Science Distributed February 4, 2020 Volume 1 Issue 4

INTEGRATED FUTURE(S)

ARTIFICIAL ILLUSIONS

Cognitive phenomena and mental shortcuts in AI

Welcome to Integrated Future(s), IF.

IF is a weekly newsletter about the future of emerging technology and the impact on science, medicine and society. We'll look at how artificial intelligence, blockchain, quantum and neuromorphic computing, neurotech and more will integrate with each other and with ***technoscience***, the technological infrastructure of science in our society.

This week we'll explore artificial intelligence (AI) and cognitive bias. Much has been written about the potential negative impact of transferring our human biases into the AI we create. This often conflates good bias and bad without tackling some critical questions.

Wait, good bias?

Yes, just as our brains have visual efficiencies that can be revealed via illusions, cognitive biases are efficiencies that can also be maladaptive or pathological. But at their core they are useful tools for humans and maybe for AI.

Please remember that the map is not the territory.

- Sean

EMBRACING COGNITIVE BIASES IN ARTIFICIAL INTELLIGENCE TO BUILD ONE SHOT LEARNERS

https://hackernoon.com/embracing-cognitive-biases-in-artificial-intelligence-to-build-one-shot-learners-8d8bd3aafb4c

When I first started thinking about the need for cognitive bias/efficiency to be incorporated into artificial intelligence a couple years ago, Daniel Jeffries article here was the only thing I found approaching this topic. This is a great and visionary piece, though maybe (still) ahead of its time as we may have overly biased ourselves against cognitive bias.

STRESSING BIAS

How can something bad also be good? When considering cognitive bias, it is helpful to think of it compared to our stress response. The stress response, the fight or flight flow of adrenaline that gets your heart racing when you're scared, is a critical life-saving response. Stress saves. However, when it shows up in the wrong context, startled by a surprise hello or that stick in the woods that looked like a snake for a second, it is maladaptive. This can be humorous or annoying. It can also be exploited (i.e. snake in a can trick). Sometimes the chronic effects of stress, post-traumatic stress disorder, can become pathological, dangerous to both the person stressed and potentially others in some cases.

LIST OF COGNITIVE BIASES

https://en.wikipedia.org/wiki/List_of_cognitive_biases

There is no authoritative source on cognitive bias, but the Wikipedia page here has a fairly comprehensive and well referenced list with brief descriptions of nearly 200 known biases. You can quickly glean that we are speaking of a wide range of cognitive phenomena when we refer to cognitive biases. We probably need better naming/taxonomy.

COGNITIVE BIAS CHEAT SHEET

https://medium.com/better-humans/cognitive-bias-cheat-sheet-55a472476b18

This Cognitive Bias Codex that Buster Benson has created, and John Manoogian has so beautifully illustrated here (I literally have a framed poster of this on my wall) is one of the finest and most useful encapsulations of information I have ever come across. The biases are categorized into four different problem areas: *What should we remember? We need to act fast. Too much information. Not enough meaning.* This should be required reading for every human and/or AI.

A SCOPING REVIEW OF COGNITIVE BIAS IN INTERNET ADDICTION AND INTERNET GAMING DISORDERS

Chia DXY, Zhang MWB

https://www.ncbi.nlm.nih.gov/pubmed/31935915

Our biases can and will be exploited, deliberately and accidentally, in many ways. Here is a scholarly review of how this happens in internet addiction and internet gaming disorders. This is just one example of study on the matter.

DISCRIMINATION BY ARTIFICIAL INTELLIGENCE IN A COMMERCIAL ELECTRONIC HEALTH RECORD—A CASE STUDY

https://www.healthaffairs.org/do/10.1377/hblog20200128.626576/full/?mc_cid=36fa9a9560&mc_eid=105181941b&

Biases that have become chronic and systematic not only can be problematic when employed by humans but can make their way into artificial intelligence/machine learning systems we develop. This can happen deliberately or unintentionally, even simply from using training data that has been created by biased humans. Here is an exploration of a recent case study where that happen in the use of AI and electronic health records.

We want to safeguard against this type of systematic bias in AI, but we also want to explore the benefits of certain types of bias (i.e. shortcuts) in creating efficient systems. In the same way neuromorphic chips and neural networks simulate brain efficiencies to improve processing performance, existing or novel cognitive biases may do the same.

That's it for this week. Next week I'll be on the road, so we will see where the simulation takes us. Share as you like. Questions and comments to seanmanion@sciencedistributed.com. Thanks for reading.

Sean T Manion PhD is a technoscientist with a focus on blockchain and other emerging tech, neuroscientist, former federal researcher/admin and bureauscientist. He is a Chief Editor at Frontiers' Blockchain for Science, a Fellow of the British Blockchain Association and co-author of the book Blockchain for Medical Research: Accelerating Trust in Healthcare with Yaël Bizouati-Kennedy (CRC Press, April 2020). He is currently performing the duties of self-appointed strategic planner for science.

Science Distributed **is a start-up recently turned non-profit focusing on improving science and its impact to society with emerging technology. More soon at sciencedistributed.com (pardon our dust).**