COURTESY OF THE ARTISTS

Intelligent Machines

A philosopher argues that an AI can't be an artist

Creativity is, and always will be, a human endeavor.

by Sean Dorrance Kelly   February 21, 2019

Portrait of Edmond Belamy (2018)
Created with AI algorithms called GANs by Parisian art collective Obvious, sold for $432,500.

O n March 31, 1913, in the Great Hall of the Musikverein concert house in Vienna, a riot broke out in the middle of a performance of an orchestral song by Alban Berg. Chaos descended. Furniture was broken. Police arrested the concert's organizer for punching Oscar Straus, a little-remembered composer of operettas. Later, at the trial, Straus quipped about the audience's frustration. The punch, he insisted, was the most harmonious sound of the entire evening. History has rendered a different verdict: the concert's conductor, Arnold Schoenberg, has gone down as perhaps the most creative and influential composer of the 20th century.

You may not enjoy Schoenberg's dissonant music, which rejects conventional tonality to arrange the 12 notes of the scale according to rules that don't let any predominate. But he changed what humans understand music to be. This is what makes him a genuinely creative and innovative artist. Schoenberg's techniques are now integrated seamlessly into everything from film scores and Broadway musicals to the jazz solos of Miles Davis and Ornette Coleman.

Creativity is among the most mysterious and impressive achievements of human existence. But what is it?
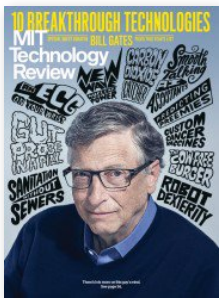
This story is part of our March/April 2019 Issue

See the rest of the issue

Subscribe

Sign up for **The Algorithm**
Artificial intelligence, demystified

| Your email | | **Sign Up** |

Stay updated on MIT Technology Review initiatives and events? ○ Yes ○ No

Creativity is not just novelty. A toddler at the piano may hit a novel sequence of notes, but they're not, in any meaningful sense, creative. Also, creativity is bounded by history: what counts as creative inspiration in one period or place might be disregarded as ridiculous, stupid, or crazy in another. A community has to accept ideas as good for them to count as creative.

As in Schoenberg's case, or that of any number of other modern artists, that acceptance need not be universal. It might, indeed, not come for years— sometimes creativity is mistakenly dismissed for generations. But unless an innovation is eventually accepted by some community of practice, it makes little sense to speak of it as creative.

Advances in artificial intelligence have led many to speculate that human beings will soon be replaced by machines in every domain, including that of creativity. Ray Kurzweil, a futurist, predicts that by 2029 we will have produced an AI that can pass for an average educated human being. Nick

Bostrom, an Oxford philosopher, is more circumspect. He does not give a date but suggests that philosophers and mathematicians defer work on fundamental questions to "superintelligent" successors, which he defines as having "intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest."

Both believe that once human-level intelligence is produced in machines, there will be a burst of progress—what Kurzweil calls the "singularity" and Bostrom an "intelligence explosion"—in which machines will very quickly supersede us by massive measures in every domain. This will occur, they argue, because superhuman achievement is the same as ordinary human achievement except that all the relevant computations are performed much more quickly, in what Bostrom dubs "speed superintelligence."

So what about the highest level of human achievement—creative innovation? Are our most creative artists and thinkers about to be massively surpassed by machines?

No.

Human creative achievement, because of the way it is socially embedded, will not succumb to advances in artificial intelligence. To say otherwise is to misunderstand both what human beings are and what our creativity amounts to.

This claim is not absolute: it depends on the norms that we allow to govern our culture and our expectations of technology. Human beings have, in the past, attributed great power and genius even to lifeless totems. It is entirely possible that we will come to treat artificially intelligent machines as so vastly superior to us that we will naturally attribute creativity to them. Should that happen, it will not be because machines have outstripped us. It will be because we will have denigrated ourselves.

Human creative achievement, because of the way it is socially embedded, will not succumb to advances in artificial intelligence.

Also, I am primarily talking about machine advances of the sort seen recently with the current deep-learning paradigm, as well as its computational successors. Other paradigms have governed AI research in the past. These have already failed to realize their promise. Still other paradigms may come in the future, but if we speculate that some notional future AI whose features we cannot meaningfully describe will accomplish wondrous things, that is mythmaking, not reasoned argument about the possibilities of technology.

Creative achievement operates differently in different domains. I cannot offer a complete taxonomy of the different kinds of creativity here, so to make the point I will sketch an argument involving three quite different examples: music, games, and mathematics.

Music to my ears

In Imaginary Landscape (2018)
Nao Tokui uses a machine-learning algorithm to create panoramas from images found in Google Street View and complements them with soundscapes created with artificial neural networks.
COURTESY OF THE ARITST

Can we imagine a machine of such superhuman creative ability that it brings about changes in what we understand music to be, as Schoenberg did?

That's what I claim a machine cannot do. Let's see why.

Computer music composition systems have existed for quite some time. In 1965, at the age of 17, Kurzweil himself, using a precursor of the pattern recognition systems that characterize deep-learning algorithms today, programmed a computer to compose recognizable music. Variants of this technique are used today. Deep-learning algorithms have been able to take as input a bunch of Bach chorales, for instance, and compose music so characteristic of Bach's style that it fools even experts into thinking it is original. This is mimicry. It is what an artist does as an apprentice: copy and perfect the style of others instead of working in an authentic, original voice. It is not the kind of musical creativity that we associate with Bach, never mind with Schoenberg's radical innovation.

So what do we say? Could there be a machine that, like Schoenberg, invents a whole new way of making music? Of course we can imagine, and even make, such a machine. Given an algorithm that modifies its own compositional rules, we could easily produce a machine that makes music as different from what we now consider good music as Schoenberg did then.

But this is where it gets complicated.

We count Schoenberg as a creative innovator not just because he managed to create a new way of composing music but because people could see in it a vision of what the world should be. Schoenberg's vision involved the spare, clean, efficient minimalism of modernity. His innovation was not just to find a new algorithm for composing music; it was to find a way of thinking about what music *is* that allows it to speak to *what is needed now*.

Some might argue that I have raised the bar too high. Am I arguing, they will ask, that a machine needs some mystic, unmeasurable sense of what is socially necessary in order to count as creative? I am not—for two reasons.

First, remember that in proposing a new, mathematical technique for musical composition, Schoenberg changed our understanding of what music is. It is only creativity of this tradition-defying sort that requires some kind of social sensitivity. Had listeners not experienced his technique as capturing the anti-traditionalism at the heart of the radical modernity emerging in early-20th-century Vienna, they might not have heard it as something of aesthetic worth. The point here is that radical creativity is not an "accelerated" version of quotidian creativity. Schoenberg's achievement is not a faster or better version of the type of creativity demonstrated by Oscar Straus or some other average composer: it's fundamentally different in kind.

Second, my argument is not that the creator's responsiveness to social necessity must be conscious for the work to meet the standards of genius. I am arguing instead that *we must be able to interpret the work as responding that way*. It would be a mistake to interpret a machine's composition as part of such a vision of the world. The argument for this is simple.

Claims like Kurzweil's that machines can reach human-level intelligence assume that to have a human mind is just to have a human brain that follows some set of computational algorithms—a view called computationalism. But though algorithms can have moral implications, they are not themselves moral agents. We can't count the monkey at a typewriter who accidentally types out *Othello* as a great creative playwright. If there is greatness in the product, it is only an accident. We may be able to see a machine's product as great, but if we know that the output is merely the result of some arbitrary act or algorithmic formalism, we cannot accept it as the expression of a vision for human good.

For this reason, it seems to me, nothing but another human being can properly be understood as a genuinely creative artist. Perhaps AI will someday proceed beyond its computationalist formalism, but that would require a leap that is unimaginable at the moment. We wouldn't just be looking for new algorithms or procedures that simulate human activity; we would be looking for new materials that are the basis of being human.

A molecule-for-molecule duplicate of a human being would be human in the relevant way. But we already have a way of producing such a being: it takes about nine months. At the moment, a machine can only do something much less interesting than what a person can do. It can create music in the style of Bach, for instance—perhaps even music that some experts think is better than Bach's own. But that is only because its music can be judged against a preexisting standard. What a machine cannot do is bring about changes in our standards for judging the quality of music or of understanding what music is or is not.

This is not to deny that creative artists use whatever tools they have at their disposal, and that those tools shape the sort of art they make. The trumpet helped Davis and Coleman realize their creativity. But the trumpet is not, itself, creative. Artificial-intelligence algorithms are more like musical instruments than they are like people. Taryn Southern, a former *American Idol* contestant, recently released an album where the percussion, melodies, and chords were algorithmically generated, though she wrote the lyrics and repeatedly tweaked the instrumentation algorithm until it delivered the results she wanted. In the early 1990s, David Bowie did it the other way around: he wrote the music and used a Mac app called Verbalizer to pseudo-

randomly recombine sentences into lyrics. Just like previous tools of the music industry—from recording devices to synthesizers to samplers and loopers—new AI tools work by stimulating and channeling the creative abilities of the human artist (and reflect the limitations of those abilities).

Games without frontiers

Much has been written about the achievements of deep-learning systems that are now the best Go players in the world. AlphaGo and its variants have strong claims to having created a whole new way of playing the game. They have taught human experts that opening moves long thought to be ill-conceived can lead to victory. The program plays in a style that experts describe as strange and alien. "They're how I imagine games from far in the future," Shi Yue, a top Go player, said of AlphaGo's play. The algorithm seems to be genuinely creative.

In some important sense it is. Game-playing, though, is different from composing music or writing a novel: in games there is an objective measure of success. We know we have something to learn from AlphaGo because we see it win.

But that is also what makes Go a "toy domain," a simplified case that says only limited things about the world.

Anna Ridler's The Fall of the House of Usher (2017)
A 12-minute animation based on Watson and Webber's 1928 silent film. Ridler created the stills using three separate neural nets: one trained on her drawings, a second trained on drawings made of the results of the first net, and a third trained on drawings made of the results of the second.
COURTESY OF THE ARITST

The most fundamental sort of human creativity changes our understanding of ourselves because it changes our understanding of what we count as good. For the game of Go, by contrast, the nature of goodness is simply not up for grabs: a Go strategy is good if and only if it wins. Human life does not generally have this feature: there is no objective measure of success in the

highest realms of achievement. <mark>Certainly not in art, literature, music, philosophy, or politics. Nor, for that matter, in the development of new technologies.</mark>

In various toy domains, machines may be able to teach us about a certain very constrained form of creativity. But the domain's rules are pre-formed; the system can succeed only because it learns to play well within these constraints. Human culture and human existence are much more interesting than this. There are norms for how human beings act, of course. But creativity in the genuine sense is the ability to change those norms in some important human domain. <mark>Success in toy domains is no indication that creativity of this more fundamental sort is achievable.</mark>

It's a knockout

A skeptic might contend that the argument works only because I'm contrasting games with *artistic* genius. There are other paradigms of creativity in the scientific and mathematical realm. In these realms, the question isn't about a vision of the world. It is about the way things actually are.

Might a machine come up with mathematical proofs so far beyond us that we simply have to defer to its creative genius?

No.

Computers have already assisted with notable mathematical achievements. But their contributions haven't been particularly creative. Take the first major theorem proved using a computer: the four-color theorem, which states that any flat map can be colored with at most four colors in such a way that no two adjacent "countries" end up with the same one (it also applies to countries on the surface of a globe).

Nearly a half-century ago, in 1976, Kenneth Appel and Wolfgang Haken at the University of Illinois published a computer-assisted proof of this theorem. The computer performed billions of calculations, checking thousands of different types of maps—so many that it was (and remains) logistically unfeasible for humans to verify that each possibility accorded with the computer's view. Since then, computers have assisted in a wide range of new proofs.
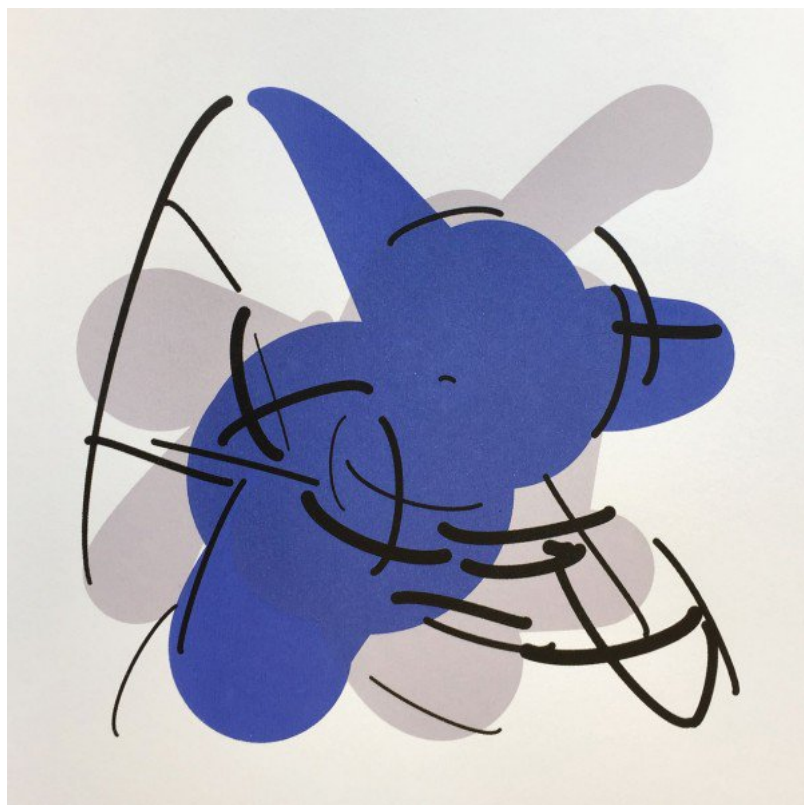
Electric Fan (2018)
Tom White uses "perception engines," algorithms that distill the data collected from thousands of photographs of common objects, to synthesize abstract shapes. He then tests and refines the results until they are recognizable by the system.
ELECTRIC FAN, COURTESY OF TOM WHITE, MAS '98, DRIB.NET

But the supercomputer is not doing anything creative by checking a huge number of cases. Instead, it is doing something boring a huge number of times. This seems like almost the opposite of creativity. Furthermore, it is so far from the kind of *understanding* we normally think a mathematical proof should offer that some experts don't consider these computer-assisted strategies mathematical proofs at all. As Thomas Tymoczko, a philosopher of mathematics, has argued, if we can't even verify whether the proof is correct, then all we are really doing is trusting in a potentially error-prone computational process.

Even supposing we do trust the results, however, computer-assisted proofs are something like the analogue of computer-assisted composition. If they give us a worthwhile product, it is mostly because of the contribution of the human being. But some experts have argued that artificial intelligence will be able to achieve more than this. Let us suppose, then, that we have the ultimate: a self-reliant machine that proves new theorems all on its own.

Could a machine like this massively surpass us in mathematical creativity, as Kurzweil and Bostrom argue? Suppose, for instance, that an AI comes up with a resolution to some extremely important and difficult open problem in mathematics.

The capacity for genuine creativity, the kind of creativity that updates our understanding of the nature of being, is at the ground of what it is to be human.

There are two possibilities. The first is that the proof is extremely clever, and when experts in the field go over it they discover that it is correct. In this case, the AI that discovered the proof would be applauded. The machine itself might even be considered to be a creative mathematician. But such a

machine would not be evidence of the singularity; it would not so outstrip us in creativity that we couldn't even understand what it was doing. Even if it had this kind of human-level creativity, it wouldn't lead inevitably to the realm of the superhuman.

Some mathematicians are like musical virtuosos: they are distinguished by their fluency in an existing idiom. But geniuses like Srinivasa Ramanujan, Emmy Noether, and Alexander Grothendieck arguably reshaped mathematics just as Schoenberg reshaped music. Their achievements were not simply proofs of long-standing hypotheses but new and unexpected forms of reasoning, which took hold not only on the strength of their logic but also on their ability to convince other mathematicians of the significance of their innovations. A notional AI that comes up with a clever proof to a problem that has long befuddled human mathematicians is akin to AlphaGo and its variants: impressive, but nothing like Schoenberg.

That brings us to the other option. Suppose the best and brightest deep-learning algorithm is set loose and after some time says, "I've found a proof of a fundamentally new theorem, but it's too complicated for even your best mathematicians to understand."

This isn't actually possible. A proof that not even the best mathematicians can understand doesn't really count as a proof. Proving something implies that you are proving it to *someone*. Just as a musician has to persuade her audience to accept her aesthetic concept of what is good music, a mathematician has to persuade other mathematicians that there are good reasons to believe her vision of the truth. To count as a valid proof in mathematics, a claim must be understandable and endorsable by some independent set of experts who are in a good position to understand it. If the experts who should be able to understand the proof can't, then the community refuses to endorse it as a proof.

For this reason, mathematics is more like music than one might have thought. A machine could not surpass us massively in creativity because either its achievement would be understandable, in which case it would not massively surpass us, or it would not be understandable, in which case we could not count it as making any creative advance at all.

The eye of the beholder

Engineering and applied science are, in a way, somewhere between these examples. There is something like an objective, external measure of success. You can't "win" at bridge building or medicine the way you can at chess, but one can see whether the bridge falls down or the virus is eliminated. These objective criteria come into play only once the domain is fairly well specified: coming up with strong, lightweight materials, say, or drugs that combat particular diseases. An AI might help in drug discovery by, in effect, doing the same thing as the AI that composed what sounded like a well-executed Bach cantata or came up with a brilliant Go strategy. Like a microscope, telescope, or calculator, such an AI is properly understood as a tool that enables human discovery—not as an autonomous creative agent.

It's worth thinking about the theory of special relativity here. Albert Einstein is remembered as the "discoverer" of relativity—but not because he was the first to come up with equations that better describe the structure of space

and time. George Fitzgerald, Hendrik Lorentz, and Henri Poincaré, among others, had written down those equations before Einstein. He is acclaimed as the theory's discoverer because he had an original, remarkable, and true understanding of what the equations *meant* and could convey that understanding to others.

For a machine to do physics that is in any sense comparable to Einstein's in creativity, it must be able to persuade other physicists of the worth of its ideas at least as well as he did. Which is to say, we would have to be able to accept its proposals as aiming to *communicate their own validity to us*. Should such a machine ever come into being, as in the parable of Pinocchio, we would have to treat it as we would a human being. That means, among other things, we would have to attribute to it not only intelligence but whatever dignity and moral worth is appropriate to human beings as well. We are a long way off from this scenario, it seems to me, and there is no reason to think the current computationalist paradigm of artificial intelligence—in its deep-learning form or any other—will ever move us closer to it.

Creativity is one of the defining features of human beings. The capacity for genuine creativity, the kind of creativity that updates our understanding of the nature of being, that changes the way we understand what it is to be beautiful or good or true—that capacity is at the ground of what it is to be human. But this kind of creativity depends upon our valuing it, and caring for it, as such. As the writer Brian Christian has pointed out, human beings are starting to act less like beings who value creativity as one of our highest possibilities, and more like machines themselves.

How many people today have jobs that require them to follow a predetermined script for their conversations? How little of what we know as real, authentic, creative, and open-ended human conversation is left in this eviscerated charade? How much is it like, instead, the kind of rule-following that a machine can do? And how many of us, insofar as we allow ourselves to be drawn into these kinds of scripted performances, are eviscerated as well? How much of our day do we allow to be filled with effectively machine-like activities—filling out computerized forms and questionnaires, responding to click-bait that works on our basest, most animal-like impulses, playing games that are designed to optimize our addictive response?

We are in danger of this confusion in some of the deepest domains of human achievement as well. If we allow ourselves to say that machine proofs we cannot understand are genuine "proofs," for example, ceding social authority to machines, we will be treating the achievements of mathematics as if they required no human understanding at all. We will be taking one of our highest forms of creativity and intelligence and reducing it to a single bit of information: yes or no.

The Butcher's Son (2018)
Mario Klingemann used two GANs, one trained on a data set of human poses and one trained on pornography, to render thousands of composite images. After evaluating each for pose and detail, he chose one to refine into the finished work.
COURTESY OF THE ARTIST

Even if we had that information, it would be of little value to us without some understanding of the reasons underlying it. We must not lose sight of the essential character of reasoning, which is at the foundation of what mathematics is.

So too with art and music and philosophy and literature. If we allow ourselves to slip in this way, to treat machine "creativity" as a substitute for our own, then machines will indeed come to seem incomprehensibly superior to us. But that is because we will have lost track of the fundamental role that creativity plays in being human.

Sean Dorrance Kelly is a philosophy professor at Harvard and coauthor of the New York Times best-selling book All Things Shining.

Learn from the humans leading the way in deep learning at EmTech Next. Register Today!
June 11-12, 2019
Cambridge, MA
Register now

Related Video

More videos