Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 190 (2021) 226–230

Procedia Computer Science

www.elsevier.com/locate/procedia

2020 Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence: Eleventh Annual Meeting of the BICA Society

Hybrid Bionic Cognitive Architecture for Artificial General Intelligence Agents

Roman V. Dushkin[b], Vladimir Y. Stepankov[a,]*

[a]*National Research Nuclear University MEPhI, Moscow 115409, Russia*
[b]*Artificial Intelligence Agency, Volkonsky 1st Lane, 15, 127473 Moscow, Russia*

Abstract

The article describes the author's proposal on cognitive architecture for the development of a general-level artificial intelligent agent («strong» artificial intelligence). New principles for the development of such an architecture are proposed — a hybrid approach in artificial intelligence and bionics. The architecture diagram of the proposed solution is given and descriptions of possible areas of application are described. Strong artificial intelligence is a technical solution that can solve arbitrary cognitive tasks available to humans (human-level artificial intelligence) and even surpass the capabilities of human intelligence (artificial superintelligence). The fields of application of strong artificial intelligence are limitless — from solving current problems facing the human to completely new problems that are not yet available to human civilization or are still waiting for their discoverer. The novelty of the work lies in the author's approach to the construction of cognitive architecture, which has absorbed the results of many years of research in the field of artificial intelligence and the results of the analysis of cognitive architectures of other researchers.

Peer-review under responsibility of the scientific committee of the 2020 Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence: Eleventh Annual Meeting of the BICA Society

Keywords: artificial intelligence; strong artificial intelligence; cognitive architecture; artificial intelligent agent; hybrid artificial intelligence; bionic approach; machine learning; multisensory integration; goal-setting; explainability.

* Corresponding author. Tel.: +7-916-577-7767.
E-mail address: vstepankov@gmail.com

1877-0509
Peer-review under responsibility of the scientific committee of the 2020 Annual International Conference on Brain-Inspired Cognitive Architectures for Artificial Intelligence: Eleventh Annual Meeting of the BICA Society
10.1016/j.procs.2021.06.028

1. Introduction

Artificial intelligence as a scientific field is already 7 decades old. But even nowadays there is an acute question — is it possible to create an artificial cognitive agent with strong learning and self-learning abilities? This question is still debatable, and there is no clear answer to it. The credo of the authors of this work is that there are no fundamental obstacles to creating a strong artificial intelligence, since we have a clear example of intelligent agents — ourselves. In other words, if there is an example of an intelligent agent, why should it be impossible to create other intelligent agents, perhaps of a different nature?

Scientists and engineers working in the field of strong artificial intelligence are currently focused on so-called cognitive architectures. This is a general presentation of principles and solutions (possibly in some aspects) that can somehow solve the problem of building a strong artificial cognitive agent. To date, a large number of cognitive architectures of strong artificial intelligence have been developed and proposed, most of which are based on the study of the properties and functionality of the human brain and central nervous system [Yates et al., 2020].

Indeed, the anthropocentric approach provides researchers with a solid foundation for trying to develop a strong artificial intelligence system [Leshchev, 2011]. Since man seems to be an intellectual being, looking at the principles of nature will allow us to find the correct direction of research at the first stage, which can later be expanded to new principles (e. g., planes, helicopters and rockets fly through the air not like birds do). Thus, the principles of bionics [Lipov, 2010] are quite suitable for the initial design of cognitive architectures for artificial general intelligence agents (strong artificial intelligence).

The purpose of this work is to present a new cognitive architecture for an artificial general intelligence agent, which is based on the long-term research in the field of artificial intelligence and, as expected, may well become the basis for further research in scientific collaborations.

2. Principles of the hybrid paradigm of artificial intelligence

Since the founding of artificial intelligence as an interdisciplinary field of research in 1956, the founding fathers have launched two paradigms of artificial intelligence — bottom-up or «dirty» and top-down or «clean». The first focuses on modeling the basic elements of the biological substrate that forms the basis of human intelligence — neurons, as well as on the study of artificial neural networks, through which it tries to reach intelligence as an emergent property. The second focuses on modeling pure cognitive processes at a high level, thereby making an attempt to create intelligence per se [Dushkin & Andronov, 2020].

Some researchers note that combining the approaches and methods of both the bottom-up and top-down paradigms into a single scheme will solve the problems of both paradigms, leaving only their best sides [Dushkin & Andronov, 2020; Tahmasebi, 2012]. The authors' position on this issue is to combine approaches in a single architecture, which generally corresponds to how perception, cognition and intelligence work at the logical level in humans. The scheme of operation of an artificial intelligent agent with a hybrid architecture is based on a cyclic repetition of the process of perception of information from the environment, its processing by sensory neural networks, decision-making using a symbolic universal output machine, and transmitting the decision to effectors (actuators) for interacting with the environment through a motor neural network [Dushkin & Andronov, 2020].

In general, this scheme really allows us to neutralize the negative aspects of both paradigms. However, to create an artificial general intelligence agent in a hybrid architecture of artificial intelligence, there is a lack of an essential characteristic that the human mind has — understanding the situation in which the agent finds itself, and understanding it taking into account the context and personal experience [Dushkin, 2020]. Updating knowledge bases «on the fly» and the ability to choose from several variants of meaning available both in personal experience and in the knowledge of all mankind — these are still unsolved problems of the hybrid paradigm in particular and of artificial intelligence in general. This problem is actual in the related areas of the creation of cybernetic organisms, where the "communication consequences of the technological expansion of the space of cognition" [Leshchev, 2014] are significant. There are also other nuances that will be discussed in the next section.

3. Bionic principles suitable for cognitive architecture

Bionics is the application of principles, approaches, and working solutions found in nature to the design and implementation of technical systems [Bionics, 1993]. Since earlier in this work an anthropocentric position was proclaimed regarding the study of intelligence in order to attempt to recreate artificial general intelligence in the form of a technical system, there is no doubt that bionic principles can help in this.

In particular, the following approaches found in living intelligent agents seem highly relevant for use in designing a new generation of cognitive architecture:

1. *Multisensory integration* [Kranowitz & Silver, 2006], [Harnad, 1990].
2. *Feedback* from the decision-making subsystems to the sensors for predicting perception [Shumsky, 2020], [Zalta, 2014].
3. *The personal memory of a cognitive agent* [Osipov, 2015].
4. *Goal setting* [Glazunov, 2011].
5. *Internal conflict resolution* [Sundas et al., 2020].

Thus, the listed (and in general, incomplete) points represent interesting bionic principles that are reasonable to use when designing the cognitive architecture of a new generation of artificial intelligent agents to move towards creating a strong artificial intelligence.

4. The architecture of artificial intelligent agent

To develop a new generation of cognitive architecture, it is necessary to consider examples of such architectures that were developed earlier to solve various problems both in the field of artificial intelligence in general, and when trying to develop strong artificial intelligence in particular. Although comparative analysis of cognitive architectures is of significant research interest in itself, this article will cover only a few relevant cognitive architectures. In particular, the author's interest is focused on the following examples:

1. Jeff Hawkins suggests a low-level architecture for the basic building blocks of a general-level cognitive agent, and they correspond to the so-called cortical columns in the human cortex [Hawkins & Blakeslee, 2005].
2. Raymond Kurzweil describes a similar architectural solution [Kurzweil, 2012], which is based on pattern recognition and prediction, and the image is broadly understood as any situation in which an intelligent agent may find itself.
3. Sergey Shumsky provides an interesting cognitive architecture for building a robot operating system [Shumsky, 2020], which simultaneously uses three pillars of modern machine learning — supervised learning, unsupervised learning, and reinforcement learning, and all these types of machine learning are used simultaneously to solve various tasks facing an intelligent agent at different levels.
4. Peter Anokhin developed the theory of functional systems [Anokhin, 1975], in which he described the cognitive architecture of a living intelligent agent.

Using the described hybrid architecture of artificial intelligence, the bionic principles and the listed examples of cognitive architectures of early researchers, it is proposed to consider a new generation of cognitive architecture, the General scheme of which is shown in the following figure.

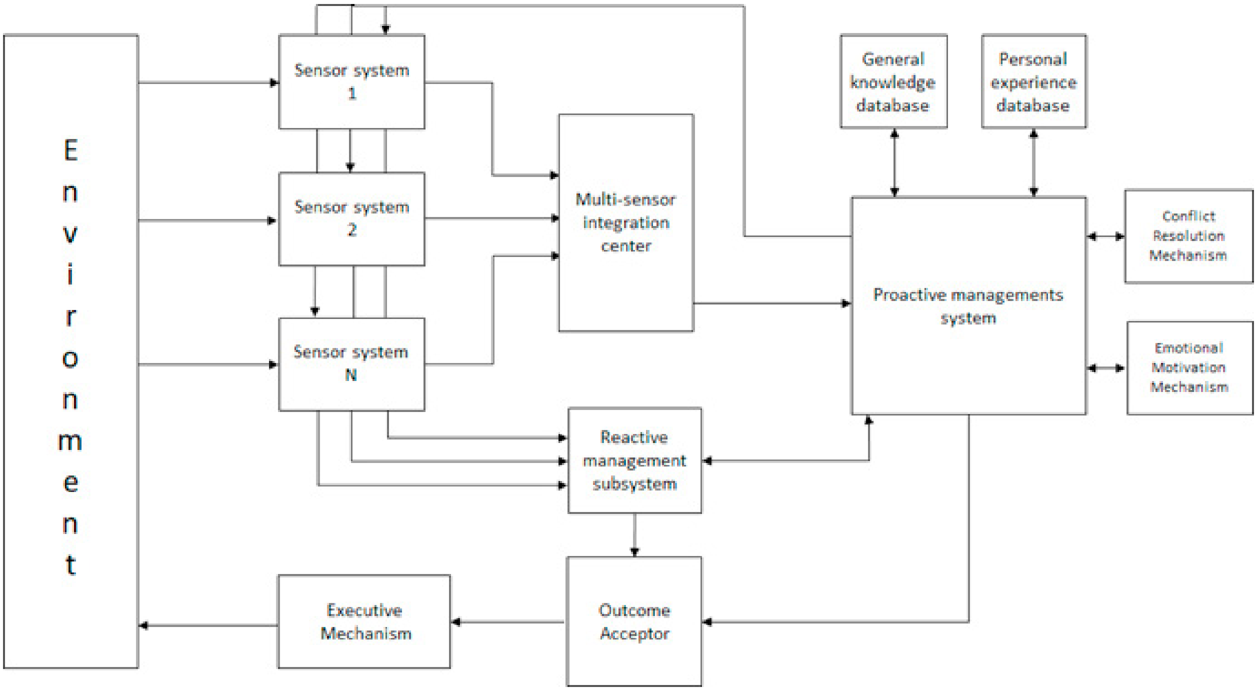

Fig. 1. General scheme of the new generation cognitive architecture.

The general process of cognition, which is defined by the presented cognitive architecture, consists of the following steps:

1. Various sensor systems of the intelligent agent receive signals from the environment, as well as predictive signals from the proactive control subsystem. Based on the actual perceived and expected images, recognizable images are formed and sent to the reactive control subsystem and the multisensory integration center.
2. A reactive (fast) control subsystem causes an instantaneous reflex reaction of an intelligent agent in cases where the rule of excitation of such a reaction is present in its structure. If there are no response rules, attention to the perceived situation is escalated to the proactive control subsystem. If a response rule exists, it is executed through the action result acceptor.
3. The multisensory integration center builds a complete description of the perceived situation in the environment of an intelligent agent and sends it to a proactive control subsystem.
4. The proactive (slow) management subsystem interacts with the general knowledge base and personal experience of the intelligent agent, getting the necessary meaning of the perceived situation from them and recording new knowledge obtained by the agent as a result of receiving information from the environment. In addition, the proactive control subsystem uses cycles of interaction with the mechanisms of emotional goal-setting and conflict resolution to make a final decision about the agent's behavior program in the perceived situation. To do this, we use a dynamic model of the environment and the agent in it, which is part of the most proactive control subsystem (not shown in the diagram). The final decision is sent to the action result acceptor for execution, and also transmitted to the reactive control subsystem for subsequent instant reaction of the agent to similar situations.
5. The action result acceptor receives the decision made and implements it through the actuators of the intelligent agent that affect the environment. In this case, the acceptor forms a mode of waiting for the result of the behavioral program execution, which is satisfied when the result is received. In this case, reinforcement

learning mechanisms should record the completed program in the agent's personal experience database as «good».

It should be noted that the described cycle repeats constantly, continuously, and even in a competitive mode with itself, since the perception of an intelligent agent must be continuous up to some level of sampling. Therefore, in the described process, it is also affected by the history of performing the same cognitive process in previous time cycles.

5. Conclusion

The article presents a general scheme of the cognitive architecture of a new generation artificial intelligence agent based on the principles of the hybrid paradigm of artificial intelligence and bionics. This cognitive architecture has absorbed some of the suggestions of other researchers in the field of artificial general intelligence. The resulting scheme is generalized and universal, which makes it possible to obtain artificial intelligent agents of various types when it is concretized by blocks for solving individual cognitive problems.

References

[1] Anokhin P. K. (1975) Essays on the physiology of functional systems. Moscow, 1975.

[2] Bionics: Nature as a Model (1993). — PRO FUTURA Verlag GmbH, München, Um-weltstiftung WWF Deutschland, 1993.

[3] Dushkin R.V. (2020) Criticism of the "Chinese Room" by J. Searle from the perspective of a hybrid model for development of artificial cognitive agents. Siberian Journal of Philosophy, 2020, Volume 18, No. 2. pp. 30-47. — DOI: 10.25205 / 2541-7517-2020-18-2-30-47.

[4] Dushkin R. V., Andronov M. G. The Hybrid Design for Artificial Intelligence Systems // In book: Arai K., Kapoor S., Bhatia R. (eds) Proceedings of the 2020 Intelligent Systems Conference (IntelliSys), Volume 1 (1250). — Springer, Cham, 2020. — P. 164-170. — ISBN 978-3-030-55179-7. — DOI: https://doi.org/10.1007/978-3-030-55180-3_13.

[5] Glazunov Yu. T. (2011) Emotional experience in the system of human goal-setting. Vestnik MGTU, volume 14, No. 1, 2011, pp. 126-140.

[6] Harnad S. (1990) The Symbol Grounding Problem. — Physica, 1990. — D 42: p. 335-346. — URL: https://clck.ru/RD7qu (Accessed 04.10.2020).

[7] Hawkins J., Blakeslee S. (2005) On Intelligence. — New York, NY: Owl Books. — ISBN 978-0-8050-7853-4.

[8] Kranowitz C. S., Silver L. B. (2006) The Out-of-Sync Child. — Penguin Books, 2006. — 352 p. — ISBN 978-0-39953-271-9.

[9] Kurzweil R. (2012) How to Create a Mind: The Secret of Human Thought Revealed. — New York: Viking Books. — ISBN 978-0-670-02529-9.

[10] Leshchev, S.V. (2014) Interfaces of social ecology from technological convergence to the Internet of Things, Filosofskie Nauki, No. 11, pp.106–113. (In Russ.)

[11] Leshchev, S.V. (2011). Cognitive Phenomenology and Artificial Intelligence: Communication, Neurophysiology, Technology. — Polygnosis. № 3-4. 2011. — Ch. 3-4, pp. 16-25. (In Russ.)

[12] Lipov A. N. (2010) At the origins of modern bionics. Bio-morphological formation in an artificial environment. — Polygnosis. № 1-2. 2010. — Ch. 1-2, pp. 126-136.

[13] Osipov G. S. (2015) Signs-Based vs. Symbolic Models // Advances in Artificial Intelligence and Soft Computing. 2015.

[14] Shumsky S. A. (2020) Machine intelligence. Essays on the theory of machine learning and artificial intelligence. Moscow, RIOR Publ., 2020. — 340 p. — ISBN: 978-5-369-01832-3.

[15] Sundas A., Bhatia A., Saggi M., Ashta J. (2020) Reinforcement Learning // In book: Machine Learning and Big Data: Concepts, Algorithms, Tools, and Applications. — John Wiley & sons, July 2020.

[16] Tahmasebi H. (2012) A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. — Computers & Geosciences. 42: p. 18-27. — Bib-code:2012CG.....42...18T. — DOI: 10.1016/j.cageo.2012.02.004.

[17] Yates S., Walker A., Kerri V. M. (2020) Artificial Intelligence, 2020.

[18] Zalta E. N. (2014) Gottlob Frege // Stanford Encyclopedia of Philosophy (Fall 2014), Edward N. Zalta (ed.).