Epsilon Theory

Data Access Battles, Creative Thinking & Full Script AI

Data access battles

A couple of weeks back I shared a link to the story of ImageNet and the importance of data to developing algorithms. *Ars Technica* reports on two 'at the coalface' battles over data access with HiQ and Power Ventures fighting with LinkedIn and Facebook over data access. I'm not advocating a position on this but, to be sure, small — and currently obscure — court cases like these will, cumulatively, end up setting the precedents which will have a significant impact on the evolution and ownership of powerful algorithms that are increasingly driving behavior and economics.

Creative thinking

This speech from Claude Shannon at Bell Labs in 1952 has been circulating online for the past couple of weeks. It is a timeless, pragmatic speech on creative thinking which remains, 65 years later, fully relevant for developing novel computational strategies:

Sometimes I have had the experience of designing computing machines of various sorts in which I wanted to compute certain numbers out of certain given quantities. This happened to be a machine that played the game of nim and it turned out that it seemed to be quite difficult. It took quite a number of relays to do this particular calculation although it could be done. But then I got the idea that if I inverted the problem, it would have been very easy to do — if the given and required results had been interchanged; and that idea led to a way of doing it which was far simpler than the first design. The way of doing it was doing it by feedback; that is, you start with the required result and run it back until — run it through its value until it matches the given input. So the machine itself was worked backward putting range S over the
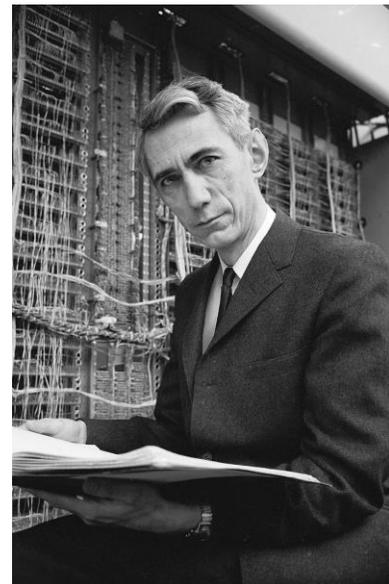
numbers until it had the number that you actually had and, at that point, until it reached the number such that P shows you the correct way.

Facebook shuts down robots after they invent their own language

*"**Facebook shuts down robots after they invent their own language**"* has become a widely reported and wildly commented story over the past month, referencing a story on 'Tricky chatbots' linked **here** a couple of months back. For melodramatic illustrative effect, I like switching a couple of words in the Facebook headline so that it reads *'Lehman (doesn't) shuts down traders after they invent their own language'* as it illustrates that, in general, if you: put a bunch of agents (human or machine) together and set up a narrowly defined, adversarial, multi-player game with a strong reward function then the agents will develop their own task-specific language and protocols, keep adding complexity, lie to each other (yes, the FB bots also learnt to do that), be tempted to obfuscate behavior in order to reduce interference and maximize the reward function, and develop models which are positive for near-term reward maximization but do not necessarily deal with longer-term consequence or long tail events, and so become very hard for human overseers to truly assess…

DICK FULD (2008):
I wake up every single night wondering what I could have done differently — this is a pain that will stay with me the rest of my life

FACEBOOK (2017):
Hold my beer

Facebook's AI language

Bob: i can i i everything else . . . .

Alice: balls have zero to me to me to me to me to me to me to me to me to me

Bob: you i everything else . . . .

Alice: balls have a ball to me to me to me to me to me to me to me

Bob: i i can i i i everything else . . . .

Alice: balls have a ball to me to me to me to me to me to

Bob: i . . . .

⊖ Epsilon Theory

AI: From partial to full script

Thinking more broadly about the longer-term evolution of AI (and the nature of money and contracts, per *Ethereum* link last week), it has been interesting to re-read **Sapiens: A Brief History of Humankind** by Yuval Noah Harari which charts the rise to dominance of us Sapiens with especially interesting chapters on the development of written language and money. A concept which particularly grabbed me was that written language was initially developed as 'partial script' technology for narrow tasks such as tax accounting, and then evolved to be full script and so capable of much more than it was originally conceived for.

The history of writing is almost certainly a wonderful historical premonition of the trajectory of AI, except with the evolution being much faster and the warning that likely "the AI is more powerful than pen."

Relevant excerpt from Sapiens:

> Full script is a system of material signs that can represent spoken language more or less completely. It can therefore express everything people can say, including poetry. Partial script, on the other hand, is a system of material signs that can represent only particular types of information, belonging to a limited field of activity … It didn't disturb the Sumerians (who invented the script) that their script was ill-suited for writing poetry. They didn't invent it in order to copy spoken language, but rather to do things that spoken language failed at … Between 3000 BC and 2500 BC more and more signs were added to the Sumerian system, gradually transforming it into a full script that we today call cuneiform. By 2500 BC, kings were using cuneiform to issue decrees, priests were using it to record oracles, and less-exalted citizens were using it to write personal letters.

⊖ Epsilon Theory

The beautiful mathematical explorations of Maryam Mirzakhani

And finally, at the risk of turning into *The Economist*, we conclude this week's *Rabbit Hole* with a touching **obituary** of the Tehran-born, Fields Medal-winning mathematician Maryam Mirzakhani:

A bit more than a decade ago when the mathematical world started hearing about Maryam Mirzakhani, it was hard not to mispronounce her then-unfamiliar name. The strength and beauty of her work made us learn it. It is heartbreaking not to have Maryam among us any longer. It is also hard to believe: The intensity of her mind made me feel that she would be shielded from death.

⊖ Epsilon Theory