Editorial: Novel methods in embodied and enactive AI and cognition

Fabio Bonsignorio[1,2]*

[1]AIFORS Lab, FER, University of Zagreb, Zagreb, Croatia, [2]Heron Robots, Genoa, Italy

Editorial on the Research Topic
Novel methods in embodied and enactive AI and cognition

1. Introduction

Many people still identify AI with GOFAI (Good Old Fashioned Artificial Intelligence), broadly speaking "functionalist" and "symbolic" AI, or Machine Learning or Deep Learning (DL) (Hinton and Salakhutdinov, 2006), mainly due to recent successes in object and speech recognition and other tasks. DL methods, and recently transformers and generative AI, text-to-image generation tools (such as MidJourney, StableDiffusion, and Dall-E), and Large Language Models (LLMs) (such as ChatGPT) have proven to be very powerful in a wide range of applications. However, they usually need to be trained on vast datasets (hence the phrase "Big Data"); they work less well and with lower effectiveness when available training data are less abundant, or when intelligent behaviors of physical agents in a physical environment require real-time reactions to a sensorimotor flow affected by agent behavior. When humans are involved, the "explainability" of AI reasoning is of great importance and is often weak. This is in part due to the fact that Deep Learning has been initially applied for the indexing of large data sets of images and sound samples, and for profiling of online customers of marketplace platforms and users of social networks. There might be deeper reasons for that. In nature, cognition and intelligence follow a completely different paradigm than in current mainstream AI and robotics. They are embedded in a physical system (a body) (Pfeifer and Bongard, 2006), emerging bottom-up from the interaction of large numbers of loosely coupled components (Bonsignorio, 2007), and they are usually associated with life. In contrast, the "mechatronics + machine learning" paradigm, which is still used to build mainstream robots, implements top-down controls and keeps the body well divided from the mind (following the views of seventeenth-century European philosopher Descartes). Frame of Reference (F-O-R) issues in AI (Dennett, 1987) also affect Deep Learning models; human "tagging" in model training or in LLM tuning *via* Reinforcement Learning by Human Feedback (RLHF) are clear examples of this. Several different approaches have been tried in order to frame the problem of designing physical agents acting in the real world in a way directly inspired by natural intelligence. Many researchers have seen in soft robotics a way to pursue approaches more closely related to what we observe in nature without questioning the widely accepted paradigms. In many cases, the models of robots make precise assumptions about the body and the environment, such as that

the body is made up of a known number of rigid objects connected to one another *via* joints of a known type, like those in the skeleton (Bongard et al., 2006; Hersch et al., 2009; Sturm et al., 2009; Koos and Mouret, 2011). Some bolder research has employed other approaches, including self-organizing maps or Kohonen-type unsupervised machine learning (Asada et al., 2009; Mori and Kuniyoshi, 2010). Probabilistic measures like Crutchfield information metrics have been used to determine topological relationships between sensors and actuators (Olsson et al., 2004; Kaplan and Hafner, 2005; Schatz and Oudeyer, 2009). New measures (Zenil et al., 2013) have been developed to help detect life (whether artificial or natural); see Terrazas et al. (2013). Kohonen's Self-Organizing Maps have been used to learn visuomotor coordination (Ritter et al., 1989). Kuiper's Spatial Semantic Hierarchy shows that intrinsic knowledge alone may be used to reconstruct an instructive and demanding sensorimotor contingency. Pierces and Kuipers concepts have been extended with even weaker assumptions through analysis of sensorimotor variable distances using statistically based information-theoretic metrics (Crutchfield, 1990; Olsson et al., 2004). Tanev et al. (2005) and Prokopenko (2013, 2014) propose to exploit information-driven self-organization. In this context, Bonsignorio (2013) recommends leveraging Lie Groups to reduce the computational load of information-driven self-organization, building on work by Chirikjian and Burdick (1994) and Chirikjian (2010). One of main missions of the Human Brain Project, a flagship EU project, is dedicated to neurorobotics, a new field aiming to merge neurosciences, AI, and robotics. Emergentist approaches to embodied/enactive AI are in principle capable of escaping the limitations of "symbolic" and "functionalist" AI and ML/DL (above all, the F-O-R issue). However, the problem of anchoring "pseudo?-symbolic" concise "pseudo?-representations" to the sensorimotor flow has still to be solved. The line between what is "pre-programmed" and "learned" in natural intelligent agents and where this should be in artificial ones has yet to be fully grasped; see Cangelosi and Asada (2022). Promising approaches leveraging system dynamics have been proposed, recently by Billard et al. (2022).

2. Contributed articles

The contributed papers cover some of the more challenging open questions in the area of Embodied and Enactive AI and propose some original approaches. Scarinzi and Cañamero argue that "artificial emotions" are a necessary tool for an agent interacting with the environment. Hernandez-Ochoa point out the potential importance and usefulness of the evo-devo approach for artificial emotional systems. The problem of anchoring a symbolic description to a neural encoding is discussed by Katz et al., who propose a "neurocomputational controller" for robotic manipulation based on a "neural virtual machine" (NVM). The NVM encodes the knowledge of a symbolic stacking system, but can then be further improved and fine-tuned by a Reinforcement Learning procedure. This is an approach attempting to bridge "symbolic descriptions" with data-driven approaches. In Hinrichs et al., the authors show via a thorough data analysis how "meaning," as it is understood by us humans in natural language, is actually

an unstable ground for symbolic representations, as it shifts from language to language. An early stage controller inspired by Piaget's schemas is proposed by Lagriffoul. The relevance of Piaget's work, which provides an insightful analysis of cognitive development in human children, has been recognized for example in Bonsignorio (2007).

We should not underestimate the importance of establishing suitable, cheap, and easily customizable testing platforms for real-world testing of scientific ideas in embodied and enactive AI and cognition. This is addressed by Stoelen et al..

3. Discussion

Despite some successes, much work still remains to be done (Ackerman and Guizzo, 2015; Yanco et al., 2015). In particular, "tabula rasa" ("blank slate") methods have shown serious limitations in terms of moving beyond "shallow understanding" to the deeper understanding that is required to achieve adaptive, resilient, and trustworthy behaviors in physical agents (Marcus and Davis, 2019). Deep Learning (DL) algorithms cannot infer high-level representations or causal links, or make strong anticipatory actions. Might more abstract approaches, reproposing hard (symbolic) modeling approaches from a system theory point of view, such as that of the Coresense project, be merged with "emergentist" data-driven pipelines? To overcome their current limitations, the fields of AI and robotics need to move from a "Cartesian" and "clock-like" mechatronic-plus-machine-learning paradigm to a radically new one, based on the reverse-engineering of animal intelligence and cognition. The new approach will need to be strongly interdisciplinary, as it will have to borrow principles and methods from—to name a few fields—AI, neuroscience, artificial life, and synthetic biology.

Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

Funding

Conflict of interest

FB is CEO and Founder of Heron Robots.

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or

claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ackerman, E., and Guizzo, E. (2015). "Darpa robotics challenge: amazing moments, lessons learned, and what's next," *IEEE Spectrum* (Piscataway, NJ).

Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., et al. (2009). Cognitive developmental robotics: a survey. *Auton. Ment. Dev. IEEE Trans.* 1, 12–34. doi: 10.1109/TAMD.2009.2021702

Billard, A., Mirrazavi, S., and Figueroa, N. (2022). *Learning for Adaptive and Reactive Robot Control - A Dynamical Systems Approach.* Cambridge, MA: MIT Press.

Bongard, J., Zykov, V., and Lipson, H. (2006). Resilient machines through continuous self-modeling. *Science* 314, 1118–1121. doi: 10.1126/science.1133687

Bonsignorio, F. (2007). "Preliminary considerations for a quantitative theory of networked embodied intelligence," in *50 Years of Artificial Intelligence - Festschrift*, eds M. Lungarella, F. Iida, and R. Pfeifer (Berlin; Heidelberg: Springer), 112–123.

Bonsignorio, F. (2013). Quantifying the evolutionary self-structuring of embodied cognitive networks. *Artif. Life* 19, 267–289. doi: 10.1162/ARTL_a_00109

Cangelosi, A., and Asada, M. (2022). *Cognitive Robotics.* Cambridge, MA: MIT Press.

Chirikjian, G. (2010). *Stochastic Models, Information Theory, and Lie Groups, Vol I and II.* Boston, MA: Birkhauser.

Chirikjian, G., and Burdick, J. (1994). A modal approach to hyper-redundant manipulator kinematics. *IEEE Trans. Robot. Autom.* 10, 343–354.

Crutchfield, J. P. (1990). "Information and its metric," in *Nonlinear Structures in Physical Systems - Pattern Formation, Chaos and Waves*, eds L. Lam and H. C. Morris (New York, NY: Springer), 119–130.

Dennett, D. (1987). "Cognitive wheels: the frame problem in AI," in *The Robots Dilemma: The Frame Problem in Artificial Intelligence*, ed Z. Pylyshyn (Norwood, NJ: Ablex), 41–64.

Hersch, M., Sauser, E., and Billard, A. (2009). Online learning of the body schema. *Int. J. Hum. Robot.* 5, 161. doi: 10.1142/S0219843608001376

Hinton, G. E., and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science* 313, 504–507. doi: 10.1126/science.1127647

Kaplan, F., and Hafner, V. V. (2005). "Mapping the space of skills: an approach for comparing embodied sensorimotor organizations," in *Development and Learning, 2005. Proceedings. The 4th International Conference on* (Piscataway, NJ: IEEE), 129–134.

Koos, S., and Mouret, J. B. (2011). "Online discovery of locomotion modes for wheel-legged hybrid robots: a transferability-based approach," in *Proceedings of CLAWAR 2011* (Singapore: World Scientific Publishing Co), 70–77.

Marcus, G., and Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust.* New York, NY: Vintage.

Mori, H., and Kuniyoshi, Y. (2010). "A human fetus development simulation: self- organization of behaviors through tactile sensation," in *Development and Learning (ICDL), 2010 IEEE 9th International Conference on* (Piscataway, NJ: IEEE), 82–87.

Olsson, L., Nehaiv, C., and Polani, D. (2004). "Information trade-offs and the evolution of sensory layouts," in *Proc. Artificial Life IX, 2004* (Cambridge, MA).

Pfeifer, R., and Bongard, J. C. (2006). *How the Body Shapes the Way We Think-A New View of Intelligence.* Cambridge, MA: MIT Press.

Prokopenko, M. (2013). "Design versus self-organization," in *Advances in Applied Self-Organizing Systems*, ed M. Prokopenko (London: Springer-Verlag), 3–21.

Prokopenko, M., Polani, D., and Ay, N. (2014). "On the cross-disciplinary nature of guided self-organization," in *Guided Self-Organization: Inception*, ed M. Prokopenko (Heidelberg; New York, NY; Dordrecht; London: Springer), 3–13.

Ritter, H., Martinetz, T., and Schulten, K. (1989). Topology-conserving maps for learning visuo-motor-coordination. *Neural Netw.* 2, 159–168.

Schatz, T., and Oudeyer, P. Y. (2009). "Learning motor dependent crutchfields information distance to anticipate changes in the topology of sensory body maps," in *Proc. of IEEE 8th International Conference on Development and Learning* (Piscataway, NJ).

Sturm, J., Plagemann, C., and Burgard, W. (2009). Body schema learning for robotic manipulators from visual self-perception. *J. Physiol. Paris* 103, 220–231. doi: 10.1016/j.jphysparis.2009.08.005

Tanev, I., Ray, T., and Buller, A. (2005). Automated evolutionary design, robustness, and adaptation of sidewinding locomotion of a simulated snake-like robot. *IEEE Trans. Robot.* 21, 632–645. doi: 10.1109/TRO.2005.851028

Terrazas, G., Zenil, H., and Krasnogor, N. (2013). Exploring programmable self-assembly in non DNA-based computing. *Nat. Comput.* 12, 499–515. doi: 10.1007/s11047-013-9397-2

Yanco, H. A., Norton, A., Ober, W., Shane, D., S. A., and Vice, J. (2015). Analysis of human-robot interaction at the darpa robotics challenge trials. *J. Field Robot.* 32, 420–444. doi: 10.1177/0278364916688254

Zenil, H., Ball, G., and Tegner, J. (2013). "Testing biological models for non-linear sensitivity with a programmability test," in *Advances in Artificial Intelligence, ECAL 2013*, eds P. Lio, O. Miglino, G. Nicosia, S. Nolfi, and M. Pavone (Cambridge, MA: MIT Press), 1222–1223.