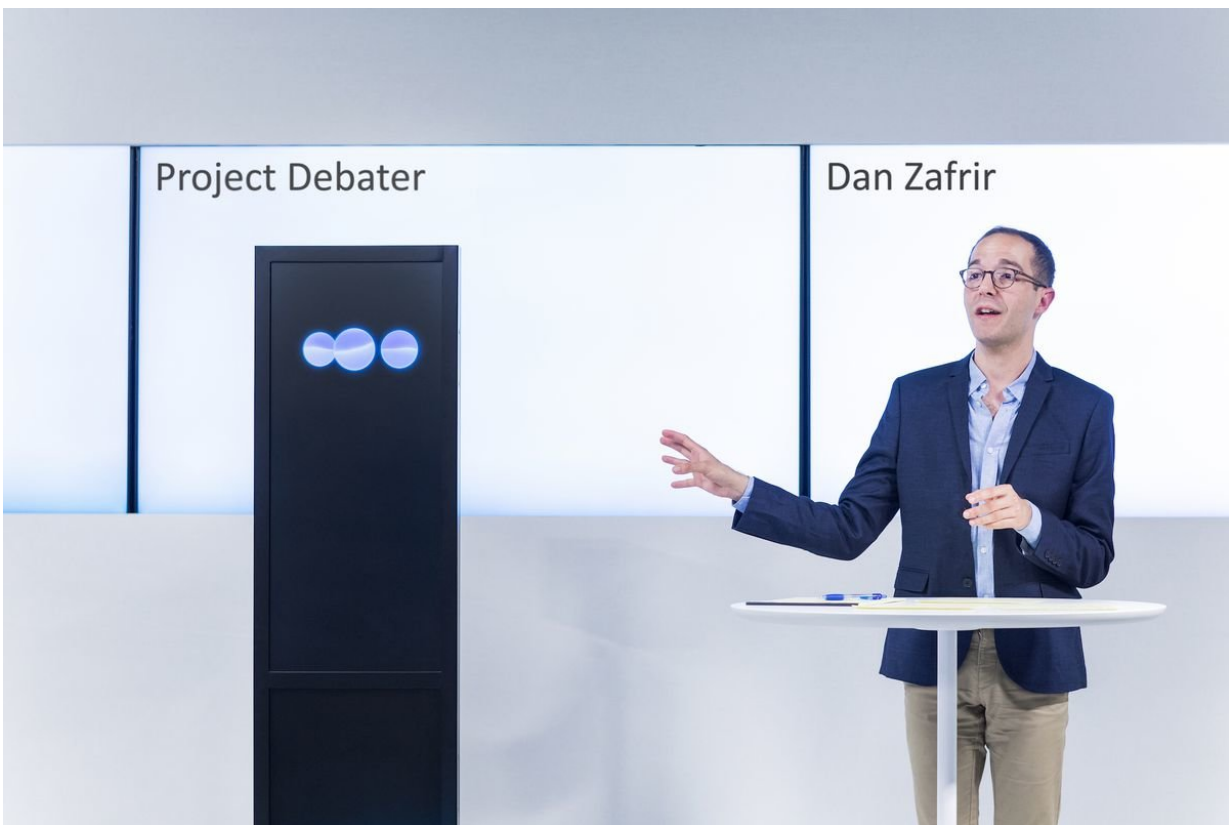


# IBM's debating computer—an AI expert's verdict

June 26 2018, by Chris Reed

---



Credit: IBM

The competition got underway when the computer's female voice, a mix of Amazon's Alexa and Stephen Hawking's communicator, spoke to its human opponent: "Hello Noa. We meet again."

I was the only academic invited into the crowded room of 50 or so journalists to witness the [recent contest](#) between the artificial intelligence of IBM's Project Debater and Israeli debate champions Noa Ovidia and Dan Zafrir. The opening gambit produced titters and eye-rolling from the audience. I was more of an eye-roller – I'm not convinced that obvious pre-scripted material really helps the cause of showcasing AI technologies. What followed though was undeniably an impressive feat of engineering—but it might be too easy to think that sci-fi AI is now just round the corner.

Project Debater follows [the announcement](#) that Google has developed an AI technology known as Duplex that can conduct natural-sounding phone conversations in order to book appointments and carry out other tasks. Both projects look like they involve AI that is nearing human-level competence, that could pass the [Turing test](#), and imminently dominate the world, perhaps. But this is an illusion borne of the careful marketing of these huge corporations. The reality is that we're still in the earliest days of understanding AI.

After the initial crowd-pleasing tactics, IBM's computer produced a four-minute speech, on the fly, on a topic selected at random from a list of 40 on which it hadn't already been trained to debate. It did this by identifying, classifying, selecting and then stitching together snippets from a library of 300m news articles. The result was largely grammatically correct, semantically on message and more or less coherent. The system was then able to listen and respond to a similar statement from its human opponent.

It's maybe worth reflecting on just how difficult these tasks are. Holding a conversation is enormously challenging once you go beyond very structured, tightly controlled domains. Deep learning systems, inspired by the human brain, are trying to map whatever the human says to a relatively small number of possible moves with a small number of

possible values. Google Duplex still works within a specific domain, like booking dinner, and so can be very robust.

Having an argument is even more demanding. It is remarkably difficult to build an algorithm to reliably determine whether a given sentence supports your position or not. On one level, the IBM team nailed it, with Project Debater producing its coherent and persuasive four-minute statement. I was also very impressed that the computer's grammatical structure was so good, especially as each sentence may have drawn from multiple articles in the library.

### **Technology still limited**

Yet as the speech went on, I got the distinct sense that the thematic structure was breaking down, with the flow flitting between topics. The machine finished bang on the four-minute mark with a nice rhetorical flourish of anticipating and attacking the opponent's argument (known [as procatalepsis](#)). But later, the computer's two-minute rebuttal to its human opponent sounded increasingly like mere repetition.

Project Debater has achieved significant new advances in areas such as searching texts for arguments ([argument mining](#)) coupled with technical solutions such as grammatical repair that involves gluing sentence parts together. But, as an orator, the computer is still making its first tiny squeaks.

The system has only the most rudimentary notion of argument structure and so often deviates from the main theme. It pays no heed to its audience, nor its opponent, and has no way of adapting its language or exploiting any of the hundreds of clever rhetorical techniques that [help win over audiences](#)

Neither IBM nor Google are claiming, or even intimating, that they've

solved all AI problems, or built machines with human-level performance. In both cases, the programmers have specific goals in mind that more or less lead directly to commercial technology.

The real value of argument technology as a whole is going to be delivered not in the debating chamber but in applications in which AI systems can contribute to human decision-making teams. Whether in the [police incident room](#), the [intelligence analysis bunker](#) or [the classroom](#), it can only be a good thing to increase the robustness of evidence-based decision making by introducing AI systems that can contribute to the conversation. They will be able to add new information or critique human reasoning.

Project Debater is a valuable step forward toward this goal, and the broader aim of building AI that can really understand and respond to us. But we are most certainly not on the verge of seeing AI systems out-debating their human counterparts. Today's AI technology is as far from these scenarios as the Romans' experiments with steam power were from the industrial revolution.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: IBM's debating computer—an AI expert's verdict (2018, June 26) retrieved 19 April 2024 from <https://phys.org/news/2018-06-ibm-debating-computeran-ai-expert.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--