Emergence of Implicit World Models from Mortal Agents

Anonymous Author(s)
Affiliation
Address
email

Life possesses agency and behaves autonomously [1, 2]. Agency refers to the ability to autonomously set goals based on intrinsic motivation (IM) and act toward achieving them. Life, by autonomously setting its own goals, is able to proactively respond to unknown situations and unpredictable events, and adjust its behavior using feedback from the environment. When attempting to mimic the agency of life, it is crucial for artificial agents to intrinsically set their own goals. Intrinsic goal setting has been explored through concepts like prediction information maximization [3, 4, 5], empowerment maximization [6], curiosity-driven learning [7, 8, 9, 10], and novelty-based learning [11, 12]. In many of these IM approaches, however, researchers explicitly design motivations, such as "novelty is good" in novelty-based learning. As a result, the complete internalization of goals within artificial agents has not yet been fully achieved, and flexible adaptation to the environment based on autonomous goal setting remains a challenge.

The most fundamental goal of life is to avoid death. Avoiding death means maintaining a state of being alive, that is, possessing homeostasis [13, 14, 15], which involves acquiring energy from external sources and keeping one's internal state within a certain range. The homeostasis is based on the objective of sustaining the very existence (being) of the self. The characterization of life based on the goal of maintaining the persistence of being was proposed as autopoiesis by Maturana and Varela [16] and later extended by Barandiaran et al. to define agency [17]. Autopoiesis is a process by which life, driven by the meta-goal of preserving its own existence (being), autonomously sets multiple internalized motivations, such as acquiring energy or escaping predators, and generates open-ended behaviors to achieve them (Appendix A) [18]. This suggests the renaissance of research stance that the existence of the agent itself and the extrinsic motivation (EM) to maintain it precedes the agent's IM, which this stance akin to the perspective of classical suggestions, such as Parisi's internal robotics [19], and Di Paolo's approach to the homeostatic adaptation using evolutionary optimizations [18].

A theoretical framework where homeostasis as the core of the EM is known as homeostatic reinforcement learning (homeostatic RL) in computational neuroscience [20, 21, 22]. By combining deep RL [23, 24], recent studies have reported the emergence of various goal-directed behaviors [25, 26]. These results suggest a possibility of the emergence of highly adaptive process of the artificial systems from our perspective, such as world models and IM (Appendix B) [27, 28, 29, 30, 31, 32, 33]. In this paper, by combining meta-RL [34, 35] and deep homeostatic RL, we hypothesize the possibility of explaining such IMs as an emergent property of homeostatic systems, together with world models. The further discussion in Appendix C suggests that including recurrent neural networks (RNN) in homeostatic RL may naturally lead to the meta-learning ability of agents. Furthermore, as reported by Wang et al. [34, 35], computational experiments have shown that, even though all of the agent architecture and optimization are carried out in *model-free*, such meta-RL agents behave like *model-based* [35]. This suggests that the agent acquires a process for implicitly constructing a model of the environment (*implicit* world model) within the unstructured network, and uses it for learning and exploration.

By conducting meta-RL based on the unified EM (homeostasis), we propose a "mortal agent" that can open-endedly generate IMs and world models according to the agent-environment coupling.

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

References

[1] Alvaro Moreno and Matteo Mossio. *Biological autonomy: A Philosophical and Theoretical Enquiry*. Springer, 2015.

[2] Ezequiel Di Paolo and Evan Thompson. The enactive approach. In *The Routledge handbook of embodied cognition*, pages 68–78. Routledge, 2014.

[3] Nihat Ay, Nils Bertschinger, Ralf Der, Frank Güttler, and Eckehard Olbrich. Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B*, 63:329–339, 2008.

[4] Jakob Hohwy. *The predictive mind*. OUP Oxford, 2013.

[5] Thomas Parr, Giovanni Pezzulo, and Karl J Friston. *Active inference: the free energy principle in mind, brain, and behavior*. MIT Press, 2022.

[6] Alexander S Klyubin, Daniel Polani, and Chrystopher L Nehaniv. Empowerment: A universal agent-centric measure of control. In *2005 ieee congress on evolutionary computation*, volume 1, pages 128–135. IEEE, 2005.

[7] Jürgen Schmidhuber. A possibility for implementing curiosity and boredom in model-building neural controllers. In *Proc. of the international conference on simulation of adaptive behavior: From animals to animats*, pages 222–227, 1991.

[8] Pierre-Yves Oudeyer, Frdric Kaplan, and Verena V Hafner. Intrinsic motivation systems for autonomous mental development. *IEEE transactions on evolutionary computation*, 11(2):265–286, 2007.

[9] Jürgen Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE transactions on autonomous mental development*, 2(3):230–247, 2010.

[10] Yuri Burda, Harri Edwards, Deepak Pathak, Amos Storkey, Trevor Darrell, and Alexei A Efros. Large-scale study of curiosity-driven learning. *arXiv preprint arXiv:1808.04355*, 2018.

[11] Jean-Baptiste Mouret and Jeff Clune. Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*, 2015.

[12] Yuri Burda, Harrison Edwards, Amos Storkey, and Oleg Klimov. Exploration by random network distillation. In *Seventh International Conference on Learning Representations*, pages 1–17, 2019.

[13] Walter Bradford Cannon. *The wisdom of the body*. Norton & Co., 1939.

[14] W Ross Ashby. *Design for a brain.* Wiley, 1952.

[15] George E Billman. Homeostasis: the underappreciated and far too often ignored central organizing principle of physiology. *Frontiers in physiology*, 11:200, 2020.

[16] Francisco G Varela, Humberto R Maturana, and Ricardo Uribe. Autopoiesis: The organization of living systems, its characterization and a model. In *Facets of systems science*, pages 559–569. Springer, 1991.

[17] Xabier E Barandiaran, Ezequiel Di Paolo, and Marieke Rohde. Defining agency: Individuality, normativity, asymmetry, and spatio-temporality in action. *Adaptive behavior*, 17(5):367–386, 2009.

[18] Ezequiel A Di Paolo. Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop. *Dynamical systems approach to embodiment and sociality*, pages 19–42, 2003.

[19] Domenico Parisi. Internal robotics. *Connection science*, 16(4):325–338, 2004.

[20] Mehdi Keramati and Boris S Gutkin. A reinforcement learning theory for homeostatic regulation. In *Advances in Neural Information Processing Systems*, pages 82–90, 2011.

[21] Mehdi Keramati and Boris Gutkin. Homeostatic reinforcement learning for integrating reward collection and physiological stability. *Elife*, 3:e04811, 2014.

[22] Oliver J Hulme, Tobias Morville, and Boris Gutkin. Neurocomputational theories of homeostatic control. *Physics of life reviews*, 31:214–232, 2019.

[23] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, 2015.

[24] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[25] Naoto Yoshida, Tatsuya Daikoku, Yukie Nagai, and Yasuo Kuniyoshi. Emergence of integrated behaviors through direct optimization for homeostasis. *Neural Networks*, 177:106379, 2024.

[26] Naoto Yoshida, Hoshinori Kanazawa, and Yasuo Kuniyoshi. Synthesising integrated robot behaviour through reinforcement learning for homeostasis. *bioRxiv*, pages 2024–06, 2024.

[27] David Ha and Jürgen Schmidhuber. Recurrent world models facilitate policy evolution. In *Advances in Neural Information Processing Systems*, pages 2451–2463, 2018.

[28] Danijar Hafner, Timothy P Lillicrap, Mohammad Norouzi, and Jimmy Ba. Mastering atari with discrete world models. In *International Conference on Learning Representations*, 2020.

[29] Danijar Hafner, Jurgis Pasukonis, Jimmy Ba, and Timothy Lillicrap. Mastering diverse domains through world models. *arXiv preprint arXiv:2301.04104*, 2023.

[30] Pierre-Yves Oudeyer and Frederic Kaplan. What is intrinsic motivation? a typology of computational approaches. *Frontiers in neurorobotics*, 1:6, 2007.

[31] Nick Haber, Damian Mrowca, Stephanie Wang, Li F Fei-Fei, and Daniel L Yamins. Learning to play with intrinsically-motivated, self-aware agents. *Advances in neural information processing systems*, 31, 2018.

[32] Ramanan Sekar, Oleh Rybkin, Kostas Daniilidis, Pieter Abbeel, Danijar Hafner, and Deepak Pathak. Planning to explore via self-supervised world models. In *ICML*, 2020.

[33] Isaac Kauvar, Chris Doyle, Linqi Zhou, and Nick Haber. Curious replay for model-based adaptation. *International Conference on Machine Learning*, 2023.

[34] Jane X Wang, Zeb Kurth-Nelson, Dhruva Tirumala, Hubert Soyer, Joel Z Leibo, Remi Munos, Charles Blundell, Dharshan Kumaran, and Matt Botvinick. Learning to reinforcement learn. *arXiv preprint arXiv:1611.05763*, 2016.

[35] Jane X Wang, Zeb Kurth-Nelson, Dharshan Kumaran, Dhruva Tirumala, Hubert Soyer, Joel Z Leibo, Demis Hassabis, and Matthew Botvinick. Prefrontal cortex as a meta-reinforcement learning system. *Nature neuroscience*, 21(6):860–868, 2018.

[36] Humberto R Maturana and Francisco J Varela. *The tree of knowledge: The biological roots of human understanding.* New Science Library/Shambhala Publications, 1987.

[37] Humberto Maturana. *Biology of cognition.* Biological Computer Laboratory, Department of Electrical Engineering . . . , 1970.

[38] Humberto R Maturana and Francisco J Varela. *Autopoiesis and cognition: The realization of the living.* Springer, 1980.

[39] Francisco J Varela. *Principles of biological autonomy.* General Systems Research. North Holland, 1979.

[40] Tom Froese and Stewart John. Life after ashby: ultrastability and the autopoietic foundations of biological autonomy. *Cybernetics and Human Knowing*, 17(4):7–50, 2010.

[41] Robert Rosen. Some realizations of (m, r)-systems and their interpretation. *The bulletin of mathematical biophysics*, 33:303–319, 1971.

[42] Ryuzo Hirota, Hayato Saigo, and Shigeru Taguchi. Reformalizing the notion of autonomy as closure through category theory as an arrow-first mathematics. In *ALIFE 2023: Ghost in the Machine: Proceedings of the 2023 Artificial Life Conference*. MIT Press, 2023.

[43] Gautier Hamon, Mayalen Etcheverry, Bert Wang-Chak Chan, Clément Moulin-Frier, and Pierre-Yves Oudeyer. Discovering sensorimotor agency in cellular automata using diversity search. *arXiv preprint arXiv:2402.10236*, 2024.

[44] Tibor Gánti. *Chemoton theory: theory of living systems*. Springer Science & Business Media, 2003.

[45] Pier Luigi Luisi. Autopoiesis: a review and a reappraisal. *Naturwissenschaften*, 90:49–59, 2003.

[46] Karl Friston. Life as we know it. *Journal of the Royal Society Interface*, 10(86):20130475, 2013.

[47] Jelle Bruineberg, Krzysztof Dołęga, Joe Dewhurst, and Manuel Baltieri. The emperor's new markov blankets. *Behavioral and Brain Sciences*, 45:e183, 2022.

[48] Takaya Araki, Tomoaki Nakamura, and Takayuki Nagai. Long-term learning of concept and word by robots: Interactive learning framework and preliminary results. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2280–2287. IEEE, 2013.

[49] Tadahiro Taniguchi, Hiroshi Yamakawa, Takayuki Nagai, Kenji Doya, Masamichi Sakagami, Masahiro Suzuki, Tomoaki Nakamura, and Akira Taniguchi. A whole brain probabilistic generative model: Toward realizing cognitive architectures for developmental robots. *Neural Networks*, 150:293–312, 2022.

A Mini Review for Autopiesis

This review provides a comprehensive overview of the development of the concept of autopoiesis since its inception 50 years ago, and its computational models. The concept of autopoiesis, first introduced by Chilean biologists Humberto Maturana and Francisco Varela in the early 1970s [36, 37, 38, 39], has had a profound impact on our understanding of life, cognition, and complex systems. Autopoiesis refers to the process by which self-maintaining and self-producing systems sustain themselves and maintain their identity through continuous interactions with their environment. The concept of autopoiesis later became connected with ideas such as enaction related to the sensorimotor loop [2] and agency [17], and biological autonomy [40, 1].

We then delve into the diverse research streams that have emerged in the field as computational model, examining the Category theory approach, Enactive approach, Synthetic Biology approach, and Bayesian approach each with its unique contributions to the understanding of autopoiesis. First, there is the approach using category theory. This begins with Rosen's (M,R) system [41], and more recently, discussions on closure have been conducted by Moreno and Mossio [1], and Hirota, Saigo and Taguchi[42]. Next is the Enactive Approach. While Di Paolo and Frose have conceptually organized the interactions between agents and their environment [2, 40], computational models like the Sensorimotor Lenia [43], which employs cellular automata, have been proposed. The third is the Synthetic Biology approach, which began with Ganti's chemoton[44] and has been further modeled by Luisi [45]. Lastly, there is the formulation of autopoietic systems using the free energy principle and Markov blankets [46, 47].Autopoiesis has developed both conceptually and computationally over the past 50 years, serving as an important guideline for constructing artificial agency.

B Graphical abstract

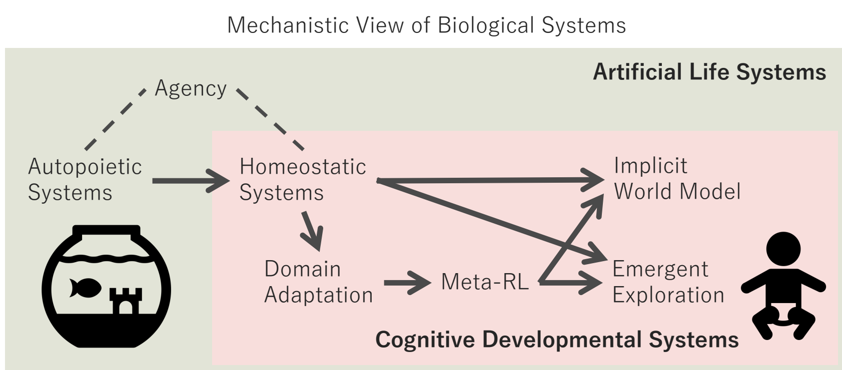

Figure 1: Relation diagram of our proposal on the emergent abilities of autonomous cognitive developmental systems, from mechanistic (= undirected, unsupervised) perspective of autonomous biological systems.

C Architecture for Mortal Agent

By combining recent meta-RL [34, 35] and deep homeostatic RL, we propose the possibility of explaining IM as an emergent property of systems adapting to a domain, together with a world model. To do this, we first focus on the possibility of mapping meta-RL and homeostasis RL (Figure 2). Specifically, the external observations x_t, latest action selection a_{t-1}, and latest reward r_{t-1} required for domain adaptation in meta-RL. These multi-modal observation are thought to correspond to exteroception x^e, proprioception x^p, and interoception x^i, in homeostatic RL [25] respectively. The multi-modal observation is common situation in studies of cognitive developmental robotics [48, 49]. Therefore, the inclusion of a recurrent neural networks (RNNs), which is essential for meta-RL, in the model architecture of homeostasis RL agent may minimally lead to the potential for a meta-learning ability.

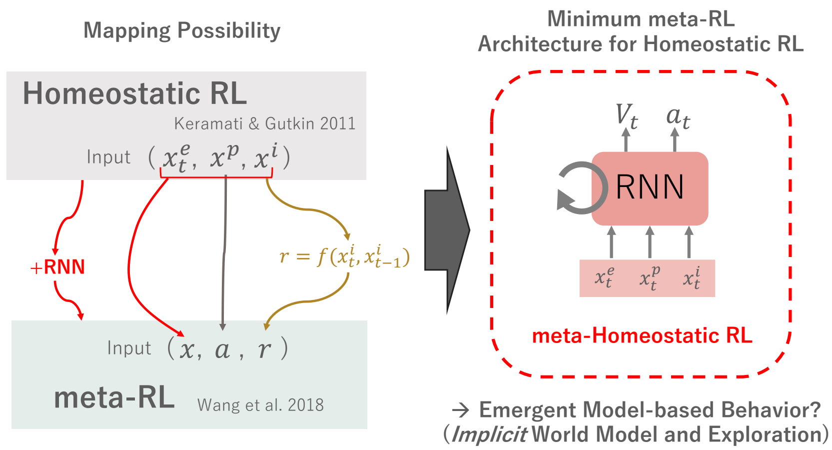

Figure 2: Implication of homeostatic extrinsic reward system combined with recurrent connection for the emergence of implicit world models and exploration.