A Critical Look Into Cognitively-Inspired Artificial Intelligence

Sara Mahmoud[1,*], Alice Plebe[2]

[1]University of Skövde, Skövde, Sweden
[2]University of Trento, Trento, Italy

Abstract
Nature has been a great source of inspiration for many inventions and theories. One of the major benefits for this inspiration is perceiving the impossible as possible. The inception of the AI field was no exception with cognitively-inspired approaches with a dream of having an intelligent system that thinks as a human. However, this journey of human intelligence into machine intelligence has been rough and more challenging that resulted in the separation of AI from cognitive studies. In this article, we highlight the main challenges and opportunities for cognitive inspiration for AI development. We then break down the source of inspiration into four abstraction levels in which the researcher may place an inspiration from. These levels then contribute into three main stages for modeling the AI system. The two dimensional mapping from cognitive levels into modeling stages and the relation between them aims to assist the process of cognitively-inspired approaches.

Keywords
Cognitively-inspired systems, cognitive abstraction levels, Artificial Intelligence

1. Introduction

Computer systems and software can be said to evolve through two main approaches; conventional and bio-inspired [1]. The former approach develops systems through the engineering methodology of system development, while the latter approach borrows inspiration from nature and applies it with the conventional approach. Cognition has been one of the origin sources for inspiration for developing artificial intelligence systems with an idealistic goal for reaching general intelligence [2] since the development of the Turing machine and setting the Turing test for determining the machine's intelligence. While the 1956 *Dartmouth Summer School on Artificial Intelligence* is considered the starting point of developing AI to simulate human intelligence presented by top researchers at the time [3].

Many conventional approaches and mathematical formulas were initiated based on theories and models of cognition and inspiration at the early stage of AI such as neural networks, natural language processing, recommender systems, deep learning and others. However, the development of these theories and applications evolved in a conventional progression method with less involvement from its source of inspiration. As a result, the association between AI

AIC 2022, 8th International Workshop on Artificial Intelligence and Cognition
*Corresponding author.
sara.mahmoud@his.se (S. Mahmoud); alice.plebe@unitn.it (A. Plebe)
0000-0003-2855-2782 (S. Mahmoud); 0000-0001-8567-0553 (A. Plebe)

techniques and cognitive inspiration started to fade gradually. AI has since become a heavily computational with the conventional approach to develop applications that does not require a cognitive equivalence. With this separation, less focus has been placed for developing clear methodologies for combining conventional approaches with bio-inspired ones [1]. Besides, the advantages of this combination started to blur in researchers' view.

1.1. From challenges to opportunities for cognitively inspired systems

Developing cognitively inspired systems is a challenging journey. It requires first an understanding of the underlying cognitive science theories. Cognitive science was originally formulated as a synergy between the fields of psychology, linguistics, computer science, philosophy, neuroscience, and anthropology [4].

Biological inspiration further requires defining the degree of faithfulness to what the scientist will replicate from the biological system. It also requires flexibility in adjusting the model from a description of a biological system to what is suitable for an artificial agent [5]. This often involves a translation into a more formal model description [6]. There is still a lack of clear methodology for transferring a biological system into a technological artifact, or phrased differently, we need a bridge between the bio-inspired approach and the conventional approach [7].

Despite these challenges, taking the route of bio-inspired approach opens opportunities that may be difficult to achieve with the conventional one. Among all the gains, taking inspiration from nature allows the researcher to be a pioneer in leading ideas of development. The initiation of the concepts of flying was inspired and tested by Ibn Firnas in the 9th century by observing birds flying [8] and set the mathematical foundation for flying which highly contributed to the invention of airplanes nowadays. Another example, at the same time period, Ibn AlHaytham studied the anatomy of the human's eye and set the foundation of capturing an image through a small hole in a dark room projecting the image into the wall called *Qoumra*. This theory was the groundwork for the invention of cameras presently [9].

In the field of AI, the impact of artificial neural networks and modern day deep learning can not be overstated [10], and it remains an excellent example of a bio-inspired technology. The origin of this technique is based on the brain neural networks forming multi-layers of neurons feeding into each other that collectively process the data and perform advanced cognitive tasks.

1.2. Levels and categorization

For the bio-inspired approach, the researcher examines the applicable cognitive theory and models. While an abundance of the literature discusses conceptual descriptions of *what* a cognitive system should do, less is discussed on *how* this translates into a formal model for an artificial system. This causes confusion for the researcher for how to take the inspiration.

One of the early works for clarifying the development of cognitive systems was by Marr and Poggio [11] who argue that a complex systems, especially the ones inspired by cognition, should be understood at three main and independent levels; the *computation* that describes the task, the *algorithm* that specifies how the task is implemented and the *hardware* that executes the implementation and data.

Newell discussed another perspective of cognitive operation levels that is based on timescale [12]. This includes *biological* (such as the neural spike) that takes nano to milliseconds, *cognitive operations* (that includes the cognitive architecture for tasks as grasping objects) that takes tenth of a second to less that ten seconds, and *rational operation* that requires reasoning (such as doing the laundry) takes few seconds to minutes. Finally, the *social behaviors* (with interactions) that may take hours to days.

An extended categorization for Newell's levels of cognition by timescale is presented by Lieto [13]. In this work, Lieto presents the relation between cognitive systems and AI historically and shows examples from the literature for cognitively inspired work as well as a view for the future relation in terms of a roadmap.

From a cognitive modeling point of view, Guest and Martin [14] propose a framework of six layers for building computational models for psychological research in which outputs each layer is fed to the following one and verifies the layer above. The layers top-down are; framework, theory, specification, implementation, hypothesis and data.

Lieder & Griffiths [15] discussed the rationality theory of human cognition as a higher level of psychological theory in a computational level (in Marr's categorization) and then they presented a resource-rational analysis as a methodology for connecting rational human cognition with AI and neuroscience.

Here, we propose a description of the transitioning from cognitive studies to AI systems as a two dimensional mapping (see Table 1). The vertical dimension defines levels of cognitive complexity that is the source of inspiration. The horizontal dimension describes the stages for modeling and developing the system.

The dimension of cognitive complexity refers to the complexity of the system that inspires the researcher. We categorize cognitive complexity in four levels, from least cognitive complexity; *the cellular, the architecture, the functional and the behavioral.*

Second, the modeling of a computational cognitive system is carried out in multiple stages as described by Guest and Martin [14] and Marr and Poggio [11]. Here we determine three main stages in which the inspiration would mostly affect. Starting from the *theoretical* stage that (framework and theory in Guest and Martin). This also includes the computational level for describing the task and the theoretical description in the algorithmic level in Marr's categorization. Then, this feeds into the implementation stage (specification and implementation in Guest and Martin) including the algorithm description in Marr's categorization. Finally, the data level includes the representation and source of the data. This includes the input/output representation from the algorithm level and the physical representation in the hardware level in Marr's categorization.

Table 1 shows the results of inspiration development to two dimensions. Inspiration at the *cellular level (A)* refers to the smallest biological component for cognition such as the neuron as the least cognitive level. (1) This inspiration feeds into the theories such as artificial neural networks of spiking neurons [16] and learning by spike-timing dependent plasticity [17, 18]. (2) Inspiration feeds into implementation with equation formulation for the spiking neurons such as Bayesian computation by spike-timing dependent plasticity [19] and deep convolutional models with spiking neurons [20]. (3) In the data stage, inspiration helps in formation and processing on low cellular level signals such as neuromorphic cameras [21].

On the *architectural level (B)* the inspiration broadens the ability to build more complex

Table 1
Examples resulting from the two dimensional mapping from cognitive inspiration to AI modeling

		Cognitive Modeling abstraction stages		
		Theory (1)	**Implementation (2)**	**Data (3)**
Cognitive complexity levels	**Cellular (A)**	Spiking neurons [16]	Bayesian computation by spike-timing dependent plasticity [19]	Signals in neuromorphic cameras [21]
	Architectural (B)	Neural semantic pointers [22]	Semantic Pointer Architecture Unified Network (SPAUN) [25]	Visual cortex sparse representations [27]
	Functional (C)	Reinforcement learning [30]	AlphaGo [31]	Data parameters for decision making [34]
	Behavioral (D)	Resource-rational analysis [15] Behavioral cloning [35]	Robots by imitation [39]	Data for human behavior cloning for driving behavior [40]

systems with large architectures including components. (1) This level feeds into theories such as neural semantic pointers [22], computational maps inspired by the visual cortex [23] and convergence-divergence zones [24]. (2) This is reflected in implementation such as Semantic Pointer Architecture Unified Network (SPAUN) [25] and highly successful object recognition deep convolutional models [26]. (3) Besides, it inspires researchers on how the data flows internally from the perception system to the rest of the processing components such as in visual cortex sparse representations [27] and Cortical magnification [28].

With the *functional level (C)* inspiration, the researcher is able to look at more complex systems with mental tasks such as perception and decision making. At this level, the physical body and the interaction with the external world is highly involved. (1) In the theory stage this inspires theories for interacting with the environment such as Local, error-driven and associative, biologically realistic algorithm (LEABRA) [29] and Reinforcement learning [30]. (2) The implementation of mental tasks inspires work as AlphaGo [31] and AlphaZero [32] that are deep implementations of reinforcement learning achieving superhuman performance in challenging combinatorial games and deep reinforcement learning for robotic dexterous manipulations [33]. (3) On the data level it inspires the flow of the data between the system and the environment as well as the form of data for interaction and decision making such as Human-like system data parameters [34].

At the *behavioral level (D)*, the scientist studies human cognition as examined by an external observer such as learning by demonstration. (1) In theory inspiration this includes resource-rational analysis [15], behavioral cloning [35], apprenticeship learning [36] and inverse reinforcement learning [37]. (2) Behavior level inspiration brings many implementation inspiration such as end-to-end deep learning networks [38] and robots by imitation [39]. (3) On the data stage, this level defines the data required and used for behavior implementation such as Behavioral cloning of human driving behaviors [40, 41]

2. Conclusion

Despite the opportunities gained by bio-inspired approaches for developing intelligent systems, it has been a challenge for researchers to translate between the two fields. In this short paper we briefly discusses the challenges and opportunities for cognitively inspired systems and a two dimensional mapping of cognitive inspiration to computation modeling stages.

We categorize cognitive complexities from the least level of *cellular* level cognition to *architectural* including brain components, then *functional* in which cognitive abilities are required for interaction with the environment, and *behavioural* cognition that includes imitation and social abilities.

The researcher then needs to understand the stages of modeling and developing a cognitive system. This includes defining the theory of the model and the required formal models and equation for implementation as well as the data and hardware for the system.

This work is in progress for further investigation of modeling a concrete artificial system based on the bio-inspired approach. The future work includes the examination of other factors for inspiration such as the degree of loyalty in copying the biological cognition. Inspiration and cognitive modeling requires a degree of adjustment from biological system modeling to formal modeling. While the exact creation of cognitive modeling could be desired but impossible, a loose copying could be argued not to be bio-inspired or mimicking the biological cognition.

References

[1] S. Ghoul, A Road Map to Bio-inspired Software Engineering, Research Journal of Information Technology 8 (2016) 75–81. URL: http://www.scialert.net/abstract/?doi=rjit.2016.75.81. doi:10.3923/rjit.2016.75.81.

[2] R. Zhu, L. Liu, M. Ma, H. Li, Cognitive-inspired computing: Advances and novel applications, 2020.

[3] J. Moor, The dartmouth college artificial intelligence conference: The next fifty years, Ai Magazine 27 (2006) 87–87.

[4] W. Bechtel, G. Graham (Eds.), A Companion to Cognitive Science, Blackwell Publishing Ltd., 1999.

[5] A. F. Morse, T. Ziemke, On the role (s) of modelling in cognitive science, Pragmatics & cognition 16 (2008) 37–56.

[6] I. van Rooij, M. Blokpoel, Formalizing verbal theories: A tutorial by dialogue., Social Psychology 51 (2020) 285.

[7] T. De Wolf, T. Holvoet, Towards a methodology for engineering self-organising emergent systems, Frontiers in Artificial Intelligence and Applications 135 (2005) 18.

[8] E. A. Jamsari, A. Nawi, A. Sulaiman, R. Sidek, Z. Zaidi, M. Zulfazdlee, Ibn firnas and his contribution to the aviation technology of the world, Advances in Natural and Applied Sciences 7 (2013) 74–78.

[9] M. Zghal, H.-E. Bouali, Z. B. Lakhdar, H. Hamam, The first steps for learning optics: Ibn sahl's, al-haytham's and young's works on refraction as typical examples, in: Education and Training in Optics and Photonics, Optical Society of America, 2007, p. ESB2.

[10] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, nature 521 (2015) 436–444.
[11] D. Marr, T. Poggio, A computational theory of human stereo vision, Proceedings of the Royal Society of London. Series B. Biological Sciences 204 (1979) 301–328.
[12] A. Newell, Unified theories of cognition, Harvard University Press, 1994.
[13] A. Lieto, Cognitive design for artificial minds, Routledge, 2021.
[14] O. Guest, A. E. Martin, How computational modeling can force theory building in psychological science, Perspectives on Psychological Science 16 (2021) 789–802.
[15] F. Lieder, T. L. Griffiths, Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources, Behavioral and Brain Sciences 43 (2020).
[16] W. Maass, Networks of spiking neurons: the third generation of neural network models, Neural Networks 10 (1997) 1659–1671.
[17] E. M. Izhikevich, Solving the distal reward problem through linkage of STDP and dopamine signaling, Cerebral Cortex 17 (2007) 2443–2452.
[18] H. Markram, W. Gerstner, P. J. Sjöström, Spike-timing-dependent plasticity: a comprehensive overview, Frontiers in Synaptic Neuroscience 4 (2012) 2.
[19] B. Nessler, M. Pfeiffer, L. Buesing, W. Maass, Bayesian computation emerges in generic cortical microcircuits through spike-timing-dependent plasticity, PLoS Computational Biology 9 (2013) e1003037.
[20] A. Tavanaei, M. Ghodrati, S. R. Kheradpisheh, T. Masquelier, A. Maida, Deep learning in spiking neural networks, Neural Networks 111 (2019) 47–63.
[21] J. Han, C. Zhou, P. Duan, Y. Tang, C. Xu, C. Xu, T. Huang, B. Shi, Neuromorphic camera guided high dynamic range imaging, in: Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, 2020, pp. 1730–1739.
[22] C. Eliasmith, C. H. Anderson, Neural Engineering Computation, Representation, and Dynamics in Neurobiological Systems, MIT Press, Cambridge (MA), 2003.
[23] K. Fukushima, Neocognitron: a hierarchical neural network capable of visual pattern recognition, Neural Networks 1 (1988) 119–130.
[24] K. Meyer, A. Damasio, Convergence and divergence in a neural architecture for recognition and memory, Trends in Neuroscience 32 (2009) 376–382.
[25] C. Eliasmith, How to build a brain: A neural architecture for biological cognition, Oxford University Press, 2013.
[26] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
[27] H. J. Kashyap, C. C. Fowlkes, J. L. Krichmar, S. Member, Sparse Representations for Object- and Ego-Motion Estimations in Dynamic Scenes, IEEE Transactions on Neural Networks and Learning Systems 32 (2021) 2521–2534.
[28] A. Plebe, J. F. Kooij, G. P. Rosati Papini, M. Da Lio, Occupancy grid mapping with cognitive plausibility for autonomous driving applications, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 2934–2941.
[29] R. O'Reilly, Six principles for biologically-based computational models of cortical cognition, Trends in Cognitive Sciences 2 (1998) 455–462.
[30] A. G. Barto, R. S. Sutton, Simulation of anticipatory responses in classical conditioning by a neuron–like adaptive element, Behavioral and Brain Science 4 (1982) 221–234.

[31] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, D. Hassabis, Mastering the game of Go with deep neural networks and tree search, Nature 529 (2016) 484–489.

[32] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. L. L. Sifre, D. Kumaran, T. G. T. Lillicrap, K. Simonyan, D. Hassabis, A general reinforcement learning algorithm that masters chess, shogi and Go through self-play, Science 362 (2018) 1140–1144.

[33] I. Akkaya, M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas, J. Schneider, N. Tezak, J. Tworek, P. Welinder, L. Weng, Q. Yuan, W. Zaremba, L. Zhang, Solving Rubik's cube with a robot hand, arXiv abs/1910.07113 (2019).

[34] P. Suresh, P. V. Manivannan, Human driver emulation and cognitive decision making for autonomous cars, in: Proceedings of 2016 International Conference on Robotics: Current Trends and Future Challenges (RCTFC), IEEE, 2017, pp. 1–6.

[35] F. Torabi, G. Warnell, P. Stone, Behavioral cloning from observation, in: International Joint Conferences on Artificial Intelligence, 2018, pp. 4950–4957.

[36] P. Abbeel, A. Y. Ng, Apprenticeship learning via inverse reinforcement learning, in: International Conference on Machine Learning, 2004.

[37] A. Y. Ng, S. J. Russell, Algorithms for inverse reinforcement learning, in: International Conference on Machine Learning, 2000.

[38] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, et al., End to end learning for self-driving cars, arXiv preprint arXiv:1604.07316 (2016).

[39] J. Tani, M. Ito, Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB, Neural Networks 17 (2004) 1273–1289.

[40] F. Codevilla, M. Müller, A. López, V. Koltun, A. Dosovitskiy, End-to-end Driving via Conditional Imitation Learning, in: IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 4693–4700.

[41] P. Hang, C. Lv, Y. Xing, C. Huang, Z. Hu, Human-Like Decision Making for Autonomous Driving: A Noncooperative Game Theoretic Approach, IEEE Transactions on Intelligent Transportation Systems 22 (2021) 2076–2087. arXiv:2005.11064.