Self-learning Symbolic AI: A critical appraisal

Hans Peter Willems
MIND|CONSTRUCT
March 2021

Abstract

In the current day and age, it is easy to forget that the scientific field of Artificial Intelligence was largely built upon the notion of "symbol manipulation". As Deep Learning is now the prevalent technology for AI, we have first stepped away from all the accomplishments in Symbolic AI, only to circle back to those accomplishments but now state them as problem areas that we need to solve with Deep Learning and other forms of Artificial Neural Nets.

The level of hype that contributed to the popularity of Deep Learning and related approaches, has also had a disastrous effect on the availability of funding for symbolic oriented projects. For those working in the field of Symbolic AI, the new AI winter has already started several years ago. Funding for Symbolic AI is close to nonexistent, while Deep Learning is "where the money is", resulting in a focus that might never yield the results we are after and ignoring technological advances that were made already, some even decades ago.

I present a critical appraisal of Symbolic AI approaches, relating to the current limitations in Deep Learning, and will show that those limitations can be solved, and in many cases already have been solved, by symbolic approaches.

Keywords: Artificial Intelligence, Deep Learning, Symbolic approach, Commonsense Knowledge, Cognitive Architectures,

Preamble

This paper is in large part inspired by the paper "Deep Learning: A Critical Appraisal" by Gary Markus. Because of that, this paper mimics the structure of Gary's paper, as to address the points that he brought up in a structured and concise way. This paper is in no way a critique on his paper, but rather a reaction to his critiques on Deep Learning, made from a Symbolic AI viewpoint.

1. Evading the Deep Learning limits: a Symbolic approach

As Gary Marcus implies (Marcus, 2017), it can be argued that Deep Learning is hitting a wall. Even now, several years after Marcus wrote his paper, there is little evidence that Deep Learning is going to crack the problems mentioned in his paper (and debated in context below), at least not within a few years from now. Artificial General Intelligence, and more specifically the concept of "commonsense knowledge" is gaining interest in scientific discussions about human-level Artificial Intelligence. However, the absence of anything relating to commonsense knowledge in Deep Learning is dominating those discussions. Yann LeCun has stated that we not only lack the technology to build AGI systems, but that we don't even have the science to be able to get there (LeCun, 2019).

Strangely enough, Symbolic Artificial Intelligence has tackled several of the problems that Gary stated in his paper. And more impressively, many of those problems were tackled many decades ago (Winograd, 1972). Going from there it seems a legitimate question why we are not working mainly with Symbolic AI technology today. The short answer can be found in the fact that the term "Machine learning" is (currently) directly linked to Artificial Neural Networks and therefor to Deep Learning. Somehow it seemed impossible to make Symbolic systems self-learning (Minsky, 2003). At least until now, as the ASTRID system, made by MIND|CONSTRUCT, has already proved (in public demonstrations: MIND|CONSTRUCT 2018), to be capable of not only self-learning but also doing this unsupervised and without the need for big datasets.

The self-learning capability in the ASTRID system is based on a symbolic approach, instead of any form of artificial neural network. It seems therefor logical to take a look at the current obstacles in Deep Learning, from the perspective of self-learning Symbolic AI.

2. What Symbolic AI is, and what it does well

Describing Symbolic AI in a few paragraphs is impossible. Every AI technology and scientific approach that does not use any form of artificial neural network, is somehow linked to the symbolic approach. This revolves largely around the Physical Symbol System Hypothesis (Newell and Simon, 1976), which basically states that the human brain processes reality in the form of symbols. Everything we humans know and understand can be described in a symbolic form (Simon, 1993).

When you start to describe reality in all its facets, in the form of symbols and the relations between those symbols, you get a knowledge structure that supports higher forms of reasoning. Causality, Planning (inverse causality), Structural dependency, are all examples of symbolic relationships. Inductive reasoning, Deductive reasoning and even Abduction where all solved in the Symbolic realm decades ago by Cognitive Architectures like Eurisko (1978), Copycat (1988), Soar (1983), MYCIN (1975) and its successor CADUCEUS (1986).

The most obvious difference between the symbolic approach and Deep Learning is that the symbolic approach centers around storing '"structured knowledge", while Deep Learning uses pattern approximation without storing any form of knowledge about the pattern. The other obvious difference (as most people still think) is that Deep Learning can "learn" the pattern while symbolic systems are not capable of learning and all "knowledge" must be manually entered into the system. This is no longer true, the ASTRID system learns commonsense knowledge fully unsupervised, but more about this below (see section 3.1).

3. Symbolic AI solutions to Deep learning limitations

Gary Marcus opens this part of his paper by stating that Deep Learning is basically a "brute force" approach. It should be clear to anyone in the field that this is hard to counter argue. The huge datasets, combined with datacenters full of servers that use an incredible amount of energy, can hardly be classified as anything else than brute force.

The symbolic approach, on the other hand, doesn't need large datasets (as I will show below) and also doesn't need large server arrays. The load on hardware for a symbolic system like ASTRID is minimal (MIND|CONSTRUCT. 2021), as it doesn't have to do the millions of calculations a Deep learning system has to do.

Following below are the ten points that Gary Marcus pointed at in his paper, but here viewed from the perspective of Symbolic-based solutions to the Deep Learning limitations.

3.1 On the need for large amounts of data

The need for large datasets, to be able to train a Deep Learning system, is well known. It should also be clear that we don't have large datasets for everything in the world. Deep Learning science tries to remedy this by generating datasets from other data. It is debatable if this is going to solve the problem of availability, but the need for large datasets is not solved by this. The training model for Deep learning stays unchanged.

A symbolic system doesn't need large datasets. It learns from sparse data the same way that humans do. New information is stored into the knowledge graph and relations to other concepts are found and added to the knowledge graph. The system actually learns by adding more knowledge, creating more relations and by that mechanism building a rich contextual representation around concepts. In human cognition this is known as knowledge integration. Problem-solving is done by traversing the learned links across the contextual representations. A symbolic system, or more specifically the ASTRID system, can learn from large amounts of information to gain large amounts of (integrated) knowledge. The amount of learned knowledge scales linear with the amount of available information. In stark contrast, a Deep learning system does not learn "more" from a larger dataset.

3.2 On the limited capability for knowledge transfer

Although researchers like Yann LeCun use the word "understanding" a lot in his descriptions of what Deep Learning systems can do, there is absolutely no understanding of any sort going on in artificial neural networks. These systems don't learn any meaning relating to the data they are trained on. This problem has been demonstrated by several parties (Marcus 2018)(Eykholt at all 2018), showing that even when only certain elements of the trained environment change, the learned task doesn't transfer to the new situation within the system's current state.

There is a bigger problem that stems from this: Deep Learning systems cannot do "transfer learning". Al available human knowledge is obviously not available in any form that can easily be transferred to a Deep Learning system. It isn't even available as datasets to be used for Deep Learning. This problem is compounded by the fact that our reality consists for a large part of abstract knowledge constructs. The Deep Learning field has no bearing on how to handle such abstractions in any way.

A symbolic approach can, and does indeed, handle abstract concepts. It can also perform transfer learning without any problems, as all human knowledge is available as symbolic notations. Learned information that is stored in a knowledge graph, can easily be transferred to another system and incorporated into that system's knowledge graph. Different contextual representations for the same base concept can be merged directly, resulting is a combined, richer context.

3.3 On the absence of Hierarchical data parsing

It can be argued that Deep Learning systems do use some form of hierarchical information processing. The way Deep Learning systems traverse image features from large features down to tiny features does indeed seem hierarchical. However, the resulting pattern does not have any hierarchical features that can be parsed afterwards. There is no relational mapping between individual features in this so-called hierarchy.

A symbolic system uses, in most cases, a knowledge graph to store learned information, which is intrinsically hierarchical in nature. Because of this fact, symbolic systems are capable of doing actual “inference”, which means they can find complex relations between features that are stored in the hierarchy. This is the basis for any form of higher reasoning, and it can therefor be stated that hierarchically stored representations are a prerequisite for higher reasoning. Advances in neural science have demonstrated that the human brain does indeed use some form of hierarchical categorization of concepts (Mastrandrea, Gabrielli, et all, 2017).

3.4 On the struggle with complex inference

This problem is a continuation on the arguments made before (sections 3.2 and 3.3). The absence of any hierarchically structured knowledge in a Deep Learning system and the resulting inability to parse its internal representations in any meaningful way, makes complex inferences impossible.

Winograd Schemas are designed to test higher reasoning in AI-systems. They actually test for the availability of commonsense knowledge in the system, as the answer to a Winograd Schema is not contained in the question itself. Handling this challenge with any level of sufficiency is a distant pipe dream for Deep Learning systems but can be handled, at least in part, by conventional Symbolic systems with ease.

3.5 On the lack of transparency

Humans are capable of meta-cognition. Although we are still far away from completely understanding how the human brain works, we can think about thinking. We are capable to follow our own “train of thought” and infer why and how we infer something about something. This is mainly possible because we use language to label our thoughts.

However, there is no meta-Deep Learning. It is not possible to follow the internal “deliberation” of a Deep Learning system. There are projects that try to solve this problem by having the Deep Learning system give additional output about the features it used to determine its outcome. This, however, does not solve the “black box” problem of Deep Learning, as we still don’t know how the system came to use those features, or even how it decided to output those features as being the important ones.

A knowledge graph in a symbolic system can be inspected and inferences can be followed and reported in human-readable form. Not only that, but a system with a sufficiently sophisticated inference engine can do meta-cognition and simply output its “train of thought”.

3.6 On the lack of cross-domain knowledge integration

One of the biggest problems in Deep Learning is the fact that a system can only be trained on one problem at the time. Stating that Artificial General Intelligence has been reached with Deep Learning because a system could be trained on different tasks without changing it (Schrittwieser, J., Antonoglou, I., Hubert, T. et al. 2020), is a perversion of reality. A system that can be trained on any single task is, albeit a generic solution for single tasks, a far cry from a system that can be trained on any combination of tasks. As soon as you try to train a Deep Learning system on two completely different tasks, or even just slightly different tasks, it will fail in epic fashion in each of the tasks.

A knowledge-graph can hold contextual information on an unlimited number of tasks and conceptual constructs. In addition, it facilitates conceptual cross-linking between different knowledge domains. There is no “catastrophic forgetting”, as new knowledge only adds to the context of any related concept. A symbolic system grows smarter, as more broad knowledge is learned by the system.

3.7 On the lack of causal reasoning and other cognitive constructs

As it is already discussed in the previous parts of this paper (and hence a trend should be obvious by now), the lack of (hierarchically) structured knowledge in a Deep Learning system is the origin of many (current) limitations in Deep Learning systems. One example is the capacity to do complex planning. Although Deep Mind’s AlphaGo and AlphaGoZero have shown impressive results in games that seemingly need complex planning, those systems don’t do anything that resembles planning in the way that humans perceive that capability. Planning needs causal, structural and temporal reasoning, none of which is happening in Deep

Learning systems. Reducing a game to a search-space with possible moves is more like a complete lack of planning and making decisions “as we go”.

A symbolic system can do all forms of complex reasoning because it has the knowledge stored in a form that supports complex reasoning. Because of this, higher level capabilities like complex planning are more or less “emergent” in symbolic systems.

3.8 On the inability to immediately handle changes in the knowledge domain

Any form of autonomous behavior in a largely dynamic environment will be impossible to attain with Deep Learning. The amounts of data needed to train such systems for every concept in the world can only be handled fast enough when the speed of CPUs grows several magnitudes faster than the current trend. Every little change in the environment needs the Deep Learning system to relearn that specific knowledge. Besides that, it is impossible to train any system (including symbolic systems) for every contingency for every situation that the system might encounter.

The symbolic solution is found in the fact that symbolic systems can use cross-domain knowledge. This opens the door to handling analogous knowledge to handle novel situations. This is also where the importance of commonsense knowledge can be seen: Commonsense knowledge creates fertile grounds for mining analogous knowledge.

3.9 On the brittleness of learned patterns

There are two major reasons why adversarial attacks have such an effect on Deep Learning systems. The first problem is the lack of a “rich contextual representation”. The well-known example of a stop sign covered by stickers being mistaken for a “well-stocked refrigerator”, is the result of lacking the knowledge that refrigerators don’t stand on a pole at a road intersection. That, of course, is commonsense knowledge.

The second problem is the fact that the Deep Learning system generalizes its inputs down to a common denominator for the learned pattern. Humans on the other hand, while using common denominators, also use “discriminant features” to distinguish between things that do somewhat look alike but are definitely not the same thing. Refrigerators not standing on a pole at road intersections is an example of a discriminant feature because a stop sign does stand on a pole at a road intersection. Deep Learning currently has no way of handling discriminant features.

Of course, to handle discriminant features you need to be able to store those as part of the contextual representation, and be able to reason about those discriminant features. These are all things that symbolic systems are already capable of.

3.10 On the opaqueness of Deep Learning technology

Deep learning technology is still hard to implement outside the obvious proof-of-concept cases. Especially with the current focus on explain-ability of AI systems, the cryptic values that make up a learned pattern in a (deep) neural network makes application in real-world scenarios debatable. Deep Learning is still a black-box approach. It is also still impossible to transfer those learned patterns to other applications, making integration into other systems really hard.

The symbolic approach, on the other hand, is based on traditional programming paradigms, making it easy to extend, integrate and optimize into several directions. Scaling a Deep Learning solution basically means more data and more processing power, while scaling a symbolic system can use all possible options that traditional Information Technology has to offer, like code optimization, optimizing compilers, removing database bottlenecks, faster programming languages, faster networking protocols, better cross-server process leveling, the list goes on.

3.11 Discussion: The elephant in the room

From the previous text we might come to the conclusion that Symbolic AI is the way to go. However, the reality is that Deep Learning is the prevalent AI technology at this moment. When we talk about systems that can be trained (without further specification), we use the term “machine-learning” which has become synonymous with Deep Learning.

The main reason the AI field largely abandoned the symbolic approach is the fact that symbolic systems proved to be difficult to make them self-learning or trainable in any form. When Deep Learning came along it was heralded as the new technology that was indeed trainable. The simple fact that this ability didn’t look in any way like how humans learn things (Hinton, 2017), was soon bulldozed over by large amounts of money for Deep Learning research. The other important fact that symbolic systems were already capable of many human-like cognitive functions, things that Deep Learning still can't do today, was also largely overlooked.

As the statements in this paper show, Symbolic AI seems to be the answer to each presented limitation in Deep Learning. That is, if a symbolic system could or would be self-learning. Despite the lack of media attention, and the choir of (mainly online) naysayers that want to believe this is impossible, an operational symbolic system capable of learning commonsense knowledge has been developed and implemented (MIND|CONSTRUCT ASTRID system, 2017). Not only has it been proven, by a working implementation, that a symbolic system can be made trainable, but it has been shown (in public demonstrations) that the system learns without supervision, from non-curated texts. The system builds, and dynamically maintains its internal world-model in real-time, from inputs over time, pretty much like humans do.

4. Excessive Deep Learning hype affecting Symbolic AI research

Research into symbolic approaches for self-learning systems has suffered immensely from the Deep Learning hype. This hype, demonstrably created by the big stakeholders in Deep Learning technology, has created a negative sentiment towards almost every other AI-technology, but strangely enough disproportionately towards symbolic approaches.

We have come to the point where research into symbolic approaches is being ridiculed as "old ways that don't work". Why people think that symbolic approaches are not worthy, can largely be subscribed to the Deep Learning hype machine. There is a fear that this Deep Learning hype might create another "AI winter", but research in the symbolic realm has already been in an AI-winter, as it is (still) very hard to find funding for this approach.

5. What would be better

Deep learning does have obvious strong points and in specific applications it will remain unbeatable. We should not try to get rid of it in the same way as we tried to get rid of symbolic approaches. Therein lies the realization that we should look seriously at what symbolic systems have to offer. Gary Marcus proposed four areas of interest. I will discuss those below, but again from a symbolic perspective.

5.1 Unsupervised and non-curated learning

Unsupervised learning is a term used in Deep Learning to point at systems that can learn from unlabeled data. An example of this approach is AlphaGoZero, which learned from playing thousands of games against itself, generating new gameplay-data in the process. However, this still doesn't negate the need for large datasets.

In symbolic systems, unsupervised learning points to the ability to use existing unlabeled knowledge to train the system (which in itself is transfer learning). The big difference with the Deep Learning side of this is that the symbolic approach handles "unlabeled" as "no matter what information is learned", while Deep Learning handles it as "no matter as long as it fits in the narrow training set".

To illustrate the point made here: The ASTRID system can read non-curated texts that are copied verbatim into the system, and figure out how to handle the information and integrate the learned knowledge with already existing knowledge. It doesn't matter if the information has nothing to do with the already previously learned knowledge.

5.2 The possibility of Hybrid solutions

It seems logical, from a Deep Learning perspective, to opt for a hybrid solution where a symbolic system augments the Deep Learning system. If possible, this would indeed help Deep Learning over some existing hurdles. I do, however, have some serious reservations about this direction.

There is a massive disconnect between the ways that Deep Learning and symbolic systems handle their internal information. This is akin to the difference between raster graphics and vector graphics. Although it is perfectly possible to convert a vector image to pixels, the other way around is problematic to say the least. And where a vector image can contain very precise measurement information, this gets lost when converted to pixels. There are pixel file formats that store additional information, but that is beside the point, as Deep Learning systems cannot store additional information relating to its learned patterns. It should be clear that, in this respect, Deep Learning is very shallow when compared to symbolic systems.

The other option is to (try to) add symbolic properties to Deep Learning systems. Unfortunately, this still doesn't solve the fact that Deep Learning systems (and artificial neural networks in general) are inherently non-symbolic in nature.

5.3 Insights from Human psychology

This is where everything comes together. Modeling the behavior of the human brain, which is proven to be at least in part symbolic in nature (Friedemann Pulvermüller, 2013), seems a more logical approach than to model the biological composition of the human brain (which ANNs are actually terrible at). The advantages and advancements of symbolic systems, even a few decades ago, won't be matched by Deep Learning systems in years to come. The symbolic approach has also crossed the threshold, set by Deep Learning, of being able to learn on its own.

A symbolic system, and the ASTRID system in particular, has many features that map directly to human behavior. Insights from many psychological domains can be used directly to model behavior in a symbolic system. Commonsense knowledge (like humans have) is another concept that fits neatly in a symbolic approach (as it does in the ASTRID system).

5.4 Bolder challenges

Several bold challenges for Deep Learning systems are yesterday's victories for Symbolic AI. Complex reasoning tasks, in need of commonsense knowledge, like Winograd Schemas, have already been tackled (at least in large part) by symbolic systems (Sharma at all, 2015), and the Turing Test will be cracked probably within one or two years. This paper clearly illustrates that the mentioned challenges for Deep Learning are no challenge at all for symbolic systems.

From a symbolic perspective, there are much bolder challenges, Machine Consciousness being the biggest one of all (Willems. 2014). While fringe hypotheses like Panpschychism are happily adopted by the Deep Learning crowd to be able to vouch for consciousness in such systems, a more logical approach is that consciousness somehow arises from cognition. The fact that symbolic systems are often categorized as "Cognitive Architectures" gives rise to the notion that symbolic systems have a much better chance to attain consciousness, than Deep Learning systems.

6. Conclusions

Looking at the current limitations in Deep Learning and the obvious fact that little to no progress has been made in the discussed areas over the last several years, we should be asking if this is the right technology to focus on for Artificial General Intelligence. Predictions on the advent of AGI lay roughly somewhere between fifty years and “never”. It seems obvious that such predictions are made from the perspective of Deep Learning being the technology to use for this quest. Obviously, Deep Learning is the hammer and AGI starts to look suspiciously like the proverbial nail.

Even hybrid-approaches that include Deep Learning, or even just some form of artificial neural networks, seem to lack any significant progress relating to cognitive structures, commonsense knowledge and human-like forms of training. At least from a distance it seems that the inclusion of ANNs (in any form) in these hybrid-approaches, will seriously impede progress towards AGI.

Symbolic AI, as contrasted to Deep Learning in this paper, doesn’t suffer any of the drawbacks discussed here. The main (historic) issue with symbolic systems, that it seemed impossible to make them self-learning, has been proven to be untrue (Hans Peter Willems, 2021). A lack of media attention and a choir of naysayers (even in the scientific realm) doesn’t make this accomplishment one to neglect. Taking this into account, the symbolic approach has already ticked most of the boxes that are seen as the building blocks for AGI. From this perspective I would argue that we are much closer to a fully operational AGI system than most scientists think.

I can’t wait to see where we get from here.

References

1. Gary Marcus (2017). Deep Learning: A Critical Appraisal. arXiv.
2. Yann LeCun (2019). Will Self-Taught, A.I. Powered Robots Be the End of Us?: World Science Festival. YouTube video. https://youtu.be/IHc5Zt7qT6o?t=1614
3. Winograd, T. (1972). Understanding Natural Language. New York: Academic Press.
4. Minsky, Marvin (2003). Why AI is Brain-dead, https://www.wired.com/2003/08/why-a-i-is-brain-dead/
5. MIND|CONSTRUCT (2018). MIND|CONSTRUCT demonstrates ASTRID's Autonomous Learning capabilities to shareholders. https://mindconstruct.com/website/timeline/1
6. Newel and Simon (1976), “Computer Science as Empirical Inquiry: Symbols and Search”, Communications of the ACM
7. Simon, Herbert A. (1993). The Human Mind: The symbolic level. Proceedings of the American Philosophical Society, Vol. 137, No. 4, 1993

8. MIND|CONSTRUCT (2021). Introducing ASTRID /Brain 5. MIND|CONSTRUCT website: https://mindconstruct.com/website/astrid

9. Gary Marcus (2018). Deep Learning: A Critical Appraisal. UBC Computer Science lecture. YouTube video: https://youtu.be/Ckogtfn6zaI?t=2038

10. Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., Prakash, A., Kohno, T., & Song, D. (2018). Robust Physical-World Attacks on Deep Learning Models. https://arxiv.org/abs/1707.08945

11. Mastrandrea, R., Gabrielli, A., Piras, F. et al. Organization and hierarchy of the human functional brain network lead to a chain-like core. Sci Rep *7, *4888 (2017). https://doi.org/10.1038/s41598-017-04716-3

12. Schrittwieser, J., Antonoglou, I., Hubert, T. et al. Mastering Atari, Go, chess and shogi by planning with a learned model. Nature *588, *604–609 (2020). https://doi.org/10.1038/s41586-020-03051-4

13. Hinton, G. (2017). Artificial intelligence pioneer says we need to start over. Axios. https://www.axios.com/ai-pioneer-advocates-starting-over-2485537027.html

14. MIND|CONSTRUCT (2017). *BRAIN 3:* ASTRID System builds its internal Complex World Model fully autonomous https://mindconstruct.com/website/timeline/3

15. Friedemann and Pulvermüller (2013). How neurons make meaning: brain mechanisms for embodied and abstract-symbolic semantics. Trends in Cognitive Sciences September 2013, Vol. 17, No. 9.

16. Sharma, A., Vo, N. H., Gaur, S., & Baral, C. (2015). An approach to solve winograd schema challenge using automatically extracted commonsense knowledge. In Logical Formalizations of Commonsense Reasoning - Papers from the AAAI Spring Symposium, Technical Report (Vol. SS-15-04, pp. 141-144). AI Access Foundation.

17. Willems H.P. (2014). Why we need Conscious Artificial Intelligence. https://mindconstruct.com/files/whyweneedconsciousai.pdf

18. Hans Peter Willems (2021). Bootstrapping Commonsense Knowledge (in preparation).