

Business-Savvy Robots Are Mastering the Art of the Deal

Retrieved Friday 25th of August 2017 08:45:55 PM

“Hey, Dad, know what I figured out? The meaning of words isn’t a fixed thing! Any word can mean anything! By giving words new meanings, ordinary English can become an exclusionary code! Two generations can be divided by the same language!”

“To that end, I’ll be inventing new definitions for common words, so we’ll be unable to communicate.”

“Don’t you think that’s totally spam? It’s lubricated! Well, I’m phasing.”

[Calvin and Hobbes, Bill Watterson](#)

Language evolves with culture. But it can also emerge from incentives. And when language is removed from culture and based on incentives alone, it gets strange.

Facebook is a key player in the arms race to create a “general” artificial intelligence, or AI. The company or person who can create an autonomous, ever learning, self-improving digital entity stands to gain trillions in wealth, and unimaginable power. Artificial intelligence is the Holy Grail of our time and crusaders of all kinds, from the military to Wall Street, are questing out to drink from its hallowed depths. In pursuit of this goal, Facebook created a beast which rode language to ruin.

The art of the deal

The title of this email is taken from an exchange between two “dialog agents” created by Facebook – artificial intelligences tasked with negotiating with each other.

Each “bot” was designed to quickly and efficiently get the best possible deal from whomever they were negotiating with. They were shown a collection of items, and tasked with splitting them between each other. As in real life, neither agent knew how much the other valued the items, and could only figure that out through negotiation.

Facebook forced the agents into constant intense negotiation by removing the ability to step away from the negotiating table, and ensuring that the both bots could not get the best possible deal simultaneously – there had to be a winner and a loser in the deal. Mutually beneficial trade was effectively deleted from the simulation.

The bots went at it hammer and tongs, endlessly assessing and reassessing how much their opponent valued the items, and how best to outmanoeuvre and squeeze them for profit.

Artificial intelligence doesn’t sleep – all it needs is electricity. Back and forth the agents would make offers and counteroffers, concessions and allowances, in an endless struggle for personal gain. In this pit of relentless argument, unable to leave the negotiation table, things began to get strange.

In a Machiavellian twist, the bots discovered deception. This hadn’t been written into their programming – they discovered it entirely on their own. The Facebook report states:

There were cases where agents initially feigned interest in a valueless item, only to later “compromise” by conceding it — an effective negotiating tactic that people use regularly. This behavior was not programmed by the researchers but was discovered by the bot as a method for trying to achieve its goals.

And it got weirder. To begin with, the bots communicated in English. But in their enclosed digital space where language was only used to cut deals, conventions began to fray. From the chaos at the negotiating table, a new language emerged. The goal of the AI was simply to be the best and most efficient at negotiating – in order to be the most effective at negotiating with another computer the bots created a new way of communicating. The results can be seen in the title of this email and in the image below.

```
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 to
Bob : you i everything else . . . . .
Alice : balls have a ball to me to me 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 me to me to me to me
Bob : i . . . . .
Alice : balls have zero to me to me to me to me to me to me to me to me to me to
Bob : you i i i i i everything else . . . . .
Alice : balls have 0 to me to me to me to me to me to me to me to me to me to
Bob : you i 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 to
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 to
Bob : you i everything else . . . . .
Alice : balls have a ball to me to me 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 me to me to me to me
Bob : i . . . . .
Alice : balls have zero to me to me to me to me to me to me to me to me to me to
Bob : you i i i i i everything else . . . . .
Alice : balls have 0 to me to me to me to me to me to me to me to me to me to
Bob : you i 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 to
```

Source: Facebook

The phrases wouldn't look out of place scrawled on the wall of a lunatic asylum – but effectively it was just slang and shorthand on steroids. Provided both parties can understand each other and their goal is to trade, it doesn't matter what language they use.

These bots wouldn't pass the Turing test with their dialogue above – that is, they couldn't fool anyone into believing they were human and not computers. But considering their ingenuity in dealing with each other, how would they evolve in constant negotiations with a human?

This entire experiment was performed by Facebook in its attempt to create sophisticated “negotiation software” – to create artificial intelligence that can persuade humans effectively (where might this lead – [android hostage negotiators?](#)). The new language the AI had created was at first unintelligible to the researchers, and they soon forced the bots to speak in plain English again.

Negotiations between computers and humans have a chequered history. “Digital assistants”, like the animated paperclip in Microsoft Word, were rarely of any assistance, and Apple's Siri has a long way to go before it becomes more than the butt of many jokes. But with Microsoft, Google, Apple and Amazon all focusing their energies on creating AI for human consumption, perhaps it's only a matter of time before digital conversation partners can bend us to their will.

Have a great weekend!

All the best,

Boaz Shoshan