From: John Brockman <[redacted]>
To: Jeffrey Epstein <jeevacation@gmail.com>
Subject: the meeting
Date: Sun, 09 Sep 2018 22:30:17 +0000

Did you tell the group of my involvement

The subject of money never came up. Nobody asked. I didn't offer.

In talking to Bob Axelrod, I told him you funded people like Seth, Nowak, Lee, and were a regular backer of EDGE, but it was not specific to the meeting. Still, there were some negative vibes, all about "where are the women?", specially from the women. And this didn't come close to anything re: Me-too. Instead of saying thank you for a spectacularly interesting weekend, Katinka and I got crapped on for out efforts. This led me to ask who did I offend?, to whom do I owe an apology. Overall, the situation is worse than ever and I don't see it getting better.

I'll be in the office from noon on. Call me. And thanks again.

JB

Here are the paragraphs for the talks. (Kahneman is quite ill and couldn't come, by the way.)

TALKS

Here are the one-paragraph statements on what each participant plans to talk about.

ROBERT AXELROD
Walgreen Professor for the Study of Human Understanding at the University of Michigan, best known for his interdisciplinary work on the evolution of cooperation; author of The Complexity of Cooperation and The Evolution of Cooperation.

Cooperation achieves its beneficial effects by improving communication, promoting gains from specialization, enhancing organizational effectiveness, and reducing the risks of harmful conflict. Members of an institutionalized academic discipline jointly benefit in all these ways. Unfortunately, members of different disciplines typically do not. The boundaries of most disciplines were largely set 100 (plus or minus 50) years ago, and efforts to redraw the boundaries (e.g. at Irvine and Carnegie Mellon) have not met with much success. I would like us to consider how the more or less fragmented research community can best respond to new opportunities (e.g. AI), new problems (e.g. climate change), new modes of education and governance, and new understandings of human behavior and values.

ROD BROOKS

Computer scientist; Panasonic Professor of Robotics, emeritus, MIT; former director, MIT Computer Science Lab; and founder, chairman, and CTO of Rethink Robotics. He is the author of Flesh and Machines.

Have we gotten into a cul-de-sac in trying to understand animals as machines from the combination of digital thinking and the crack cocaine of computation uber alles that Moore's law has provided us? What revised models of brains might we be looking at to provide new ways of thinking and studying the brain and human behavior? Did the Macy conferences get it right? Is it time for a reboot?

DAVID CHALMERS
University Professor of Philosophy and Neural Science, and Co-Director of the Center for Mind, Brain, and Consciousness at New York University; Distinguished Professor of Philosophy at the Australian National University; best known for his work on consciousness, including his formulation of the "hard problem" of consciousness.

Would every possible mind face a mind-body problem? Once we develop AI systems that can reflect on themselves and reason, will they invariably report that their minds seem from the inside to be more than a collection of circuits? These questions are closely tied to the meta-problem of consciousness: give an algorithmic explanation of why we're bothered by the problem of consciousness. That's a tractable project for philosophy/psychology/AI that may just shed light on the mind-body problem itself

FREEMAN DYSON
Professor of physics at the Institute for Advanced Study in Princeton who has worked on nuclear reactors, solid-state physics, ferromagnetism, astrophysics, and biology, looking for problems where elegant mathematics could be usefully applied; his books include Disturbing the Universe, Weapons and Hope, Infinite in All Directions, and Maker of Patterns.

I am asking whether our brains might be quantum analog computers. I believe this possibility was first suggested by Richard Feynman. The brain might be an amplifier, sensitive to the quantum states of memory molecules, and amplifying the molecular information until it becomes a signal strong enough to drive motor neurons to action. Quantum jumps in the memory, unpredictable according to quantum mechanics, would control executive decisions. Philosophers would continue to argue whether this gives us free will. Experimenters must learn how to observe in detail what goes on inside the head of a three-month-old baby. ow does that little head sort out the neural inputs from eyes and ears, recognize faces and voices, master grammar and syntax, know the difference between nouns and verbs, and learn how to exploit the weaknesses of grown-ups?

GEORGE DYSON
Historian of science and technology; author of Baidarka: the Kayak; Darwin Among the Machines, Project Orion, and Turing's Cathedral.

Nature's response to those who believe they can build machines to control everything will be to allow them to build a machine that controls them instead. We are stuck in this digital mindset—believing that machine learning, deep learning, etc. is just a way to cultivate better algorithms that the usual actors can domesticate and sell. Analog computing has no algorithms. They aren't there. People who think they are hidden and waiting to be understood and explained (and controlled) are just fooling themselves. Nature discovered this on her own and so will machines, whether we recognize it or not.

PETER GALISON
Science historian; Joseph Pellegrino University Professor and co-founder of the Black Hole Initiative at Harvard University; author of Einstein's Clocks and Poincaré's Maps: Empires of Time.

For centuries, scientists and scientifically-minded philosophers have argued over what it is that science wants—or should want. Is it a causal account of the world? Is it prediction? Is it understanding? For many branches of scientific inquiry Newton's inverse square law was not only a profound insight into tides, planets, moons, comets, galaxies, it was also a model for laws more generally. Darwin's epochal theory joined and explained like nothing the biological sciences had ever seen—but the impact of the account did not aim for the kind of prediction Newton sought when he focused on the moons of Jupiter When Monte Carlo simulations entered and later prospered throughout the physical (and other) sciences, they were simultaneously celebrated and derided for pushing aside an absolute focus on fundamental laws—prediction might do very well to shield the operator of a nuclear power plants even absent a latter-day 1/r2. Maybe our absolutist arguments about AI and science aren't what we need: perhaps we really are perfectly happy to put all stress on prediction when it comes to the design of efficacious drugs or decent-enough foreign language translations… and yet want something very different in grasping the confluence of quantum field theory and general relativity, or in understanding why we sentence different people to different sentences for similar crimes. Systematic reason has never wanted one thing across its precincts: Maybe in the age of AI we don't want and don't need one unified set of virtues now.

NEIL GERSHENFELD
Physicist; Director of MIT's Center for Bits and Atoms; founder of the global fab lab network; author of FAB, co-author (with Alan Gershenfeld & Joel Cutcher-Gershenfeld) of Designing Reality.

I consider computer science to be one of the worst things to have happened to either computers or science, because it's based on a fiction that digital is not physical. It's had a good run, but cracks appearing in that matrix are indications of the need for a do-over. Embracing rather than avoiding the reality that information is physical highlights a number of prevailing false dichotomies. One is that digital and analog are in opposition; analog degrees of freedom can be used to solve digital problems more effectively, in a way that could be understood in an introductory course on optimization but would confound a neurobiologist. Another is the segregation of computation from fabrication, which merge in what is literally the mother of all algorithms, morphogenesis.

ALISON GOPNIK
Developmental psychologist at UC Berkeley; her books include The Philosophical Baby and, most recently, The Gardener and the Carpenter: What the New Science of Child Development Tells Us About the Relationship Between Parents and Children.

Back in 1950, Turing argued that for a genuine AI we might do better by simulating a child's mind than an adult's. This insight has particular resonance given recent work on "life history" theory in evolutionary biology —the developmental trajectory of a species, particularly the length of its childhood, is highly correlated with adult intelligence and flexibility across a wide range of species. This trajectory is also reflected in brain development, with its distinctive transition from early proliferation to later pruning. I've argued that this developmental pattern reflects a distinctive evolutionary way of resolving explore-exploit tensions that bedevil artificial intelligence. Childhood allows for a protected period of broad, high-temperature search through the space of solutions and hypotheses, before the requirements of focused, goal-directed planning set in. This distinctive exploratory childhood intelligence, with its characteristic playfulness, imagination and variability, may be the key to the human ability to innovate creatively yet intelligently, an ability that is still far beyond the purview of AI. More generally, a genuine understanding of intelligence requires a developmental perspective.

TOM GRIFFITHS
Henry R. Luce Professor of Information, Technology, Consciousness, and Culture at Princeton University; co-author (with Brian Christian) of Algorithms to Live By.

The success of deep learning has largely been a consequence of the availability of increasingly large data sets and increasingly more computational resources for processing those data sets. But maybe more data and more computation are leading us down the wrong track for building systems that display the same kind of general intelligence as people. When we look at what makes human learning impressive, it's often the ability to learn from small amounts of data. Likewise, our limited onboard computational resources have forced humans to develop a sophisticated ability to make decisions about how to efficiently deploy those resources: when faced with new problems, we are able to develop strategies and heuristics that allow us to use the skills and knowledge we have already acquired to our best ability. Thinking about how to do more with less might be the path to the next AI revolution.

W. DANIEL HILLIS
Inventor, entrepreneur, and computer scientist; pioneer of the concept of parallel computers that is now the basis for most supercomputers, as well as the RAID disk array technology; The Judge Widney Professor of Engineering and Medicine at USC; author of The Pattern on the Stone: The Simple Ideas That Make Computers Work.

While be we have all been distracted by the hypothetical emergence of computer-based intelligence's we have missed noticing that technology-enabled super-human intelligences have already emerged. We originally created this these intelligences as corporations, NGO, and nation states to serve us, but in many ways, they have grown more powerful than us, and they have goals of their own. We don't notice these super intelligences as such because they live on the substrate of humans and technology that is us: to us fish, they are water. I believe that most of the concerns raised about AIs could be better address by refocusing them from the hypothetical to these actual examples.

CAROLINE A. JONES
Professor of art history in the Department of Architecture at MIT; author of Eyesight Alone: Clement Greenberg's Modernism and the Bureaucratization of the Senses; Machine in the Studio: Constructing the Postwar American Artist; and The Global Work of Art.

Let's question together what we mean when we refer to "Intelligence" in machines. Rather than what Scientific American ridiculed as the "Oz behind the curtain" model of brain the master-controller, we need a much more distributed notion that goes well beyond the sacred cranium and may not be bounded by our skin. The world of art (with which I am most familiar) has been critiquing the cranial paradigm for decades now. Multi-sensorial kinds of art involving new media and immersive installations have led visitors to acknowledge that they "know" the art through infrasonic vibrations in viscera (sound art), or through a diffused haptic response that seems to involve a highly distributed mirror system, or through non-verbal proxemics. In parallel to this humbling proliferation of non-cranial and non-visual aesthetic forms are astonishing new insights from biology that confound the boundaries between mind and body or brain and viscera. The gut-brain axis (in which mental health relies on partnerships with xenobacterial microbiota), or revelations about the "immune brain" distributed through our lymph system, should give the narrow definers of intelligence extreme pause. The adaptive learning system of our "meat machines" (e.g., the immune system) that gains and recalls knowledge about friends and foes every time we put something in our mouth, take it into a cut in our shin, or breathe it in our nasal mucosa is a general intelligence—one that seems to be completely independent of our conscious thought. This adaptive and responsive wetware, and its dependence on a larger living ecosystem, is something I recommend we try to understand more fully before claiming that it is "intelligence" we've produced in our machines.

DANIEL KAHNEMAN
Nobel Laureate in Economic Sciences (2002); Eugene Higgins Professor of Psychology Emeritus at Princeton University, Professor of Psychology and Public Affairs Emeritus at the Woodrow Wilson School; winner of the 2013 Presidential Medal of Honor; author of Thinking, Fast and Slow.

My late teacher Yehoshua Bar-Hillel was once asked, in the 1950's, whether computers would ever understand language. He answered unhesitatingly "Never" and immediately clarified that by "Never" he meant "at least 50 years." I am puzzled by the number of references to what AI "is" and what it "cannot do" when in fact the new AI is less than ten years old and is moving so fast that references to it in the present tense are dated almost before they are uttered. The statements that AI doesn't know what it's talking about or is not enjoying itself are trivial if they refer to the present and undefended if they refer to the medium-range future—say 30 years. Hype is bad, but the debunkers should remember that the AI Winter was brought about by two brilliant people proving what a one-layer Perceptron could not do. My optimistic question would be "Where will the next breakthrough in AI come from—and what will it easily do that deep learning is not good at?"

SETH LLOYD
Theoretical physicist at MIT; Nam P. Suh Professor in the Department of Mechanical Engineering; external professor at the Santa Fe Institute; pioneer in quantum computation, quantum communication and quantum biology, including proposing the first technologically feasible design for a quantum computer; author of Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos.

I am working on problems of quantum machine learning, developing novel protocols that allow quantum systems to find patterns in nature that can't be revealed by classical machine learning algorithms. I am also working with experimentalists to implement these quantum machine learning protocols on photonic and superconducting quantum computers. As part of this research, my collaborators and I are trying to develop a general theory of how quantum systems obtain information about the world, and how they use that information to harvest free energy from their environment. We are applying this theory to understand how pre-biotic systems —before the existence of DNA, RNA, or self-reproduction—compete to gather energy, and how that competition gives rise to complex structures of energy harvesting and information processing. That is, we are investigating how the universe begins to compute.

IAN MCEWAN

Novelist whose works have earned him worldwide critical acclaim; recipient, the Man Booker Prize for Amsterdam (1998), the National Book Critics' Circle Fiction Award and the Los Angeles Times Prize for Fiction for Atonment (2003). His most recent novel is On Chesil Beach.

I would like to set aside the technological constraints in order to imagine how an embodied artificial consciousness might negotiate the open system of human ethics—not how people think they should behave, but how they do behave. For example, we may think the rule of law is preferable to revenge but matters get blurred when the cause is just and we love the one who exacts the revenge. A machine incorporating the best angel of our nature might think otherwise. The ancient dream of a plausible artificial human might be scientifically useless but culturally irresistible. At the very least, the quest so far has taught us just how complex we (and all creatures) are in our simplest actions and modes of being. There's a semi-religious quality to the hope of creating a being less cognitively flawed than we are.

FRANK WILCZEK

Nobel Laureate in Physics (2004); Herman Feshbach Professor of Physics at MIT; Director Quantum Chinal; author of A Beautiful Question: Finding Nature's Deep Design.

Fundamental comparison of the capabilities of natural and existing artificial intelligence shows that both have striking advantages. This is reflected in their performance. The present relevant disadvantages of artificial intelligence plausibly can be addressed by a new kind of engineering which features self-reproducing units and winnowing of massively connected networks, as in the development and learning of human brains. This will take considerable time to accomplish though, so I foresee what physicists call a crossover, rather than a singularity. That is fortunate, since it will allow AI morality and social behavior to evolve using input from practical experience. The contrary idea, that one could program morality and social behavior, in the style of expert systems, seems to me to fly in the face of experience teaching AI how to do things we humans do without knowing how we do them (e.g. image processing, locomotion, language).

STEPHEN WOLFRAM
Scientist, inventor, and the founder and CEO of Wolfram Research; creator of the symbolic computation program Mathematica and its programming language, Wolfram Language, as well as the knowledge engine Wolfram|Alpha; author of A New Kind of Science.

I've spent several decades creating a computational language that aims to give a precise symbolic representation for computational thinking, suitable for use by both humans and machines. I'm interested in figuring out what can happen when a substantial fraction of humans can communicate in computational language as well as human language. It's clear that the introduction of both human spoken language and human written language had important effects on the development of civilization. What will now happen (for both humans and AI) when computational language spreads?