

generate double-strand breaks (DSBs) in mtDNA. They used RNA sequencing to analyse changes in gene expression in cells treated with mTLNs, and found increased transcription of nuclear genes involved in the innate immune response; these included interferon-response genes, which are typically involved in combating viral infections. The authors also found that the transcription factor STAT1 was modified by phosphate groups and relocated to the nucleus – a key part of the interferon response⁸.

Breaks in mtDNA that occur through other means, such as treatment with toxic, DNA-damaging agents or errors in replication, often lead to compromised organelle function⁹. But Tigano and colleagues found that the mTLN treatment reduced the number of mtDNAs by only around 60%, which did not seem to have an immediate impact on mitochondrial function. The group observed no changes in key indicators of normal mitochondrial function, such as morphology, the gradient of protons (H⁺ ions) across the membrane, and the generation of reactive oxygen species. These data indicate that mtDNA cleavage is a key trigger of antiviral responses.

Next, Tigano *et al.* set out to identify the signalling molecules that relay the message of mtDNA instability to the nucleus. Although the mTLN-treated cells had intact mitochondrial function and were not apoptotic, the group showed that BAK–BAX pores did form on the membrane, consistent with mitochondrial herniation. The authors found that mtRNA – but not mtDNA – accumulated in the cytosol of these cells. The mtRNA molecules were detected by an RNA-sensing protein called RIG-I, which is better known as a sensor of viral RNA in the cytosol¹⁰. Working with its adaptor protein on the mitochondrial outer membrane, dubbed mitochondrial antiviral signalling (MAVS), RIG-I triggers a signalling pathway that activates interferon-response genes in the nucleus¹⁰. These findings point to a framework by which cells engage mitochondrial signalling molecules in immune-surveillance mechanisms (Fig. 1).

DNA-damaging agents such as radiation, which is used to treat cancer, elicit a systemic immune response that is thought to be driven by DNA damage in the nucleus. Tigano and colleagues found that radiation depleted mtDNA numbers by 40% and elicited the same immune response as mTLNs, suggesting that DSBs occur in mtDNA as well as in nuclear DNA following irradiation. Strikingly, induction of the interferon response during irradiation was nearly completely abrogated in cells lacking mtDNA. This observation indicates that mtDNA damage caused by radiation can be a driver of interferon responses. Of note, the induction of several other innate immune responses still occurred in cells lacking mtDNA, suggesting that depletion of mtDNA

specifically impairs the interferon response.

The study highlights an immunostimulatory role for mtRNA. However, questions remain. For instance, mtRNA molecules are highly unstable in nature¹¹ – how are mtRNAs stabilized so that they accumulate in the cytosol, as was observed in the current study? Another avenue for further investigation is the factors that stimulate the formation of BAK–BAX pores following mtDNA breaks. It would be of broad interest to study whether drugs that inhibit this pore formation can suppress an inflammatory immune response. The discovery of a mechanism by which cells recognize self-RNAs from mitochondria to initiate an immune response also raises the question of whether this pathway might be involved in autoimmune disease. Finally, it would be exciting to explore whether artificially induced mtDNA damage could be used to increase the efficacy of targeted immunotherapies for cancer.

Artificial intelligence

Argument technology for debating with humans

Chris Reed

A fully autonomous computer system has been developed that can take part in live debates with people. The findings hint at a future in which artificial intelligence can help humans to formulate and make sense of complex arguments. **See p.379**

The study of arguments has an academic pedigree stretching back to the ancient Greeks, and spans disciplines from theoretical philosophy to computational engineering. Developing computer systems that can recognize arguments in natural human language is one of the most demanding challenges in the field of artificial intelligence (AI). On page 379, Slonim *et al.*¹ report an impressive development in this field: Project Debater, an AI system that can engage with humans in debating competitions. The findings showcase how far research in this area has come, and emphasize the importance of robust engineering that combines different components, each of which handles a particular task, in the development of technology that can recognize, generate and critique arguments in debates.

Less than a decade ago, the analysis of human discourse to identify the ways in which evidence is adduced to support conclusions – a process now known as argument mining² – was firmly beyond the capabilities of state-of-the-art AI. Since then, a combination of technical advances in AI and increasing

Nandhitha Uma Naresh and **Cole M. Haynes** are in the Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA.

e-mail: cole.haynes@umassmed.edu

1. Shpilka, T. & Haynes, C. M. *Nature Rev. Mol. Cell Biol.* **19**, 109–120 (2018).
2. Guo, X. *et al.* *Nature* **579**, 427–432 (2020).
3. Tigano, M., Vargas, D. C., Tremblay-Belzile, S., Fu, Y. & Sfeir, A. *Nature* **591**, 477–481 (2021).
4. McArthur, K. *et al.* *Science* **359**, ea06047 (2018).
5. White, M. J. *et al.* *Cell* **159**, 1549–1562 (2014).
6. West, A. P. *et al.* *Nature* **520**, 553–557 (2015).
7. Dhir, A. *et al.* *Nature* **560**, 238–242 (2018).
8. Majoros, A. *et al.* *Front. Immunol.* **8**, 29 (2017).
9. Alexeyev, M., Shokolenko, I., Wilson, G. & LeDoux, S. *Cold Spring Harb. Perspect. Biol.* **5**, a012641 (2013).
10. Rehwinkel, J. & Gack, M. U. *Nature Rev. Immunol.* **20**, 537–551 (2020).
11. Borowski, L. S., Dziembowski, A., Hejnowicz, M. S., Stepień, P. P. & Szczesny, R. J. *Nucleic Acids Res.* **41**, 1223–1240 (2013).

This article was published online on 24 February 2021.

maturity in the engineering of argument technology, coupled with intense commercial demand, has led to rapid expansion of the field. More than 50 laboratories worldwide are working on the problem, including teams at all the large software corporations.

One of the reasons for the explosion of work in this area is that direct application of AI systems that can recognize the statistical regularities of language use in large bodies of text has been transformative in many applications of AI (see ref. 3, for example), but has not, on its own, been as successful in argument mining. This is because argument structure is too varied, too complex, too nuanced and often too veiled to be recognized as easily as, say, sentence structure. Slonim *et al.* therefore decided to initiate a grand challenge: to develop a fully autonomous system that can take part in live debates with humans. Project Debater is the culmination of this work.

Project Debater is, first and foremost, a tremendous engineering feat. It brings together new approaches for harvesting and interpreting argumentatively relevant



Figure 1 | Project Debater takes on a human opponent. Slonim *et al.*¹ have developed Project Debater, an AI system that can take part in debating competitions with humans.

material from text with methods for repairing sentence syntax (which enable the system to redeploy extracted sentence fragments when presenting its arguments; the role of their syntax-repair technology is modestly underplayed by the authors). These components of the debater system are combined with information that was pre-prepared by humans, grouped around key themes, to provide knowledge, arguments and counter-arguments about a wide range of topics. This knowledge base is supplemented with ‘canned’ text – fragments of sentences, pre-authored by humans – that can be used to introduce and structure a presentation during a debate.

Project Debater is extraordinarily ambitious, both as an AI system and as a grand challenge for AI as a field. As with almost all AI research that sets its sights so high, a key bottleneck is in acquiring enough data to be able to compute an effective solution to the set challenge⁴. Project Debater has addressed this obstacle using a dual-pronged approach: it has narrowed its focus to 100 or so debate topics; and it harvests its raw material from data sets that are large, even by the standards of modern language-processing systems.

In a series of outings in 2018 and 2019, Project Debater took on a range of talented, high-profile human debaters (Fig. 1), and its performance was informally evaluated by the audiences. Backed by its argumentation techniques and fuelled by its processed data sets, the system creates a 4-minute speech that opens a debate about a topic from its repertoire, to which a human opponent responds. It then reacts to its opponent’s points by producing a second 4-minute speech. The opponent replies with their own 4-minute rebuttal, and the debate concludes with both participants giving a 2-minute closing statement.

Perhaps the weakest aspect of the system

is that it struggles to emulate the coherence and flow of human debaters – a problem associated with the highest level at which its processing can select, abstract and choreograph arguments. Yet this limitation is hardly unique to Project Debater. The structure of argument is still poorly understood, despite two millennia of research. Depending on whether the focus of argumentation research is language use, epistemology (the philosophical theory of knowledge), cognitive processes or logical validity, the features that have been proposed as crucial for a coherent model of argumentation and reasoning differ wildly⁵.

Models of what constitutes good argument are therefore extremely diverse⁶, whereas models of what constitutes good debate amount to little more than formalized intuitions (although disciplines in which the goodness of debate is codified, such as law and, to a lesser extent, political science, are ahead of the game on this front). It is therefore no wonder that Project Debater’s performance was evaluated simply by asking a human audience whether they thought it was “exemplifying a decent performance”. For almost two thirds of the debated topics, the humans thought that it did.

A final challenge faced by all argument-technology systems is whether to treat arguments as local fragments of discourse influenced by an isolated set of considerations, or to weave them into the larger tapestry of societal-scale debates. To a large degree, this is about engineering the problem to be tackled, rather than engineering the solution. By placing *a priori* bounds on an argument, theoretical simplifications become available that offer major computational benefits. Identifying the ‘main claim’, for example, becomes a well-defined task that can be performed almost as reliably by machine as by

humans^{7,8}. The problem is that humans are not at all good at that task, precisely because it is artificially engineered. In open discussions, a given stretch of discourse might be a claim in one context and a premise in another.

Moreover, in the real world, there are no clear boundaries that delimit an argument: discourses that happen beyond debating chambers are not discrete, but connect with a web of cross-references, analogy, exemplification and generalization. Ideas about how such an argument web might be tackled by AI have been floated in theory⁹ and implemented using software – a system called DebateGraph (see go.nature.com/30g2ym4), for example, is an Internet platform that provides computational tools for visualizing and sharing complex, interconnected networks of thought. However, the theoretical challenges and socio-technical issues associated with these implementations are formidable: designing compelling ways to attract large audiences to such systems is just as difficult as designing straightforward mechanisms that allow them to interact with these complex webs of argument.

Project Debater is a crucial step in the development of argument technology and in working with arguments as local phenomena. Its successes offer a tantalizing glimpse of how an AI system could work with the web of arguments that humans interpret with such apparent ease. Given the wildfires of fake news, the polarization of public opinion and the ubiquity of lazy reasoning, that ease belies an urgent need for humans to be supported in creating, processing, navigating and sharing complex arguments – support that AI might be able to supply. So although Project Debater tackles a grand challenge that acts mainly as a rallying cry for research, it also represents an advance towards AI that can contribute to human reasoning – and which, as Slonim *et al.* put it, pushes far beyond the comfort zone of current AI technology.

Chris Reed is at the Centre for Argument Technology, University of Dundee, Dundee DD1 4HN, UK.
e-mail: c.a.reed@dundee.ac.uk

1. Slonim, N. *et al.* *Nature* **591**, 379–384 (2021).
2. Lawrence, J. & Reed, C. *Comput. Linguist.* **45**, 765–818 (2020).
3. Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. Preprint at <https://arxiv.org/abs/1810.04805> (2018).
4. Feigenbaum, E. A. *Proc. 5th Int. Joint Conf. Artif. Intell.* 1014–1029 (Morgan Kaufman, 1977).
5. van Eemeren, F. H. *et al.* *Handbook of Argumentation Theory* (Springer, 2014).
6. Hahn, U. *Trends Cogn. Sci.* **24**, 363–374 (2020).
7. Levy, R., Yonatan, B., Hershcovich, D., Aharoni, E. & Slonim, N. *Proc. COLING 2014, 25th Int. Conf. Comput. Linguist. Tech. Pap.* 1489–1500 (2014).
8. Trautmann, D., Daxenberger, J., Stab, C., Schütze, H. & Gurevych, I. *Proc. AAAI Conf. Artif. Intell.* **34**, 9048–9056 (2020).
9. Rahwan, I., Zablith, F. & Reed, C. *Artif. Intell.* **171**, 897–921 (2007).