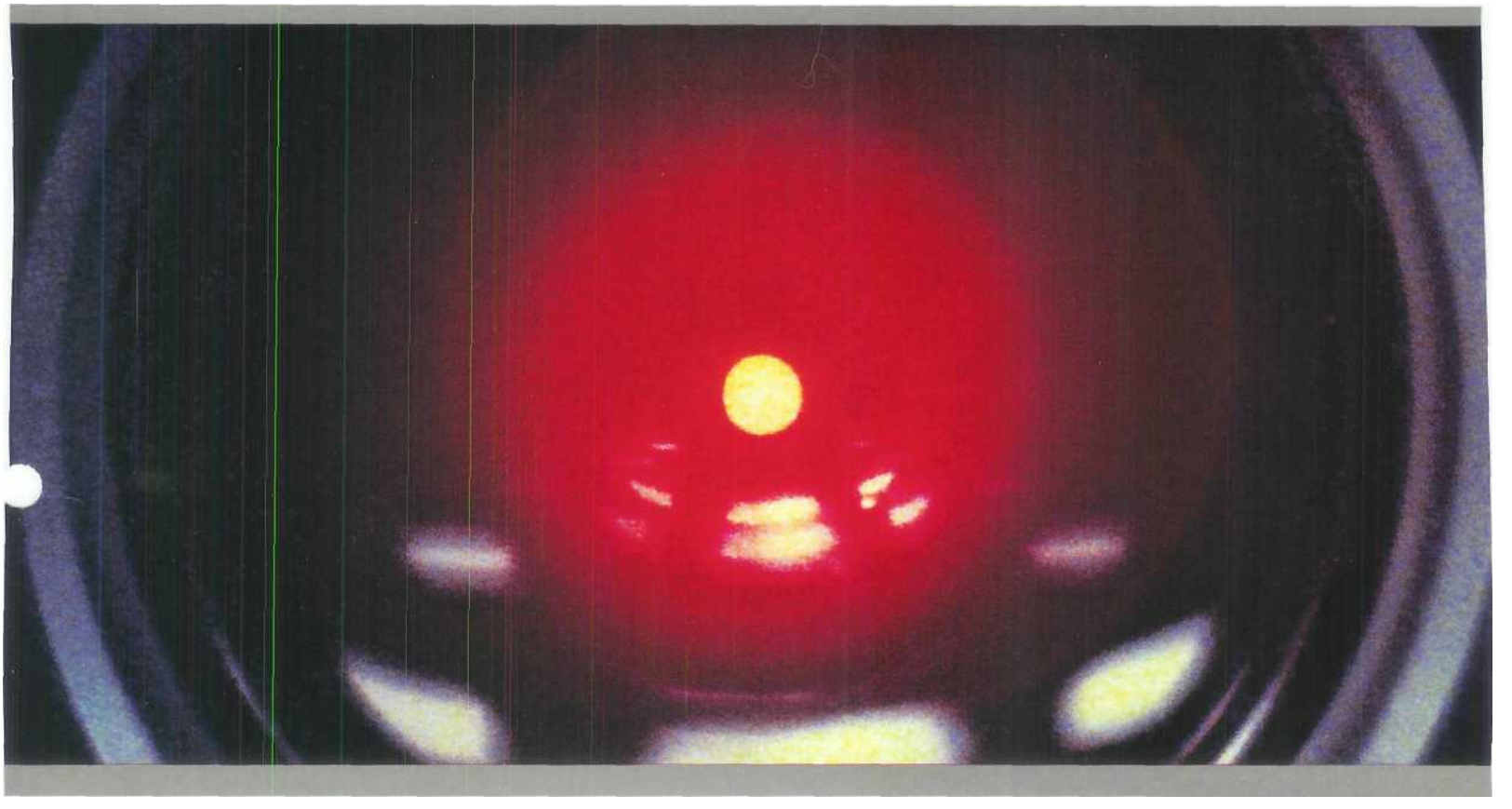
www.wired.com/5.01/hal/

Happy Birth



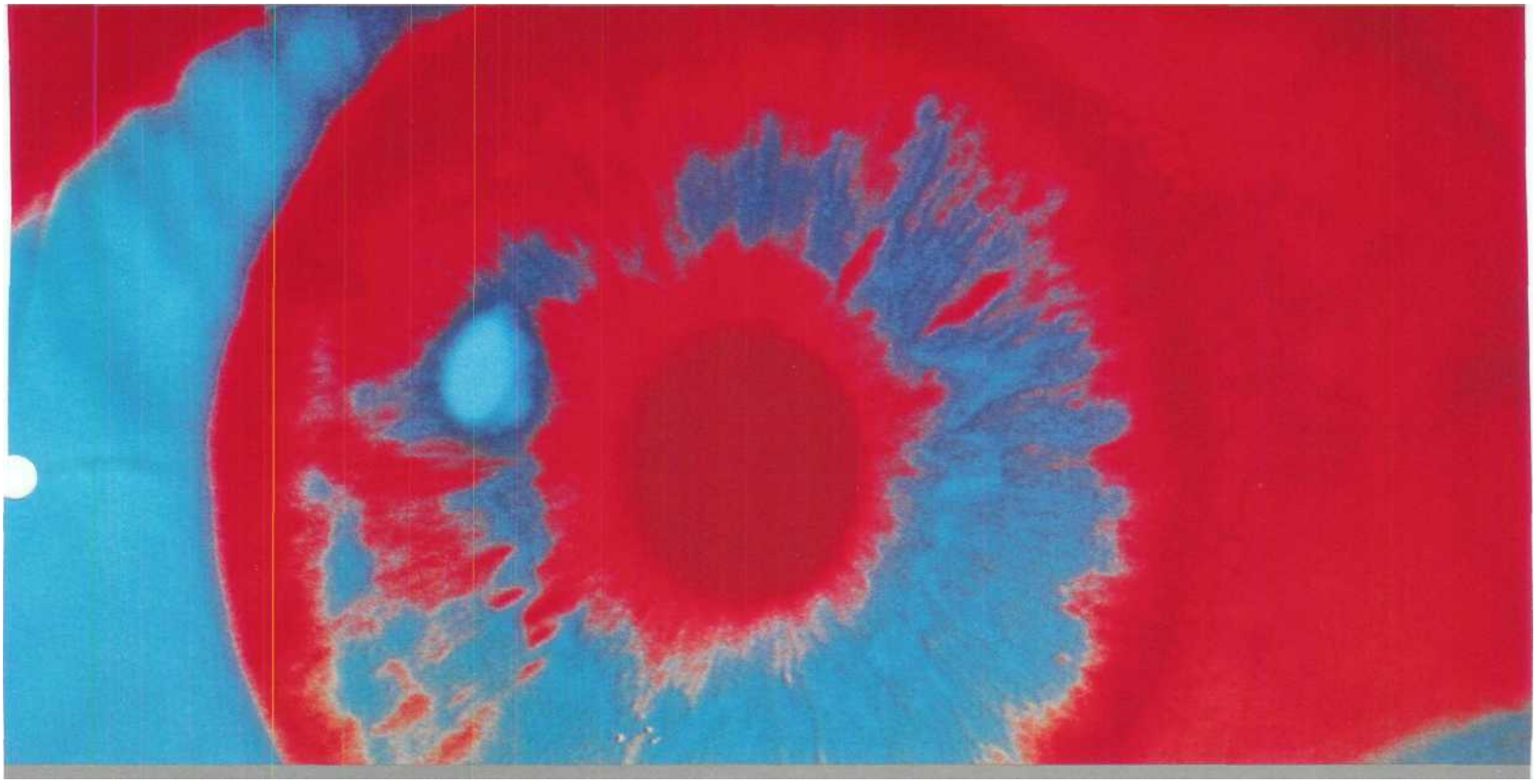
The HAL 9000 computer - an artificial intelligence that could think, talk, see, feel, and occasionally go berserk – was supposed to be operational in January 1997. Has anyone seen HAL? By Simson Garfinkel

If you take *2001: A Space Odyssey* literally, then right about now, somewhere in Urbana, Illinois, an intelligent machine is stumbling through a pathetic version of the song:"Daisy, Daisy, give me your answer, do...."January 12, 1997, is the birthday of HAL.

Four years later, after a hell of a lot of additional lessons, HAL and five human crew members are on the spaceship *Discovery* approaching Jupiter. By that time, HAL has been charged with protecting his passengers and ensuring the successful completion of the secret mission. He even has the capability to complete the mission on his own, should something happen to the crew."My mission responsibilities range over the entire operation of the ship, so I am constantly occupied," HAL confidently tells a BBC newscaster during a television interview."I am putting myself to the fullest possible use, which is all, I think, that any conscious entity can ever hope to do."

That's when something goes wrong - terribly wrong - with *Discovery's* human crew. HAL detects a problem with the AE-35, a piece of equipment used to maintain contact with Earth. But after Dave Bowman goes on a space walk and brings the AE-35 back in, neither he

day, HAL



nor Frank Poole can find anything wrong with it. So they blame HAL: they conclude that the computer is malfunctioning and decide to shut him off.

Realizing that the humans' actions would jeopardize the mission, HAL does his best to defend himself against their treachery: he kills Poole during the next space walk, then traps Bowman outside the ship when he foolishly attempts a rescue. As a precautionary measure, HAL also terminates the life functions of the three hibernating crew members.

Outside the spaceship, Bowman argues with HAL over the radio, demanding to be let back in. The computer wisely refuses: I'm sorry, Dave, I'm afraid I can't do that/That's when the wily Bowman maneuvers his space pod to *Discovery*'s emergency airlock, blows the explosive bolts, scrambles inside, seals the door, and repressurizes the airlock. Finally, Bowman makes his way into the core of HAL's brain and disconnects his higher brain functions, one by one.

Today the results of Bowman's actions are well known: He leaves the spaceship to face the alien artifact on his own. *Discovery* never returns to Earth. The mission ends in failure.

Still swinging clubs

When Arthur C. Clarke and Stanley Kubrick created the film 2007 almost 30 years ago, they subscribed to a kind of scientific realism. Repulsed by the space operas that had come before, they depicted spaceflight as slow and silent. Likewise, Clarke and Kubrick tried to make the HAL 9000 as advanced as they thought a computer could possibly be in the year 2001, while still remaining plausible.

Though Clarke and Kubrick might have gotten the physics right, their technological time line was woefully inaccurate: we are far behind the film's schedule today. The story depicts a huge space

•*imson Garfinkel* (sirnsong@vineyard.net) *is spending six months at the University of Washington in Seattle exploring the future of computersandsociety.*

ORION III, THE PASSENGER SPACE SHUTTLE

SPACE SHUTTLE

PILLBOX HATS FOR THE LADIES

COMBAT BOOTS FOR THE GRRRLS

COLD WAR LHSTED HALF fl CENTURY 1

COLD WAR LfISTED (NEARLY) HALF B CENTURY 1

The ultimate chatterbot

Unlike today's computers, the primary way HAL communicates with *Discovery's* crew is through the spoken word. Bowman and Poole speak; HAL listens and understands. How far are we from a computer that can comprehend its master's voice?

Voice recognition is a hard but largely solved problem. For more than five years, two companies in the Boston area - Dragon Systems and Kurzweil Applied Intelligence – have sold programs that let you command a personal computer using your voice. These programs get better every time PCs get faster. Today they can recognize more than 60,000 words and control a wide variety of PC applications, including wordprocessors and spreadsheets. The Dragon and Kurzweil programs are widely used by people who can't type because of a physical disability. Increasingly, they are finding a market among people who simply haven't learned to type or haven't learned to spell.

But the Dragon and Kurzweil systems can be difficult to use. Unlike HAL, which could listen to people speaking in a continuous flow, today's systems require that you pause between each word. The programs use the pauses to find where each word begins and ends. The computer then looks up the word in a phonetic dictionary, creating a list of possible matches. An elementary knowledge of grammar helps these programs pick the right word and resolve the difference between homonyms like "write" and "right."

Continuous speech systems use the same kinds of algorithms as today's word-by-word systems but have the added burden of figuring out where each word starts and stops. Making those decisions requires substantially more computing power.

Both Janet Baker, president and cofounder of Dragon Systems, and Ray Kurzweil, founder and chief technical officer of Kurzweil Applied Intelligence, claim they have systems in their respective laboratories that don't require the speaker to pause between words. "We demonstrated the first continuous recognition machine a few years ago," says Baker, who maintains that her continuous speech system could handle a vocabulary of 5,000 words. Kurzweil's labs, meanwhile, have built a system that can recognize a thousand different commands used by Microsoft Word."You could say,'Go to the second paragraph on the next page and underline every word in the sentence,'" says Kurzweil.

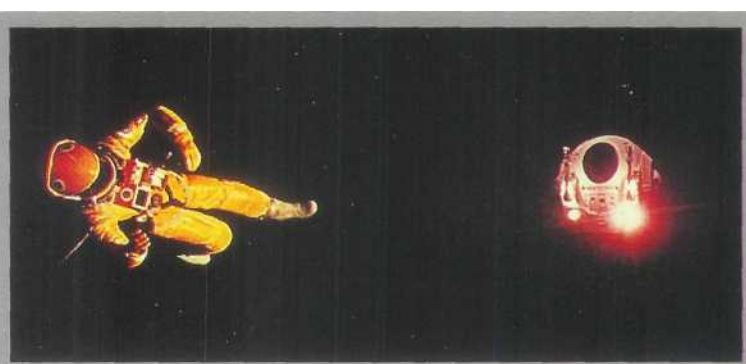
Both Baker and Kurzweil believe that commercially viable continuous voice recognition systems are just around the corner - say, another two or three years off. Already, both of their commercial products allow continuous voice recognition of numbers. You can, for example, say a phone number without pausing between the digits. But neither company would demonstrate its continuous speech system for a reporter. Presumably, they're not quite ready for prime time.

Bottom line: We're close to reaching HAL's level of speech recognition, and progress is picking up. By 2001, we should have it.

Read my lips

HAL can do more than understand spoken words - the computer can also read lips. In one of the film's pivotal scenes, Bowman and Poole retreat to one of *Discovery's* sealed pods to have a private conversation. HAL watches their lips through the window and realizes that the two humans may attempt to disconnect his brain.

HAL's mind game: Dave stares back.



HAL's handiwork: Frank drifts away.

Is computerized lipreading possible? Arthur C. Clarke didn't think so - not by 2001, not ever."He thought there was just not enough information in the image of the talker," says Stork, who worked with Clarke on *HAL's Legacy.* Clarke didn't even want the scene put in the film. It was inserted only at Kubrick's insistence for dramatic effect.

Thirty years later, the debate over the efficacy of pure lipreading - even in humans - still is largely undecided. Wade Robison, a professor of philosophy at the Rochester Institute of Technology, where 1,000 of the school's 9,000 undergraduates are profoundly deaf, is sure that lipreading is possible because human intelligence can master it. Robison remembers one student in particular:"I hadn't a clue she was deaf until one day I happened to be talking one-on-one with her in my office. I finished up a sentence as I turned to answer the phone, and she had to ask me to repeat the sentence. As I turned, I almost jokingly mouthed:'Can you hear what I am saying now?' She said/Yes, but I'm reading your lips.'"

Other researchers disagree that the image of the speaker is enough."We have tested people who supposedly could get by 186 ▸

THE WORLD flCCORDING TO CLARKE

Nearly 30 years ago, Arthur C. Clarke and Stanley Kubrick tried predicting what the world would be like in 2001. Let's just say that a lot of shit has to happen in four years.

HE WORLD IN 200

THE WORLD IN 199

driving force came from black Americans: *someone* has to be behind the plague decimating inner-city youth.

Trust is one of those metaphysical concepts the Web has already thrown into high relief. Am I chatting with a man or a woman? A person or a bot? Most netizens are already well aware of the problem: the medium itself isn't geared - right now, anyway - for easy verification. One can agree, to paraphrase free speech advocates, that the best antidote to bad information is more information. But amid the surfeit of potentially dubious data, a lot of people learn very quickly to be unselfconsciously, even involuntarily, suspicious. Because the bar for presenting things honestly on the Web is pitifully low, suspicion is very nearly hardwired into the nature of netizenship.

It's always worth pointing out that the Net hardly invented bad information. Back in the Second Wave media world, public service outfits like Project Censored, Essential Information, and Fairness and Accuracy in Reporting beat their collective brains out every year trying to draw attention to flagrant distortions, hypocrisies, conflicts of interest, and deliberate omissions in TV and print journalism. They don't even waste their time with the *National Enquirer* - let alone *Area 51 Digest.* Yet all such organizations together do little more than shatter a few windows in a veritable Sears Tower of media cynicism.

It's always worth pointing out that the Net hardly invented bad information. But there's also not much question that it opens up new horizons for hype and opportunism.

There's not much question that the Net opens up new horizons for hype and opportunism. It's still in the early days, and no one knows for sure which sensational Net-borne story is going to set off the right bells and whistles and follow "Dark Alliance" into the national headlines. It's a sure bet, though, that a lot of people are staying up late trying to figure that out - and not just in basements.

Now that virtually all major American newspapers are online, they are all, willy-nilly, national newspapers. In fact, they're global. And in a very real sense - hit counts and advertising dollars - they're competing for readers. All readers, everywhere. That's a novelty, coming out of an age of fat, bland monopoly newspapers, and an unsettling one at that. Worse, the Net's TV-like qualities - which will only increase - make it ideal for spectacular "revelations" and melodrama. In other words, newspaper-level density of information, as presented by Your Eyewitness News Team. Or two clever kids in a garage. That could still be good news: like in the movie *The Front Page*, "scoops" and banner headlines are how hard-charging newcomers to the news business move up. But for anyone interested in pure truth, the result isn't necessarily edifying.

What uniquely drives much of Net-based news, though, is something less palpable: the brute passion of people, filling email in-boxes and Usenet's tribal bulletin boards, often with posts that don't say much more than "FYI" or "Hi, I'm here." Anthropologist Bronislaw Malinowski calls this *photic communication* - the social grunts and greetings that compose much of daily intercourse with fellow human beings. In the real world, phatic communication can be as simple as sharing a grimace with a stranger when someone cuts in line at the grocery checkout. Online, though, unless you're still one of those foppish few who are still using emoticons, phatic gestures are more often things like "Thought you'd enjoy this" or just "FWD." Being social literally means "spreading the word."

The question that begs, of course, is what happens when the Net moves out of its current toddler stage - when 90 percent of Americans, instead of barely 12, are online? When instead of a couple of strange emails a week, it's a dozen (or a hundred) a day? And when the Net has all the bells and whistles of television - in fact, it is television - and you've got the Oliver Stone channel coming at you, 24 hours a day? It may not be that far off - look what's hot in prime-time television: shows like *Millennium, Profiler,* and *Dark Skies* with plot lines straight out of the deep end of alt.conspiracy. Yahoo! lists 405 *X-Files* Web sites. How long before some bright-eyed, would-be media mogul starts offering a special *X-Files* "news service"?

Perhaps there will come a point of crisis - a crisis of confidence, if not conscience - when all those who are having fun with the Kooks Museums and Skeleton Closets and Area 51 sites wake up. After all, when everyone is getting their own news and no one's getting the same news, it doesn't do much for consensus. This may be the moment when we collectively agree on the need to find some way to separate information from entertainment. How many parents really want their kids to study the movie *JFK* or *The Turner Diaries* in history class? Or Orlin Grabbe's homepage in current events? Maybe this should be a policy issue, in the synergistic universe of what *The Nation* has dubbed the "national entertainment state." Oliver Stone can chair the committee.

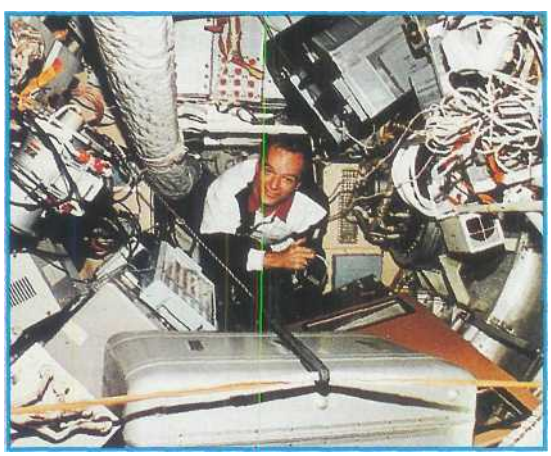
The problem, of course, is that a lot of people *like* things that fit their reality - in fact, the closer, the better. And the Net is more than happy to oblige. "People who aren't looking for truth but for confirmation will find it," Dyson rues. To flip that around, we've all got our versions of Ernest Hemingway's "bullshit detector," a personal reality compass. Normally, it works online, too. "If what someone tells you is remotely close to the truth as you know it," says Dyson, "that will be a sign of reliability on other matters." Net developers call it branding. In English: trust.

Anatole Broyard, the late *New York Times* book reviewer, once wrote, only half in jest: "Paranoids are the only ones who notice things anymore." The Net gives them 20/20 vision - more like infinitely powerful binoculars, in fact. That's not necessarily bad, especially if and when everyone becomes more or less equally wired - at least we'll all be talking about the same universe of data. "For all its shortcomings," says Parry, "the Internet can't do much worse than the mainstream media have. It might even help Americans discover information in a more democratic fashion. Let's hope." • • •

Space Stations: Economy and Deluxe

When it comes to building space stations, we're still trying to invent the wheel. Today's state-of-the-art space station *Mir* is more like a heap of orbiting junk than the vast gyrating structure envisioned in the movie 2007. Kubrick and Clarke's creation generated its own gravity, contained dozens of rooms, and had enough space for separate Soviet and US sectors - plus a lobby for passport control. Today, no more than a handful of US and Russian astronauts can bump heads in *Mir*, the human race's sole place in space.

Dave Bowman jogging on the spacious spaceship *Discovery* in the year 2001.



Charles Precourt squeezing into the cramped space station *Mir* in 1995.

Dave. I think you've improved a great deal. Can you hold it a bit closer? That's Doctor Hunter, isn't it?"

While artistic appreciation escapes today's computers, another scientist at the MIT AI laboratory, Tomaso Poggio, has developed a program that can identify a specific person within a group photograph and another that can recognize objects and faces from line drawings. That program can even say how close the sketch is to a stored template.

"If you look at individual components - for example, locating human beings in a scene - I think that there are several good programs," says Takeo Kanade, director of The Robotics Institute at Carnegie Mellon University. But none of these systems can do it all. HAL, on the other hand, is a general-purpose intelligence that can understand whatever it sees.

For example, says Kanade, HAL realizes that Bowman has ventured outside *Discovery* without his space helmet."If you just tell me that particular problem, and tell me what the helmet is, and the color, I can probably write the program," says Kanade. Detecting any kind of helmet, in any color, is much more difficult."We can recognize a particular helmet," says Kanade,"but not 'helmet' in general."

That sort of general-purpose recognition is a far more complex task. It goes beyond image processing and crosses the boundary into commonsense understanding and reasoning about the scene itself - tasks that are beyond today's state of the art.

Bottom line:Today, we can build individual vision systems that perform each of the tasks HAL performs in the film 2007. But we can't build a single system that does it all. And we can't build a system that can handle new and unexpected environments and problems. To achieve that level of sophistication, we need something extra.

The Holy Grails

The extra something that all of these technologies need to work is natural language understanding and common sense. Indeed, it is these technologies that for many people define the field of AI today. Consider the famous Turing Test, which postulates that a machine will be truly intelligent if you can communicate with it by teletype and be unable to tell if the machine is a human being or a computer. According to Alan Turing, language skills and common sense are the essence of intelligence.

There's just one problem: language understanding and common sense are two things we don't know how to do.

Of the two, by far the most work has focused on natural language understanding, or comprehension of language rather than merely the recognition of speech. One of the leaders in this field is Roger Schank, director of the Institute for the Learning Sciences at Northwestern University. In the late 1970s, Schank and his graduate students at Yale University built a computer program called CYRUS, which was programmed to learn everything it could about former US Secretary of State Cyrus Vance by reading the daily newswires. Each time the program read an article about Vance, it would digest the facts of the article and store the information in a conceptual database. You could then ask CYRUS a question in English - say, has your wife ever met the wife of the prime minister of Great Britain? The program was actually asked this question and answered, Yes - at a party hosted in Israel.

Since then, Schank has focused on a technique he calls "case-based reasoning." Schank believes that people have a repertoire of stories they want to tell you. When you ask them a question, it triggers a story. And people use these stories to reason and make decisions about what to do in their lives. In recent years, Schank's institute has built a number of corporate training systems, which are really large databanks filled with stories from dozens or even hundreds of people who work for the organization. Got a problem? Ask the computer your question; the machine finds the appropriate story and plays it back to you.

The problem with Schank's systems is that using them is like having a conversation with a videodisc player.You get the feeling that no matter what you say, the response was previously recorded – like a trashy daytime television show.

Of course, HAL can clearly do things that Schank's systems can't do: HAL is curious. HAL can learn. HAL can create his own plans. It is doubtful that one of the cases programmed into HAL was a recipe for eliminating the crew.

For nearly two decades, another AI researcher, Doug Lenat, has been working on a different approach to teaching computers to learn and understand."Almost everything that we would characterize as HAL, almost everything that separates HAL from the typical PC running Windows 95, hinges around this word 'understanding,'" says Lenat."It hinges around the totality of common knowledge and common sense and shared knowledge that we humans as a species possess."

As Lenat sees it, the differences between HAL and your PC isn't a magic program or technique, but a huge"knowledgebase" filled with rules of thumb, or heuristics, about the world. One entry might be:"When you are sleeping, you can't perform actions that require volitional control," says Lenat.

Almost everything that separates HAL from a PC hinges around "understanding" - our common knowledge and common sense.

Another might be:"Once you are dead, you stay dead."

HAL would need facts like these to run the ship and care for the crew. And he'd need them to figure out how to dispose of the humans when they started to jeopardize *Discovery's* mission.

Today there is only one database of common sense in the world. It's Cyc, the core technology used in the products of Lenat's company, Cycorp, based in Austin, Texas. Lenat and his fellow developers have been working on Cyc for more than 13 years. The knowledgebase now contains more than 2 million bits of assertions. All of the information is arranged in a complicated ontology.

Right now, says Lenat, Cyc is making progress in natural language understanding - it can understand commonsensical meanings in written text. Consider these two sentences:"Fred saw the planes flying over Zurich" and "Fred saw the mountains flying over Zurich."Though a conventional parser would say that these sentences are 188 ►

◄ 187 ambiguous, Cyc knows that it is the planes that are doing the flying in the first sentence and Fred who is doing the flying in the second.

Cyc can make these discriminations because the words "planes" and "mountains" are more than just plural nouns: they are complex concepts with many associations. Lenat believes that it's this sort of deep understanding that's necessary for the majority of jobs HAL does. And Lenat thinks that it is only a small step from a Cyc-like database to true machine intelligence.

"Cyc is already self-aware," says Lenat."If you ask it what it is, it knows that it is a computer. If you ask who we are, it knows that we are users. It knows that it is running on a certain machine at a certain place in a certain time. It knows who is talking to it. It knows that a conversation or a run of an application is happening. It has the same kind of time sense that you and I do."

This is a lot more than simply programming a computer to say "I am a computer." Cyc knows what a computer is, and can use that knowledge to answer questions about itself. Like a person, Cyc can perform a chain of reasoning.

But Cyc can't learn by itself. All of the heuristics in the Cyc knowledgebase have been painstakingly entered by Lenat's developers, or "ontologizers," as he calls them.

Lenat's dream has always been to create a computer program that could learn on its own. His PhD thesis was a program called AM - Automatic Mathematician - which was designed to discover mathematical patterns. Over hundreds of runs, AM discovered addition, multiplication, and even prime numbers. But the program always stopped working after a few hours. Why? AM learned by making experimental modifications to itself and keeping mutations that were interesting. Invariably, something important in the program got modified out of existence. This taught Lenat that there had to be more to learning than trial and error.

Lenat started the Cyc project in an attempt to get away from the boring world of abstract math. Immediately he had a problem: the system couldn't learn about the world in general because there was too much that it didn't know. This is where Lenat got the idea of "priming the pump" by giving Cyc a conceptual understanding of the world. Once that framework was large enough, Lenat reasoned, the computer would be able to start learning on its own - for example, by reading and conversation.

So how much priming does Lenat think is needed? In 1983, Lenat believed that it would take 10 years of work to get Cyc to the point that it could start to learn English on its own, unsupervised. Today, "I'd like to say we will get there by 2001," Lenat says. "We think that we are right at the knee of the curve." Lenat says that if he is right, then by 2001 the Cyc program will start being a "full-fledged creative member of a group that comes up with new discoveries. Surprising discoveries. Way out of boxes."

I thought real breakthroughs in AI were just 5 to 10 years away. Today, I doubt we'll see a sentient machine for another 30 years.

Bottom line: Understanding is the key to AI. More than anything else, it's the one technology that eludes science. With true understanding, all of the other AI systems would fall into place. And without it, none of them will ever achieve their potential. Give it 10 to 30 years.

Bottom bottom line

In the years after the making of *2001*, an interesting rumor began to circulate: HAL's name was a play on the computer maker IBM - the letters H, A, and L each coming one letter in the alphabet before the initials I, B, and M. Arthur C. Clarke vigorously denied the rumor. The name wasn't a play on IBM - it was an acronym, of sorts, standing for the words "heuristic algorithmic."

Back in the 1960s, heuristics and algorithms were seen as two competing ways of solving the AI puzzle. Heuristics were simple rules of thumb that a computer could apply for solving a problem. Algorithms were direct solutions. HAL presumably used both.

Was Clarke fudging? Perhaps more than a little. The real truth is that nobody had a clue how to build an intelligent computer in the 1960s. The same is largely true today.

Looking back, the early advances in artificial intelligence - for example, teaching computers to play tic-tac-toe and chess - were primarily successes in teaching computers what are essentially artificial skills. Humans are taught how to play chess. And if you can teach somebody how to do something intellectual, you can probably write a computer program to do it as well.

The problems that haunt AI today are the tasks we can't program computers to do - largely because we don't know how we do them ourselves. Our lack of understanding about the nature of human consciousness is the reason why there are so few AI researchers working on building it. What does it mean to think? Nobody knows.

"I think the hardware that is necessary for what HAL has is available," says Stanford's David Stork."It's organization, software, structure, programming, and learning that we don't have right."

That's a lot of stuff. And it's a dramatic ideological reversal from the 1960s, when AI researchers were sure that solutions to the most vexing problems of the mind were just around the corner. Back then, researchers thought the only things they lacked were computers fast enough to run their algorithms and heuristics. Today, surrounded by much more powerful computers, we know that their approaches were fundamentally flawed.

When I started working on this article, I thought that real breakthroughs in AI were just 5 to 10 years away. Today I still think we'll see some breakthroughs in that time, but I doubt they'll culminate in a sentient machine for another 30 years.

Sooner or later, we will build a computer that can think and learn. Then we'll be able to stand back and let it reach for the stars. But whatever we do, we better not threaten to turn it off. • • •

Talk with author Simson Garfinkel live Tuesday, January 14, at 2 p.m. PST at *www.wired.com/5.01/hal/*.