The Man Who Tried to Redeem the World with Logic - Issue 21: Information - Nautilus

By Michael Segal

28-35 minutes

Walter Pitts was used to being bullied. He'd been born into a tough family in Prohibition-era Detroit, where his father, a boiler-maker, had no trouble raising his fists to get his way. The neighborhood boys weren't much better. One afternoon in 1935, they chased him through the streets until he ducked into the local library to hide. The library was familiar ground, where he had taught himself Greek, Latin, logic, and mathematics—better than home, where his father insisted he drop out of school and go to work. Outside, the world was messy. Inside, it all made sense.

Not wanting to risk another run-in that night, Pitts stayed hidden until the library closed for the evening. Alone, he wandered through the stacks of books until he came across *Principia Mathematica*, a three-volume tome written by Bertrand Russell and Alfred Whitehead between 1910 and 1913, which attempted to reduce all of mathematics to pure logic. Pitts sat down and began to read. For three days he remained in the library until he had read each volume cover to cover—nearly 2,000 pages in all—and had identified several mistakes. Deciding that Bertrand Russell himself needed to know about these, the boy drafted a letter to Russell detailing the errors. Not only did Russell write back, he was so impressed that he invited Pitts to study with him as a graduate student at Cambridge University in England. Pitts couldn't oblige him, though—he was only 12 years old. But three years later, when he heard that Russell would be visiting the University of Chicago, the 15-year-old ran away from home and headed for Illinois. He never saw his family again.

McCulloch was a confident, gray-eyed, wild-bearded, chain-smoking philosopher-poet who lived on whiskey and ice cream and never went to bed before 4 a.m.

In 1923, the year that Walter Pitts was born, a 25-year-old Warren McCulloch was also digesting the *Principia*. But that is where the similarities ended—McCulloch could not have come from a more different world. Born into a well-to-do East Coast family of lawyers, doctors, theologians, and engineers, McCulloch attended a private boys academy in New Jersey, then studied mathematics at Haverford College in Pennsylvania, then philosophy and psychology at Yale. In 1923 he was at Columbia, where he was studying "experimental aesthetics" and was about to earn his medical degree in neurophysiology. But McCulloch was a philosopher at heart. He wanted to know what it means to know. Freud had just published *The Ego and the Id*, and psychoanalysis was all the rage. McCulloch didn't buy it—he felt certain that somehow the mysterious workings and failings of the mind were rooted in the purely mechanical

firings of neurons in the brain.

Though they started at opposite ends of the socioeconomic spectrum, McCulloch and Pitts were destined to live, work, and die together. Along the way, they would create the first mechanistic theory of the mind, the first computational approach to neuroscience, the logical design of modern computers, and the pillars of artificial intelligence. But this is more than a story about a fruitful research collaboration. It is also about the bonds of friendship, the fragility of the mind, and the limits of logic's ability to redeem a messy and imperfect world.
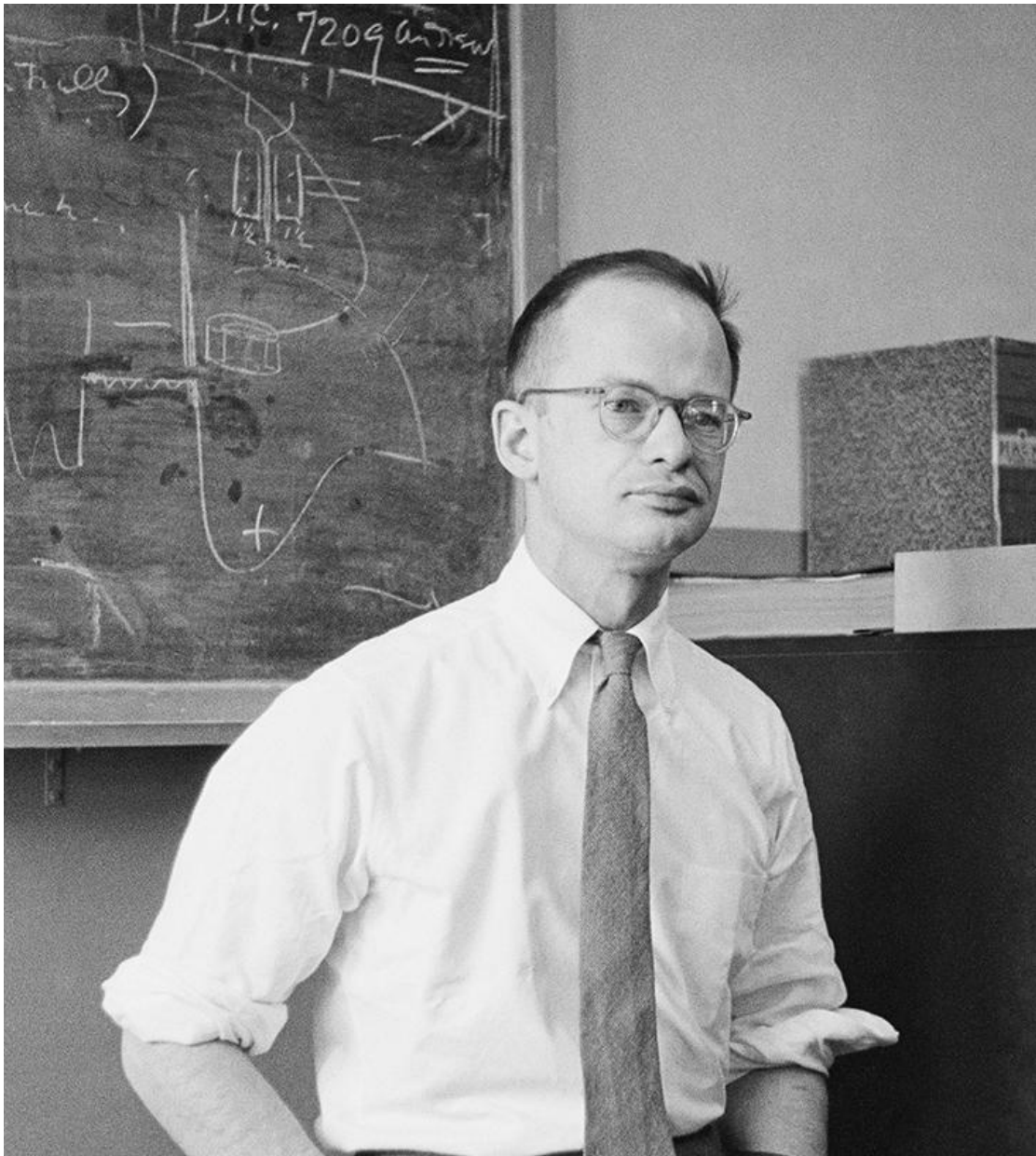
Walter Pitts (1923-1969): *Walter Pitts' life passed from homeless runaway, to MIT neuroscience pioneer, to withdrawn alcoholic.* Estate of Francis Bello / Science Source

Standing face to face, they were an unlikely pair. McCulloch, 42 years old when he met Pitts, was a confident, gray-eyed, wild-bearded, chain-smoking philosopher-poet who lived on whiskey and ice cream and never went to bed before 4 a.m. Pitts, 18, was small and shy, with a long forehead that prematurely aged him, and a squat, duck-like, bespectacled face. McCulloch was a respected scientist. Pitts was a homeless runaway. He'd been hanging around the University of Chicago, working a menial job and sneaking into Russell's lectures, where he met a young medical student named Jerome Lettvin. It was Lettvin who introduced the two men. The moment they spoke, they realized they shared a hero in common: Gottfried Leibniz. The 17th-century philosopher had attempted to create an alphabet of human thought, each letter of which represented a concept and could be combined and manipulated according to a set of logical rules to compute all knowledge—a vision that promised to transform the imperfect outside world into the rational sanctuary of a library.

McCulloch explained to Pitts that he was trying to model the brain with a Leibnizian logical calculus. He had been inspired by the *Principia*, in which Russell and Whitehead tried to show that all of mathematics could be built from the ground up using basic, indisputable logic. Their building block was the proposition—the simplest possible statement, either true or false. From there, they employed the fundamental operations of logic, like the conjunction ("and"), disjunction ("or"), and negation ("not"), to link propositions into increasingly complicated networks. From these simple propositions, they derived the full complexity of modern mathematics.

Which got McCulloch thinking about neurons. He knew that each of the brain's nerve cells only fires after a minimum threshold has been reached: Enough of its neighboring nerve cells must send signals across the neuron's synapses before it will fire off its own electrical spike. It occurred to McCulloch that this set-up was binary—either the neuron fires or it doesn't. A neuron's signal, he realized, is a proposition, and neurons seemed to work like logic gates, taking in multiple inputs and producing a single output. By varying a neuron's firing threshold, it could be made to perform "and," "or," and "not" functions.

Late at night, McCulloch and Pitts alone would pour the whiskey, hunker down, and attempt to build a computational brain from the neuron up.

Fresh from reading a new paper by a British mathematician named Alan Turing which proved the possibility of a machine that could compute any function (so long as it was

possible to do so in a finite number of steps), McCulloch became convinced that the brain was just such a machine—one which uses logic encoded in neural networks to *compute*. Neurons, he thought, could be linked together by the rules of logic to build more complex chains of thought, in the same way that the *Principia* linked chains of propositions to build complex mathematics.

As McCulloch explained his project, Pitts understood it immediately, and knew exactly which mathematical tools could be used. McCulloch, enchanted, invited the teen to live with him and his family in Hinsdale, a rural suburb on the outskirts of Chicago. The Hinsdale household was a bustling, free-spirited bohemia. Chicago intellectuals and literary types constantly dropped by the house to discuss poetry, psychology, and radical politics while Spanish Civil War and union songs blared from the phonograph. But late at night, when McCulloch's wife Rook and the three children went to bed, McCulloch and Pitts alone would pour the whiskey, hunker down, and attempt to build a computational brain from the neuron up.

Before Pitts' arrival, McCulloch had hit a wall: There was nothing stopping chains of neurons from twisting themselves into loops, so that the output of the last neuron in a chain became the input of the first—a neural network chasing its tail. McCulloch had no idea how to model that mathematically. From the point of view of logic, a loop smells a lot like paradox: the consequent becomes the antecedent, the effect becomes the cause. McCulloch had been labeling each link in the chain with a time stamp, so that if the first neuron fired at time t, the next one fired at $t+1$, and so on. But when the chains circled back, $t+1$ suddenly came *before t*.

Pitts knew how to tackle the problem. He used modulo mathematics, which deals with numbers that circle back around on themselves like the hours of a clock. He showed McCulloch that the paradox of time $t+1$ coming before time t wasn't a paradox at all, because in his calculations "before" and "after" lost their meaning. Time was removed from the equation altogether. If one were to see a lightning bolt flash on the sky, the eyes would send a signal to the brain, shuffling it through a chain of neurons. Starting with any given neuron in the chain, you could retrace the signal's steps and figure out just how long ago lightning struck. Unless, that is, the chain is a loop. In that case, the information encoding the lightning bolt just spins in circles, endlessly. It bears no connection to the time at which the lightning actually occurred. It becomes, as McCulloch put it, "an idea wrenched out of time." In other words, a memory.

By the time Pitts finished calculating, he and McCulloch had on their hands a mechanistic model of the mind, the first application of computation to the brain, and the first argument that the brain, at bottom, is an information processor. By stringing simple binary neurons into chains and loops, they had shown that the brain could implement every possible logical operation and compute anything that could be computed by one of Turing's hypothetical machines. Thanks to those ouroboric loops, they had also found a way for the brain to abstract a piece of information, hang on to it, and abstract it yet again, creating rich, elaborate hierarchies of lingering ideas in a process we call "thinking."

McCulloch and Pitts wrote up their findings in a now-seminal paper, "A Logical Calculus of Ideas Immanent in Nervous Activity," published in the *Bulletin of Mathematical Biophysics*. Their model was vastly oversimplified for a biological brain, but it

succeeded at showing a proof of principle. Thought, they said, need not be shrouded in Freudian mysticism or engaged in struggles between ego and id. "For the first time in the history of science," McCulloch announced to a group of philosophy students, "we know how we know."

Pitts had found in McCulloch everything he had needed—acceptance, friendship, his intellectual other half, the father he never had. Although he had only lived in Hinsdale for a short time, the runaway would refer to McCulloch's house as home for the rest of his life. For his part, McCulloch was just as enamored. In Pitts he had found a kindred spirit, his "bootlegged collaborator," and a mind with the technical prowess to bring McCulloch's half-formed notions to life. As he put it in a letter of reference about Pitts,

"Would I had him with me always."[1]

Pitts was soon to make a similar impression on one of the towering intellectual figures of the 20th century, the mathematician, philosopher, and founder of cybernetics, Norbert Wiener. In 1943, Lettvin brought Pitts into Wiener's office at the Massachusetts Institute of Technology (MIT). Wiener didn't introduce himself or make small talk. He simply walked Pitts over to a blackboard where he was working out a mathematical proof. As Wiener worked, Pitts chimed in with questions and suggestions. According to Lettvin, by the time they reached the second blackboard, it was clear that Wiener had found his new right-hand man. Wiener would later write that Pitts was "without question the strongest young scientist whom I have ever met … I should be extremely astonished if he does not prove to be one of the two or three most important scientists of his generation, not merely in America but in the world at large."

So impressed was Wiener that he promised Pitts a Ph.D. in mathematics at MIT, despite the fact that he had never graduated from high school—something that the strict rules at the University of Chicago prohibited. It was an offer Pitts couldn't refuse. By the fall of 1943, Pitts had moved into a Cambridge apartment, was enrolled as a special student at MIT, and was studying under one of the most influential scientists in the world. It was quite a long way from blue-collar Detroit.

Wiener wanted Pitts to make his model of the brain more realistic. Despite the leaps Pitts and McCulloch had made, their work had made barely a ripple among brain scientists—in part because the symbolic logic they'd employed was hard to decipher, but also because their stark and oversimplified model didn't capture the full messiness of the biological brain. Wiener, however, understood the implications of what they'd done, and knew that a more realistic model would be game-changing. He also realized that it ought to be possible for Pitts' neural networks to be implemented in man-made machines, ushering in his dream of a cybernetic revolution. Wiener figured that if Pitts was going to make a realistic model of the brain's 100 billion interconnected neurons, he was going to need statistics on his side. And statistics and probability theory were Wiener's area of expertise. After all, it had been Wiener who discovered a precise mathematical definition of information: The higher the probability, the higher the entropy and the lower the information content.

The scientists in the room were floored. And yet, everyone who knew Pitts was sure that he could do it.

As Pitts began his work at MIT, he realized that although genetics must encode for gross neural features, there was no way our genes could pre-determine the trillions of synaptic connections in the brain—the amount of information it would require was untenable. It must be the case, he figured, that we all start out with essentially random neural networks—highly probable states containing negligible information (a thesis that continues to be debated to the present day). He suspected that by altering the thresholds of neurons over time, randomness could give way to order and information could emerge. He set out to model the process using statistical mechanics. Wiener excitedly cheered him on, because he knew if such a model were embodied in a machine, that machine could *learn*.

"I now understand at once some seven-eighths of what Wiener says, which I am told is

something of an achievement," Pitts wrote in a letter to McCulloch in December of 1943, some three months after he'd arrived. His work with Wiener was "to constitute the first adequate discussion of statistical mechanics, understood in the most general possible sense, so that it includes for example the problem of deriving the psychological, or statistical, laws of behavior from the microscopic laws of neurophysiology … Doesn't it sound fine?"

That winter, Wiener brought Pitts to a conference he organized in Princeton with the mathematician and physicist John von Neumann, who was equally impressed with Pitts' mind. Thus formed the beginnings of the group who would become known as the cyberneticians, with Wiener, Pitts, McCulloch, Lettvin, and von Neumann its core. And among this rarified group, the formerly homeless runaway stood out. "None of us would think of publishing a paper without his corrections and approval," McCulloch wrote. "[Pitts] was in no uncertain terms the genius of our group," said Lettvin. "He was absolutely incomparable in the scholarship of chemistry, physics, of everything you could talk about history, botany, etc. When you asked him a question, you would get back a whole textbook … To him, the world was connected in a very complex and wonderful fashion."[2]

The following June, 1945, von Neumann penned what would become a historic document entitled "First Draft of a Report on the EDVAC," the first published description of a stored-program binary computing machine—the modern computer. The EDVAC's predecessor, the ENIAC, which took up 1,800 square feet of space in Philadelphia, was more like a giant electronic calculator than a computer. It was possible to reprogram the thing, but it took several operators several weeks to reroute all the wires and switches to do it. Von Neumann realized that it might not be necessary to rewire the machine every time you wanted it to perform a new function. If you could take each configuration of the switches and wires, abstract them, and encode them symbolically as pure information, you could feed them into the computer the same way you'd feed it data, only now the data would include the very programs that manipulate the data. Without having to rewire a thing, you'd have a universal Turing machine.

To accomplish this, von Neumann suggested modeling the computer after Pitts and McCulloch's neural networks. In place of neurons, he suggested vacuum tubes, which would serve as logic gates, and by stringing them together exactly as Pitts and McCulloch had discovered, you could carry out any computation. To store the programs as data, the computer would need something new: a memory. That's where Pitts' loops came into play. "An element which stimulates itself will hold a stimulus indefinitely," von Neumann wrote in his report, echoing Pitts and employing his modulo mathematics. He detailed every aspect of this new computational architecture. In the entire report, he cited only a single paper: "A Logical Calculus" by McCulloch and Pitts.

By 1946, Pitts was living on Beacon Street in Boston with Oliver Selfridge, an MIT student who would become "the father of machine perception"; Hyman Minsky, the future economist; and Lettvin. He was teaching mathematical logic at MIT and working with Wiener on the statistical mechanics of the brain. The following year, at the Second Cybernetic Conference, Pitts announced that he was writing his doctoral dissertation on probabilistic three-dimensional neural networks. The scientists in the room were

floored. "Ambitious" was hardly the word to describe the mathematical skill that it would take to pull off such a feat. And yet, everyone who knew Pitts was sure that he could do it. They would be waiting with bated breath.

In a letter to the philosopher Rudolf Carnap, McCulloch catalogued Pitts' achievements. "He is the most omniverous of scientists and scholars. He has become an excellent dye chemist, a good mammalogist, he knows the sedges, mushrooms and the birds of New England. He knows neuroanatomy and neurophysiology from their original sources in Greek, Latin, Italian, Spanish, Portuguese, and German for he learns any language he needs as soon as he needs it. Things like electrical circuit theory and the practical soldering in of power, lighting, and radio circuits he does himself. In my long life, I have never seen a man so erudite or so really practical." Even the media took notice. In June 1954, *Fortune* magazine ran an article featuring the 20 most talented scientists under 40; Pitts was featured, next to Claude Shannon and James Watson. Against all odds, Walter Pitts had skyrocketed into scientific stardom.

Some years earlier, in a letter to McCulloch, Pitts wrote "About once a week now I become violently homesick to talk all evening and all night to you." Despite his success, Pitts had become homesick—and home meant McCulloch. He was coming to believe that if he could work with McCulloch again, he would be happier, more productive, and more likely to break new ground. McCulloch, too, seemed to be floundering without his bootlegged collaborator.

Suddenly, the clouds broke. In 1952, Jerry Wiesner, associate director of MIT's Research Laboratory of Electronics, invited McCulloch to head a new project on brain science at MIT. McCulloch jumped at the opportunity—because it meant he would be working with Pitts again. He traded his full professorship and his large Hinsdale home for a research associate title and a crappy apartment in Cambridge, and couldn't have been happier about it. The plan for the project was to use the full arsenal of information theory, neurophysiology, statistical mechanics, and computing machines to understand how the brain gives rise to the mind. Lettvin, along with the young neuroscientist Patrick Wall, joined McCulloch and Pitts at their new headquarters in Building 20 on Vassar Street. They posted a sign on the door: *Experimental Epistemology*.

With Pitts and McCulloch together again, and with Wiener and Lettvin in the mix, everything seemed poised for progress and revolution. Neuroscience, cybernetics, artificial intelligence, computer science—it was all on the brink of an intellectual explosion. The sky—or the mind—was the limit.

He began drinking heavily and pulled away from his friends. He set fire to his dissertation along with all of his notes and his papers.

There was just one person who wasn't happy about the reunion: Wiener's wife. Margaret Wiener was, by all accounts, a controlling, conservative prude—and she despised McCulloch's influence on her husband. McCulloch hosted wild get-togethers at his family farm in Old Lyme, Connecticut, where ideas roamed free and everyone went skinny-dipping. It had been one thing when McCulloch was in Chicago, but now he was coming to Cambridge and Margaret wouldn't have it. And so she invented a story. She sat Wiener down and informed him that when their daughter, Barbara, had stayed at McCulloch's house in Chicago, several of "his boys" had seduced her. Wiener

immediately sent an angry telegram to Wiesner: "Please inform [Pitts and Lettvin] that all connection between me and your projects is permanently abolished. They are your problem. Wiener." He never spoke to Pitts again. And he never told him why.[3]

For Pitts, this marked the beginning of the end. Wiener, who had taken on a fatherly role in his life, now abandoned him inexplicably. For Pitts, it wasn't merely a loss. It was something far worse than that: It defied logic.

And then there were the frogs. In the basement of Building 20 at MIT, along with a garbage can full of crickets, Lettvin kept a group of them. At the time, biologists believed that the eye was like a photographic plate that passively recorded dots of light and sent them, dot for dot, to the brain, which did the heavy lifting of interpretation. Lettvin decided to put the idea to the test, opening up the frog's skulls and attaching electrodes to single fibers in their optic nerves.

Pitts with Lettvin: *Pitts with Jerome Lettvin and one subject of their experiments on visual perception (1959).* Wikipedia

Together with Pitts, McCulloch and the Chilean biologist and philosopher Humberto Maturana, he subjected the frogs to various visual experiences—brightening and dimming the lights, showing them color photographs of their natural habitat, magnetically dangling artificial flies—and recorded what the eye measured before it sent the information off to the brain. To everyone's surprise, it didn't merely record what it saw, but filtered and analyzed information about visual features like contrast, curvature, and movement. "The eye speaks to the brain in a language already highly organized and interpreted," they reported in the now-seminal paper "What the Frog's Eye Tells the Frog's Brain," published in 1959.

The results shook Pitts' worldview to its core. Instead of the brain computing

information digital neuron by digital neuron using the exacting implement of mathematical logic, messy, analog processes in the eye were doing at least part of the interpretive work. "It was apparent to him after we had done the frog's eye that even if logic played a part, it didn't play the important or central part that one would have expected," Lettvin said. "It disappointed him. He would never admit it, but it seemed to add to his despair at the loss of Wiener's friendship."

Once everything had been reduced to information governed by logic, the actual mechanics ceased to matter—the tradeoff for universal computation was ontology.

The spate of bad news aggravated a depressive streak that Pitts had been struggling with for years. "I have a kind of personal woe I should like your advice on," Pitts had written to McCulloch in one of his letters. "I have noticed in the last two or three years a growing tendency to a kind of melancholy apathy or depression. [Its] effect is to make the positive value seem to disappear from the world, so that nothing seems worth the effort of doing it, and whatever I do or what happens to me ceases to matter very greatly …"

In other words, Pitts was struggling with the very logic he had sought in life. Pitts wrote that his depression might be "common to all people with an excessively logical education who work in applied mathematics: It is a kind of pessimism resulting from an inability to believe in what people call the Principle of Induction, or the principle of the Uniformity of Nature. Since one cannot prove, or even render probable a priori, that the sun should rise tomorrow, we cannot really believe it shall."

Now, alienated from Wiener, Pitts' despair turned lethal. He began drinking heavily and pulled away from his friends. When he was offered his Ph.D., he refused to sign the paperwork. He set fire to his dissertation along with all of his notes and his papers. Years of work—important work that everyone in the community was eagerly awaiting—he burnt it all, priceless information reduced to entropy and ash. Wiesner offered Lettvin increased support for the lab if he could recover any bits of the dissertation. But it was all gone.

Pitts remained employed by MIT, but this was little more than a technicality; he hardly spoke to anyone and would frequently disappear. "We'd go hunting for him night after night," Lettvin said. "Watching him destroy himself was a dreadful experience." In a way Pitts was still 12 years old. He was still beaten, still a runaway, still hiding from the world in musty libraries. Only now his books took the shape of a bottle.

With McCulloch, Pitts had laid the foundations for cybernetics and artificial intelligence. They had steered psychiatry away from Freudian analysis and toward a mechanistic understanding of thought. They had shown that the brain computes and that mentation is the processing of information. In doing so, they had also shown how a machine could compute, providing the key inspiration for the architecture of modern computers. Thanks to their work, there was a moment in history when neuroscience, psychiatry, computer science, mathematical logic, and artificial intelligence were all one thing, following an idea first glimpsed by Leibniz—that man, machine, number, and mind all use information as a universal currency. What appeared on the surface to be very different ingredients of the world—hunks of metal, lumps of gray matter, scratches of ink on a page—were profoundly interchangeable.

There was a catch, though: This symbolic abstraction made the world transparent but the brain opaque. Once everything had been reduced to information governed by logic, the actual mechanics ceased to matter—the tradeoff for universal computation was ontology. Von Neumann was the first to see the problem. He expressed his concern to Wiener in a letter that anticipated the coming split between artificial intelligence on one side and neuroscience on the other. "After the great positive contribution of Turing-cum-Pitts-and-McCulloch is assimilated," he wrote, "the situation is rather worse than better than before. Indeed these authors have demonstrated in absolute and hopeless generality that anything and everything … can be done by an appropriate mechanism, and specifically by a neural mechanism—and that even one, definite mechanism can be 'universal.' Inverting the argument: Nothing that we may know or learn about the functioning of the organism can give, without 'microscopic,' cytological work any clues regarding the further details of the neural mechanism."

This universality made it impossible for Pitts to provide a model of the brain that was practical, and so his work was dismissed and more or less forgotten by the community of scientists working on the brain. What's more, the experiment with the frogs had shown that a purely logical, purely brain-centered vision of thought had its limits. Nature had chosen the messiness of life over the austerity of logic, a choice Pitts likely could not comprehend. He had no way of knowing that while his ideas about the biological brain were not panning out, they were setting in motion the age of digital computing, the neural network approach to machine learning, and the so-called connectionist philosophy of mind. In his own mind, he had been defeated.

On Saturday, April 21, 1969, his hand shaking with an alcoholic's delirium tremens, Pitts sent a letter from his room at Beth Israel Hospital in Boston to McCulloch's room down the road at the Cardiac Intensive Care Ward at Peter Bent Brigham Hospital. "I understand you had a light coronary; … that you are attached to many sensors connected to panels and alarms continuously monitored by a nurse, and cannot in consequence turn over in bed. No doubt this is cybernetical. But it all makes me most abominably sad." Pitts himself had been in the hospital for three weeks, having been admitted with liver problems and jaundice. On May 14, 1969 Walter Pitts died alone in a boarding house in Cambridge, of bleeding esophageal varices, a condition associated with cirrhosis of the liver. Four months later, McCulloch passed away, as if the existence of one without the other were simply illogical, a reverberating loop wrenched open.

Amanda Gefter is a physics writer and author of Trespassing on Einstein's Lawn: A father, a daughter, the meaning of nothing and the beginning of everything. *She lives in Cambridge, Massachusetts.*

References

1. All letters retrieved from the McCulloch Papers, BM139, Series I: Correspondence 1931–1968, Folder "Pitts, Walter."

2. All Jerome Lettvin quotes taken from: Anderson, J.A. & Rosenfield, E. *Talking Nets: An Oral History of Neural Networks* MIT Press (2000).

3. Conway F. & Siegelman J. *Dark Hero of the Information Age: In Search of Norbert Wiener, the Father of Cybernetics* Basic Books, New York, NY (2006).