Topic #: 2 “Does AI need more innate machinery?” (Yann LeCun, Gary Marcus)

Rahul Basu (rb622) and Indira Akkiraju (ian8)

Nature vs nurture has been debated for many centuries in philosophy and psychology. Now, the debate has moved to AI where critics are beginning to wonder who has the upper hand; innate machinery or better learning algorithms. On one side Yann LeCun explains that the very best learning is learning from scratch. This would need innate machinery that is made up of simple components which can work together at a very fundamental level. Gary Marcus explains that learning has limitations that can only be overcome by using the very best innate machinery. He criticizes the neural networks approach by predicting that the only way to match human cognitive capability is to have new kinds of functionality which can only be brought forth by more advanced innate machinery.

Yann LeCun is no stranger to this dispute. In his childhood days his favorite movie was “2001 a Space Odyssey” where technology meets human intelligence in a science fiction world. He was always intrigued by the evolution of intelligence and technology. He studied electrical engineering at Ecole Superieure d'Ingénieur en Electrotechnique et Electronique (ESIEE), Paris in 1983, where he got interested in chip design. It was also here that he stumbled upon the nature vs nurture theories of Jean Piaget and Noam Chomsky. His interest in philosophy later blossomed into groundbreaking advances in the fields of AI during his PhD research work. In 1987 he received his PhD in Computer Science from Université Pierre et Marie Curie where he worked on various projects involving neural networks despite the field being relatively dull at the time. Later in his PhD thesis, he was one of the first people to propose a form of the back-propagation learning algorithm which was further developed a decade later, to become the cornerstone of computer vision.

Back in the 1970s, the pivotal question haunting Yann was how to train neural networks with multiple layers (these are known today as the hidden layers of neural nets). Stemming from Yann’s previous interests and insights, he delved further inside the perceptron model, which to him, seemed like the perfect meeting point of intelligence and technology. The perceptron model is a type of linear classifier algorithm from machine learning for supervised learning of binary classifiers. This essentially would be the backbone of how a computer can make decisions, and take appropriate action. Yann focused on the challenge of developing multi layered neural nets, which he was convinced would be more capable then simple linear classifiers. It was during this time in France that he met Terry Sejnowski, already famous for the Boltzmann machine, and enthusiastically discussed his early thoughts about back propagation before he was even working on his PhD thesis. Terry secretly knew of a small group of researchers in America who were also working on the same idea of backpropagation and told his friend Geoffrey Hinton of Bell Labs that a young man in France was trying to prove this very same theory. A few months later in 1985, Geoffrey Hinton came to France on a conference and met with Yann. Shortly after, in 1986, Geoffrey Hinton and Ronald J Williams co-authored the very famous backpropagation paper "Learning representations by back-propagating errors" published in Nature, where Yann’s work was acknowledged and referenced. Yann finished his thesis and graduated in in 1987.

The mathematical framework for backpropagation consists of vector matrices paired with appropriate weights for every feature. The problem with training neural networks with multiple layers was that the hidden layers were not part of the input or output, therefore very difficult to decipher. Yann’s innovative “backwards” approach to this problem involved having all the training data be used to calculate these weights. Gradient descent using the cost functions was used to adjust the weights accordingly in somewhat of a stepwise manner. This was a better alternative to the previous perceptron-convergence procedure because it creates useful new features in each step of deciphering the hidden layers. By this time Yann had really impressed Terry Sejnowski and built a strong relationship with Geoffrey Hinton’s group at Bell Labs so he decided to come to America for his post-doctoral studies under Geoffrey Hinton at Bell Labs.

It was here that Yann struck gold with his work in developing LeNet 1 which was used to publish “handwritten character recognition using neural network architectures”. It was this project that used data collected by the USPS and eventually lead to later inventing convolutional neural networks (ConvNets) and the graph transformer networks which he applied to improve handwriting recognition and OCR (optical character recognition). By this time he was well established in the Adaptive Systems research department at Bell Labs. Despite the innovation of Bell Labs and his seniority in the lab, it was rather difficult to share his works with other fields. For instance, although there was email, there was no internet. Despite having the processing power of Sun Microsystems’ top computers, the best at the time, there was no open source code in a user friendly language like python, or strong analytical support like in tensorflow. To help spread the technology of ConvNets and LeNet, Yann worked with researchers to create applications similar to CRF (conditional random field which is based on hidden markov model) that can interpret sequences of characters and eventually make this technology accessible to other fields. This was eventually deployed into major French banks to automate reading cheques, and it was used in America to read 10% of all checks during the late 1990s and early 2000s.

In 1996 AT&T broke up, and as a result Yann was promoted to department head at the Image-processing research department. It was also around this time that the internet slowly was becoming popular and one problem that persisted was how to get all the paper knowledge on to the internet. Yann’s team developed DjVu which helped disseminate academic research papers into online available documents. This project used Yann’s previous work of ConvNets and also file compressing to scan and clearly display printed material onto the web. Throughout these years while he was working with DjVu, the belief that neural nets could further develop learning in AI never left Yann. Due to the breakup of AT&T, the patent on ConvNets was given to another spinoff company, who did not further develop it. It was not until 2007 that the patent expired and Yann was once again able to work on ConvNets again. During this time in the scientific community, however, ConvNets was not popular yet. In 2012 ConvNets made a major come back through ImageNet which recaptured the scientific community and propagated many researchers to apply this to their specific field.

Currently Yann is working as the founding director of NYU Center for Data Science and also the director of Facebook AI Research in NYC.

Gary Marcus is somewhat of an expert in the field of Nativism. As a teenager in high school he created a program that translated Latin into English. This brought him to the conclusion that “one cannot build programs within machines that understand language without understanding how people can understand language”. After graduating high school, he went to Hampshire College where he studied Cognitive Science. It was here that he began working with human reasoning. He continued on to pursue graduate studies at MIT where he worked under advisor Steven Pinker.

Much of Marcus’s research work is centered around theories of language and the mind, as well as, the meeting point of biology and psychology. He is a strong proponent of innatism which states that the mind is not born as a blank slate but born with abilities and knowledge that is not drawn from experiences and the senses. In 2001 Marcus published his first book “The Algebraic Mind: integrating Connectionism and Cognitive Science”(MIT Press 2001). Challenging popular belief, he wrote that the workings of the mind involved the use of symbolism and “neurons can be put together to build circuits in order to do things such as process rules or process structured representations”. In his view, humans are able to learn a lot in their lifetime for the simple reason that our ancestors, over many millennia have evolved machinery for representing things like space, time and enduring objects. Their evolved machinery eventually passed down to us and we spend the majority of our mental computation time simply using it and a very small portion of time further evolving it for the future generations.

In 2012 Marcus published the book “Guitar Zero” which was a New York Times best seller. This lead to Marcus writing for The New Yorker and also in 2012 his famous ‘News Desk’ piece titled “Is ‘Deep Learning’ a revolution in Artificial intelligence” was published. It was one of the most read and most shared articles of the year. In this article he writes,

“… Realistically, deep learning is only part of the larger challenge of building intelligent machines. Such techniques lack ways of representing casual relationships…and are likely to face challenges in acquiring abstract ideas…They have no obvious ways of performing logical inferences, and they are also still a long way from integrating abstract knowledge…”

“to paraphrase an old parable, [deep learning is] a better ladder; but a better ladder doesn’t necessarily get you to the moon.”

Marcus criticized the deep learning hype and challenged the success of using deep learning to solve complex problems. In his article he also says in deep learning, “…initial approximations are largely random, but the human brain doesn’t function that way.” Marcus firmly believes that deep learning and neural nets alone are not enough to make progress in some of today’s tough big data problems.

As his popularity grew and his skeptics transformed to curious audiences, Gary Marcus felt it was time to enter the commercial market with his insights. In 2014, Doug Bemis, Gary Marcus, Ken Stanley, and Zoubin Gharamani founded the company Geometric Intelligence. It was one of the first ‘algorithmic support as a service’ companies to utilize machine learning and use Gary’s nativist perspective. The edge of their patent-pending techniques came from not only Gary’s nativist perspective but also from the ability to extract more learning from relatively less data then compared to using deep learning.

In 2016 Geometric Intelligence was acquired by Uber to create Uber AI Labs. Uber’s Chief Product Officer Jeff Holden claims that “despite which side of the business you look at “negotiating the real world’ that represents a ‘high-order intelligence problem’” is always complex. Being able to make AI systems work with smaller data input and smarter capabilities will “quickly helping ramp the effectiveness of products at Uber that don’t necessarily have an equivalent data set to draw upon”. Often times this could be the difference in making a business decision during the beta version of an Uber product when the stakes are lower, as opposed to waiting to gather enough customer data months into the fiscal year.

Gary’s most notable and controversial works came in 2017 when he published this paper with NYU titles “Deep learning: A Critical Appraisal”. In this paper, Gary relentlessly lists out 10 problems with deep learning and includes examples of why these are potential pitfalls. This paper immediately ignites the scientific and mathematical community and the birth of the debate “Does AI need more innate machinery” was officially here. Gary complains “AI and ML has almost never included nativism into their designs and is usually ignored”. His prediction is that the gap in progress made by the current AI and ML can be bridged once innate machinery is incorporated into it. “Innate constraints often play an unacknowledged yet vital role in getting neural nets to actually work” Gary claims.

Gary and Yann met for a very interesting debate in October 2017 at NYU to discuss if more innate machinery was the answer to achieving more effective General AI. They both seem to agree on what effective AI truly means. They both mentioned that AI is in its infancy – most of its applications that have been adapted for business are much the same and rely on similar methods for training and validation. The rate of research in academia and industry is exponential. Each year sees its former best improved in some significant way, as is the wider technological availability and growing interest. The state of AI in only a few years is completely unpredictable and is bounded only by human imagination.

They both seem to find ML as a fundamental necessity in reaching strong AI – ML today has become synonymous with Artificial Intelligence, even though there’s fields of Artificial Intelligence that have nothing to do with ML. This is due to its recent effectiveness and the advancements in the few years have as a result of implementing more narrow machine learning - Deep Learning or otherwise. Hence Deep Learning will always remain as a powerful technique for Machine Learning and algorithmic advancement. But on its own, Deep Learning is not sufficient.

Both Yann and Gary recognize that Deep Learning needs to be supplemented with additional tools in order to better achieve general purpose AI. They have different ideas of how to go about this, but they both realize that achieving human-like cognitive abilities would require delving outside the purview of Deep Learning alone. Along the same lines, Model-free RL is not the end all solution – Even though Deep Reinforcement Learning has been responsible for a large number of advancement in game-AI, it’s not something that would be effective in real world (and real-time) scenarios simply because of the vast number of trials it requires to achieve an efficient level of performance.

A recurring theme in both their research is how AI systems still need better internal forward models – Forward models refer to a forward algorithm, which calculates 'belief state': the probability of a state at a certain time, given the history of evidence. This is something that is closely related to the Viterbi algorithm. These are popularly used for speech and pattern recognition, and they both agree that there’s scope for much better implementations for these.

To both of their disappointment AI continues to fall short in general purpose tasks, despite the steady increase in academic and industry interest in this field. Common sense reasoning remains fundamentally unsolved and AI is constantly outperformed by toddlers in this regard. Where the state-of-the-art deep learning model would need somewhere close to 2 million epochs to learn 53 words, a baby can pick those up with very few trials, and sometimes even by passively hearing conversations around them. The same goes for building analogies, planning, logical reasoning and utilizing memory. The only area that seems to have been greatly aided by deep learning and AI seems to be perception. There is a high amount of corpus data for a few common examples and little data for less common examples. This makes common problems easy to solve and rare issues difficult to analyze. This is analogous to the relative genetic testing data available for a common form of cancer vs the relative genetic testing data available for a rare form of cancer. Both studies are important but the common form of cancer will most certainly have faster advances in its studies than the less common form simply due to the lack of data available.

Yann and Gary both agree that for a successful implementation of AI there needs to be systems with more thorough representations of the world. In his research, Gary expands on the capabilities of animals and humans at birth and how they begin their learning with these complex innate abilities. Yann agrees that the current structural and learning models of deep learning is no match for this type of inbuilt machinery that animals and humans start off with. Where they differ, is precisely how much intelligence should be provided in the starting stage of learning. Gary believes that cleverer and more detailed innate machinery is needed for strong AI. However, Yann believes that the best AI often has the simplest structure which, can in turn facilitate more versatile in the learning than more detailed innate machinery.

Yann agrees that symbol manipulation and abstract representations of real world constraints are needed but sees these aspects as being part of a larger elaborate learning process rather than a part of the starting innate machinery. By having less complex machinery and structure as a built-in starting point, the efficiency of learning increases to mold more tightly around the data provided. The generalization of certain smaller details are not overlooked. One example of Yann’s research that supports this belief is his work with handwriting recognition. When working with the USPS library to build LeNet 1, he found that the hidden layer of the ConvNet was most successful when including simplistic and generalized characters. Even though these general characters, like a line or a circle, could be a part of more than one USPS character, eliminating them as part of the learning process step by step from general features like these and to more specific features was more effective then to begin directly with the more subtle features.

Gary’s research with baby animals and children, however, have yielded different results from what Yann observed in his work. In the famous debate at NYU, Gary uses his daughter getting stuck in a chair as an example to show that although she has never encountered such a problem before (no training data to help learning), she was able to go through a very small number of trial-and-error steps, learn from these trial-and-error steps, and finally solve the problem. None of this would have been possible without her advanced innate machinery of skills like memory, tactile responsiveness, spatial intuition, etc. She did not take time to develop these skills as part of her learning, but rather just used them to solve the problem at hand.

Another point of difference is their interpretation of how AI should model human cognitive abilities. According to Yann, good predictive models capable of inferring from existing information in an environment would be enough for successful AI. Predictive modelling provides insight into any question and allows users to infer and predict based on what they observe and what they know from previous experience. It is also an attempt at modelling common sense in humans, which Yann describes as the ability to fill in the blanks. As a result of this ability in humans, our brain fills in the visual field at the retinal blind spot and also gives us an intuitive idea of what occluded images and missing segments in speech represent: essentially completing the information that is missing based on a model, created on the information present, and past information that is known.

Gary Marcus believes that intelligent machines that have mechanisms to match human cognitive capabilities can better achieve the same. These cognitive capabilities, as he outlines them are as follows - Representation of objects, operations over variables, type-token distinction, capacity to represent sets, locations, paths, trajectories and enduring individuals, a means of representing the affordances of objects, spatiotemporal contiguity / conservation of mass, causality, translational invariance; and a capacity for cost-benefit analysis. He believes that unsupervised learning can perform better with a richer system of primitives and representations.

As in the realm of psychology, there is no one correct approach. In the debate of Yann LeCun and Gary Marcus, there is no right or wrong. The most AI efficiency can be expected from using both aspects, advanced innate machinery paired with comprehensive learning. Often times the situation and environment of the problem determine which approach should take the upper hand. A speech recognition problem poses different challenges than a 2 year old being stuck in a chair. As Albert Einstein once said many years before either Yann LeCun or Gary Marcus, “The formulation of the problem is often more essential than its solution, which may be merely a matter of mathematical or experimental skill.”

Other resources:

https://en.wikipedia.org/wiki/Nature\_versus\_nurture

https://en.wikipedia.org/wiki/Self-domestication

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

“Does AI need more innate machinery” – NYU 2017 Debate, Gary Marcus and Yann LeCun

https://www.smithsonianmag.com/history/gobekli-tepe-the-worlds-first-temple-83613665/

LeNet 1 - https://www.youtube.com/watch?v=FwFduRA\_L6Q&feature=youtu.be

https://upclosed.com/people/gary-marcus/

https://psychologydictionary.org/connectionism/

https://www.newyorker.com/news/news-desk/is-deep-learning-a-revolution-in-artificial-intelligence

https://www.technologyreview.com/s/544606/can-this-man-make-AI-more-human/

https://www.crunchbase.com/organization/geometric-intelligence#section-overview

https://techcrunch.com/2016/12/05/uber-acquires-geometric-intelligence-to-create-an-AI-lab/

The AlgebrAIc Mind – Gary Marcus (MIT Press 2001)

“Does AI need more innate machinery” – NYU 2017 Debate, Gary Marcus and Yann LeCun

https://www.youtube.com/watch?v=vdWPQ6iAkT4