



Editorial: Probabilistic Perspectives on Brain (Dys)function

Thomas Parr^{1*}, Dimitrije Marković², Maxwell James D. Ramstead^{1,3,4,5}, Ryan Smith⁶, Casper Hesp^{1,7} and Karl Friston¹

¹Wellcome Centre for Human Neuroimaging, Queen Square Institute of Neurology, University College London, London, United Kingdom, ²Department of Psychology, Technische Universität Dresden, Dresden, Germany, ³Division of Social and Transcultural Psychiatry, Department of Psychiatry, McGill University, Montreal, QC, Canada, ⁴Spatial Web Foundation, Los Angeles, CA, United States, ⁵Nested Minds Network, London, United Kingdom, ⁶Laureate Institute for Brain Research, Tulsa, OK, United States, ⁷Amsterdam Brain and Cognition Center, University of Amsterdam, Amsterdam, Netherlands

Keywords: neuroscience, artificial intelligence, computational psychiatry, Bayesian inference, generative models

Editorial on the Research Topic

Probabilistic Perspectives on Brain (Dys)function

While observations in neurobiology provide inspiration for methods in artificial intelligence and machine learning—most famously, in the development of artificial neural networks (McCulloch and Pitts 1943; Rosenblatt 1958; Smolensky 1986)—the reciprocal relationship has also proved fruitful. Put simply, many of the problems that machine learning is designed to solve have already been solved by the brain. When we have a good understanding of how the brain deals with a problem, we can draw inspiration from this solution in other domains. When we have a poor understanding of aspects of brain function, we can look to how these functions are performed in machine learning. If natural selection has arrived at the same optimum, we hypothesize that brain architectures support analogous procedures. Perhaps the most obvious example of this translation is the Bayesian brain hypothesis (Knill and Pouget 2004; Doya 2007), and recent extensions of this idea (Ramstead et al., 2018). This perspective treats the brain as a statistician who makes use of a probabilistic model of the world to make sense of sensory input. It has been central to the development of theories of brain function—like predictive coding (Srinivasan et al., 1982; Rao and Ballard 1999; Friston and Kiebel 2009; Bastos et al., 2012). This research topic was designed to showcase the application of contemporary probabilistic methods to understanding how the brain works, and how it can go awry in psychiatric disorders.

Broadly, the applications of probabilistic methods to the brain fall into two camps. The first applies these methods to neurobiological or psychophysical data to draw better inferences about the brain. The second assumes the brain itself makes use of these methods and engages in inference about the data it gathers from receptors in the eyes, ears, and other sensory organs. Both approaches are usefully illustrated by Feltgen and Daunizeau. Their focus is on refinement of the estimation procedure for drift-diffusion models (Ratcliff and McKoon, 2008). While drift-diffusion dynamics may be seen as a metaphor for evidence accumulation in the brain, the estimation procedure advocated by the authors represents a means of drawing inferences about cognition from psychophysical measurements.

A related perspective on evidence accumulation is offered by Heins et al., who show the emergence of drift-diffusion like dynamics in belief updating under a deep temporal model (Friston et al., 2017). This introduces an active aspect, in which we must decide how to sample our sensory data, over multiple timescales, to ensure we assimilate the most informative data (Mirza et al., 2016). The neural realization of this assimilation process was probed by Loued-Khenissi and Preuschoff in a functional imaging experiment in which participants engaged in a probabilistic gambling task. The task allowed the authors to disambiguate neural correlates of the confidence with which an outcome was predicted from the information gain when it is observed.

OPEN ACCESS

Edited and reviewed by:

Thomas Hartung,
Johns Hopkins University,
United States

*Correspondence:

Thomas Parr
thomas.parr.12@ucl.ac.uk

Specialty section:

This article was submitted to
Medicine and Public Health,
a section of the journal
Frontiers in Artificial Intelligence

Received: 15 May 2021

Accepted: 24 May 2021

Published: 07 June 2021

Citation:

Parr T, Marković D, Ramstead MJD,
Smith R, Hesp C and Friston K (2021).
Editorial: Probabilistic Perspectives on
Brain (Dys)function
Front. Artif. Intell. 4:710179.
doi: 10.3389/frai.2021.710179

Chen et al. exploit the same active inferential formalism as Heins et al., but apply it to understand how the brain might optimize the space of hypotheses it entertains. Specifically, the authors employ Bayesian model reduction (Friston et al., 2016; Friston et al., 2018)—a technique originally developed to compare dynamic causal models in neuroimaging—to prune the set of behavioral policies a creature can select between. Policies here are alternative sequences (of actions) over time. These could be sequences of saccadic eye movements, or steps through a maze (Kaplan and Friston, 2018). Such sequences are ubiquitous in planning and decision-making problems.

Temporal sequences of this sort are central to two other contributions to this Research Topic. Frölich et al. review the generation of sequences in neural systems in the form of robust and reproducible activation patterns and argue for their central role in probabilistic and predictive information processing. FitzGerald et al. complement this by considering the role of retrospective (postdictive) inference; through the perspective of Bayesian filtering (prospective) and smoothing (prospective and retrospective). The authors propose a middle ground between the two by limiting the number of past time-steps over which retrospective inference is performed—curtailing the computational cost accrued in modeling long sequences—and demonstrate the success of the resulting scheme on a probabilistic reversal learning task.

At a more conceptual level, Safron provides a broad overview of active inference and its relationship to other influential theories of brain and consciousness, including the global neuronal workspace theory (Baars, 1993) and integrated information theory (Tononi et al., 2016). Gershman adds an interesting novel perspective to this through proposing a generative adversarial theory of brain function. This is based upon the widely used deep learning networks of the same name (Goodfellow et al., 2014). Generative adversarial networks learn a generative model of the data they are exposed to. Their objective is to generate new data that are indistinguishable from the original inputs. Gershman highlights how human brain architectures could support the generative and discriminative parts of such networks.

A key area of application for theoretical neurobiology is in computational psychiatry (Montague et al., 2012). This interdisciplinary field is well-represented by the contributions from Leptourgos and Corlett and Mehlretter et al. The former

set out a theory for the distortions in the sense of agency experienced by some people with schizophrenia. They do so through assuming the brain makes use of two distinct predictive hierarchies that deal with the feeling of, and the judgment of, agency, respectively. This dual hierarchy allows them to incorporate features of prominent theories of passivity phenomena (Blakemore and Frith 2003; Synofzik et al., 2008). Mehlretter et al. take a different perspective on computational psychiatry and make use of deep learning methods in feature selection to predict remission of symptoms in patients taking antidepressants. Their focus is on the important challenge of interpretability for such analyses.

The papers outlined above offer a snapshot of the exciting work at the interface of neuroscience and probabilistic reasoning and the enduring symbiotic relationship between the two fields.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

DM was funded by the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft), SFB 940/2, 543 project A9. KF was a Wellcome Principal Research Fellow (Ref: 088130/Z/09/Z). RS is supported by the William K. Warren Foundation, the Stewart G. Wolf Fellowship, and a Center Grant from the National Institute of General Medical Sciences (P20GM121312). Postdoctoral Fellowship from the Social Sciences and Humanities Research Council of Canada (Ref: 756-2020-0704) (MR).

ACKNOWLEDGMENTS

We are grateful to the authors who contributed their work to this special issue, and to the peer reviewers for their invaluable assistance in evaluating the submissions.

REFERENCES

- Baars, B. J. (1993). *A Cognitive Theory of Consciousness*. Cambridge, United Kingdom: Cambridge University Press.
- Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., and Friston, K. J. (2012). Canonical Microcircuits for Predictive Coding. *Neuron* 76 (4), 695–711. doi:10.1016/j.neuron.2012.10.038
- Blakemore, S.-J., and Frith, C. (2003). Self-awareness and Action. *Curr. Opin. Neurobiol.* 13 (2), 219–224. doi:10.1016/s0959-4388(03)00043-6
- Doya, K. (2007). *Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge, MA: MIT press.
- Friston, K. J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., van Wijk, B. C. M., et al. (2016). Bayesian Model Reduction and Empirical Bayes for Group (DCM) Studies. *NeuroImage* 128 (Suppl. C), 413–431. doi:10.1016/j.neuroimage.2015.11.015
- Friston, K. J., Rosch, R., Parr, T., Price, C., and Bowman, H. (2017). Deep Temporal Models and Active Inference. *Neurosci. Biobehavioral Rev.* 77, 388–402. doi:10.1016/j.neubiorev.2017.04.009
- Friston, K., and Kiebel, S. (2009). Predictive Coding under the Free-Energy Principle. *Phil. Trans. R. Soc. B* 364 (1521), 1211–1221. doi:10.1098/rstb.2008.0300
- Friston, K., Parr, T., and Zeidman, P. (2018). “Bayesian Model Reduction.” arXiv preprint arXiv:1805.07092.
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., et al. (2014). “Generative Adversarial Networks.” arXiv preprint arXiv:1406.2661.
- Kaplan, R., and Friston, K. J. (2018). *Planning and Navigation as Active Inference*. Biological Cybernetics.
- Knill, D. C., and Pouget, A. (2004). The Bayesian Brain: the Role of Uncertainty in Neural Coding and Computation. *Trends. Neurosci.* 27 (12), 712–719. doi:10.1016/j.tins.2004.10.007

- McCulloch, W. S., and Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bull. Math. Biophys.* 5 (4), 115–133. doi:10.1007/bf02478259
- Mirza, M. B., Adams, R. A., Mathys, C. D., and Friston, K. J. (2016). Scene Construction, Visual Foraging, and Active Inference. *Front. Comput. Neurosci.* 10 (56), 1–16. doi:10.3389/fncom.2016.00056
- Montague, P. R., Dolan, R. J., Friston, K. J., and Dayan, P. (2012). Computational Psychiatry. *Trends Cogn. Sci.* 16 (1), 72–80. doi:10.1016/j.tics.2011.11.018
- Ramstead, M. J. D., Badcock, P. B., and Friston, K. J. (2018). Variational Neuroethology: Answering Further Questions. *Phys. Life Rev.* 24, 59–66. doi:10.1016/j.plrev.2018.01.003
- Rao, R. P. N., and Ballard, D. H. (1999). Predictive Coding in the Visual Cortex: a Functional Interpretation of Some Extra-classical Receptive-Field Effects. *Nat. Neurosci.* 2 (1), 79–87. doi:10.1038/4580
- Ratcliff, R., and McKoon, G. (2008). The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks. *Neural Comput.* 20 (4), 873–922. doi:10.1162/neco.2008.12-06-420
- Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychol. Rev.* 65 (6), 386–408. doi:10.1037/h0042519
- Smolensky, P. (1986). “Information Processing in Dynamical Systems: Foundations of harmony Theory,” in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (foundations: MIT Press), Vol. 1, 194–281.
- Srinivasan, M. V., Laughlin, S. B., Dubs, A., and Horridge, G. A. (1982). Predictive Coding: a Fresh View of Inhibition in the Retina. *Proc. R. Soc. Lond. B. Biol. Sci.* 216 (1205), 427–459. doi:10.1098/rspb.1982.0085
- Synofzik, M., Vosgerau, G., and Newen, A. (2008). Beyond the Comparator Model: A Multifactorial Two-step Account of agency. *Conscious. Cogn.* 17 (1), 219–239. doi:10.1016/j.concog.2007.03.010
- Tononi, G., Boly, M., Massimini, M., and Koch, C. (2016). Integrated Information Theory: from Consciousness to its Physical Substrate. *Nat. Rev. Neurosci.* 17 (7), 450–461. doi:10.1038/nrn.2016.44

Conflict of Interest: MR was employed by the company Spatial Web Foundation and Nested Minds Network.

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

Copyright © 2021 Parr, Marković, Ramstead, Smith, Hesp and Friston. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.