INNATENESS, INDUCTIVE BIAS AND AI

(WORK IN PROGRESS)

Scott O'Hara

Metrowest Developers Machine Learning Group

06/02/2021

REFERENCES

- **“Innateness, AlphaZero, and Artificial Intelligence,” Gary Marcus, 2018**

  https://arxiv.org/abs/1801.05667

- **“Induction, Inductive Biases, and Infusing Knowledge into Learned Representations,” Sam Finlayson, 2020**

  https://sgfin.github.io/2020/06/22/Induction-Intro/

AN ANCIENT DEBATE, UPDATED FOR THE MACHINE AGE

- Nature vs Nurture
- Plato: ideas are innate
- Chomsky, Pinker: Language is innate
- Locke: Tabula rasa, empiricism
- Skinner: Behaviorism

AN ANCIENT DEBATE, UPDATED FOR THE MACHINE AGE [2]

'Virtually all modern observers would concede that genes and experience work together; it is "nature *and* nurture", not "nature versus nurture'. – Marcus

'Nativists … have argued for a view in which a significant part of a creature's innate armamentarium consist not of specific knowledge but of learning mechanisms, a form of innateness that *enables* learning.' - Marcus

E.g., Chomsky's Universal Grammar proposal.

AN ANCIENT DEBATE, UPDATED FOR THE MACHINE AGE [3]

'… there is ample reason to believe that humans and many other creatures are born with significant amounts of innate machinery.'

The … question … is whether *artificially* intelligent systems ought similarly to be endowed with significant amounts of innate machinery, or whether, in virtue of the powerful learning systems that have recently been developed, it might suffice for such systems to work in a more bottom up, *tabula rasa* fashion

[Marcus 2018]

INNATENESS IN BIOLOGICAL CREATURES

Over 90% of our genes are expressed in the development of the brain (Miller et al., 2014; Bakken et al., 2016; Kang et al., 2011), and a significant number of those are expressed selectively, in a way that allows the brain to self-assemble, even, to some non-trivial degree in the absence of experience.

Mechanisms such as cell division, cell differentiation, cell migration, cell death, and axon guidance combine to self-assemble a rich first draft of the human brain, even prior to experience. Even in the absence of synaptic transmission, the primary mechanism by which experience is conveyed to the brain, the basic structure of the newborn brain is preserved (Verhage et al., 2000).'

[Marcus 2018]

INNATENESS IN BIOLOGICAL CREATURES [2]

'… there is considerable evidence from the psychological literature that children are endowed early in life with what Spelke has called a "core knowledge" of domains like physics (Spelke & Kinzler, 2007).'

'… as Spelke (1994) has noted, it is difficult to see how knowledge of abstractions like objects, sets, and places could arise through (e.g.,) associative learning.'

[Marcus 2018]

INNATENESS IN BIOLOGICAL CREATURES [3]

‘Elsewhere in the animal kingdom are many precocial animals with a capacity to walk and (to some degree) navigate obstacles within moments of birth. Newborn chicks appear to be able to recognize faces (Wood & Wood, 2015) and to distinguish biological motion from nonbiological motion (Mayer, Rosa-Salva, Morbioli, & Vallortigara, 2017).’

[Marcus 2018]

INNATENESS IN MACHINE

- cognition = f(a, r, k, e), where
- a = innate algorithms (domain specific or domain-general)
- r = innate representational forms (domain specific or domain-general)
- k = innate knowledge (domain specific or domain-general)
- e = experience

A BLANK SLATE

- k = 0, r = 0
- a = extremely minimal value (e.g., an operation for adjusting weights relative to reinforcement signals.
- Leave the rest to e (experience)

A BLANK SLATE [2]

In the words of John Locke (1694):

All ideas come from sensation or reflection.

Let us then suppose the mind to be, as we say, white paper, void of all characters, without any ideas:—How comes it to be furnished? ....

Whence has it all the materials of reason and knowledge?

To this I answer, in one word, from EXPERIENCE.

A BLANK SLATE MACHINE

Deep learning pioneer ***Yann LeCun*** similarly appears to believe that both *a*, *r* and *k* should approach zero, as he made clear at an October 5, 2017 debate with me at NYU [tinyurl.com/lecunmarcusdebate].

In particular, in my own remarks I proposed a list of ten candidate elements that I felt would be important for AI, mostly on the representational side (see section 5).

When questioned about that list by the moderator (David Chalmers) LeCun took a strong empiricist position, very much in the spirit of John Locke, suggesting that none of those 10 elements needed to be innate for AI systems.

[Marcus 2018]

DEEPMIND'S ALPHA ZERO

Perhaps the strongest argument for keeping the values of a, r, and k small, while relying on a high value of e, comes from DeepMind's groundbreaking work on playing classical board games through reinforcement learning, masterfully presented by Demis Hassabis at December 7, 2017 NIPS Symposium on Kinds of Intelligence, and in a series of three papers.

The first, published in *Nature*, introduced **AlphaGo** (Silver et al., 2016);

the second, also published in *Nature*, focused on a more powerful successor, **AlphaGo Zero** (Silver et al., 2017a);

the third focused on **Alpha Zero**, a still-more powerful variation on the theme that played Go, chess and shogi at unprecedented levels, published on *arXiv*, December 2017 (Silver et al., 2017b).

[Marcus 2018]

DEEPMIND'S NATURE PAPER: ALPHAGO IS A BLANK SLATE

DeepMind's 2017 Nature paper frames their results, throughout, as an implicit, and at times explicit, argument for a strong version of empiricism. Their strongly antinativist framing began with title of the paper, which purported to show that they had demonstrated "Mastering the game of Go without human knowledge".

The abstract similarly claimed that the system they presented achieved its undeniably impressive results by "starting tabula rasa."

The conclusions report that the paper had shown that "a pure reinforcement learning approach is fully feasible, even in the most challenging of domains: it is possible to train to superhuman level, without human.

[Marcus 2018]

MARCUS: NOT SO FAST!

'… I don't view **AlphaStar** as a relatively unstructured "end-to-end" deep learning system, of the sort one typically finds in image recognition, in which one might expect nothing other than a deep network with many layers, with pixels on the input, and move choices on the output.

Rather, **AlphaStar** is something much closer to the sort of thing I have been advocating: a deeply structured hybrid, making important use of deep learning, but also reliant on rich integration with more traditional symbolic techniques like tree search. It's a system in which deep learning is a fundamental tool, but embedded in a symbolic context. In that respect it is to some degree closer to what I was advocating architecturally.'

[Marcus 2018]

MARCUS: ALPHAGO NOT A BLANK SLATE. A ≠ 0; R ≠ 0; K ≠ 0.

Reinforcement learning is supplement[ed] with other techniques, and human knowledge did in fact enter the system. Most of the paper's seventeen authors were deeply familiar with Go. One, Fan Hui, is a four-time European go champion…

…like virtually all computer Go systems of the last decade, the system built in **Monte Carlo tree search**, a technique, most often used in games, for evaluating moves and countermoves, with intermediate results accumulated and tested statistically over tree structures.

Similarly, **artfully placed convolutional layers** allow the system to recognize that many patterns on the board are translation invariant.

[Marcus 2018]

MARCUS: ALPHAGO NOT A BLANK SLATE. A ≠ 0; R ≠ 0; K ≠ 0. [2]

Crucially, **the Monte Carlo tree structure apparatus was not learned** from the data, by pure reinforcement learning. Rather, it was built in innately, into each iteration of AlphaStar, by DeepMind's programmers.

Likewise, **the convolutional layers were structured in a precise way**, not induced purely via reinforcement learning, with parameters **appropriate for playing Go.**

AlphaGo Zero also **included a special sampling algorithm for dealing with** the **symmetries** (eg reflections and rotations) of Go boards…

[Marcus 2018]

MARCUS: THE RIGHT STUFF

… rather being an illustration of the power of tabula rasa learning, **AlphaGo is actually an illustration of** the opposite: of **the power of building in the right stuff to begin with.**

- With the right initial algorithms and knowledge, complex problems are learnable (or learnable given some sort of real-world constraints on compute and data).
- Without the the right initial algorithms, representations and knowledge, many problems remain out of reach.

Convolution is the prior that has made the field of deep learning work.

tree search has been vital for game playing.

AlphaZero has combined the two.

TEN PROPOSED INNATE PRIMITIVES

1. Representations of objects
2. Structured, algebraic representations
3. Operations over variables
4. A type-token distinction
5. A capacity to represent sets, locations, paths, trajectories, obstacles and enduring individuals.
6. A way of representing the affordances of objects.
7. Spatiotemporal contiguity
8. Causality
9. Translational invariance
10. Capacity for cost-benefit analysis

TWO RESEARCH METHODOLOGIES FOR AI

The Reductive Strategy: One approach to discerning how much innateness might be required for AI would be to create synthetic agents that do difficult tasks, with some initial degree of innateness, achieve state of the art performance with those tasks, and then iterate, reducing as much innateness as possible, ultimately converging on some (putatively) minimal amount of innate machinery. In the AlphaStar series of papers, DeepMind has essentially followed exactly this strategy.

The Top-Down Strategy: An alternative approach might start from the top down, examining properties of (e.g.) adult or child cognitive systems, seeking clues as to what might be innate (based on behavioral grounds), or conceptually necessary.

[Marcus 2018]

TWO RESEARCH METHODOLOGIES FOR AI [2]

The **reductive approach** is about distilling a set of cognitive primitives, by successively factoring out needless complexity… .

The **top-down approach** is about using what we independently know about cognition in order constrain what we might as starting points for AI.

[Marcus 2018]

TWO RESEARCH METHODOLOGIES FOR AI [3]

Reductive approach example: In the AlphaStar series of papers, DeepMind has essentially followed exactly this strategy.

Top-down approach examples: … as Kant argued with respect to space and time, in his *Critique of Pure Reason* (Kant, 1781). Chomsky's argument for tree structure extended from empirical considerations (through the study of multiple languages); [Marcus'] arguments [in *The Algebraic Mind*] stemm[ing] primarily from empirical considerations about the nature of human cognition.

[Marcus 2018]

INNATENESS, INDUCTIVE BIAS AND AI

(WORK IN PROGRESS)

Scott O'Hara

Metrowest Developers Machine Learning Group

06/02/2021