

## Augmenting Humans: IBM's Project Debater AI gives human debating teams a hand at Cambridge

Two teams, sparring on a controversial topic — whether artificial intelligence would bring more harm than good — the Thursday night debate in front of 300-strong audience seemed rather typical for Cambridge Union, the world's oldest debating society.

**London 26th November 2019** - Two teams, sparring on a controversial topic — whether artificial intelligence would bring more harm than good — the Thursday night debate in front of 300-strong audience seemed rather typical for Cambridge Union, the world's oldest debating society.

Except it wasn't.

It was the first-of-its-kind for this ornate, 150-year-old Debating Chamber that once hosted British prime minister Margaret Thatcher, US President Theodore Roosevelt, the current Dalai Lama, theoretical physicist Stephen Hawking — and even, to the delight of the TV series Baywatch fans, American actress Pamela Anderson.

Because this time, the humans taking part in the debate were augmented.

“AI will not be able to make morally correct decisions because morality is unique to humans.” The soft, pleasant female voice was strangely omnipresent, like fog enveloping the public. And then, after a pause, the voice spoke again: “AI has a lower rate of error than humans... It will be a great advantage as it will free up more time from having to do mundane and repetitive tasks.” It seemed to have completely changed its stance.

Under the Chamber's wooden balcony decorated with the crests of all of the Cambridge colleges, rivals sat in burgundy red leather seats on either side of a three-step podium, and the society's President — in a larger, throne-like chair in its center. Typically, the audience's attention would gravitate towards the President as the highest point in the room. Now, though, all eyes were glued on the black 2m-tall monolithic block with blue lights in front of her.

When the voice spoke, the blue lights flickered as if the voice was coming from the machine. An AI innovation — the creation of IBM Research, more than six years in the making, dubbed [Project Debater](#). The AI technology gave two teams of humans, each made up of a professor and an expert debater, a hand (not that it has hands) — in a debate questioning its very own purpose. “I'm really intrigued and excited to be working with a machine,” one of the participants, Sharmila Parmanand, an expert debater on the against-AI

team, said ahead of the event— eyeing the stage with the black monolith.

And, unlike any traditional debates, this time the audience also was part of the discussion — as the AI sifted through more than a thousand arguments submitted by the public in advance, pro and against AI in society, to develop its arguments.

But it was not just about the debate. Tonight’s demonstration of IBM’s technology was also very different from its previous showcases. Until now, the AI competed against a human —like champion debater Harish Natarajan in February 2019 at [IBM’s Think conference](#). Natarajan made a return at tonight’s event.

Remembering his first encounter with the machine, he said that “after about a minute of the strangeness of debating against AI, it became just another debate.” [\*\(read in-depth Q&A with Natarajan here\)\*](#)

Back then, though, it was mostly about the AI’s listening comprehension skills and the ability to construct a meaningful response to an argument, using 400 million entries in its database, drawn from newspaper and magazine articles. Those skills were already way beyond the abilities of a typical digital assistant in your smart home or mobile device, which can listen to one-sentence commands like “book a place at a restaurant” and execute them.

Instead, the AI inside Project Debater is as if “you crowd-sourced thousands of pro and con restaurant reviews and had them summarized in a few paragraphs,” says one of the scientists behind the research, IBM computer scientist Dan Lahav. He himself is a renown debater, and was crowned the best speaker in the world in 2018.

Back in February, despite the media frenzy around Natarajanbeating the software, the experiment was not about whether the AI could outwit the human or vice versa. It was about showing the very different strengths of both.

“Humans are significantly better at understanding the audience and getting context based upon our cultures and traditions, but machines have the ability to just go through large amounts of data and use it to enrich our knowledge, to make much more specific decisions, to make us confound the other side and remove some of the biases that we as humans have,” says Lahav.

Tonight’s debate was showing exactly that as well — and thus, that an intelligent machine can work alongside people — and assist them.

“In 2016, Project Debater was at a level of a toddler, in 2019, we’ve reached university level, ” says Noam Slonim, the principal investigator of the project and IBM’s distinguished engineer. And, he adds, “Project Debater can help humans in many ways, including by reducing inherent human bias — by presenting both

sides of a controversial topic.”

Indeed, helping humans navigate their messy existence has always been the primary role of the most useful tools we’ve ever built. As far as AI goes, Project Debater took IBM’s own achievements of the Deep Blue supercomputer defeating in 1997 the then-top world chess player Garry Kasparov and IBM Watson outwitting two champions of US television quiz show Jeopardy! in 2011 to a whole new level. “If you’re playing chess, you have a very clear winning function — once you’ve done checkmate, you’ll know exactly how it looks on the board. And if you make a move that no one understands, it may still pay off 10 or 15 moves down the line of the game,” says Lahav. “In debating, everything is very subjective. There are no assigned scores, and if you’re going to make a move that no one from the audience understands, if the machine is incomprehensible, no one will be able to follow.”

Project Debater “comes to life” as a monolith with the voice of female actress Eliza Foss. Its role is simply to be a visual aide, so that the audience and the debaters can pin their eyes onto something when the AI speaks.

“Embodying AI helps us to reason about it, we are not used to disembodied intelligences, but in reality it is a disembodied intelligence and we should think about it differently from other humans,” says Neil Lawrence, professor of computer science at Cambridge University and one of the debaters participating in the event. “It’s a trade-off between increasing our comfort levels as humans and correctly representing technologies.”

The evening started with the Project Debater making two opening statements: first, arguing in favor of AI, and then, with a brief pause, against. The software was a player on both teams, and each opposing team had to come up straight away with a rebuttal. But what was unique is how the machine actually put together these opening statements — coherent, argumentative and analytical speeches lasting about five minutes.

To do so, it didn’t pull the data out of thin air, Isaac Asimov’s robot novels, or its own database — instead, it combed through more than [1,100](#) arguments submitted online prior to the debate and fed to it at the start of the event. These were short statements of no more than 36 words arguing either for or against artificial intelligence, such as:

*As machines’ ability to sense, learn, interact naturally and act autonomously increases, they will blur the distinction between the physical and the digital world. AI systems will interconnect to predict to our human needs and emotions.*

*In 2030, advancing AI will not leave most people better off than they are today, because our global digital mission is not strong enough and not principled enough to assure that no one is left behind.*

The idea was to tweak the traditional debate style where the audience is mute and get the public to express

their views too. This would then make the narrative in the opening speeches much more balanced, compelling those on stage to engage with the opinions of the audience and making the debate much more inclusive.

But the voice of the machine in the Debating Chamber on this chilly November night didn't read these hundreds of statements to the teams and the audience, which would have probably dealt with anyone's worst insomnia. Instead, the IBM researchers applied their Speech by Crowd technology — an AI platform based on the core AI behind Project Debater for crowd-sourcing decision support, able to collect free-text arguments from large audiences and automatically construct persuasive viewpoints to support or contest a topic. In about a minute, the tech analyzed every single entry and assessed its stance and quality, filtering out irrelevant and redundant ones. It then composed a coherent narrative based on the main and most important arguments made by the public — one pro and the other against... well, itself. It effectively summarized the opinions of the public based on the opposing stances.

IBM created a machine able to analyze information and deliver a comprehensible data-driven speech. Human debates have long been one of the most intellectually challenging tasks that exist. Some actions, though, such as understanding whether an argument is for or against, are intuitive to us — but incredibly complex to a machine. Then there's the challenge of judging the argument to be of high enough quality to persuade others. And finally, the AI also has to arrange the arguments in a narrative way understandable by humans.

As the evening progressed, the audience and the human contenders alike seemed to get much more relaxed around the machine. The team arguing against AI started the rebuttals — with Parmanand voicing the arguments, looking just as much at the public in the room and her rivals as she was at the machine.

When Sylvie Delacroix, a professor of law and ethics at the University of Birmingham and a member of the opposing team took to the stage, she started her speech by jokingly calling the Project Debater 'Debbie'. "AI as a tool matters because of the sheer speed it's transforming us — the way we work and the way we date," she said, then asking people in the room trusting technology to choose their partner for life.

After the rebuttals, it was time to deliver the closing statements. Both teams tried to be as persuading as they could. It did somewhat look like the machine was on trial, its fate being decided by a bunch of humans in red leather chairs. Taking turns, the teams spoke, then fell silent.

"A junior doctor now has access to far more information than ever before and can save far more lives than ever before," argued Natarajan of the pro-AI team, and I couldn't help but imagine the monolith trying to hide a smile. "We should believe that AI will do us harm because it's the only way to prevent those harms," argued Lawrence from the opposing team, and at this point in an Asimov novel, the machine would probably gasp, its artificial brain calculating all the possible consequences of a ruling based on these conclusions.

For Lawrence, the demonstration's main goal was to encourage people to think about how they're interacting with AI — how “we can ensure our interactions with machines enhance us rather than control us.” Was the goal achieved? “The jury is out,” he says, smiling.

Project Debater didn't take part in the rebuttals or the closing statements — although it had the ability to do so. For Lahav, the experiment was a success — “a very cool combination that I think demonstrates two things in one: first, it's about augmenting people, and second, the machine also gives them the ability to see what everyone in the room is thinking, so they're going to be able to better engage with that.”

Natarajan agrees — and he's arguably the best person to judge, having already matched wits with Project Debater. “A machine's thought process will differ from that of a human, but that is a feature more than a bug. Machines are able to find statistically valid connections that humans may either miss or ignore. This human failing may be due to our propensity to overuse heuristics, default to biases, or limits in our ability to process information,” he says. On the other hand, “humans have the ability to use common sense, to make decisions that are consistent with deeply held principle of justice and may provide the emotional intelligence to communicate a decision.” A perfect team.

As the debate ended, it was time for the audience to pick a winner. Instead of casting votes, the room suddenly filled with the *whoosh* of hundreds of people getting up at once. The audience had three doors to choose from to go through — a “ayes” door in support of the proposition, a “noes” door in support of the opposition, and “abstain” door for those who were wavering. The narrow majority crowded in front of the noes door — meaning that they voted in favor of AI (*the final tally: 48.17% ayes, 51.22% noes and 0.61% abstention*).

Debates in Cambridge apart though, what could the technology be used for outside the walls of the society's Victorian building? Even before the Cambridge Union was set up in 1815, debating pervaded most human activities. It is present in any job that requires data-taking and making a decision based on it, be it in law or medicine or sales. A company could use the technology to better understand its employees, or the market. “You can think about it as a research engine versus a search engine,” says Lahav. “Imagine in a political discourse, it'll give the ability to elevate the level of discussion, because you would get the sense of what the audience is saying, you'll give them a platform to engage. And lawmakers can be better at seeing the other side because the machine can generate arguments at both sides.”

And then there's the issue of bias. Humans in any debate, be it journalists, political advisors or psychologists, have some degree of personal bias, despite often pledging to be completely impartial. A programmer writing a code can inadvertently introduce bias just as well. But, says Lahav, with thousands of different arguments, introducing systematic bias is very hard. The quality filters are done using a large group of people chosen

from many different platforms.

“I don’t think it is feasible to build a technology that doesn’t introduce any bias, whether it’s the bias of the creator or the bias of the data. What I think is important is to see whether the trend the technology introduces helps us to reduce the level of bias that would have existed without it,” says Lahav. In other words, if the Thursday debate was with just six humans, each with their individual biases, the audience would just have to sit there and listen. With the AI, the audience was able to take active part — and from the get-go, at least in the opening statements, some of the bias was eliminated. “I think that proper framing of this question would be whether the technology helps to reduce bias in a world in which it exists and I think the answer is a clear yes — and that’s why I have high hopes that it’s actually a way to elevate public discourse, rather than harming it by a new biases,” says Lahav.

Parmanand says that short of romanticizing algorithms because they too can entrench certain biases, it’s true that a machine’s “scope and range and capacity to absorb information is a lot wider than us. I think there are a lot of really interesting applications like basic fact-checking and much more.”

So will AI ever replace humans? “At this stage, at least, it looks likely that the real future of AI is to augment humans rather than to fully replace us,” says Natarajan. And after all, the fear of a dystopian future where machines rule the world may not be logical — or justified. Smartphones can already play chess just as well as IBM’s Deep Blue. So who knows, maybe we’ll be able to do most things with our own personal AI software, and then the fears of mass inequality or deprivation are overstated. “I really don’t know what the future of AI is,” says Natarajan. “It certainly will have a large impact, but we can’t say what that is.”

**Read and hear the full transcript from project debater [here](#).**

**Watch the full debate [here](#).**

## **Press Contacts**

Emma Bennet

UKI External Communications

07976 839109

---