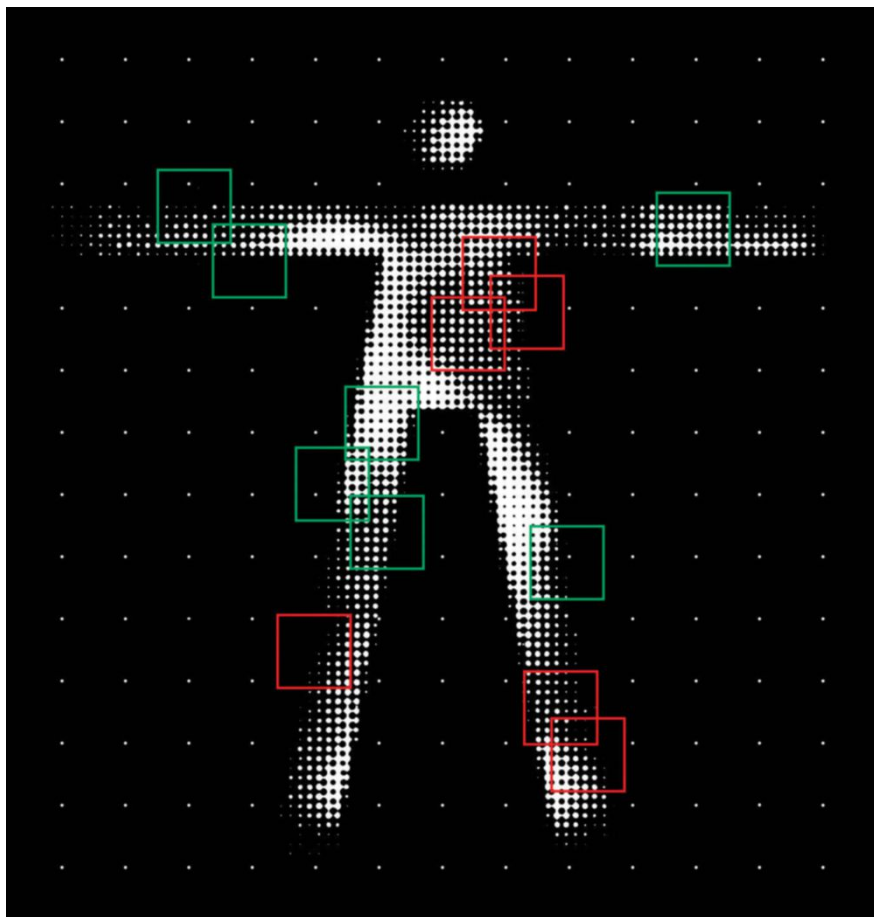# THE ALGORITHM WILL SEE YOU NOW

When it comes to diagnosis, will A.I. replace the M.D.?

BY SIDDHARTHA MUKHERJEE

In some trials, "deep learning" systems have outperformed human experts.

ONE EVENING LAST November, a fifty-four-year-old woman from the Bronx arrived at the emergency room at Columbia University's medical center with a grinding headache. Her vision had become blurry, she told the E.R. doctors, and her left hand felt numb and weak. The doctors examined her and ordered a CT scan of her head.

A few months later, on a morning this January, a team of four radiologists-in-training huddled in front of a computer in a third-floor room of the hospital. The room was windowless and dark, aside from the light from the screen, which looked as if it had been filtered through seawater. The residents filled a cubicle, and Angela Lignelli-Dipple, the chief of neuroradiology at Columbia, stood behind them with a pencil and pad. She was training them to read CT scans.

"It's easy to diagnose a stroke once the brain is dead and gray," she said. "The trick is to diagnose the stroke before too many nerve cells begin to die." Strokes are usually caused by blockages or bleeds, and a neuroradiologist has about a forty-five-minute window to make a diagnosis, so that doctors might be able to intervene—to dissolve a growing clot, say. "Imagine you are in the E.R.," Lignelli-Dipple continued, raising the ante. "Every minute that passes, some part of the brain is dying. Time lost is brain lost."

She glanced at a clock on the wall, as the seconds ticked by. "So where's the problem?" she asked.

Strokes are typically asymmetrical. The blood supply to the brain branches left and right and then breaks into rivulets and tributaries on each side. A clot or a bleed usually affects only one of these branches, leading to a one-sided deficit in a part of the brain. As the nerve cells lose their blood supply and die, the tissue swells subtly. On a scan, the crisp borders between the anatomical structures can turn hazy. Eventually, the tissue shrinks, trailing a parched shadow. But that shadow usually appears on the scan several hours, or even days, after the stroke, when the window of intervention has long closed. "Before that," Lignelli-Dipple told me, "there's just a hint of something on a scan"—the premonition of a stroke.

The images on the Bronx woman's scan cut through the skull from its base to the apex in horizontal planes, like a melon sliced from bottom to top. The residents raced through the layers of images, as if thumbing through a flipbook, calling out the names of the anatomical structures: cerebellum, hippocampus, insular cortex, striatum, corpus callosum, ventricles. Then one of the residents, a man in his late twenties, stopped at a picture and motioned with the tip of a pencil at an area on the right edge of the brain. "There's something patchy here," he said. "The borders look hazy." To me, the whole image looked patchy and hazy—a blur of pixels—but he had obviously seen something unusual.

"Hazy?" Lignelli-Dipple prodded. "Can you describe it a little more?"

The resident fumbled for words. He paused, as if going through the anatomical structures in his mind, weighing the possibilities. "It's just not uniform." He shrugged. "I don't know. Just *looks* funny."

Lignelli-Dipple pulled up a second CT scan, taken twenty hours later. The area pinpointed by the resident, about the diameter of a grape, was dull and swollen. A series of further scans, taken days apart, told the rest of the story. A distinct wedge-shaped field of gray appeared. Soon after the woman got to the E.R., neurologists had tried to open the clogged artery with clot-busting

drugs, but she had arrived too late. A few hours after the initial scan, she lost consciousness, and was taken to the I.C.U. Two months later, the woman was still in a ward upstairs. The left side of her body—from the upper arms to the leg—was paralyzed.

I walked with Lignelli-Dipple to her office. I was there to learn about learning: How do doctors learn to diagnose? And could machines learn to do it, too?

MY OWN INDUCTION into diagnosis began in the fall of 1997, in Boston, as I started my clinical rotations. To prepare, I read a textbook, a classic in medical education, that divided the act of diagnosis into four tidy phases. First, the doctor uses a patient's history and a physical exam to collect facts about her complaint or condition. Next, this information is collated to generate a comprehensive list of potential causes. Then questions and preliminary tests help eliminate one hypothesis and strengthen another—so-called "differential diagnosis." Weight is given to how common a disease might be, and to a patient's prior history, risks, exposures. ("When you hear hoofbeats," the saying goes, "think horses, not zebras.") The list narrows; the doctor refines her assessment. In the final phase, definitive lab tests, X-rays, or CT scans are deployed to confirm the hypothesis and seal the diagnosis. Variations of this stepwise process were faithfully reproduced in medical textbooks for decades, and the image of the diagnostician who plods methodically from symptom to cause had been imprinted on generations of medical students.

But the real art of diagnosis, I soon learned, wasn't so straightforward. My preceptor in medical school was an elegant New Englander with polished loafers and a starched accent. He prided himself on being an expert diagnostician. He would ask a patient to demonstrate the symptom—a cough, say—and then lean back in his chair, letting adjectives roll over his tongue. "Raspy and tinny," he might say, or "base, with an ejaculated thrum," as if he were describing a vintage bottle of Bordeaux. To me, all the coughs sounded exactly the same, but I'd play along—"Raspy, yes"—like an anxious impostor at a wine tasting.

The taxonomist of coughs would im-mediately narrow down the diagnostic possibilities. "It sounds like a pneumonia," he might say, or "the wet rales of congestive heart failure." He would then let loose a volley of questions. Had the patient experienced recent weight gain? Was there a history of asbestos exposure? He'd ask the patient to cough again and he'd lean down, listening intently with his stethoscope. Depending on the answers, he might generate another series of possibilities, as if strengthening and weakening synapses. Then, with the élan of a roadside magician, he'd proclaim his diagnosis—"Heart failure!"—and order tests to prove that it was correct. It usually was.

A few years ago, researchers in Brazil studied the brains of expert radiologists in order to understand how they reached their diagnoses. Were these seasoned diagnosticians applying a mental "rule book" to the images, or did they apply "pattern recognition or non-analytical reasoning"?

Twenty-five such radiologists were asked to evaluate X-rays of the lung while inside MRI machines that could track the activities of their brains. (There's a marvellous series of recursions here: to diagnose diagnosis, the imagers had to be imaged.) X-rays were flashed before them. Some contained a single pathological lesion that might be commonly encountered—perhaps a palm-shaped shadow of a pneumonia, or the dull, opaque wall of fluid that had accumulated behind the lining of the lung. Embedded in a second group of diagnostic images were line drawings of animals; within a third group, the outlines of letters of the alphabet. The radiologists were shown the three types of images in random order, and then asked to call out the name of the lesion, the animal, or the letter as quickly as possible while the MRI machine traced the activity of their brains. It took the radiologists an average of 1.33 seconds to come up with a diagnosis. In all three cases, the same areas of the brain lit up: a wide delta of neurons near the left ear, and a moth-shaped band above the posterior base of the skull.

"Our results support the hypothesis that a process similar to naming things in everyday life occurs when a physician promptly recognizes a characteristic and previously known lesion," the researchers concluded. Identifying a lesion was a process similar to naming the animal. When you recognize a rhinoceros, you're not considering and eliminating alternative candidates. Nor are you mentally fusing a unicorn, an armadillo, and a small elephant. You recognize a rhinoceros in its totality—as a pattern. The same was true for radiologists. They weren't cogitating, recollecting, differentiating; they were seeing a commonplace object. For my preceptor, similarly, those wet rales were as recognizable as a familiar jingle.

IN 1945, THE British philosopher Gilbert Ryle gave an influential lecture about two kinds of knowledge. A child knows that a bicycle has two wheels, that its tires are filled with air, and that you ride the contraption by pushing its pedals forward in circles. Ryle termed this kind of knowledge—the factual, propositional kind—"knowing that." But to learn to ride a bicycle involves another realm of learning. A child learns how to ride by falling off, by balancing herself on two wheels, by going over potholes. Ryle termed this kind of knowledge—implicit, experiential, skill-based—"knowing how."

The two kinds of knowledge would seem to be interdependent: you might use factual knowledge to deepen your experiential knowledge, and vice versa. But Ryle warned against the temptation to think that "knowing how" could be reduced to "knowing that"—a playbook of rules couldn't teach a child to ride a bike. Our rules, he asserted, make sense only because we know how to use them: "Rules, like birds, must live before they can be stuffed." One afternoon, I watched my seven-year-old daughter negotiate a small hill on her bike. The first time she tried, she stalled at the steepest part of the slope and fell off. The next time, I saw her lean forward, imperceptibly at first, and then more visibly, and adjust her weight back on the seat as the slope decreased. But I hadn't taught her rules to ride a bike up that hill. When her daughter learns to negotiate the same hill, I imagine, she won't teach her the rules, either. We pass on a few precepts about the universe but leave the brain to figure out the rest.

Some time after Lignelli-Dipple's session with the radiology trainees, I

spoke to Steffen Haider, the young man who had picked up the early stroke on the CT scan. How had he found that culprit lesion? Was it "knowing that" or "knowing how"? He began by telling me about learned rules. He knew that strokes are often one-sided; that they result in the subtle "graying" of tissue; that the tissue often swells slightly, causing a loss of anatomical borders. "There are spots in the brain where the blood supply is particularly vulnerable," he said. To identify the lesion, he'd have to search for these signs on one side which were not present on the other.

I reminded him that there were plenty of asymmetries in the image that he had ignored. This CT scan, like most, had other gray squiggles on the left that weren't on the right—artifacts of movement, or chance, or underlying changes in the woman's brain that preceded the stroke. How had he narrowed his focus to that one area? He paused as the thought pedalled forward and gathered speed in his mind. "I don't know—it was partly subconscious," he said, finally.

"That's what happens—a clicking together—as you grow and learn as a radiologist," Lignelli-Dipple told me. The question was whether a machine could "grow and learn" in the same manner.

In January, 2015, the computer scientist Sebastian Thrun became fascinated by a conundrum in medical diagnostics. Thrun, who grew up in Germany, is lean, with a shaved head and an air of comic exuberance; he looks like some fantastical fusion of Michel Foucault and Mr. Bean. Formerly a professor at Stanford, where he directed the Artificial Intelligence Lab, Thrun had gone off to start Google X, directing work on self-learning robots and driverless cars. But he found himself drawn to learning devices in medicine. His mother had died of breast cancer when she was forty-nine years old—Thrun's age now. "Most patients with cancer have no symptoms at first," Thrun told me. "My mother didn't. By the time she went to her doctor, her cancer had already metastasized. I became obsessed with the idea of detecting cancer in its

earliest stage—at a time when you could still cut it out with a knife. And I kept thinking, Could a machine-learning algorithm help?"

Early efforts to automate diagnosis tended to hew closely to the textbook realm of explicit knowledge. Take the electrocardiogram, which renders the heart's electrical activity as lines on a page or a screen. For the past twenty years, computer interpretation has often been a feature of these systems. The programs that do the work tend to be fairly straightforward. Characteristic waveforms are associated with various conditions—atrial fibrillation, or the blockage of a blood vessel—and rules to recognize these waveforms are fed into the appliance. When the machine recognizes the waveforms, it flags a heartbeat as "atrial fibrillation."

In mammography, too, "computer-aided detection" is becoming commonplace. Pattern-recognition software highlights suspicious areas, and radiologists review the results. But here again the recognition software typically uses a rule-based system to identify a suspicious lesion. Such programs have no built-in mechanism to learn: a machine that has seen three thousand X-rays is no wiser than one that has seen just four. These limitations became starkly evident in a 2007 study that compared the accuracy of mammography before and after the implementation of computer-aided diagnostic devices. One might have expected the accuracy of diagnosis to have increased dramatically after the devices had been implemented. As it happens, the devices had a complicated effect. The rate of biopsies shot up in the computer-assisted group. Yet the detection of small, invasive breast cancers—the kind that oncologists are most keen to detect—*decreased*. (Even later studies have shown problems with false positives.)

Thrun was convinced that he could outdo these first-generation diagnostic devices by moving away from rule-based algorithms to learning-based ones—from rendering a diagnosis by "knowing that" to doing so by "knowing how." Increasingly, learning algorithms of the kind that Thrun works with involve a

computing strategy known as a "neural network," because it's inspired by a model of how the brain functions. In the brain, neural synapses are strengthened and weakened through repeated activation; these digital systems aim to achieve something similar through mathematical means, adjusting the "weights" of the connections to move toward the desired output. The more powerful ones have something akin to layers of neurons, each processing the input data and sending the results up to the next layer. Hence, "deep learning."

Thrun began with skin cancer; in particular, keratinocyte carcinoma (the most common class of cancer in the U.S.) and melanoma (the most dangerous kind of skin cancer). Could a machine be taught to distinguish skin cancer from a benign skin condition—acne, a rash, or a mole—by scanning a photograph? "If a dermatologist can do it, then a machine should be able to do it as well," Thrun reasoned. "Perhaps a machine could do it even better."

Traditionally, dermatological teaching about melanoma begins with a rule-based system that, as medical students learn, comes with a convenient mnemonic: ABCD. Melanomas are often asymmetrical ("A"), their borders ("B") are uneven, their color ("C") can be patchy and variegated, and their diameter ("D") is usually greater than six millimetres. But, when Thrun looked through specimens of melanomas in medical textbooks and on the Web, he found examples where none of these rules applied.

Thrun, who had maintained an adjunct position at Stanford, enlisted two students he worked with there, Andre Esteva and Brett Kuprel. Their first task was to create a so-called "teaching set": a vast trove of images that would be used to teach the machine to recognize a malignancy. Searching online, Esteva and Kuprel found eighteen repositories of skin-lesion images that had been classified by dermatologists. This rogues' gallery contained nearly a hundred and thirty thousand images—of acne, rashes, insect bites, allergic reactions, and cancers—that dermatologists had categorized into nearly two thousand diseases. Notably, there was a set of two thousand lesions that had also been biopsied and examined by
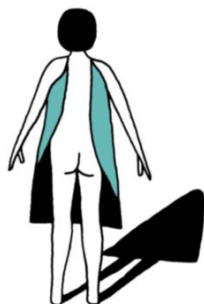
pathologists, and thereby diagnosed with near-certainty.

Esteva and Kuprel began to train the system. They didn't program it with rules; they didn't teach it about ABCD. Instead, they fed the images, and their diagnostic classifications, to the neural network. I asked Thrun to describe what such a network did.

"Imagine an old-fashioned program to identify a dog," he said. "A software engineer would write a thousand if-then-else statements: *if* it has ears, and a snout, *and* has hair, and *is not* a rat . . . and so forth, ad infinitum. But that's not how a child learns to identify a dog, of course. At first, she learns by seeing dogs and being told that they are dogs. She makes mistakes, and corrects herself. She thinks that a wolf is a dog—but is told that it belongs to an altogether different category. And so she shifts her understanding bit by bit: this is 'dog,' that is 'wolf.' The machine-learning algorithm, like the child, pulls information from a training set that has been classified. Here's a dog, and here's not a dog. It then extracts features from one set versus another. And, by testing itself against hundreds and thousands of classified images, it begins to create its own way to recognize a dog—again, the way a child does." It just knows *how* to do it.

In June, 2015, Thrun's team began to test what the machine had learned from the master set of images by presenting it with a "validation set": some fourteen thousand images that had been diagnosed by dermatologists (although not necessarily by biopsy). Could the system correctly classify the images into three diagnostic categories—benign lesions, malignant lesions, and noncancerous growths? The system got the answer right seventy-two per cent of the time. (The actual output of the algorithm is not "yes" or "no" but a probability that a given lesion belongs to a category of interest.) Two board-certified dermatologists who were tested alongside did worse: they got the answer correct sixty-six per cent of the time.

Thrun, Esteva, and Kuprel then widened the study to include twenty-five dermatologists, and this time they used a gold-standard "test set" of roughly two thousand biopsy-proven images. In almost every test, the machine was more sensitive than doctors: it was less likely to miss a melanoma. It was also more specific: it was less likely to call something a melanoma when it wasn't. "In

every test, the network outperformed expert dermatologists," the team concluded, in a report published in *Nature*.

"There's one rather profound thing about the network that wasn't fully emphasized in the paper," Thrun told me. In the first iteration of the study, he and the team had started with a totally naïve neural network. But they found that if they began with a neural network that had already been trained to recognize some unrelated feature (dogs versus cats, say) it learned faster and better. Perhaps our brains function similarly. Those mind-numbing exercises in high school—factoring polynomials, conjugating verbs, memorizing the periodic table—were possibly the opposite: mind-sensitizing.

When teaching the machine, the team had to take some care with the images. Thrun hoped that people could one day simply submit smartphone pictures of their worrisome lesions, and that meant that the system had to be undaunted by a wide range of angles and lighting conditions. But, he recalled, "In some pictures, the melanomas had been marked with yellow disks. We had to crop them out—otherwise, we might teach the computer to pick out a yellow disk as a sign of cancer."

It was an old conundrum: a century ago, the German public became entranced by Clever Hans, a horse that could supposedly add and subtract, and would relay the answer by tapping its hoof. As it turns out, Clever Hans was actually sensing its handler's bearing. As the horse's hoof-taps approached the correct answer, the handler's expression and posture relaxed. The animal's neural network had not learned arithmetic; it had learned to detect changes in human body language. "That's the bizarre thing about neural networks," Thrun said. "You cannot tell what they are picking up. They are like black boxes whose inner workings are mysterious."

The "black box" problem is endemic in deep learning. The system isn't guided by an explicit store of medical knowledge and a list of diagnostic rules; it has effectively taught itself to differentiate moles from melanomas by making vast numbers of internal adjustments—something analogous to strengthening and weakening synaptic connections in the brain. Exactly how did it determine that

THE MIND-BODY PROBLEM

Get up.

No.
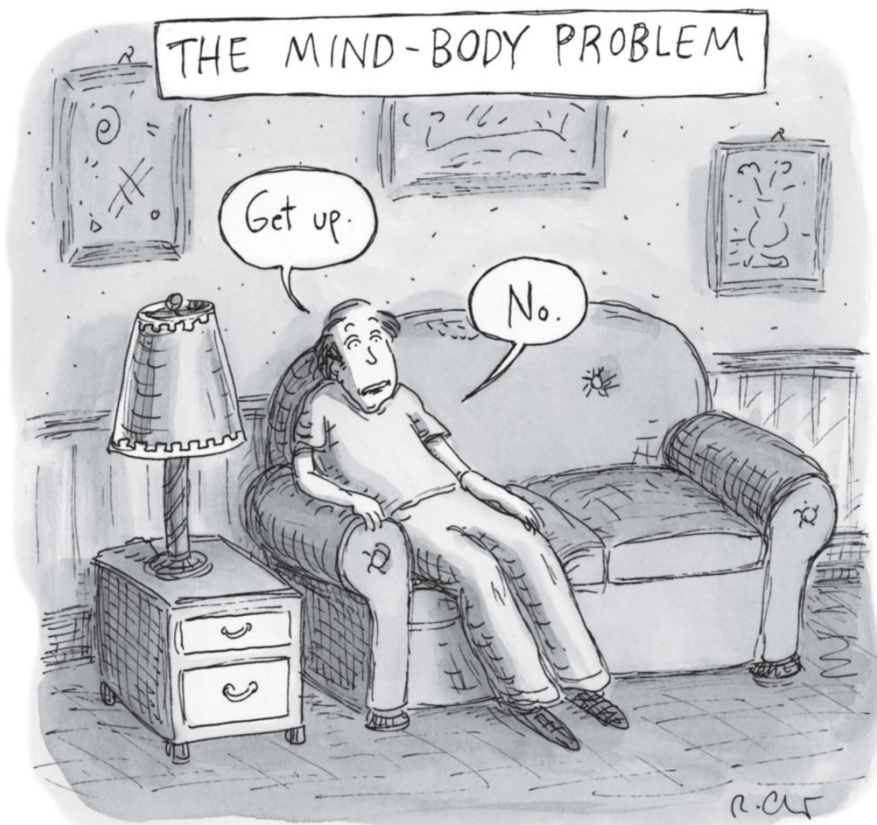
a lesion was a melanoma? We can't know, and it can't tell us. All the internal adjustments and processing that allow the network to learn happen away from our scrutiny. As is true of our own brains. When you make a slow turn on a bicycle, you lean in the opposite direction. My daughter knows to do this, but she doesn't know *that* she does it. The melanoma machine must be extracting certain features from the images; does it matter that it can't tell us which? It's like the smiling god of knowledge. Encountering such a machine, one gets a glimpse of how an animal might perceive a human mind: all-knowing but perfectly impenetrable.

Thrun blithely envisages a world in which we're constantly under diagnostic surveillance. Our cell phones would analyze shifting speech patterns to diagnose Alzheimer's. A steering wheel would pick up incipient Parkinson's through small hesitations and tremors. A bathtub would perform sequential scans as you bathe, via harmless ultrasound or magnetic resonance, to determine whether there's a new mass in an ovary that requires investigation. Big Data would watch, record, and evaluate you: we would shuttle from the grasp of one algorithm to the next. To enter Thrun's world of bathtubs and steering wheels is to enter a hall of diagnostic mirrors, each urging more tests.

It's hard not to be seduced by this vision. Might a medical panopticon that constantly scans us in granular—perhaps even cellular—detail, comparing images day by day, enable us to catch cancer at its earliest stages? Could it provide a breakthrough in cancer detection? It sounds impressive, but there's a catch: many cancers are destined to be self-limited. We die *with* them, not *of* them. What if such an immersive diagnostic engine led to millions of unnecessary biopsies? In medicine, there are cases where early diagnosis can save or prolong life. There are also cases where you'll be worried longer but won't live longer. It's hard to know how much you want to know.

"I'm interested in magnifying human ability," Thrun said, when I asked him about the impact of such systems on human diagnosticians. "Look, did industrial farming eliminate some forms of farming? Abso-

lutely, but it amplified our capacity to produce agricultural goods. Not all of this was good, but it allowed us to feed more people. The industrial revolution amplified the power of human muscle. When you use a phone, you amplify the power of human speech. You cannot shout from New York to California"—Thrun and I were, indeed, speaking across that distance—"and yet this rectangular device in your hand allows the human voice to be transmitted across three thousand miles. Did the phone replace the human voice? No, the phone is an augmentation device. The cognitive revolution will allow computers to amplify the capacity of the human mind in the same manner. Just as machines made human muscles a thousand times stronger, machines will make the human brain a thousand times more powerful." Thrun insists that these deep-learning devices will not replace dermatologists and radiologists. They will *augment* the professionals, offering them expertise and assistance.

G EOFFREY HINTON, a computer scientist at the University of Toronto, speaks less gently about the role that learning machines will play in clinical medicine. Hinton—the great-great-grandson of George Boole, whose Boolean algebra is a keystone of digital computing—has sometimes been called the father of deep learning; it's a topic he's worked on since the mid-nineteen-seventies, and many of his students have become principal architects of the field today.

"I think that if you work as a radiologist you are like Wile E. Coyote in the cartoon," Hinton told me. "You're already over the edge of the cliff, but you haven't yet looked down. There's no ground underneath." Deep-learning systems for breast and heart imaging have already been developed commercially. "It's just completely obvious that in five years deep learning is going to do better than radiologists," he went on. "It *might* be ten years. I said this at a hospital. It did not go down too well."

Hinton's actual words, in that hospital talk, were blunt: "They should stop training radiologists now." When I

brought up the challenge to Angela Lignelli-Dipple, she pointed out that diagnostic radiologists aren't merely engaged in yes-no classification. They're not just locating the embolism that brought on a stroke. They're noticing the small bleed elsewhere that might make it disastrous to use a clot-busting drug; they're picking up on an unexpected, maybe still asymptomatic tumor.

Hinton now qualifies the provocation. "The role of radiologists will evolve from doing perceptual things that could probably be done by a highly trained pigeon to doing far more cognitive things," he told me. His prognosis for the future of automated medicine is based on a simple principle: "Take any old classification problem where you have a lot of data, and it's going to be solved by deep learning. There's going to be *thousands* of applications of deep learning." He wants to use learning algorithms to read X-rays, CT scans, and MRIs of every variety—and that's just what he considers the near-term prospects. In the future, he said, "learning algorithms will make pathological diagnoses." They might read Pap smears, listen to heart sounds, or predict relapses in psychiatric patients.

We discussed the black-box problem. Although computer scientists are working on it, Hinton acknowledged that the challenge of opening the black box, of trying to find out exactly what these powerful learning systems know and how they know it, was "far from trivial—don't believe anyone who says that it is." Still, it was a problem he thought we could live with. "Imagine pitting a baseball player against a physicist in a contest to determine where a ball might land," he said. "The baseball player, who's thrown a ball over and over again a million times, might not know any equations but knows exactly how high the ball will rise, the velocity it will reach, and where it will come down to the ground. The physicist can write equations to determine the same thing. But, ultimately, both come to the identical point."

I recalled the disappointing results from older generations of computer-assisted detection and diagnosis in mammography. Any new system would need
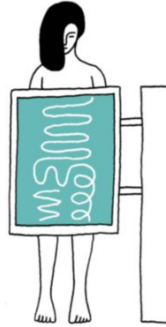
to be evaluated through rigorous clinical trials, Hinton conceded. Yet the new intelligent systems, he stressed, are designed to learn from their mistakes—to improve over time. "We could build in a system that would take every missed diagnosis—a patient who developed lung cancer eventually—and feed it back to the machine. We could ask, What did you miss here? Could you refine the diagnosis? There's no such system for a human radiologist. If you miss something, and a patient develops cancer five years later, there's no systematic routine that tells you how to correct yourself. But you could build in a system to teach the computer to achieve exactly that."

Some of the most ambitious versions of diagnostic machine-learning algorithms seek to integrate natural-language processing (permitting them to read a patient's medical records) and an encyclopedic knowledge of medical conditions gleaned from textbooks, journals, and medical databases. Both I.B.M.'s Watson Health, headquartered in Cambridge, Massachusetts, and DeepMind, in London, hope to create such comprehensive systems. I watched some of these systems operate in pilot demonstrations, but many of their features, especially the deep-learning components, are still in development.

Hinton is passionate about the future of deep-learning diagnosis, in part, because of his own experience. As he was developing such algorithms, his wife was found to have advanced pancreatic cancer. His son was diagnosed with a malignant melanoma, but then the biopsy showed that the lesion was a basal-cell carcinoma, a far less serious kind of cancer. "There's much more to learn here," Hinton said, letting out a small sigh. "Early and accurate diagnosis is not a trivial problem. We can do better. Why not let machines help us?"

O N AN ICY March morning, a few days after my conversations with Thrun and Hinton, I went to Columbia University's dermatology clinic, on Fifty-first Street in Manhattan. Lindsey Bordone, the attending physician, was scheduled to see forty-nine patients that day. By ten o'clock, the waiting room was filled with people. (Identifying details have been changed.) A bearded man, about sixty years old, sat in the corner concealing a rash on his neck with a woollen scarf. An anxious couple huddled over the *Times*.

Bordone saw her patients in rapid succession. In a fluorescent-lit room in the back, a nurse sat facing a computer and gave a one-sentence summary—"fifty years old with no prior history and new suspicious spot on the skin"—and then Bordone rushed into the examining room, her blond hair flying behind her.

A young man in his thirties had a scaly red rash on his face. As Bordone examined him, the skin flaked and fell off his nose. Bordone pulled him into the light and looked at the skin carefully, and then focussed her handheld dermoscope on it.

"Do you have dandruff in your hair?" she asked.

The man looked confused. "Sure," he said.

"Well, this is facial dandruff," Bordone told him. "It's a particularly bad case. But the question is why it appeared now, and why it's getting worse. Have you been using some new product in your hair? Is there some unusual stress in your family?"

"There's definitely been some stress," he said. He had lost his job recently, and was dealing with the financial repercussions.

"Keep a diary," she advised. "We can determine if there's a link." She wrote a prescription for a steroidal cream, and asked him to return in a month.

In the next room, there was a young paralegal with a spray of itchy bumps on his scalp. He winced as Bordone felt his scalp. "Seborrheic dermatitis," she said, concluding her exam.

The woman in another room had undressed and donned a hospital gown. In the past, she had been diagnosed with a melanoma, and she was diligent about getting preventive exams. Bordone pored over her skin, freckle by freckle. It took her twenty minutes, but she was thorough and comprehensive, running her fingers over the landscape of moles and skin tags and calling out diagnoses as she moved. There were nevi and keratoses, but no melanomas or carcinomas.

"Looks all good," she said cheerfully at the end. The woman sighed in relief.

And so it went: Bordone came; she saw; she diagnosed. Far from Hinton's coyote, she seemed like a somewhat manic roadrunner, trying to keep pace with the succession of cases that treadmilled beneath her. As she wrote her notes in the back room, I asked her about Thrun's vision for diagnosis: an iPhone pic e-mailed to a powerful off-site network marshalling undoubted but inscrutable expertise. A dermatologist in full-time practice, such as Bordone, will see about two hundred thousand cases during her lifetime. The Stanford machine's algorithm ingested nearly a hundred and thirty thousand cases in about three months. And, whereas each new dermatology resident needs to start from scratch, Thrun's algorithm keeps ingesting, growing, and learning.

Bordone shrugged. "If it helps me make decisions with greater accuracy, I'd welcome it," she said. "Some of my patients could take pictures of their skin problems before seeing me, and it would increase the reach of my clinic."

That sounded like a reasonable response, and I remembered Thrun's reassuring remarks about augmentation. But, as machines learn more and more, will humans learn less and less? It's the perennial anxiety of the parent whose child has a spell-check function on her phone: what if the child stops learning how to spell? The phenomenon has been called "automation bias." When cars gain automated driver assistance, drivers may become less alert, and something similar may happen in medicine. Maybe Bordone was a lone John Henry in a world where the steam drills were about to come online. But it was impossible to miss how her own concentration never wavered and how seriously she took every skin tag and mole that she ran her fingers over. Would that continue to be true if she partnered with a machine?

I noticed other patterns in Bordone's interactions with her patients. For one thing, they almost always left feeling better. They had been touched and scrutinized; a conversation took place. Even the naming of lesions—"nevus," "keratosis"—was an emollient: there was something deeply reassuring about the process. The woman who'd had the skin exam left looking fresh and unburdened, her anxiety exfoliated.

There was more. The diagnostic moment, as the Brazilian researchers might have guessed, came to Bordone in a flash of recognition. As she called out "dermatitis" or "eczema," it was as if she were identifying a rhinoceros: you could almost see the pyramid of neurons in the lower posterior of her brain spark as she recognized the pattern. But the visit did not end there. In almost every case, Bordone spent the bulk of her time investigating causes. Why had the symptoms appeared? Was it stress? A new shampoo? Had someone changed the chlorine in the pool? Why now?

The most powerful element in these clinical encounters, I realized, was not knowing that or knowing how—not mastering the facts of the case, or perceiving the patterns they formed. It lay in yet a third realm of knowledge: knowing *why*.

"Pretty good. The ending was a bit predictable."

• • •

E XPLANATIONS RUN SHALLOW and deep. You have a red blister on your finger because you touched a hot iron; you have a red blister on your finger because the burn excited an inflammatory cascade of prostaglandins and cytokines, in a regulated process that we still understand only imperfectly. Knowing why—*asking* why—is our conduit to every kind of explanation, and explanation, increasingly, is what powers medical advances. Hinton spoke about baseball players and physicists. Diagnosticians, artificial or human, would be the baseball players—proficient but opaque. Medical researchers would be the physicists, as removed from the clinical field as theorists are from the baseball field, but with a desire to know "why." It's a convenient division of responsibilities— yet might it represent a loss?

"A deep-learning system doesn't have any explanatory power," as Hinton put it flatly. A black box cannot investigate cause. Indeed, he said, "the more powerful the deep-learning system becomes, the more opaque it can become. As more features are extracted, the diagnosis becomes increasingly accurate. Why these features were extracted out of millions of other features, however, remains an unanswerable question." The algorithm can solve a case. It cannot build a case.

Yet in my own field, oncology, I couldn't help noticing how often advances were made by skilled practitioners who were also curious and penetrating researchers. Indeed, for the past few decades, ambitious doctors have strived to be at once baseball players *and* physicists: they've tried to use diagnostic acumen to understand the pathophysiology of disease. Why does an asymmetrical border of a skin lesion predict a melanoma? Why do some melanomas regress spontaneously, and why do patches of white skin appear in some of these cases? As it happens, this observation, made by diagnosticians in the clinic, was eventually linked to the creation of some of the most potent immunological medicines used clinically today. (The whitening skin, it turned out, was the result of an immune reaction that was also turning against the melanoma.) The chain of discovery can begin in the clinic. If more and more clinical practice were relegated to increasingly opaque learning machines, if the daily, spontaneous intimacy between implicit and explicit forms of knowledge—knowing how, knowing that, knowing why—began to fade, is it possible that we'd get better at doing what we do but less able to reconceive what we *ought* to be doing, to think outside the algorithmic black box?

I spoke to David Bickers, the chair of dermatology at Columbia, about our automated future. "Believe me, I've tried to understand all the ramifications of Thrun's paper," he said. "I don't understand the math behind it, but I do know that such algorithms might change the practice of dermatology. Will dermatologists be out of jobs? I don't think so, but I think we have to think hard about how to integrate these programs into our practice. How will we pay for them? What are the legal liabilities if the machine makes the wrong prediction? And will it diminish our practice, or our self-image as diagnosticians, to rely on such algorithms? Instead of doctors, will we end up training a generation of technicians?"

He checked the time. A patient was waiting to see him, and he got up to leave. "I've spent my life as a diagnostician and a scientist," he said. "I know how much a patient relies on my capacity to tell a malignant lesion from a benign one. I also know that medical knowledge emerges from diagnosis."

The word "diagnosis," he reminded me, comes from the Greek for "knowing apart." Machine-learning algorithms will only become better at such knowing apart—at partitioning, at distinguishing moles from melanomas. But knowing, in all its dimensions, transcends those task-focussed algorithms. In the realm of medicine, perhaps the ultimate rewards come from knowing together. ♦