Physics of Life Reviews 47 (2023) 271–272

Contents lists available at ScienceDirect

Physics of Life Reviews

journal homepage: www.elsevier.com/locate/plrev

A guided tour through the spaces of particular "minds": A Comment on "Path integrals, particular kinds, and strange things"

Ali Rahmjoo, Daniel Ari Friedman *

Active Inference Institute, Davis, CA 95616, United States

In this comment, we provide an assemblage of reflections arising from the "Path integrals, particular kinds, and strange things" by Friston et al. [1].

In 1984 Aaron Sloman published his influential paper "The Structure of the Space of Possible Minds" [2] calling for scholars from various disciplines to begin to construct a rigorous map of "mindspace." It didn't attract much attention at the time, but due to the explosion of AI research in recent years and the subsequent urgent need for developing an increasingly rigorous taxonomy of minds and specifying their characteristic dimensions, Sloman's project is recently taken up by researchers such as Yampolskiy [3] and Shanahan [4] among others. Yampolskiy has even gone so far as proposing a new field of study named *intellectology* devoted to investigating taxonomies of minds [5].

Viewed within this context, the typology of particular kinds proposed in the target paper [1] can be regarded as a potential contribution to such a project. Specifically by observing the nature of a given particle's dynamics and the structure of its generative model (which together determine its particular kind), it would be possible to place it within state spaces or on manifolds of sentience. In this light, Active Inference provides a framework for measuring, modeling, and implementing such minds, while the Free Energy Principle (FEP) provides a first-principles theoretical grounding.

While perspectivism is seemingly unavoidable, many attempts to construct a map of possible minds are resolutely anthropocentric [4,6,7]. By placing an encultured adult human mind as the exemplar or apotheosis of sentience against which all properties of "other" minds should be compared and analyzed, such attempts fail to sufficiently account for the Umwelt and subjectivity of the kinds of sentience that are possibly far removed from the grasp of even the most imaginative human minds [8,9].

By moving away from such an anthropocentric perspective and rather examining the dynamic behavior of a given particle with regards to its interactions through blanket states (a behavior which may be read as manifesting a kind of "agency" for strange particles and the ones with deeper generative models), FEP enables a more comprehensive approach towards mapping the space of possible minds and their interactions in "ecosystems of shared intelligence" [10]. This exciting direction, which synthesizes approaches from Bayesian Statistics and Machine Learning with the analytical rigor of Category Theory, is known as "compositional cognitive cartography" [11] – alluding to the exploratory and perspectival map-like relation [12] between cognitive models and their target system of interest.

It remains to be seen to what extent it can be reconciled with the other taxonomies of minds. But some tentative signposts for possible future research are as follows:

- Gray et al.'s two-dimensional mindspace is constituted by "agency" and "experience" axes [6]. As the particular kinds (Figure 2 in [1]) move from inert to active to conservative and to strange particles, one can envisage them as moving towards higher values of

* Corresponding author.
E-mail address: Daniel@ActiveInference.Institute (D.A. Friedman).

https://doi.org/10.1016/j.plrev.2023.11.001
Received 2 November 2023; Accepted 6 November 2023
Available online 7 November 2023

the "agency" axis. And as the hierarchical depth of generative models increases (hence granting the agents with the capability of forming beliefs about beliefs), sentient subjectivity ("experience") can be seen as ever increasing.

- The emergence of apparent or autonomous agency in strange particles can also be traced via Terrence Deacon's notion of *teleodynamics* and the corresponding *significant information* it produces [13]. More precisely, *homeodynamics* drives the behavior of inert particles, whereas the dynamics of coupled conservative particles can be characterized through their *morphodynamics*. Only through the mutual constraint of two morphodynamic systems the self-referential convolution of the given teleodynamic system emerges which can be read as indicating the emergence of a sense of agency or "selfhood" in strange particles, i.e., belief about one's beliefs.
- Eliezer Yudkowsky describes the map of mind design space as a nested hierarchy of the space of all human minds, within the space of all biological minds, within a much larger space of "minds-in-general," all embedded within a vast space of optimization processes [14]. Viewing optimization processes as the unifying thread among all possible kinds of minds is reminiscent of the first-principles approach of FEP, albeit less rigorously. Hence, the typology of particular kinds can serve both as an exemplar and an underlying scheme for conceiving or modifying various taxonomies of mind in a much more precise and principled way.

On an educational note, the authors of the present commentary have been involved with learning and teaching the first Active Inference textbook of 2022 [15], across 5 cohorts of textbook groups at the Active Inference Institute [16]. We have found the scope and visual representations of this paper to be almost uniquely relevant in these educational settings, where learners vary greatly in terms of familiarity with the Active Inference Ontology [17] and burgeoning literature [18]. Specifically, this target paper codified how (maps of) strange minds and their ecosystems can be specified and composed in an arbitrary and seemingly open-ended fashion. As each modeled *thing* is modeled according to a particular partition under the FEP (Figure 1 in [1]), there is formal continuity among inert, active, and strange systems (Figure 2 in [1]). From an implementation or application perspective, this axiomatic approach towards cognitive modeling may help us realize more applicable, interpretable, resilient, efficient, and adaptive synthetic intelligences.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Friston K, Da Costa L, Sakthivadivel DAR, Heins C, Pavliotis GA, Ramstead M, et al. Path integrals, particular kinds, and strange things. Phys Life Rev 2023;47:35–62. https://doi.org/10.1016/j.plrev.2023.08.016.
[2] Sloman A. The structure of the space of possible minds. 1984 [cited 23 Oct 2023]. Available: https://www.cs.bham.ac.uk//research/projects/cogaff/sloman-space-of-minds-84.html.
[3] Yampolskiy RV. The space of possible mind designs. Artificial general intelligence. Springer International Publishing 2015:218–27. https://doi.org/10.1007/978-3-319-21365-1_23.
[4] Shanahan M. Beyond humans, what other kinds of minds might be out there?. In: Aeon [Internet] Aeon Magazine 19 Oct 2016 [cited 23 Oct 2023]. Available: https://aeon.co/essays/beyond-humans-what-other-kinds-of-minds-might-be-out-there.
[5] Yampolskiy R.V. The Universe of Minds. arXiv [cs.AI]. 2014. Available: http://arxiv.org/abs/1410.0369.
[6] Gray HM, Gray K, Wegner DM. Dimensions of mind perception. Science 2007;315:619. https://doi.org/10.1126/science.1134475.
[7] Koch C. Consciousness: confessions of a romantic reductionist. MIT Press; 2012. https://doi.org/10.7551/mitpress/9367.001.0001.
[8] Hacker PMS. Is there anything it is like to be a bat? Philosophy 2002;77:157–74. https://doi.org/10.1017/S0031819102000220.
[9] Massimi M.Perspectival Realism. Oxford University Press; 2022. doi:10.1093/oso/9780197555620.001.0001.
[10] Friston K.J., Ramstead M.J.D., Kiefer A.B., Tschantz A., Buckley C.L., Albarracin M., et al. Designing ecosystems of intelligence from first principles. arXiv [cs.AI]. 2022. Available: http://arxiv.org/abs/2212.01354.
[11] St Clere Smithe T. Mathematical foundations for a compositional account of the bayesian brain. arXiv [q-bio.NC]. 2022. Available: http://arxiv.org/abs/2212.12538.
[12] Ramstead M.J.D., Sakthivadivel D.A.R., Friston K.J. On the map-territory fallacy fallacy. arXiv [physics.hist-ph]. 2022. Available: http://arxiv.org/abs/2208.06924.
[13] Deacon T.W. Incomplete nature. 2011. Available: https://en.wikipedia.org/w/index.php?title=Incomplete_Nature&oldid=1181630358.
[14] Yudkowsky E. Artificial Intelligence as a positive and negative factor in global risk. 2008 [cited 30 Oct 2023]. doi:10.1093/oso/9780198570509.003.0021.
[15] Parr T., Pezzulo G., Friston K.J. Active inference: the free energy principle in mind, brain, and behavior. MIT Press; 2022. Available: https://mitpress.mit.edu/books/active-inference.
[16] Active Inference Institute, Aguirre A, Boik J., Burian L., Brown M., Cordes R.J., et al. The Active Inference Institute and Active Inference Ecosystem. 2023. doi:10.5281/zenodo.8266281.
[17] Active Inference Institute. ActiveInferenceInstitute/Active_Inference_Ontology: v5. 2023. doi:10.5281/zenodo.7972289.
[18] Knight V.B., Cordes R.J., Friedman D. The free energy principle & active inference: a systematic literature analysis. 2022. doi:10.5281/zenodo.7449368.