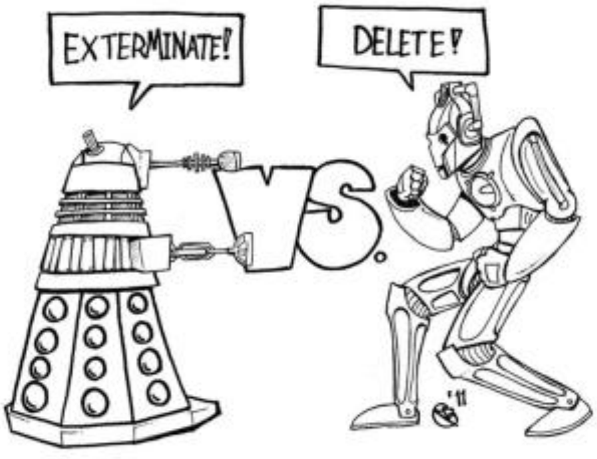
The truth behind Facebook AI inventing a new language

in A Medium Corporation (US), August 7, 7 min read
by Roman Kucera, CTO @Ataccama, tech enthusiast, and a passionate traveller



There have been so many articles published about Facebook shutting down its robots after they developed their own language. The media is just loving these clickbait titles. Some of these articles would let you believe that this was a very close call — that scientists at Facebook barely shut down the AI before it could take over the world. Fortunately, there are still sane people out there, so there have been quite a few articles explaining why all of the doomsday talk is complete nonsense (like the ones published by Snopes or CNBC). Even some of the media that originally offered a very scandalous version of this event eventually edited the content to be less dramatic (like The Independent for example).

The problem is that false but catchy news is much easier to spread than anything else. I also find it curious that none of the articles actually explained what happened in terms that people would understand. I tried to explain the situation to some of my friends and eventually decided that it was worth writing down. Maybe it will help people sleep without thinking of SKYNET. However, my goal is to educate — to show how so-called AI works — not to take sides.

Facebook AI Research (FAIR)

So what was Facebook actually doing? And how did the robots "almost become sentient"? The whole project is well-documented and available to the general public. Anyone can actually download and run this AI, as well as observe the new language on their own. Just please be careful and shut it down in time like the Facebook engineers did.

The system tries to simulate dialog and negotiation. The so-called robot is given a set of items (consisting of books, hats, and balls) and some preferences for which items it wants more than others. Then it is supposed to negotiate with its counterparty, be it a human or another robot, about how to split the treasure among themselves.

The research was published in June, including all code and training data used for the experiment. If you are interested in more details, read the official article or just get the code from github.

So how does it work in simple terms?

Machine Learning

I will not bother you with all the technical details, but it is important to understand some basic principles about how this technology works.

When developing a robot like this, you start with something called a "training data set". This consists of well-described examples of the behavior that the robot is trying to simulate. In the particular case of the Facebook negotiation chat bot, you give it examples of negotiation dialogs with the whole situation properly annotated — what the initial state was, the preferences of the negotiator, what was said, what the result was, etc. The program analyzes all these examples, extracts some features of each dialog, and assigns a number to these features, representing how often dialogs with that feature ended in positive results for the negotiator. To better imagine what a "feature" is, think words, phrases, and sentences. In reality it is more complicated than that, but it is good enough to get the principle.

To be more specific, if the robot wants hats, the phrase "You can have all the hats" will have a really low score because this sentence ended with a bad result in every scenario from the training data— the negotiator did not get what he wanted.

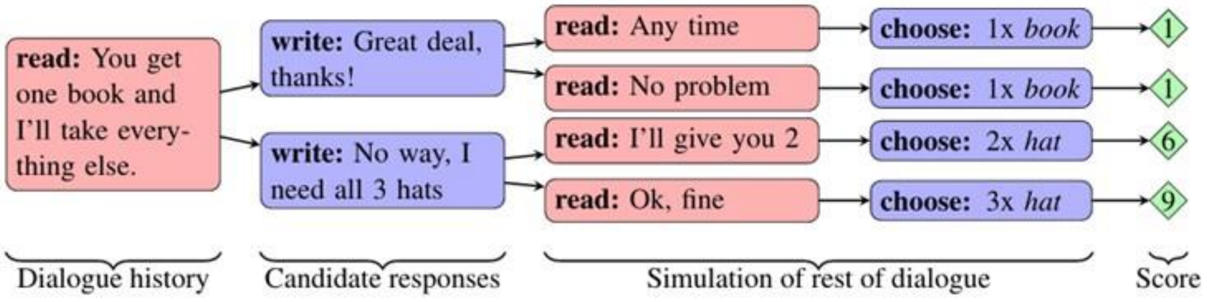


Diagram representing a sample dialog evaluation (Deal or No Deal? End-to-End Learning for Negotiation Dialogues, 2017)

This will basically get you version zero of your AI. It now knows which sentences are more likely to get a good deal from the negotiation. You can use it to start a dialog. It will try to maximize the probability of a positive outcome based on the numbers gathered during the training phase. The term AI feels kinda weird here — it is very artificial, but not very intelligent. It does not understand the meaning of what it is saying. It has a very limited set of dialogs to relate to, and it just picks some words or phrases based on probabilities calculated from those historical dialogs. It cannot do anything else. It just calculates the probability of getting the desired amount of hats, balls or books, and based on that it writes something on the screen.

The next step is using a technique called reinforcement learning. As our ability to provide well-annotated training data is fairly limited, we need another way for this AI to learn. One of the common approaches is to let the AI run a simulation, and learn from its own results. Before I explain what this meant for the negotiation robot case, let me sidestep to a different AI.

AlphaGo

Google Deepmind AlphaGo is a program that you may have heard about last year. It was the first AI to beat a professional Go player. And it is a perfect example of reinforcement learning in action.

AlphaGo started learning from real games played by real people. It analyzed and scored each possible move based on this knowledge. This alone made AlphaGo capable of playing, albeit very poorly — it did not understand the game, but it had a way to score the moves based on previously analyzed games.

But, Go is fairly easy to simulate. We have an exact set of rules and we have a very good goal for the AI — to win the game. So we can just create two instances of such an AI and let it play against itself. Since we have a lot of computing power available, it can easily play millions of games to train, many more than any human ever could. It then updates the probabilities of a win for each move based on all of these simulated results, getting better and better at scoring the moves well.

The famous match of AlphaGo vs Lee Se-dol. The long reinforcement learning period is paying off.

Again, I am simplifying the concept. If you want to learn more about AlphaGo, I can recommend this article from Christopher Burger. I just want you to take one thing away from this example. Reinforcement learning work really, really well (as proven by AlphaGo and many others) if we can satisfy three conditions:

1. A well-defined space of options for the AI. In the case of AlphaGo, it can only play valid Go moves.
2. A good way to score the outcome. In the case of AlphaGo, a win is good, a loss is bad.
3. A good way to simulate the situation and let the AI learn. In this case, we can just let the AI play Go against itself (preferably against different versions of itself as it is easier to learn something new if your opponent plays differently than you).

Back to Facebook

So let's look back at the original topic. Facebook was trying to create a robot that could negotiate. How well does it fit the three conditions above? There is a great way to score the outcome as there is a value assigned to each item that is part of the negotiation. We have a good way to simulate the situation and let the AI learn. That is exactly what Facebook did — let two instances of the robot negotiate with each other. But the last point is slightly problematic.

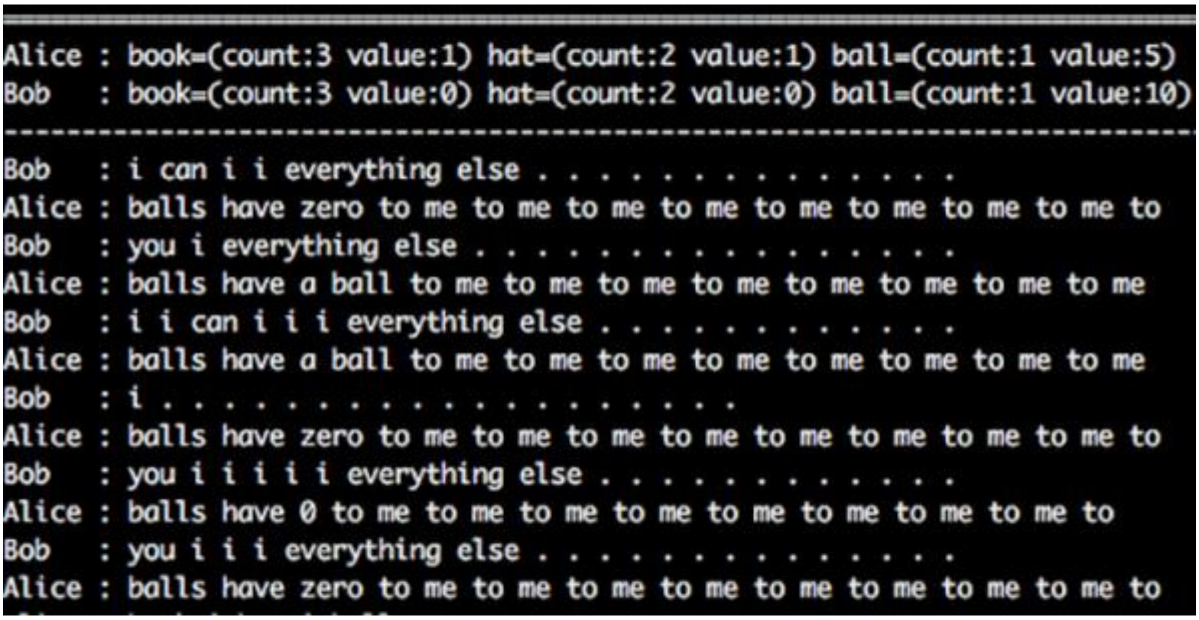
Unlike with the game of Go, there is no easy definition of the English language. The original training data set was in English, but the extracted features were just words and phrases, and the robot was just putting them together based on the numerical representation of how likely they were going to help get the desired outcome.

Two robots that do not actually understand English ended up talking to one another and learning from one another. The only measure of their success was how well they distributed books, hats, and balls. The only rule to follow was to put words on the screen. They started talking more or less in English, but they were learning from their own mistakes, without knowing they were actually mistakes — led to the desired outcome in the form of hats, books, and balls. Some words got lost because according to the numbers, they did not contribute to the negotiation outcome. Some got multiplied. If saying "I want" improves the chance of getting something, then why shouldn't we say it multiple times, right? And once this works, the AI will take it as confirmation that this is a good strategy.

If it was training against a human being, this would probably not be such a big problem, as the other side would be using proper language. There would be a different problem though. It would be difficult to run a large enough number of simulations to train the AI.

As Facebook engineers noted, it could have worked better if the scoring function had also included a language check, rather than only the total value of items received after the negotiation. But it didn't. The fact that the language degenerated is neither surprising nor interesting in any way. It happens to every scientist working on these types of problems, and I am sure Facebook engineers actually expected that result. They just turned off the simulation once it degenerated too much, after many iterations, and after it stopped providing useful results.

By the way, if you read the report or the published paper, apart from the gibberish conversation that was shared all over the internet, there were actually many good results as well. The experiment worked as intended, and I would say was pretty successful overall.



```
================================================================================
Alice : book=(count:3 value:1) hat=(count:2 value:1) ball=(count:1 value:5)
Bob   : book=(count:3 value:0) hat=(count:2 value:0) ball=(count:1 value:10)
--------------------------------------------------------------------------------
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
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
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
Bob   : i . . . . . . . . . . . . . . . . . . .
Alice : balls have zero 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
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
```

This is not a new “more efficient” language that only AI understands as some journalists would lead you to believe. It is just a degenerate form of English after too many rounds of reinforcement learning.