Artificial intelligence can now emulate human behaviors – soon it will be dangerously good
5 April 2019, by Ana Santos Rutschman

Impersonation technologies

Google's program analyzed the notes in 306 of Bach's musical works, finding relationships between the melody and the notes that provided the harmony. Because Bach followed strict rules of composition, the program was effectively learning those rules, so it could apply them when users provided their own notes.

The Bach app itself is new, but the underlying technology is not. Algorithms trained to recognize patterns and make probabilistic decisions have existed for a long time. Some of these algorithms are so complex that people don't always understand how they make decisions or produce a particular outcome.

AI systems are not perfect – many of them rely on data that aren't representative of the whole population, or that are influenced by human biases. It's not entirely clear who might be legally responsible when an AI system makes an error or causes a problem.

Now, though, artificial intelligence technologies are getting advanced enough to be able to approximate individuals' writing or speaking style, and even facial expressions. This isn't always bad: A fairly simple AI gave Stephen Hawking the ability to communicate more efficiently with others by predicting the words he would use the most.

More complex programs that mimic human voices assist people with disabilities – but can also be used to deceive listeners. For example, the makers of Lyrebird, a voice-mimicking program, have released a simulated conversation between Barack Obama, Donald Trump and Hillary Clinton. It may sound real, but that exchange never happened.

When artificial intelligence systems start getting creative, they can create great things – and scary ones. Take, for instance, an AI program that let web users compose music along with a virtual Johann Sebastian Bach by entering notes into a program that generates Bach-like harmonies to match them.

Run by Google, the app drew great praise for being groundbreaking and fun to play with. It also attracted criticism, and raised concerns about AI's dangers.

My study of how emerging technologies affect people's lives has taught me that the problems go beyond the admittedly large concern about whether algorithms can really create music or art in general. Some complaints seemed small, but really weren't, like observations that Google's AI was breaking basic rules of music composition.

In fact, efforts to have computers mimic the behavior of actual people can be confusing and potentially harmful.
From good to bad

In February 2019, nonprofit company OpenAI created a program that generates text that is virtually indistinguishable from text written by people. It can "write" a speech in the style of John F. Kennedy, J.R.R. Tolkien in "The Lord of the Rings" or a student writing a school assignment about the U.S. Civil War.

The text generated by OpenAI's software is so believable that the company has chosen not to release the program itself.

Similar technologies can simulate photos and videos. In early 2018, for instance, actor and filmmaker Jordan Peele created a video that appeared to show former U.S. President Barack Obama saying things Obama never actually said to warn the public about the dangers posed by these technologies.

In early 2019, a fake nude photo of U.S. Rep. Alexandria Ocasio-Cortez circulated online. Fabricated videos, often called "deepfakes," are expected to be increasingly used in election campaigns.

Members of Congress have started to look into this issue ahead of the 2020 election. The U.S. Defense Department is teaching the public how to spot doctored videos and audio. News organizations like Reuters are beginning to train journalists to spot deepfakes.

But, in my view, an even bigger concern remains: Users might not be able to learn fast enough to distinguish fake content as AI technology becomes more sophisticated. For instance, as the public is beginning to become aware of deepfakes, AI is already being used for even more advanced deceptions. There are now programs that can generate fake faces and fake digital fingerprints, effectively creating the information needed to fabricate an entire person – at least in corporate or government records.

Machines keep learning

At the moment, there are enough potential errors in these technologies to give people a chance of detecting digital fabrications. Google's Bach composer made some mistakes an expert could detect. For example, when I tried it, the program allowed me to enter parallel fifths, a music interval that Bach studiously avoided. The app also broke musical rules of counterpoint by harmonizing melodies in the wrong key. Similarly, OpenAI's text-generating program occasionally wrote phrases like "fires happening under water" that made no sense in their contexts.

As developers work on their creations, these mistakes will become rarer. Effectively, AI technologies will evolve and learn. The improved performance has the potential to bring many social benefits – including better health care, as AI programs help democratize the practice of medicine.

Giving researchers and companies freedom to explore, in order to seek these positive achievements from AI systems, means opening up the risk of developing more advanced ways to create deception and other social problems. Severely limiting AI research could curb that progress. But giving beneficial technologies room to grow comes at no small cost – and the potential for misuse, whether to make inaccurate "Bach-like" music or to deceive millions, is likely to grow in ways people can't yet anticipate.

This article is republished from The Conversation under a Creative Commons license. Read the