

From: "jeffrey E." <jeevacation@gmail.com>

To: Joscha Bach <[REDACTED]>

Subject: Re: Mechanisms for learning

Date: Sun, 10 Jul 2016 11:50:13 +0000

what i like is the idea that ubermensch could be the melding of humans, put together in one brain an increased motor system, , an increased awareness , an increased status calculator. , (blacks jews, women) ubermensch could be the combination of the best of humans not the best of a specic race or gender. fun idea. we can call it super intelligence. or uberintelligence

. the image generator combines with watson, for inference space , might produce module and concept . but I dont think so. . look at melody, it has a compicated INTRA relationship as well as external. . more like a coherent shape. . that must be seen from outside. that being said there is a body of work on is the image real or fake. pixel by pixel analysis to see dirrences and ratio. if i am a note , can i tell within my dimension that i am part of a melody. or just random sounds. if i am a pixel can i tell i am part of a photo? , can a computer say , yes that is a melody , or not. how do we know in a mozart piece when heard for even the first time that the performer hit a " bad " note. . it did nt belong in the series. , what did that mean? it was a note in a different key. out of timing , out of harmony. ? beethoven was like in your drug state, bypassing ego and writing the music of his brain.

On Sun, Jul 10, 2016 at 12:42 AM, Joscha Bach <[REDACTED]> wrote:

Dear Jeffrey,

thank you for your support and encouragement, even where I fail.

Sorry for being such an embarrassment today. I will spell out today's argument a bit better and cohesive when I get to it. Also, I should have recognized that the main point I tried to make would trigger Noam (who was as always very generous, patient, kind and humble on the personal level, even though he did not feel like conceding anything on the conceptual one). Almost all of Noam's work focused on the idea that humans have very specific circuits or modules (even when most people in his field began to have other ideas), and his frustration is that it is so hard to find or explain them.

I found Noam's hypothesis very compelling in the past. I still think that the idea that language is somehow a cultural or social invention of our species is wrong. But I think that there is a chance (we don't know that, but it seems to most promising hypothesis IMHO) that the difference between humans and apes is not a very intricate special circuit, but genetically simple developmental switches. The bootstrapping of cognition works layer by layer during the first 20 years of our life. Each layer takes between a few months and a few years to train in humans. While a layer is learned, there is not much going on in the higher layers yet, and after the low level learning is finished, it does not change very much. This leads to the characteristic bursts in child development, that have famously been described by Piaget.

The first few layers are simple perceptual stuff, the last ones learn social structure and self-in-society. The switching works with something like a genetic clock, very slowly in humans, but much more quickly in other apes, and very fast in small mammals. As a result, human children take nine months before their brains are mature enough to crawl, and more than a year before they can walk. Many African populations are quite a bit faster. In the US, black children outperform white children in motor development, even in very poor and socially disadvantaged households, but they lag behind (and never catch up) in cognitive development even after controlling for family income.

Gorillas can crawl after 2 months, and build their own nests after 2.5 years. They leave their mothers at 3-4 years. Human children are pretty much useless during the first 10-12 years, but during each phase, their brains have the opportunity to encounter many times as much training data as a gorilla brain. Humans are literally smarter on every level, and because the abilities of the higher levels depend on those of the lower levels, they

can perform abstractions that mature gorillas will never learn, no matter how much we try to train them.

The second set of mechanisms is in the motivational system. Motivation tells the brain what to pay attention to, by giving reward and punishment. If a brain does not get much reward for solving puzzles, the individual will find mathematics very boring and won't learn much of it. If a brain gets lots of rewards for discovering other people's intentions, it will learn a lot of social cognition.

Language might be the result of three things that are different in humans:

- extended training periods per layer (after the respective layer is done, it is difficult to learn a new set of phonemes or the first language)
- more layers
- different internal rewards. Perhaps the reward for learning grammatical structure is the same that makes us like music. Our brains may enjoy learning compositional regular structure, and they enjoy making themselves understood, and everything else is something the universal cortical learning figures out on its own.

This is a hypothesis that is shared by a growing number of people these days. In humans, it is reflected for instance by the fact that races with faster motor development have lower IQ. (In individuals of the same group, slower development often indicates defects, of course.)

Another support comes from machine learning: we find that the same learning functions can learn visual and auditory pattern recognition, and even end-to-end-learning. Google has built automatic image recognition into their current photo app:

<http://blogs.wsj.com/digits/2015/07/01/google-mistakenly-tags-black-people-as-gorillas-showing-limits-of-algorithms/>

The state of the art in research can do better than that: it can begin to "imagine" things. I.e. when the experimenter asks the system to "dream" what a certain object looks like, the system can produce a somewhat compelling image, which indicates that it is indeed learning visual structure. This stuff is something nobody could do a few months ago:

<http://www.creativeai.net/posts/Mv4WG6rdzAerZF7ch/synthesizing-preferred-inputs-via-deep-generator-networks>

A machine learning program that can learn how to play an Atari game without any human supervision or hand-crafted engineering (the feat that gave DeepMind 500M from Google) now just takes about 130 lines of Python code.

These models do not have interesting motivational systems, and a relatively simple architecture. They currently seem to mimic some of the stuff that goes on in the first few layers of the cortex. They learn object features, visual styles, lighting and rotation in 3d, and simple action policies. Almost everything else is missing. But there is a lot of enthusiasm that the field might be on the right track, and that we can learn motor simulations and intuitive physics soon. (The majority of the people in AI do not work on this, however. They try to improve the performance for the current benchmarks.)

Noam's criticism of machine translation mostly applies to the Latent Semantic Analysis models that Google and others have been using for many years. These models map linguistic symbols to concepts, and relate concepts to each other, but they do not relate the concepts to "proper" mental representations of what objects and processes look like and how they interact. Concepts are probably one of the top layers of the learning hierarchy, i.e. they are acquired *after* we learn to simulate a mental world, not before. Classical linguists ignored the simulation of a mental world entirely.

It seems miraculous that purely conceptual machine translation works at all, but that is because concepts are shared between speakers, so the structure of the conceptual space can be inferred from the statistics of language use. But the statistics of language use have too little information to infer what objects look like and how they interact.

My own original ideas concern a few parts of the emerging understanding of what the brain does. The "request-confirmation networks" that I have introduced at a NIPS workshop in last the December are an attempt at modeling how the higher layers might self-organize into cognitive programs.

Cheers!

Joscha

--

please note

The information contained in this communication is confidential, may be attorney-client privileged, may constitute inside information, and is intended only for the use of the addressee. It is the property of JEE

Unauthorized use, disclosure or copying of this communication or any part thereof is strictly prohibited and may be unlawful. If you have received this communication in error, please notify us immediately by return e-mail or by e-mail to jeevacation@gmail.com, and destroy this communication and all copies thereof, including all attachments. copyright -all rights reserved