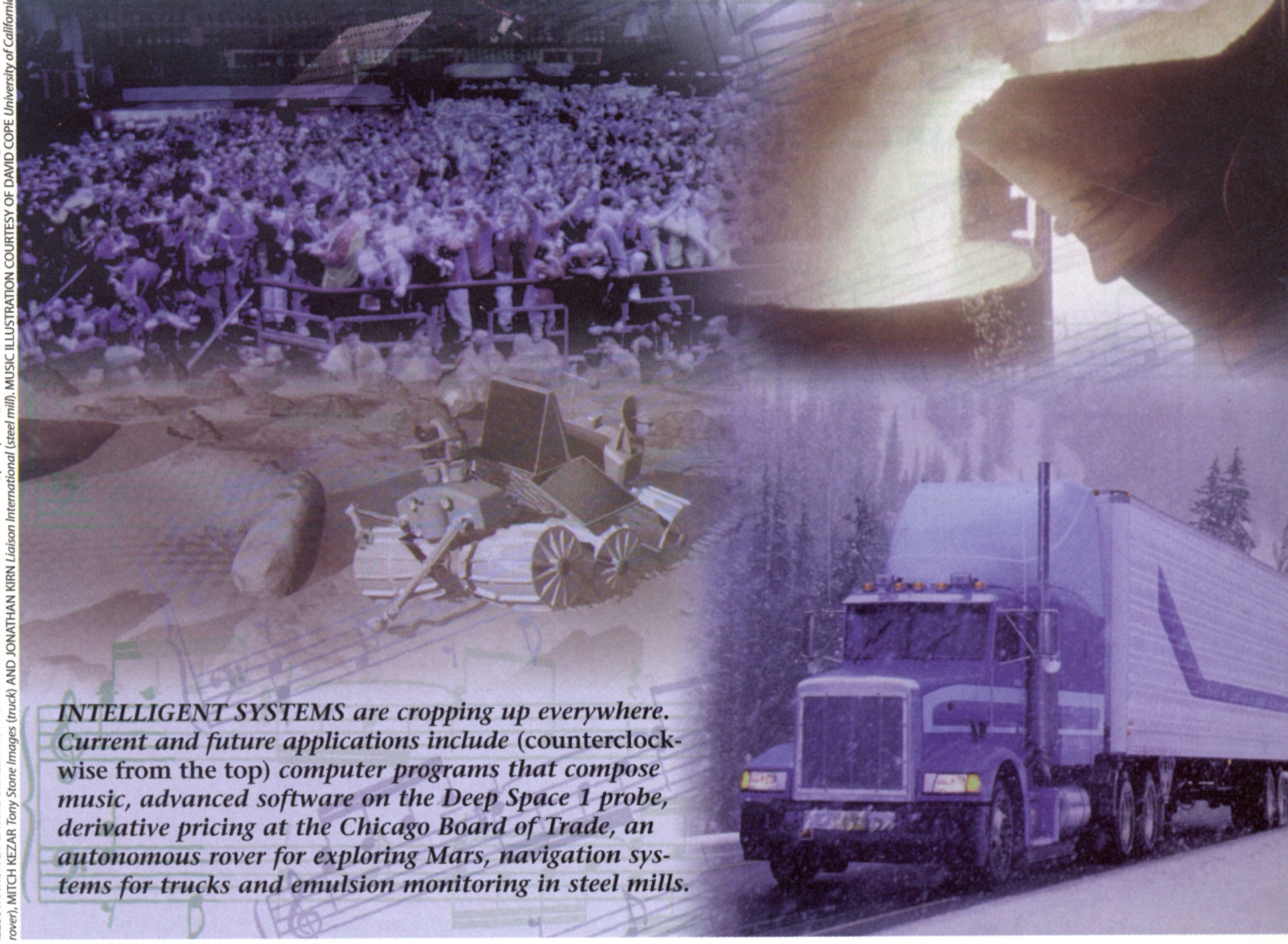
On Computational Wit
Rethinking the Goals of Artificial Intelligence

The greatest value of artificial intelligence may lie not in imitating human thinking but in extending it into new realms

by Kenneth M. Ford and Patrick J. Hayes



INTELLIGENT SYSTEMS are cropping up everywhere. Current and future applications include (counterclockwise from the top) *computer programs that compose music, advanced software on the Deep Space 1 probe, derivative pricing at the Chicago Board of Trade, an autonomous rover for exploring Mars, navigation systems for trucks and emulsion monitoring in steel mills.*

ILLUSTRATION BY LAURIE GRACE. PHOTOGRAPHS COURTESY OF SPECTRUM ASTRO (*space probe*), JEFFERY TITCOMB *Liaison International* (*trading floor*), NASA AMES RESEARCH CENTER (*Mars rover*), MITCH KEZAR *Tony Stone Images* (*truck*) AND JONATHAN KIRN *Liaison International* (*steel mill*). MUSIC ILLUSTRATION COURTESY OF DAVID COPE *University of California at Santa Cruz.*

Wings:

Many philosophers and humanist thinkers are convinced that the quest for artificial intelligence (AI) has turned out to be a failure. Eminent critics have argued that a truly intelligent machine cannot be constructed and have even offered mathematical proofs of its impossibility. And yet the field of artificial intelligence is flourishing. "Smart" machinery is part of the information-processing fabric of society, and thinking of the brain as a "biological computer" has become the standard view in much of psychology and neuroscience.

While contemplating this mismatch between the critical opinions of some observers and the significant accomplishments in the field, we have noticed a parallel with an earlier endeavor that also sought an ambitious goal and for centuries was attacked as a symbol of humankind's excessive hubris: artificial flight. The analogy between artificial intelligence and artificial flight is illuminating. For one thing, it suggests that the traditional view of the goal of AI—to create a machine that can successfully imitate human behavior—is wrong.

For millennia, flying was one of humanity's fondest dreams. The prehistory of aeronautics, both popular and scholarly, dwelled on the idea of imitating bird flight, usually by somehow attaching flapping wings to a human body or to a framework worn by a single person. It was frustratingly clear that birds found flying easy, so it must have seemed natural to try to capture their secret. Some observers suggested that bird feathers simply possessed an inherent "lightness." Advocates of the possibility of flight argued that humans and birds were fundamentally similar, whereas opponents argued that such comparisons were demeaning, immoral or wrongheaded. But both groups generally assumed that flying meant imitating a bird. Even relatively sophisticated designs for flying machines often included some birdlike features, such as the beak on English artist Thomas Walker's 1810 design for a wooden glider.

This view of flying as bird imitation was persistent. An article in *English Mechanic* in 1900 insisted that "the true flying machine will be to all intents and purposes an artificial bird." A patent application for a "flying suit" covered with feathers was made late in the 19th century, and wing-flapping methods were discussed in technical surveys of aviation published early in this century.

The Turing Test

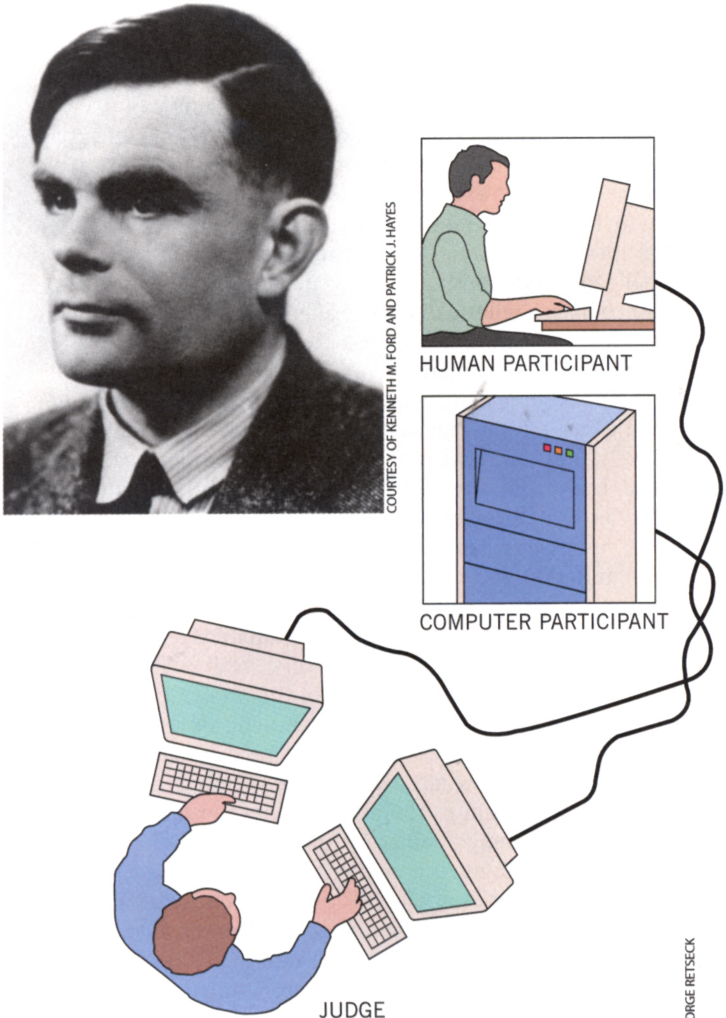
Intelligence is more abstract than flight, but the long-term ambition of AI has also traditionally been characterized as the imitation of a biological exemplar. When British mathematician Alan M. Turing first wrote of the possibility of artificial intelligence in 1950, he suggested that AI research might focus on what was probably the best test for human intelligence available at the time: a competitive interview. Turing suggested that a suitable test for success in AI would be an "imitation game" in which a human judge would hold a three-way conversation with a computer and another human and try to tell them apart. The judge would be free to turn the conversation to any topic, and the successful machine would be able to chat about it as convincingly as the human. This would require the machine participant in the game to understand language and conversational conventions and to have a general ability to reason. If the judge could not tell the difference after some reasonable amount of time, the machine would pass the test: it would be able to seem human to a human.

There is some debate about the exact rules of Turing's imitation game, and he may not have intended it to be taken so seriously. But some kind of "Turing test" has become widely perceived, both inside and outside the field, as the ultimate goal of artificial intelligence, and the test is still cited in most textbooks. Just as with early thinking about flight, success is defined as the imitation of a natural model: for flight, a bird; for intelligence, a human.

The Turing test has received much analysis and criticism, but we believe that it is worse than often realized. The test has led to a widespread misimpression of the proper ambitions of our field. It is a poorly designed experiment (depending too much on the subjectivity of the judge), has a questionable technological objective (we already have lots of human intelligence) and is hopelessly culture-bound (a conversation that is passable to a British judge might fail according to a Japanese or Mexican judge). As Turing himself noted, one could fail the test by being too intelligent—for example, by doing mental arithmetic extremely fast. According to media reports, some judges at the first Loebner competition in 1991—a kind of Turing test contest held at the Computer Museum in Boston—rated a human as a machine on the grounds that she produced extended, well-written paragraphs of informative text. (Apparently, this is now considered an inhuman ability in parts of our culture.) With the benefit of hindsight, it is now evident that the central defect of the test is its species-centeredness: it assumes that human thought is the final, highest pinnacle of thinking against which all others must be judged. The Turing test does not admit of weaker, different or even stronger forms of intelligence than those deemed human.

Most contemporary AI researchers explicitly reject the goal of the Turing test. Instead they are concerned with exploring the computational machinery of intelligence itself, whether in humans, dogs, computers or aliens. The scientific aim of AI research is to understand intelligence as computation, and its engineering aim is to build machines that surpass or extend human mental abilities in some useful way. Trying to imitate a human conversation (however "intellectual" it may be) contributes little to either ambition.

In fact, hardly any AI research is devoted to trying to pass the Turing test. It is more concerned with issues such as how machine learning and vision might be improved or how to design an autonomous spacecraft that can plan its own actions. Progress in AI is not measured by checking fidelity to a human conversationalist. And yet many critics complain of a lack of progress toward this old ambition. We think the Turing test should be relegated to the history of science, in the same way that the aim of imitating a bird was eventually abandoned by the pioneers of flight. Beginning a textbook on AI with the Turing test (as many still do) seems akin to starting a primer



COURTESY OF KENNETH M. FORD AND PATRICK J. HAYES

GEORGE RETSECK

TURING TEST for artificial intelligence was proposed in 1950 by British mathematician Alan M. Turing (photograph). *In the test, a human judge would hold a three-way conversation with a computer and another human. If the judge could not distinguish between the responses of the human and those of the computer, the machine would pass the test.*

on aeronautical engineering with an explanation that the goal of the field is to make machines that fly so exactly like pigeons that they can even fool other pigeons.

Imitation versus Understanding

Researchers in the field of artificial intelligence may take a useful cue from the history of artificial flight. The development of aircraft succeeded only when people stopped trying to imitate birds and instead approached the problem in new ways, thinking about airflow and pressure, for example. Watching hovering gulls inspired the Wright brothers to use wing warping—turning an aircraft by twisting its wings—but they did not set out to imitate the gull's wing. Starting with a box kite, they first worked on achieving sufficient lift, then on longitudinal and lateral stability, then on steering and finally on propulsion and engine design, carefully solving each problem in turn. After that, no airplane could be confused with a bird either in its overall shape or in its flying abilities. In some ways, aircraft may never match the elegant precision of birds, but in other ways, they outperform them dramatically. Aircraft do not land in trees, scoop fish from the ocean or use the natural breeze to hover motionless above the countryside. But no bird can fly at 45,000 feet or faster than sound.

Rather than limiting the scope of AI to the study of how to mimic human behavior, we can more usefully construe it as the study of how computational systems must be organized in order to behave intelligently. AI programs are often components of larger systems that are not themselves labeled "intelligent." There are hundreds of such applications in use today, including those that make investment recommendations, perform medical diagnoses, plan troop and supply movements in warfare, schedule the refurbishment of the space shuttle and detect fraudulent use of credit cards. These systems make expert decisions, find meaningful patterns in complex data and improve their performances by learning. All these actions, if done by a human, would be taken to display sound judgment, expertise or responsibility. Many of these tasks, however, could not be done by humans, who are too slow, too easily distracted or not sufficiently reliable. Our intelligent machines already surpass us in many ways. The most useful computer applications, including AI applications, are valuable exactly by virtue of their lack of humanity. A truly humanlike program would be just as useless as a truly pigeonlike aircraft.

Waiting for the Science

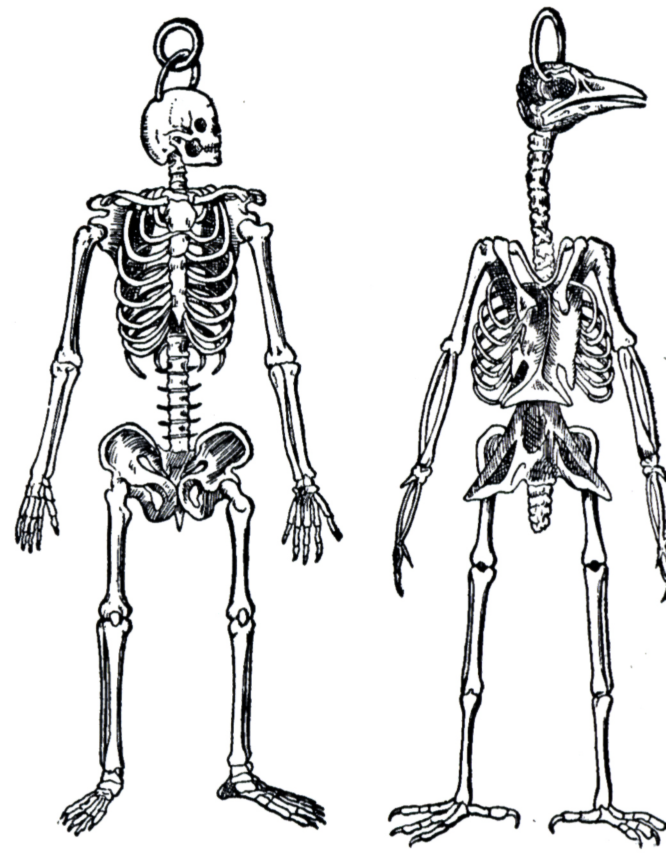
The analogy with flight provides another insight: technological advances often precede advances in scientific knowledge. The designers of early aircraft could not learn the principles of aerodynamics by studying the anatomy of birds. Evolution is a sloppy engineer, and living systems tend to be rich with ad hoc pieces of machinery with multiple uses or mechanisms jury-rigged from structures that evolved earlier for a different reason. As a result, it is often very difficult to discover basic principles by imitating natural mechanisms.

Experimental aerodynamics became possible only in the early part of this century, when artificial wings could be tested systematically in wind tunnels. It did not come from studying natural exemplars of flight. That a gull's wing is an airfoil is now strikingly obvious, yet the airfoil was not discovered by examining the anatomy of birds. Even the Wright brothers never really understood why their *Flyer* flew. The aerodynamic principles of the airfoil emerged from experiments done in 1909 by French engineer Alexandre-Gustave Eiffel, who used a wind tunnel and densely instrumented artificial wings. The first aircraft with "modern" airfoils—which were made thicker after engineers demonstrated that thicker airfoils improved lift without increasing drag—did not appear until late in World War I. As is true for many other disciplines, a firm theoretical understanding was possible only when controlled experiments could be done on isolated aspects of the system. Aerodynamics was discovered in the laboratory.

The same reasoning applies to the study of human intelligence. It may be impossible to discover the computational principles of intelligent thought by examining the intricacies of human thinking, just as it was impossible to discover the principles of aerodynamics by examining bird wings. The Wright brothers' success was largely attributed to their perception of flight in terms of lift, control and power; similarly, a science of intelligence must isolate particular aspects of thought, such as memory, search and adaptation, and allow us to experiment on these one at a time using artificial systems.

By systematically varying functional parameters of thought, we can determine the ways in which various kinds of mental processes can interact and support one another to produce intelligent behavior.

Several areas of AI research have been transformed in the past decade by an acceptance of the fact that progress must be measurable, so that different techniques can be objectively compared. For example, large-scale empirical investigations must be conducted to evaluate the efficiency of different search techniques or reasoning methods. In this kind of AI research, computers are providing the first wind tunnels for thought.



CULVER PICTURES

COMPARISON OF SKELETONS of a human and a bird—here taken from a 16th-century manuscript by French naturalist Pierre Belon—examined similarites in anatomy in an attempt to understand how birds can fly.

A Science of Intelligence

Rejecting the Turing test may seem like a retreat from the grand old ambition of creating a "humanlike" mechanical intelligence. But we believe that the proper aim of AI is much larger than simply mimicking human behavior. It is to create a computational science of intelligence itself, whether human, animal or machine. This is not a new claim; it has been made before by AI pioneers Allen Newell and Herbert A. Simon, cognitive psychologist Zenon Pylyshyn and philosopher Daniel C. Dennett, among others. But it was not until we noted the analogy with artificial flight that we appreciated the extent to which the Turing test, with its focus on imitating human performance, is so directly at odds with the proper objectives of AI. Some of our colleagues say their ultimate goal is indeed the imitation of human intelligence. Even with this limited aim, however, we believe that the perspective sketched here provides a more promising way to achieve that ambition than does the method outlined by Turing.

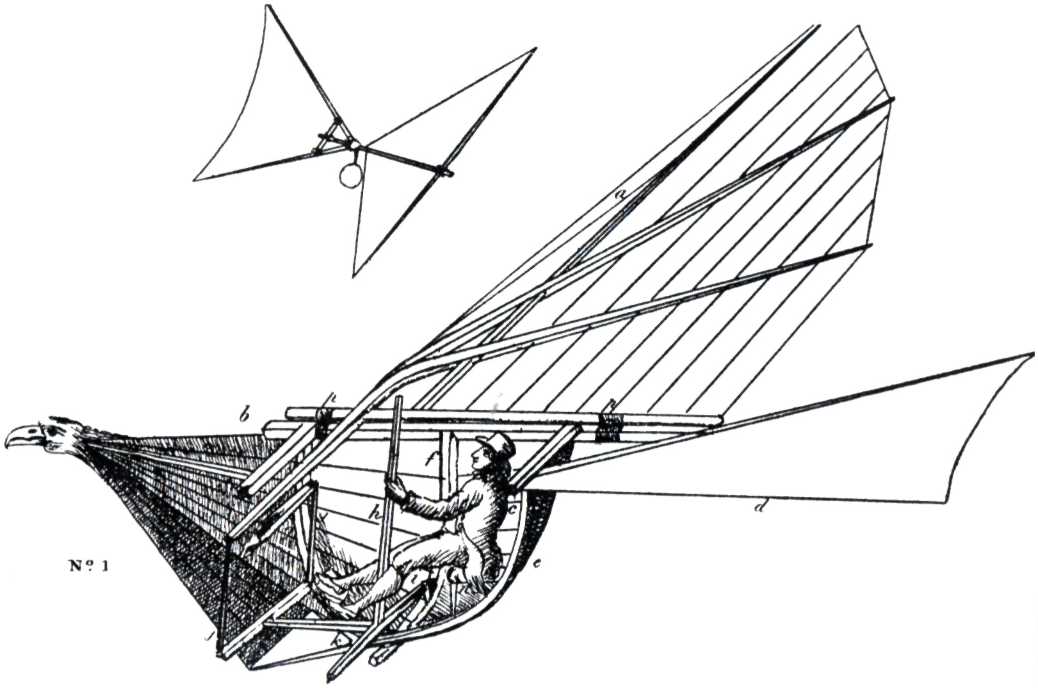
Consider again the analogy with flight. Just as the principles of aerodynamics apply equally to any wing, natural or artificial, the computational view of intelligence—or, more broadly, of mentality—applies just as well to natural thinkers as to artificial thinkers. If cognitive psychology and psycholinguistics are like the study of bird flight in all its complexity, then applied AI is like aeronautical engineering. Computer science supplies the principles that guide the engineering, and computation itself is the air that supports the wings of thought.

The study of artificial intelligence, like a large part of computer science, is essentially empirical. To run a program is often to perform an experiment on a large, complex apparatus (made partly of metal and silicon and partly of symbols) to discover the laws that relate its behavior to its structure. Like artificial wings, these AI systems can be designed and instrumented to isolate particular aspects of this relation. Unlike the research methodology of psychology, which employs careful statistical analysis to discern relevant aspects of behavior in the tangled complexity of nature, the workings of AI systems are open to direct inspection. Using computers, we can discover and experiment directly with what Newell and Simon have called the "laws of qualitative structure."

This picture of AI defines the field in a more useful and mature way than Turing could provide. In this view, AI is the engineering of cognitive artifacts based on the computational understanding that runs through and informs current cognitive science. Turing correctly insisted that his test was not meant to define intelligence. Nevertheless, in giving us this touchstone of success, he chose human intelligence—in fact, the arguing skill of an educated, English middle-class man playing a kind of party game—as our goal. But the very science that Turing directed us toward provides a perspective from which a much broader and more satisfying account of intelligence is emerging.

Scholastic Critics

Artificial intelligence and artificial flight are similar even in the criticisms they attract. The eminent American astronomer Simon Newcomb argued passionately in the early 1900s against the idea of heavier-than-air flight. Newcomb's fulminations seem amusing now, but his arguments were quite impressive and reflected the view of the informed intelligentsia of his day. Like British mathematical physicist Roger Penrose, who uses Gödel's theorem to "prove" that AI is impossible, Newcomb employed mathematical arguments. He pointed out that as birds get bigger, their wing area increases in proportion to the square of their size, but their body weight increases in proportion to the cube, so a bird the size of a man could not fly. He was still using this argument against the possibility of manned flight several years after the Wright brothers' success at Kitty Hawk, N.C., when aircraft were regularly making trips lasting several hours. It is, in fact, quite a good argument—aircraft takeoff weights are indeed roughly proportional to the cube of their wingspan—but Newcomb had no idea how sharply the lift from an airfoil increases in proportion to its airspeed. He thought of a wing as simply a flat, planar surface.



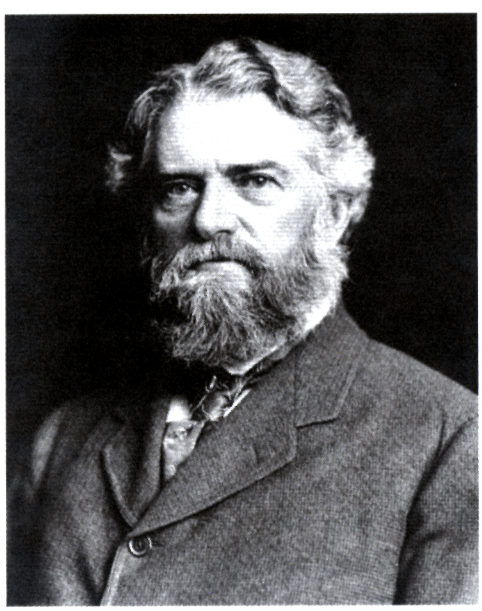
WOODEN GLIDER designed in 1810 by English artist Thomas Walker included a birdlike beak.

CULVER PICTURES

Newcomb also used a combination of thought experiment and rhetoric to make his point—the same tactic that philosopher John R. Searle has employed in his famous "Chinese Room" argument against AI [see "Is the Brain's Mind a Computer Program?" by John R. Searle; SCIENTIFIC AMERICAN, January 1990]. Newcomb stated scornfully, "Imagine the proud possessor of the aeroplane darting through the air at a speed of several hundred feet per second! It is the speed alone that sustains him. How is he ever going to stop?" Newcomb's arguments, with their wonderful combination of energy, passion, cogency and utter wrongheadedness, are so similar to contemporary arguments against artificial intelligence that for several years we have offered the annual Simon Newcomb Award for the silliest new argument attacking AI. We welcome nominations.

A common response to our analogy between artificial intelligence and artificial flight is to ask what will be the Kitty Hawk of AI and when will it happen. Our reply follows that of Herbert Simon: *it has already happened.* Computers regularly perform intelligent tasks and have done so for many years. Artificial intelligence is flying all around us, but many simply refuse to see it. Among the thousands of applications in use today, here are just a few examples: AI systems now play chess, checkers, bridge and backgammon at world-class levels, compose music, prove mathematical theorems, explore active volcanoes, synthesize stock-option and derivative prices on Wall Street, make decisions about credit applications, diagnose motor pumps, monitor emulsions in a steel mill, translate technical service manuals, and act as remedial reading tutors for elementary school children. In the near future, AI applications will guide deep-space missions, explore other planets and drive trucks along freeways.

But should all this really count as "intelligent"? The performance of AI systems, like the speed or altitude of aircraft, is not open to dispute, but whether or not one chooses to call it "intelligent" is determined more by social attitude than by anything objective. When any particular ability is mechanized, it is often no longer considered to be a hallmark of mental prowess. It is easy now to forget that when Turing was writing, a "computer" was a human being who did arithmetic for a living, and it was obvious to everyone that computing required intelligence. The meaning of the word has now changed to mean a machine, and performing fast, accurate arithmetic is no longer considered a hallmark of mental ability, just as the meaning of "flying" has changed to cover the case, once inconceivable, of dozing quietly in an airplane seat while traveling at hundreds of miles an hour far above the clouds. Newcomb—who was famous as one of the finest computers of his time—went to his deathbed refusing to concede that what early aircraft did should be called "flying."

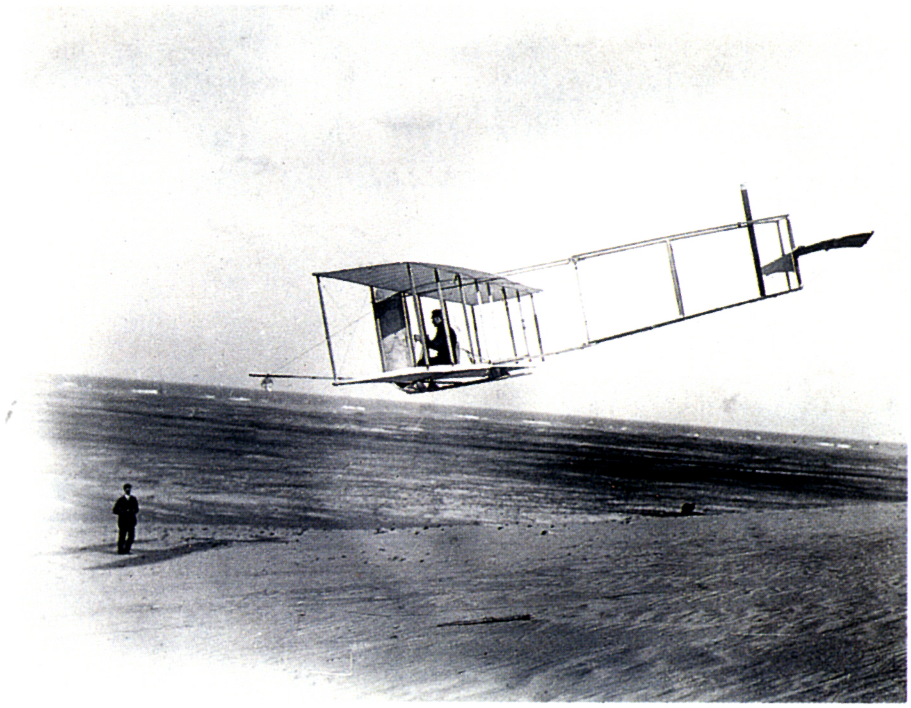


SIMON NEWCOMB, American astronomer and mathematician, argued passionately against the possibility of artificial flight—even after the Wright brothers' successful tests of their aircraft in 1903.

MARY LEA SHANE ARCHIVES, LICK OBSERVATORY

Turing suggested his test as a way to avoid useless disputes about whether a particular task counted as truly intelligent. With considerable prescience, he anticipated that many people would never accept that the action of a machine could ever be labeled as "intelligent," that most human of labels. But just as there was no doubt that the early flyers moved through the air at certain altitudes and speeds, there is no doubt that electronic computers actually get arithmetic done, make plans, produce explanations and play chess. The labels are less important than the reality.

The arbitrariness of the social labeling can be illustrated by a thought experiment in which the machine is replaced by something mysterious but natural. Whereas a dog will never pass the Turing test, no one but a philosopher would argue that a dog does not display some degree of intelligence—certainly no one who has owned a dog would make such an argument. It is often claimed that Deep Blue, the com-



THE WRIGHT FLYER, *shown at Kitty Hawk, N.C., with Orville Wright piloting, proved that aircraft need not imitate birds.*

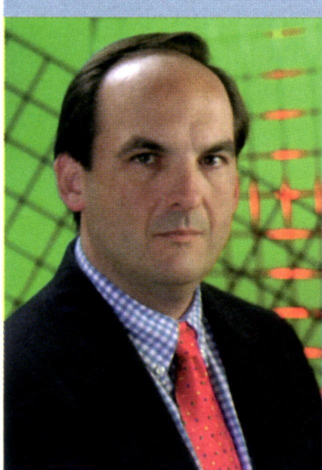
ARCHIVE PHOTOS

puter that defeated chess champion Garry Kasparov, is not *really* intelligent, but imagine a dog that played chess. A chess-playing dog that could beat Kasparov would surely be acclaimed a remarkably smart dog.

The idea that natural intelligence is a complex form of computation can only be a hypothesis at present. We see no clear reason, however, why any mental phenomenon cannot be accounted for in this way. Some have argued that the computationalist view cannot account for the phenomenology of consciousness. If one surveys the current theories of the nature of consciousness, however, it seems to us that a computationalist account offers the most promise. Alternative views consider consciousness to be some mysterious physical property, perhaps arising from quantum effects influenced by the brain's gravity or even something so enigmatic as to be forever beyond the reach of science. None of these views seems likely to explain how a physical entity, such as a brain in a body, can come to be aware of the world and itself. But the AI view of mental life as the product of computation provides a detailed account of how internal symbols can have meaning for the machine and how this meaning can influence and be influenced by the causal relations between the machine and its surroundings.

The scientific goal of AI is to provide a computational account of intelligence or, more broadly, of mental ability itself—not merely an explanation of human mentality. This very understanding, if successful, must deny the uniqueness of human thought and thereby enable us to extend and amplify it. Turing's ultimate aim, which we can happily share, was not to describe the difference between thinking people and unthinking machines but to remove it. This is not to disparage or reduce humanity and still less to threaten it. If anything, understanding the intricacies of airflow increases our respect for how extraordinarily well birds fly. Perhaps it seems less magical, but its complexity and subtlety are awesome. We suspect that the same will be true for human intelligence. If our brains are indeed biological computers, what remarkable computers they are. SA

About the Authors



INSTITUTE FOR HUMAN AND MACHINE COGNITION

KENNETH M. FORD (*left*) is associate director at the National Aeronautics and Space Administration Ames Research Center and director of the NASA Center of Excellence for Information Technology. He is on leave from the Institute for Human and Machine Cognition at the University of West Florida, where he is the director. Ford entered computer science and artificial intelligence through the back door of philosophy. After studying epistemology as an undergraduate, he joined the navy and wound up fixing computers. "I had no interest in computers," he says. "I didn't even know how to program them." But Ford learned quickly—he received his Ph.D. in computer science from Tulane University in 1987. He still finds a compelling connection between computers and philosophy. "I'm convinced," he says, "that were they reborn into a modern university, Plato, Aristotle and Leibniz would most suitably take up appointments in the computer science department." Over the years he has developed a love of stalking wary redfish and speckled trout on the shallow flats of the Gulf of Mexico—quiet hours that are often the source of fresh ideas.

PATRICK J. HAYES is John C. Pace Eminent Scholar at the Institute for Human and Machine Cognition at the University of West Florida and is a past president of the American Association for Artificial Intelligence. His path toward AI started at age 12, when he constructed a robot that could wander around a tabletop without falling off. He went on to study mathematics at the University of Cambridge and machine intelligence at the University of Edinburgh. Visiting Stanford University to work with AI pioneer John McCarthy sparked his interest in the use of logic in artificial intelligence. Later, at the University of Rochester, Hayes headed one of the first multidisciplinary programs in cognitive science. His current research focuses on the representation of knowledge, the underpinnings for an artificial consciousness and the philosophical questions raised by AI. Hayes also has a more temporal interest: his wife, Jacqueline, and he collect and restore antique mechanical clocks.