Machines and Meta-Perspectives

Molly Graham
April 2018

In order to create a machine capable of general intelligence, knowledge and concepts must be constructed from the computer's point of view rather than ours. GOFAI made the mistake of assuming objects and relations can be explicitly programmed as a way to produce complex behaviour. However, this is not how skill nor intelligence is accomplished in nature. Instead, computer scientists and philosophers ought to be reviewing the findings from neuroscience, psychology, and biology for inspiration about how to uncover crucial aspects of cognition which have not been previously considered. Not only would this create new directions for the field of artificial intelligence, it would also lead to advancements in our current theories surrounding the ontogenesis of intelligence.

The development of human cognitive abilities begins even before birth. The interactions between DNA and the environment, which in this case encompasses both the mother's body and the world she inhabits, provides a crucial foundation for learning to begin via adaptation. Though the term 'learn' often colloquially refers to explicit knowledge, the ability to implicitly learn about the world must also be considered, despite having little ability to perform actions to mitigate this process. However, evidence from developmental psychology suggests a child's knowledge begins to accumulate and become strengthened even before leaving the womb. (Hepper 95; van Heteren 1169). After birth, adequate infantile development creates a set of conditions which enables learning to occur, where an individual's first few years are spent strengthening perceptual abilities by interacting with people and simple objects. (Spelke 432) The child's caregivers play an important role in this stage by acting as reinforcement, where the feedback they provide helps the child learn more efficiently than if it were absent. Eventually, cultural knowledge and facts about the world emerge to become the new foundation for understanding broader concepts. As the child matures, their existing knowledge facilitates the integration of new information and serves as the basis for the ability to reason and create abstract concepts. (Tomasello 650) Thus, childhood can be seen as a period characterized by the emergence of intellectual life, provided the necessary conditions for both adequate brain development and knowledge acquisition are present.

Although a complete explanation of how the brain achieves and accomplishes certain behaviours is still in progress, experimental evidence has identified features of cognition which contribute to learning. While most agree that memory is a key aspect of knowledge, attention plays a large role as well, since it guides which aspects in the environment are represented in the mind. Furthermore, our attentional processes are also able to uncover features of our own minds and bodies by having access to sensory information and perceptual representations, which become connected with the contexts in which they arise. Over time, we bind these representations together to create a cohesive image about ourselves and how we fit and interact with the world. These ideas are stored in memory as a way to guide behaviour in novel situations requiring some level of reasoning to determine the best course of action. As our understanding of the world develops, behavioural and mental abilities operating on this knowledge expands as well, and eventually supports the complex thinking and reasoning associated with intelligence.

Deep learning systems serve as a promising method for creating intelligent machines since they are comprised of neural networks, which mimic the way associations are formed between individual neurons in the brain. These associations produce a specific pattern of activity which results in further activation throughout the network. If the resulting action is deemed appropriate, it can be reinforced to increase the likelihood of reactivation in the future. These associative networks have the potential to create simple representations based on the modality of a given input. For example, neural networks have been implemented to learn about features of images, speech and audio, as well as natural language. These four components may be all that is required to create intelligence, since humans seem to mostly rely on their sight, hearing, and language capabilities to discern features of the world. Although there are many other important faculties for learning, they might not be *necessary* for intelligence. Theoretically, these systems could be coordinated to relay information as a way to create associations between representations, giving rise to richer concepts over time and through experience. This notion might be explained better by way of an example. Say a computer vision system learns to recognize cars and trees, and eventually reaches a point where one may feel confident stating these visual representations are relatively stable within the system. However, a separate system would be dedicated to speech detection, and reinforcement learning could create representations for the sounds of the words 'car' and 'tree' with a high degree of accuracy across a broad range of phonemic variations. Furthermore, a different deep learning system can learn how the words 'car' and 'tree' are used in sentences, associating each term with related words based on how they are used in contemporary language. Finally, another deep learning system would be able to recognize when sounds are similar to another, allowing for associations to form between sound bites. This system would be able to sort and categorize specific sounds, such as a car engine or the sound of rustling leaves. Once each sensory system had formed some level of consistent representation, the outputs of each could be interconnected to form a richer concept of what it means to be a 'car' or 'tree.' This process would be performed by a separate deep learning system, which for the sake of simplicity, can be called an *attentional system*.

The attentional system is analogous to a person’s sense of awareness, or the active processing which directs thought and manipulates knowledge. It would act as an interface between the world and the representations which emerged as a result of modality-specific deep learning, allowing for connections between each representation to be strengthened. Through repetition and error correction, the attentional process would begin to form higher level concepts by connecting and associating specific features of particular objects. Supplying this attentional process with different types of inputs, such as pictures or sounds, would provide cues for retrieving all the relevant information associated with that input, similar to how we test children in school. For example, providing an audio file of a car driving by would result in the attentional system to parse its stored associations to find the most likely cause for this noise, or display an assortment of related content. A human could then flag the relevant results as correct, allowing the attentional system to learn which behaviours or suggestions are appropriate. Through many iterations of reinforcement learning using large datasets, the attentional system would organize the knowledge it has access to in order to form robust representations of the world. It would also require the ability to reflect upon its own knowledge when it wasn’t receiving input, in order to identify any patterns or inconsistencies between pieces of information. This continual self-reflection would serve as a method for knowledge consolidation and integration, further solidifying abstract concepts over time. As various inputs become increasingly complex, the computer develops its own representations by continuously working with examples from its past.

As concepts become increasingly intricate and robust, they can be further interconnected to handle larger or more abstracted contexts. To demonstrate this by way of an example, imagine the sound of a moving car versus the sound of a car driving through a puddle. Although there are commonalities associated with each sound bite, there is an extra feature within the latter sound, namely that of moving water, which could indicate a subset of environmental features to suggest a particular context. Eventually, as this context becomes learned by associating it with other items sharing similar characteristics, such as the presence of water, it can be abstracted to form a different concept, such as ‘raining’ or ‘is wet.’ Through continual association and abstraction, a system can formalize its own knowledge into different layers of representation and build a network of associated concepts. By creating neural network systems based on the strengthening of associations and dissociations, knowledge develops gradually over a series of interactions using a variety of data. Since these neural networks were missing in GOFAI systems, the associations it was able to make were limited to the contents decided upon by developers. Furthermore, this inadequate list of concepts would restrict a computer’s proficiency for understanding new ideas beyond those programmed into its software. Without a way to dynamically learn about new information by relating it to its own knowledge, it is unlikely that intelligent behaviour will emerge later on.

The reason why deep learning architectures are a preferred method to GOFAI is because it attempts to create intelligence from the bottom up, rather than a top down perspective. Hubert Dreyfus discusses this idea in *What Computers Can’t Do*, arguing against the conception of GOFAI as a model of general intelligence. He examines the biological, psychological, epistemological, ontological assumptions made by GOFAI to provide support for his argument. Biologically, he argues the brain not like a digital computer because the materials which allow neurons to communicate are comprised of a variety of electrochemical signals which sends graded information to and fro neurons. (74) This serves as evidence against the usage of concepts or formalized information as building-blocks for intelligence, suggesting a new approach is required. Neural networks and deep learning use a mathematical system which governs how associations are reinforced, and is fundamentally designed in the same manner as the neuronal structures which support cognitive functions. This is the most important aspect with respect to attempting to create a general artificial intelligence, since the assumptions of psychology partially rest on the correct implementation of cellular systems. However, neuroscience was not a prominent field of study when GOFAI machines were created, and our understanding of the brain to guide intelligence design was inaccessible at the time. Moreover, psychology and cognitive science have also made impressive contributions to theories of intelligence, providing new ways and additional ideas for conceptualizing an artificial mind. Although a complete understanding of cognition and the related neural substrates is yet to be determined, we may have sufficient resources to guide the implementation of a general intelligence system.

The epistemological and ontological assumptions made by GOFAI also differ from those made by neural networks. Epistemologically speaking, Dreyfus argues that information cannot be formalized and programmed into a computer to produce behaviour, because this is not how knowledge is generated in humans. (101) Although our own mental heuristics may be useful for guiding implementation, these operations are not formal, logical, nor scientific laws which can be distilled and copied to produce analogous mechanical functionality. (118) Rather, intelligence emerges by interacting with the world and its objects, which are subject to change based on context. The epistemological argument Dreyfus makes also relies on an ontological assumption suggesting the world is not sliced into separate conceptual pieces. Dreyfus argues that computers use discrete forms while the brain does not, creating a problem in translating how we assume the mind operates versus how it actually does, and its role for producing intelligence. (118, 137) Although heuristics and behavioural principles arise from the mind, they are not explicitly built into the mind itself as a way to govern behaviour. Instead, reinforcement learning serves as the foundation for the epistemological argument of intelligence, and by taking the opposite approach relative to GOFAI, better results will be achieved.

In sum, computer science may benefit from designing artificial intelligence in a way which is analogous to how intelligence arises in biology, rather than generating explicit procedures. By analyzing the findings from fields studying the brain and human life, we may be able to model intelligence with better outcomes than previous attempts. Although there is still much to learn about how the brain operates, science and technology will continue to lead us closer to replicating this functionality in machines. By studying a variety of disciplines related to human development and cognition, eventually our machines will have minds of their own.

References

Hepper, Peter G. "An examination of fetal learning before and after birth." *The Irish Journal of Psychology* 12.2 (1991): 95-107.

Spelke, Elizabeth. "Initial knowledge: Six suggestions." *Cognition* 50.1 (1994): 431-445.

Tomasello, Michael. "Cultural learning redux." *Child development* 87.3 (2016): 643-653.

van Heteren, Cathelijne F., et al. "Fetal learning and memory." *The Lancet* 356.9236 (2000): 1169-1170.