Debating with Robots: IBM Project Debater and the Advent of Augmentive Automated Argumentation

Mitchell, Gordon R.
University of Pittsburgh, Pittsburgh, PA, USA

First came "Deep Blue" vs. Kasparov (1997), then "Watson" on Jeopardy! (2011). IBM's latest artificial intelligence "grand challenge" unfolded in summer 2018, when the company's "Project Debater" unit squared off in a series of public debates against human debating champions. Although that spectacle sparked widespread conversation about whether robots would soon be eclipsing human debate talent, a follow-on event at the University of Cambridge on November 21, 2019 has drawn less notice. That debate, held on the motion, "This House believes AI will bring more harm than good," featured two teams, each paired with two humans and one Project Debater robot. Using newly unveiled "Speech by Crowd" technology, Project Debater gave the opening speech on each side of the motion, developing arguments based on crowdsourced material submitted by humans to an online portal weeks prior to the event. IBM touted the unique format as a successful demonstration of how Project Debater can work as a support tool to augment (rather than replace) human argumentation. This paper deploys Aakhus and Jackson's "argumentation by design" perspective to reconstruct the "design hypotheses" inchoate in the format of the 2019 Cambridge Union debate, then tests those hypotheses through rhetorical analysis of the debate transcript and crowdsourced arguments contributed via the "Speech by Crowd" portal. Such analysis stands to contribute insight regarding the evolution of AI technology, IBM's artificial intelligence business model, and how the prospects of "automated argumentation" implicate argumentation pedagogy, practice, and scholarship.

1. INTRODUCTION

IBM's "grand challenges" pace the corporation's technological innovation and dramatize rollout of new products, particularly in the area of artificial intelligence (AI). One memorable grand challenge took place in 1996-1997, when IBM's "Deep Blue" program defeated chess world champion Garry Kasparov. Following the spectacle, dramatic headlines such as "Big Blue's hand of God" (Levy, 1997) framed the event as a key moment in the epochal contest of "man versus machine" (Goodman & Keene, 1997). As years passed and IBM's AI initiatives grew more sophisticated, more difficult grand challenges were arranged, as in 2011, when IBM's "Watson" artificial intelligence platform competed successfully against human participants in an episode of the quiz show *Jeopardy!*. Again, headlines such as "Computer finishes off human opponents" (Hanna, 2011) captured public imagination and fueled speculation about what

human faculty computers might conquer next.

Cue to 2018, when IBM's "Project Debater" program sought to bring AI to the realm of argumentation, facing off in a series of formal debates against human counterparts. Features of these events were structured to make the AI task for Project Debater especially challenging, as topics were not announced until minutes prior to the event, and winners were determined by a vote of humans watching in a live audience. Although Project Debater performed impressively, the human debate champions selected for the grand challenge held their own (winning some of the debates in the eyes of the live human audience members), prompting *Vanity Fair's* Kenzie Bryant (2019) to quip, "the robot takeover has been held off another day."

A "grand challenge" gestures toward the concept of a scientific "crucial experiment," where a single experimental result is framed as a litmus test for a scientific hypothesis, or even an entire scientific paradigm (see Holton, 1969;

Dumitru, 2013). Did the 2018 Project Debater demonstrations mark such an inflection point in the science and technology of AI? This broad question raises ontological and epistemological issues regarding the nature of human argumentation, artificial intelligence and boundaries between human and machine learning, best left for more extended treatment. Following the trajectory of IBM's Project Debater rollout, a different, more subtle, set of questions emerge.

In November 2019, IBM collaborated with Cambridge University to convene a public debate at the Cambridge Union, one of the world's most venerable debating chambers. Unlike the earlier series of Project Debater grand challenge debates, the Cambridge debate eschewed the design principle of pitting human versus machine, utilizing instead a format that formed two competing three-person teams, each composed of one Project Debater machine and two human debaters. These teams squared off to debate the motion, "This House believes AI will bring more harm than good," in a parliamentary style debate conducted in the Cambridge Union.

Another design twist in the Cambridge debate provided a vehicle for IBM to highlight its "Speech by Crowd" application. Whereas in the initial grand challenge, Project Debater generated arguments by drawing from a digital corpus of several million curated news articles on myriad topics, in the Cambridge debate it crowd-sourced content for its arguments. This crowd-sourcing was enabled by contributions of over 1,000 users, who were invited to submit short arguments on either side of the motion to an online portal opened several weeks prior to the event. Using content from this user-generated argument corpora, Project Debater extracted what it determined to be key themes and fashioned them into high quality arguments on both sides of the motion.

With Project Debater positioned as the first speaker for each team, the debate opened with one IBM speech in favor of the motion, "This House believes AI will bring more harm than good," followed by a second IBM speech against that motion. Thus, before even turning to human speech, the debate format provided audience members with an automated stereophonic *dissoi logoi,* an airing of what the machine selected as the strongest arguments on each side of the question.

IBM's public statements and promotional materials touting integration of its "Speech by Crowd" application with Project Debater strike quite a different tone in comparison to the common "machine triumphs again over humans" tropes that circulated following the initial series of grand challenge debates in 2018. With "Speech by Crowd," the script was tweaked to "AI augments human decision-making" and "AI can help human collectivities escape their filter bubbles." In part this pivot highlighted the machine's role in supporting human cognition, rather than supplanting it, and was enabled by the fact that Project Debater was serving as something of a stenographer in selecting and tailoring human-generated arguments tailored specifically for the debate and contributed via the Speech by Crowd portal.

Reflection on the form and content of the 2019 Cambridge Union debate promises to yield insight regarding the evolution of debating in a world increasingly transformed by machine learning, artificial intelligence, and the corporate platforms that develop and market such technologies. In what follows, part one reflects on how the pragma-dialectical and argumentation as design approaches provide a useful theoretical scaffolding to support analysis of the structured public debate. Parts two through four examine, in turn, the confrontation, opening, argumentation, and concluding stages of the debate. Reflection on findings and implications of the analysis are offered in a final section.

2. THE DESIGN PERSPECTIVE

The 2018 Project Debater demonstrations showcased results from IBM's AI research program, including advances in machine listening comprehension (Lavee, et al., 2019), natural language processing (Shachar, et al., 2018), and argumentation mining of large datasets (Levy, et al., 2017). Progress in these areas was particularly notable, because such machine capabilities mimicked the talents of top human debaters who exhibit quick-draw refutation and are skilled in *kairos*—the ability to find just the right words to use in a timely way.

Yet different Project Debater capabilities were on display in the 2019 Cambridge Union debate, an event that featured a substantially different format, recasting the tenor of the debate. Specifically, these features included IBM's "Speech by Crowd" AI platform for crowd-sourcing decision-support, and a format wrinkle that pitted Project Debater against itself, arguing on both sides of the debate motion. This latter

feature might be understood as *automated antilogic,* drawing from the ancient Greek sophist Protagoras' term to describe the principle that "Two accounts [logoi] are present about every 'thing,' opposed to each other" (Schiappa, 1991).

These unique design features may reveal even more about Project Debater than the nuts and bolts of speech recognition and argumentation invention capabilities, in that they speak to the system's broader functionality. This is particularly relevant for the present study, which focuses on these more general design questions (as opposed to the specifics of coding). Such an angle of inquiry directs attention to the *telos,* or broader purpose of Project Debater. A robust literature on design in argumentation helps elucidate these dimensions.

Aakhus and Jackson (2005) have elaborated a research program that views argumentation through the prism of design: "The work central to a design enterprise involves creating techniques, procedures, and devices that make forms of communicative activity possible that were once impossible or that realize an improved form of communicative practice" (Aakhus & Jackson 2005, p. 416; see also Aakhus 2007, 2003; Jackson 1998, 2015; Greco, 2018). Adapting nomenclature from the field of architecture, they distinguish between "natural" (pre-designed) and "built" (new) communication.

It can be useful to view public debates from a design perspective, because such events incorporate both natural and built elements. On the one hand, public debates are "built"—each is designed with unique format features. On the other hand, this construction comes on top of "natural" edificies formed by debate history, which stretches back for millennia. Contemporary public debate grows out of an ancient tradition that can be traced back to Protagoras, the Greek teacher of oratory who championed the art of *dissoi logoi*, or "contrasting arguments" (see Schiappa, 1991). Through structured exercises, Protagoras taught Athenians to use the art of debating as a way to measure the strength of competing positions and inform judgments on questions of civic import. Later, the Romans would develop this tradition through a method of instruction they called *in utramque partem*—Cicero's term for arguing "on both sides of the case" (see Mendelson, 2002, pp. 173-203).

Designed public debates "build" on this "natural" edifice by inheriting the basic foundational infrastructure of back-and-forth argumentation, then inflecting the exchange through deliberate design choices regarding topic wording, format, speaker selection, incorporation of technology, and other design features.

Aakhus and Jackson stipulate that each design feature of communication contains an inchoate hypothesis. In the case of IBM's 2019 Cambridge demonstration, that hypothesis could be: Project Debater augments, rather than supplants, human decision-making. Testing this hypothesis calls for interpretation and judgment. Pragma-dialectical argumentation theory can be useful in this respect, as the approach is concerned with how disagreements are normatively structured and how they play out in practice. This theory can be a useful reference point for exploring the extent to which the design hypothesis implicit in the Cambridge demonstration held up.

In key respects, a structured public debate is designed to resemble an ideal model for critical discussion (see Table 1), with discrete format phases (topic formulation, opening speeches, question and answer, rebuttal speeches) mapping onto the phases of a critical discussion (confrontation stage, opening stage, argumentation stage, and concluding stage) (van Eemeren & Grootendorst, 2004, 59-62) "The ideal model of a critical discussion does not represent a utopia," stipulates van Eemeren (2018), "but a theoretically motivated idealization . . . suitable to serve as a point of reference in analysing and evaluating oral and written argumentative discourse" (p. 35).

Critical Discussion Phase	Description
Confrontation Stage	☐ Difference of opinion presents itself; disagreement arises.
Opening Stage	☐ Protagonist and antagonist identify their initial commitments and standpoints.
Argumentation Stage	☐ Rounds of argumentation as the protagonist responds to critical responses of the antagonist.
Concluding Stage	☐ Determination of whether the protagonist's standpoint has been successfully defended.

Table 1. Four stages of critical discussion in pragma-dialectical argumentation theory (adapted from Van Eemeren, Grootendorst, and Snoeck Henkemans 1996), pp. 281-282.

Pragma-dialectics' ideal model of a critical discussion does not seamlessly match the typical the "built environment" of a designed public debate, but the fit is close, and subtle variations can be instructive. For example, according to pragma-dialectical theory, in the confrontation stage of a critical discussion, the protagonist and antagonist locate grounds for disagreement. Although this element is also present in structured public debates, a third party (typically the organizer or moderator) plays an important role in isolating the points of disagreement and framing the scope of debate. "Public debate propositions do not simply serve to limit the discussion and define the sides of the debate," observe Broda-Bahm, Kempf and Driscoll (2004, 125); "they also play an important role in gaining attention and communicating the purpose of the debate."

Similar observations could be made about the opening, argumentation, and concluding stages of a critical discussion, which correlate roughly to phases and features of a designed public debate. Mapping features of the Project Debater Cambridge Demonstration onto the pragma-dialectical critical discussion model yields the following breakdown (see Table 2).

The following analysis considers these format features as they relate to pragma-dialectical critical discussion phases and explores the content of argumentation advanced in each stage. In this case, the fact that the design hypothesis and topic relate synecdochically provides a unique opportunity to generate insight about the event. In other words, the debate motion, "This house believes AI will bring more harm than good," lays groundwork for speakers to address the debate's design hypotheses reflexively, as they advance standpoints regarding Project Debater during the debate.

Critical Discussion Stage	Debate Format Element	Speaker
Confrontation Stage	☐ Pre-debate topic and format formulation ☐ Crowd by Speech argument sourcing ☐ Event introduction	☐ Noah Slonim (IBM)
Opening Stage	☐ First proposition speech ☐ First opposition speech ☐ Floor speeches	☐ Project Debater ☐ Project Debater ☐ Cambridge Union students
Argumentation Stage	☐ Second proposition speech ☐ Second opposition speech ☐ Floor speeches ☐ Third proposition speech ☐ Third opposition speech	☐ Sharmila Parmanand (Cambridge) ☐ Sylvie Delacroix (U. Birmingham) ☐ Cambridge Union students ☐ Neil Laurence (Cambridge) ☐ Harish Natarajan (AKE Int'l)
Concluding Stage	☐ Audience Q&A ☐ Voting	☐ Noah Slonim (IBM) and Cambridge Union students

Table 2. 2019 IBM-Cambridge Union public debate format mapped on pragma-dialectical ideal model for critical discussion

3. ANALYSIS

3.1 Confrontation Stage

In pragma-dialectical argumentation theory, the confrontation stage of a critical discussion involves the "initial situation" where interlocutors come together to assess whether their difference of opinion warrants an attempt to resolve it through critical discussion (van Eemeren, 2018, 36).

In the case of the IBM-Cambridge demonstration, the choice of topic wording—"This House believes AI will bring more harm than good"—reflects the organizers' intention to shape the critical discussion according to the norms of British parliamentary debate (hence, "This House . . ."), with the debate convened in the venerable Cambridge Union, home to thousands of previous events featuring a basic similar structure. That structure typically includes alternating pro/con ("proposition" and "opposition"), time-limited speeches, delivered by invited guest speakers and Cambridge Union student members, with an opportunity for audience members to participate through "floor speeches" and voting on considered motions.

Whereas the inaugural 2018 Project Debater demonstration featured a one-on-one, machine vs. human format, the 2019 Cambridge event expanded this format, placing three speakers on each side. It is possible that all three speaking roles on one side could have been assigned to Project Debater, with the opposing side being comprised of human debaters (such a format would have largely replicated the dynamic of the 2018 demonstration). Yet organizers of the Cambridge demonstration chose a different approach, one that carried significant design implications. That approach entailed placing, in the confrontation stage, one Project Debater unit on each side of motion, with each machine joined by two human debaters, forming opposing sides of three speakers (one machine and two humans on each side). This design created a dynamic in which Project Debater would be debating against itself (with added intrigue, given the subject matter of the motion).

According to Jackson (2015), a design hypothesis in argumentation is "some notion, theoretical or intuitive, about how argumentation works to achieve its purpose or how it might be conducted to better achieve its purpose" (250).

For the 2018 inaugural Project Debater demonstration, a plausible design hypothesis could be reconstructed as: "AI can hold its own debating against a human opponent." In comparison, design of the 2019 IBM-Cambridge demonstration reflected a different hypothesis, along the lines of: "AI can augment human learning in a debate context." Other design features in the confrontation stage reinforce this subtle, yet significant, shift. In the 2018 demonstration, Project Debater generated its initial standpoints in a compressed confrontation stage, crunching through millions of news articles and other information sources from its library, after being given the specific motion for debate, "We should subsidize space exploration," just minutes before the event.

By way of contrast, the confrontation stage in the 2019 IBM Cambridge demonstration was extended for several weeks, as IBM's "Speech by Crowd" platform crowdsourced arguments on both sides of the proposition from hundreds of human contributors who logged onto a dedicated IBM portal prior to the event (see Figure 1).
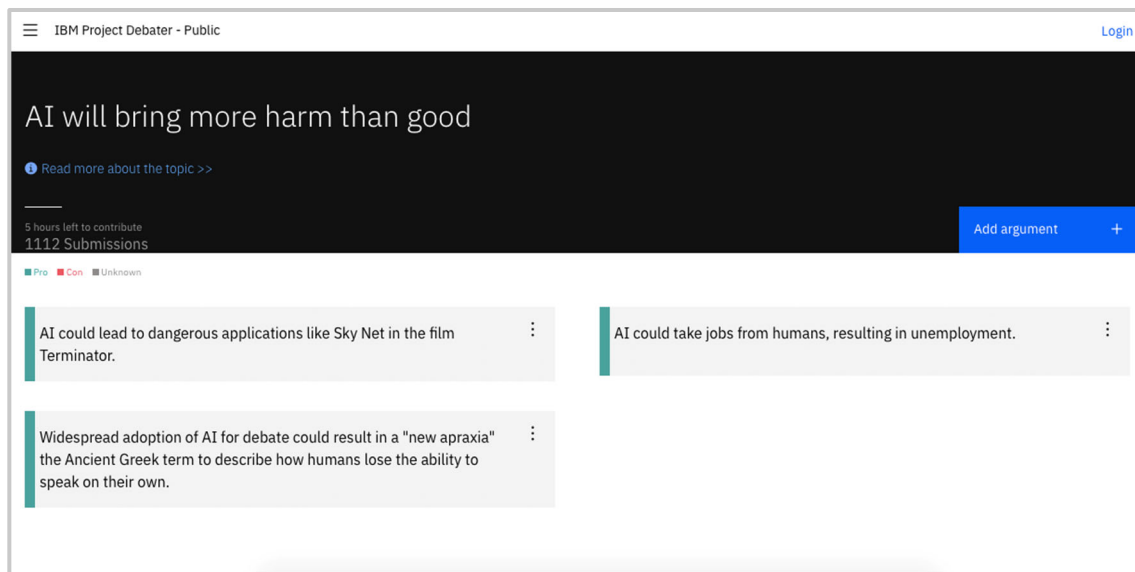

Figure 1. Example of arguments contributed through the IBM "Speech by Crowd" portal (permission pending).

In all, the Speech by Crowd platform received 1,100 arguments that people submitted to IBM through a website in the week prior to the debate. It categorized 570 comments as being in favor of the idea that AI would cause more harm than good and 511 comments as being opposed. It discarded some comments as irrelevant to the debate (Kahn, 2019).

3.2 Opening Stage

Project Debater's subsequent speeches exhibited the sort of "defining, specifying and amplifying" typical of the type of argumentation appropriate for this stage in pragma-dialectical argumentation theory (Van Eemeren, 2018, 42). During its opening presentation for the proposition, the machine advanced the following five standpoints, thus establishing the protagonist position in the critical discussion:

- Since AI is not human, its capability for moral decision-making will be limited.
- Data sets that train AI contain bias, which will be amplified in discriminatory AI applications.
- AI will create unemployment by displacing human workers.
- AI will ruin society by instilling human laziness and removing the human element from almost everything we do.
- AI will magnify the power of rogue actors to do harm.

This opening speech covered substantial argumentative ground, although each standpoint was developed cosmetically, and often somewhat haltingly, as illustrated in the following example of Project Debater's rendering of the proposition argument regarding AI's tendency to displace human workers:

Let's move to employment. While my job at IBM is secure, at least I hope so, I know this issue is quite pertinent to our discussion today. AI will make lots of people lose their jobs. It will bring more harm than good in that it will displace a lot of workers and cause employment problems. We risk creating a workforce that puts people out of employment. Jobs involving vehicles such as travel is one of the biggest employers and those jobs will be lost because of AI. (Project Debater, 2019)

Clearly the most engaging aspect of the above sample of argumentation is the use of humor. Where did the joke about IBM come from, and how did Project Debater know to deploy humor in this way? An answer to that question emerges later in the debate, but for now it may be useful to reflect on the fact that the cosmetic nature of the serious content in this passage perhaps reflects limitations placed on contributions to the Speech by Crowd portal, where each of the some 1,100 arguments submitted were constrained to a Twitter-type text box holding only several hundred characters. There were no options, for example, for contributors to submit footnotes, hyperlinks, images, or sounds as supporting evidence.

Indeed, the opening speech by Project Debater for the opposition side reflects similar dynamics, as the machine advanced the following standpoints to initiate antagonist argumentation in the confrontation stage:

- AI will relieve humans from the drudgery of repetitive tasks and reduce human errors.
- AI will open up more opportunities for human leisure time and entertainment.
- AI will create new jobs for humans in certain economic sectors.
- AI will improve medicine, transportation, and even inspire new forms of music.
- AI will general enhance the quality of human life, as fundamentally, programmed machines are governed by the laws of humanity.

Closer scrutiny of Project Debater's argument regarding AI's potential to spur technological advances reveals a curious parallel to the standpoint it advanced in the previous speech; the argument begins with a joke, then develops with *logos*-based reasoning, albeit unspooled with a few inelegant turns of phrase:

Let's move to an issue close to my artificial heart—technology. AI will enable technology to advance and further medical research, which will save lives. It will enable us to develop more and more impressive technology. While regulation and serious consideration of the concerns are in place, the benefits of AI technology are enormous and are way beyond the over-exaggerated potential harms. Autonomous vehicles are prime examples of how artificial intelligence is impacting the automotive industry. A large segment of autonomous vehicles are connected, and thus able to share the learning with each other. (Project Debater, 2015)

At the end of the opening stage, audience members and human debate participants were left to ponder an argumentative tableau crafted by Project Debater: Two mirror-image speeches, each covering five major standpoints, backed by *logos*-centric reasoning, with the exception of a single joke sprinkled in. With protagonist and antagonist standpoints established in the opening stage by Project Debater, next participants turned to the argumentation stage, as explored in the following section.

3.3 Argumentation Stage
In pragma-dialectical argumentation theory, moves made by interlocutors in the argumentation stage are tied to standpoints established in the previous, opening stage of a critical discussion. A similar convention holds in academic parliamentary debate, where the first speakers establish their side's interpretation of the motion and build an opening case that sets parameters for subsequent argumentation. The IBM-Cambridge public debate is especially notable in this light, as the first speaker for each side in the opening stage was an AI machine. How would human speakers, in the argumentation stage, work with the material handed to them by their machine partners? Transcript analysis reveals extensive co-ordination between human speakers and machine, with 17 total references to Project Debater advanced in the argumentation stage. Six of these references mentioned specific argumentative content introduced by Project Debater. There were six instances of human speakers making observations about Project Debater's role in the debate, and five times human speakers deployed Project Debater as a rhetorical synecdoche, pointing to specific dimensions of its presence

and performance to make a wider point about the general AI motion up for debate. Review of these instances provides vivid texture of the dynamics at play in the argumentation stage.

At times, human speakers would call attention to Project Debater's role in the debate, offering a window into how the participants were perceiving the experience of including a machine in their distinctively human interaction. For example, in opening the second proposition speech, Sharmila Parmanand (2018) explained, "My role here is to support and extend the arguments of my teammate, Project Debater, and also to respond to what has been raised by my opponent, [chuckle] Project Debater." Parmanand's chuckle underscored the double game going on—the debaters were willing to treat Project Debater like a human partner, yet doing so entailed verbal contortions like acknowledging the same speaker arguing simultaneously on both sides of the motion. Parmanand wove a similar reference into the end of her speech, closing with, "So, we on our side are very happy to be with Project Debater, but in general, a bit concerned about AI, so we are proud to propose."

On the antagonist side, second opposition speaker Sylvie Delacroix (2019) began her speech with a charitable gesture toward the machine, saying, "Actually, first I think it is embarrassing that we still haven't given a name. I mean, don't you think 'Project Debater' is not a very good name? 'Debbie' was given during dinner—I think I'm going to call you Debbie, unless there is any objection." Having thus anthropormorphized Project Debater, Delacroix continued to identify with the machine, reflecting, for example, how her debate preparation habits bore resemblance to the machine's search techniques modeled in the 2018 demonstrations: "Just like Debbie, before I came here tonight, I did go and browse the web. Why? Well first, because I wanted to try and anticipate what Debbie might say, because Debbie is very good at browsing the web." Later in the second opposition speech, Delacroix referred again to Project Debater's freshly-minted human name, in the process emphasizing her intention to build the opposition side's case using more than just instrumental patterns of reasoning:

> I don't want to win this debate on the basis of instrumental reasons alone. Sorry Debbie. I mean, you have done a great job at helping

here, but I don't want to win this debate on the basis of instrumental considerations. Why? Well, because we would lose sight of a very important aspect, a very important consideration. And what is it? Well, again, no offense Debbie, but this debate is not so much about you, the AI, but about us—who we are, and who we want to become.

Responding indirectly to Delacroix's move to "Debbify" Project Debater, third proposition speaker Neil Laurence pointed to the tendency of humans to anthropomorphize non-human objects. This tendency, according to Laurence, stood as a poignant marker of fundamentally different forms of human and machine cognition: "Our own method of computation is, because we're so limited, is to use our powerful computation in our head to think about the motivations of all around us and to and to anthropomorphasise the things we communicate [with] and we do that to these machines that's why we like to give them names but in reality they don't have names."

These passages clearly indicate that the figure of Project Debater cast a long shadow over the Cambridge debate—indeed the looming black obelisk in the middle of the Cambridge Union debating chamber was hard to miss (see Figure 2).

Yet Project Debater influenced the course of the debate in another register, as human speakers referenced argumentative standpoints generated by the machine in the debate's first two speeches. For example, as a preface to an argument about AI and the labor market in the second proposition speech, Parmanand (2018) stated, "So first, let's talk about the displacement of labor on a massive scale, and this is something that my teammate discussed at length, right?" Later in the same speech, Parmanand (2018) built on her machine partner's earlier argument AI's tendency to stultify humanity:

> My teammate was correctly concerned about humans losing things like creativity, staying sharp, staying adaptive, our evolutionary instincts becoming more dull when we outsource everything to robots. I was very concerned when my [chuckle] AI opponent said that maybe we will have machines replacing teachers in the classroom. The quality of education that is likely to ensue won't be as good because nothing can replace the kind of emotional intimacy that is necessary in a classroom setting, for example.

Figure 2. IBM Project Debater (right, obelisk) during 2019 Cambridge Union public debate. Photo: IBM (permission pending).

Parmanand (2018) also drew from her machine partner's earlier argumentation to bolster the proposition side's standpoint regarding AI and bias:

> So if you listened to how my opponent explained why AI is better than humans, there was this assertion that AI doesn't replicate the errors that humans do—right—AI reduces human error—that is precisely why it is going to be very hard for us, as a society, to deal with the biases that AI will entrench, because there is this perception that it eliminates human bias. We just instinctively think if it is math, it is fair. But that is not actually the case.

On the opposition side, Delacroix pointed out how her argument regarding AI's potential for economic stimulus countered the standpoint that AI would cause unemployment, advanced in Project Debater's first proposition speech: "This economic and political power is, by far, the most disrupting, and promising aspect of AI. And Debbie, by the way, that means a lot of new jobs."

Adopting a meta-view in the third opposition speech, Natarajan observed that Project Debater was able to generate impressive content on both sides of the motion in the debate's first two speeches: "So I don't think it escaped anyone's interest that at the heart of the debate is this: That a piece of technology can simultaneously be both terrifying and awe-inspiring," he said. "I think for Project Debater—on both sides, for my teammate Project Debater, and from both teams, we got elucidation of what some of those risks are."

Specifically, Natarajan highlighted how Project Debater's mirror-imaged argumentation in the debate's first two speeches underscored his point about the transformative effects of AI on the labor market:

> I think this is a realistic problem which many people have identified, in different forms, throughout this debate, starting with Project Debater on the side of the proposition, and my own partner, Project Debater, giving you the opposite side of it, which is this: The economic system that we live under changes massively when we have artificial intelligence doing jobs.

Human speakers also utilized Project Debater in a third way through deployment of the rhetorical figure of *synecdoche.* As a strategy of persuasion that invokes relationships between part and whole to make a point, the synecdoche can be a powerful tool of argumentation in debates that unfold on multiple levels. In the Cambridge-IBM public debate, the motion (regarding artificial intelligence), coupled with Project Debater's participation in the debate (as an instantiation of AI), afforded rhetorical resources for human speakers to invent synecdochic argumentative appeals.

For example, in the third proposition speech, Laurence introduced the story of Jean Dominique Bauby, former editor-in-chief of *Elle* magazine, whose tragic stroke at age 43 rendered him speechless, able only to "dictate" letters by signaling with his left eye: "The remarkable thing about Bauby is we know his story because he

wrote a book. And it took him, I think, 7 months of four hours a day to write this book," said Laurence. "I think when we think about that we all think about what it would be like to be in that state, and the first important point is [that] relative to our friend Project Debater, we are all in that state. A locked in state." The structure of Laurence's appeal was synecdochic—the specific relation between Bauby and Project Debater is deployed to underwrite a larger point about the potential danger of AI. Adding granularity to this line of argument, Laurence invoked information theorist Claude Shannon: "Shannon also estimated the entropy of the English language … and I can tell you that I'm roughly communicating to you at a rate of 2000 bits per minute. Our friend Project Debater is communicating, when it desires to do so, at a rate of around 60 billion bits per minute." Providing a counterpoint to Delacroix's move to humanize Project Debater by naming it "Debbie," Laurence drove home the upshot of his standpoint: "So Sylvie gave Project Debater a name, she called her, him, it … Debbie. I'm going to try the name 'Cybertronia the All-Knowing' because in some sense that's more representative of what we're dealing with."

3.4 Concluding Stage

At the end of a critical discussion, according to pragma-dialectical argumentation theory, participants reflect on whether the content of the exchange has led to the protagonist upholding or failing to support their standpoints offered in the opening stage. Correlates in designed public debates come in the form of adjudicated decisions and/or audience votes. For example, it is a Cambridge Union tradition to gauge audience opinion at the end of a debate by inviting audience members to exit the venue through a certain door, corresponding to their final vote in the debate. In the case of the IBM-Cambridge demonstration, this process yielded a mixed result: "Votes were split almost equally for and against the motion, with the team who argued in favor of AI garnering 51% of votes" (Ziady, 2019). Departing from the "human versus machine" narrative invited by the earlier Project Debater demonstrations, design of the IBM-Cambridge debate steered audience members away from viewing the motion as a referendum on the Project Debater technology, and more as a demonstration of how the automated argumentation could be viewed as augmenting human critical thinking.

Also in the concluding stage, leading IBM engineer Noah Slonim fielded questions from the audience, pulling back the proverbial designer's curtain to provide deeper perspective on some of the key moments in the debate. For example, several audience members were intrigued by Project Debater's attempts at humor, and their exchange with Slonim yielded important insight regarding this aspect of the machine's design:

> **Slonim:** I think I heard the question, actually: "How does the system make jokes"? So, it's a good question. So the system is not inventing jokes; it has a bank of, I would say more colorful or humoristic comments that it tries to use in the right timing. This is, by itself, is challenging. The system, also, you know, it lacks tact. So sometimes it will make a humoristic comment at, you know, in the wrong moment, which, again, could be amusing but not in the exact way that we planned it. But also, that said, I think it is interesting to point out that the type of humor that the system is using, where the subtext is really about: I am a machine. Alright, so this is the subtext of what this humor is really trying to convey—that this is a machine, not trying to replace humans, but actually to accompany them.
>
> **Audience member:** But I feel like it might reinforce the image of the machine being conscious, or whatever, and like talking to humans about, "Hey, I'm a machine, but I'm talking to you about being a machine," which requires some consciousness . . .
>
> **Slonim: . . .** Yes, so just to be clear, the machine is not conscious, okay. Alright. So yes, but again, the machine is trying to do it's best to be more engaging. I think that humor, at the end of the day, is a rhetorical tool that sometimes we use in debates. So ignoring this aspect while developing this machine is wrong, so this is why we added this capability. And again, in some debates, it works well, and in some debates it does not work well, okay?

In clarifying that Project Debater's humor—perhaps the most "human" element of its performance—was pre-scripted in a "joke bank," Slonim revealed how IBM's engineers ventriloquized their own argumentation through Project Debater's speech. Future research might explore how such a maneuver entails use of *praeteritio*—the rhetorical figure

of pointing to something by saying you are not. Slonim explained that nearly all of the jokes scripted into Project Debater's bank involve self-deprecating jokes that poke fun at the limitations of AI (recall from the opening stage, Project Debater wisecracking, "While my job at IBM is secure, at least I hope so.") Of course, the dramatic element of such humor is that it invites audience amazement at the fact that a machine could generate such sophisticated humor "on its own." Slonim's exchange with the Cambridge Union audience highlights blurriness of the human/machine boundary and serves as a reminder that some of the most dazzling displays of apparently spontaneous machine intelligence may be more the product of purely human invention than we realize.

4. CONCLUSION

The 2019 IBM-Cambridge demonstration debate showcased IBM's Project Debater technology in an innovative format designed to demonstrate how the AI platform is able to augment human decision-making through argumentation. The preceding analysis has explored how design of the debate, and content of the argumentation in the event, bear on IBM's "design hypothesis" regarding this issue. Such analysis may have enduring salience, given IBM's commitment to integrate Project Debater into its commercial suite of AI applications.

Study limitations include the fact that robust generalizations may be difficult to generate from qualitative analysis of a single event. Indeed, future projects might usefully explore other instances where IBM's Project Debater and Speech by Crowd platforms have been demonstrated, such as the effort to deploy machine-assisted crowd-sourcing to catalyze public discussion on the value of autonomous vehicles in the Swiss city of Lugano (Curioni, 2019). And when it comes to automated argumentation, IBM is not the only game in town—there are also collaborative efforts by Scottish and Dutch scholars to build comparable platforms (see, e.g. Visser, Lawrence, Wagemans, and Reed, 2019). How do such platforms compare, and how might the emergence of automated argumentation shape the human experience of using *dissoi logoi* to inform critical judgments and learn alternative perspectives? When future scholars look back on the next 20 years of the Tokyo Argumentation Conference, they may spot trends in which such questions move to the fore of the argumentation studies research agenda.

REFERENCES

Aakhus, M. (2003). Neither naïve nor critical reconstruction: Dispute mediators, impasse, and the design of argumentation. *Argumentation 17*, pp. 265-290.

Aakhus, M. (2007). Communication as design. *Communication Monographs 74*(1), pp. 112–117.

Aakhus, M., and S. Jackson. (2005). Technology, interaction, and design. In K. L. Fitch & R. E. Sanders (Eds.), *Handbook of Language and Social Interaction* (pp. 411–435). Mahwah, NJ: Lawrence Erlbaum Associates.

Bar-Haim, R., D. Krieger, O. Todelo-Ronen, L. Edelstein, Y. Bilu, A. Halfon, Y. Katz, A. Menczel, R. Aharanov, and N. Slonim. (2019). From surrogacy to adoption; from bitcoin to cryptocurrency: Debate topic expansion. In A. Korhonen, D.R. Traum, and L/ Marquez (Eds.), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 977-990). Stroudsburg, PA: Association for Computational Linguistics.

Bryant, K. (2019). A human beat A.I. in a debate tournament, and robot takeover has been held off another day. *Vanity Fair,* February 13.

Curioni, A. (2019). IBM project debater demonstrates the future of democracy in Switzerland. IBM Research Blog, September 24. https://www.ibm.com/blogs/research/2019/09/speech-by-crowd-switzerland/

Delacroix, S. (2019). Statement. Public debate on "This house believes AI will bring moreharm than good." Cambridge Union, Cambridge University. London. November 24. https://www.youtube.com/watch?v=lWkvopzAvoQ

Dumitru, C. (2013). Crucial instances and crucial experiments in Bacon, Boyle, and Hooke. *Society and Politics 7,* 45-61.

Eemeren, F.H. van, R. Grootendorst, and F.S. Henkemans, et al. (1996). *Fundamentals of argumentation theory.* Mahwah, NJ: Lawrence Erlbaum Associates.

Eemeren, F.H. van, R. Grootendorst. (2004). *A systematic theory of argumentation: The pragma-dialectical approach.* Cambridge: Cambridge University Press.

Eemeren, F.H. van. 2018. *Argumentation theory: A pragma-dialectical perspective.* Cham, Switzerland: Springer.

Gleize, M., E. Shnarch, L. Chosehen, L. Dankin, G. Moshkowich, R. Aharonov, and N. Slonim. Are you convinced? Choosing the more convincing evidence with a Siamese network. In A. Korhonen, D. R. Traum, and L. Marquez (Eds.), *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 967-976). Stroudsburg, PA: Association for Computational

Linguistics.

Goodman, D., and R. Keene. (1997). *Man versus machine: Kasparov versus Deep Blue.* Cambridge, MA: H3 Press.

Greco, S. (2018). Designing dialogue: Argumentation as conflict management in social interaction. *Travaux Neuchâtelois de Linguistique* 68, 7-15

Hanna, J. (2011). Computer finishes off human opponents on 'Jeopardy!' *CNN,* February 17.

Holton, G. (1969). Einstein, Michelson, and the 'crucial' experiment. *Isis* 60, 132-197.

Jackson, S. (2015). Design thinking in argumentation theory and practice. *Argumentation* 29, 243-263.

Kahn, J. (2019). IBM showcases A.I. that can parse arguments in Cambridge Union debate. *Fortune,* November 22, https://fortune.com/2019/11/22/ibm-ai-debate-arguments-cambridge-union/.

Lavee, T., M. Orbach, L. Kotlerman, Y. Kantor, S. Gretz, L. Dankin, S. Mirkin, M. Jacovi, Y. Bilu, R. Aharonov, and N. Slonim. (2019). Towards effective rebuttal: Listening comprehension using corpus-wide claim mining. In P. Nakov and A. Palmer (Eds.), *Proceedings of the 6th Conference on Argument Mining* (pp. 967-976). Stroudsburg, PA: Association for Computational Linguistics.

Levy, R., S. Gretz, B. Sznajder, S. Hummel, R. Aharonov, and N. Slonim. (2017). Unsupervised corpus-wide claim detection. In I. Habernal, et al. (Ed.), *Proceedings of the 4th Workshop on Argument Mining* (pp. 79-84). Copenhagen: Association for Computational Linguistics.

Levy, S. (1997). Big Blue's hand of God. *Newsweek,* May 18.

Laurence, N. (2019). Statement. Public debate on "This house believes AI will bring moreharm than good." Cambridge Union, Cambridge University. London. November 24. https://www.youtube.com/watch?v=lWkvopzAvoQ

Mirkin, S., M. Jacovi, T. Lavee, H.K. Kuo, S. Thomas, L. Sager, L. Kotlerman, E. Venezian, and N. Slonim. (2018). A recorded debating dataset. In N. Calzolari (Ed.), *Proceedings of the Eleventh International Conference on Language Resources and Evaluation,* Miyazaki, Japan: European Language Resources Association.

Natarajan, H. (2019). Statement. Public debate on "This house believes AI b more harm than good." Cambridge Union, Cambridge University. London. November 24. https://www.youtube.com/watch?v=lWkvopzAvoQ

Parmanand, S. (2019). Statement. Public debate on "This house believes AI will bring moreharm than good." Cambridge Union, Cambridge University. London. November 24. https://www.youtube.com/watch?v=lWkvopzAvoQ

Project Debater. (2019). Statement. Public debate on "This house believes AI will bring more harm than good." Cambridge Union, Cambridge University. London. November 24. https://www.youtube.com/watch?v=lWkvopzAvoQ

Schiappa, E. (1991). *Protagoras and logos: A study in Greek philosophy and rhetoric.* Columbia: University of South Carolina Press.

Slonim, N. (2019). Statement. Public debate on "This house believes AI will bring moreharm than good." Cambridge Union, Cambridge University. London. November 24. https://www.youtube.com/watch?v=lWkvopzAvoQ

Visser, J., J. Lawrence, J. Wagemans, and C. Reed. (2019). An annotated corpus of argument schemes in US election debates. In B. Garssen, D. Godden, G.R. Mitchell, & J. Wagemans (Eds.), *Proceedings of the Ninth Conference of the International Society for the Study of Argument* (pp. 1101-1111). Amsterdam: SicSat.

Ziady, H. (2019). IBM computer struggles in Cambridge debate on the dangers of AI. *CNN Business.* November 22. https://www.cnn.com/2019/11/21/tech/ai-cambridge-university-debate/index.html