Editorial

Applying the Free Energy Principle to Complex Adaptive Systems

Paul B. Badcock 1,2,* , Maxwell J. D. Ramstead 3,4 , Zahra Sheikhbahaee 5 and Axel Constant 6

1 Centre for Youth Mental Health, The University of Melbourne, Melbourne, VIC 3010, Australia
2 Orygen, Parkville, VIC 3052, Australia
3 VERSES Research Lab and the Spatial Web Foundation, Los Angeles, CA 90016, USA;
maxwell.d.ramstead@gmail.com
4 Wellcome Centre for Human Neuroimaging, University College London, London WC1E 6BT, UK
5 David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada;
zsheikhb@uwaterloo.ca
6 Charles Perkins Centre, The University of Sydney, John Hopkins Drive, Camperdown, NSW 2006, Australia;
axel.constant.pruvost@gmail.com
* Correspondence: pbadcock@unimelb.edu.au

The free energy principle (FEP) is a formulation of the adaptive, belief-driven behaviour of self-organizing systems that gained prominence in the early 2000s as a unified model of the brain [1,2]. Since then, the theory has been applied to a wide range of biotic phenomena, extending from single cells and flora [3,4], the emergence of life and evolutionary dynamics [5,6], and to the biosphere itself [7]. For our part, we have previously proposed that the FEP can be integrated with Tinbergen’s seminal four questions in biology to furnish a multiscale ontology of living systems [8]. We have also explored more specific applications, e.g., to the evolution and development of human phenotypes [9–11], socio-cultural cognition, behaviour, and learning [12,13], as well as the dynamic construction of environmental niches by their denizens [14,15].

Despite such contributions, the capacity of the FEP to extend beyond the human brain and behaviour, and to explain living systems more generally, has only begun to be explored. This begs the following questions: Can the FEP be applied to any organism? Does it allow us to explain the dynamics of all living systems, including large-scale social behaviour? Does the FEP provide a formal, empirically tractable theory of any complex adaptive system, living or not? With such questions in mind, the aim of this Special Issue was to showcase the breadth of the FEP as a unified theory of complex adaptive systems, biological or otherwise. Instead of concentrating on the human brain and behaviour, we welcomed contributions that applied the FEP to other complex adaptive systems, with the hope of exemplifying the extent of its explanatory scope.

For the uninitiated, it is worth briefly outlining what the FEP is. Variational free energy refers to an information theoretic quantity that places an upper limit on the entropy of a system’s observations, relative to a generative model instantiated by an agent. (In this context, entropy is defined as the time-average of ‘surprise’ or the negative log probability of the agent’s sensory data.) Generative models harness probabilistic mappings from hidden causes in the environment to observed consequences (i.e., sensory data), and state transitions inherent to the environment [2]. Under the FEP, an organism is modelled as implicitly ‘expecting’ to find itself within a limited range of phenotypic states; as such, deviations from these states elicit a type of ‘phenotypic surprise’ (i.e., the deviation between actual and phenotypically preferred states). Consequently, organisms remain alive by acting in ways that minimize this type of surprise (e.g., a fish avoiding the ‘state’ of being out of water). In other words, and more heuristically, free energy scores the discrepancy between desired and sensed data; and the FEP states that the imperative of all self-organizing
There are two main ways for a self-organizing system to minimize free energy. The first is by changing its perception of the world. Previously, this has been explored through reference to human neural processing. The FEP appeals to a view of the brain as a hierarchi-
cal “inference machine”, which instantiates a hierarchy of hypotheses about the world (i.e., a generative model) that enables an organism to minimize free energy (and therefore keep entropic dissipation at bay, at least locally) by reducing discrepancies between incoming sensory inputs and top-down predictions (i.e., prediction errors) [2]. Neurobiologically, expectations about sensory data are thought to be encoded by deep pyramidal cells (i.e., representation units) at every level of the cortical hierarchy, which carry predictions downward to suppress errors at the level below, whereas prediction errors themselves are encoded by superficial pyramidal cells and are carried forward to revise expectations at the level above [16]. The relative influence of ascending (error) and descending (representation) signals is weighted by their inverse variance or precision (e.g., a high precision on ascending error signals lowers confidence in descending predictions), which is mediated by neuro-modulation and reflected psychologically by attentional selection and sensory attenuation. In short, the recursive neural dynamics described here enable us to minimise free energy (resp. prediction error) by updating our internal models (i.e., perception).

Second, an organism can reduce surprise directly by acting upon the world in order to fulfill its expectations and generate unsurprising sensations. This process of ‘active inference’ describes how an organism reduces free energy through self-fulfilling cycles of action and perception [17]. Active inference models implement action selection as the minimization of expected free energy, which is the free energy expected under beliefs about possible courses of action, or policies. By selecting actions that are expected to minimize free energy, the organism can maintain itself within preferred, phenotypically unsurprising states. Thus, survival mandates that an organism must not only reduce free energy from moment to moment; it must also reduce the expected free energy associated with the future outcomes of action [18,19].

Having briefly outlined the rudiments of the FEP, let us turn briefly to complex adaptive systems (CAS). This concept is synonymous with complexity science and has its roots in evolutionary systems theory, which assumes a dynamic, inextricable relationship between generalised selection and self-organization [11]. Broadly speaking, a CAS refers to any multi-component, self-organising system that adapts to its environment through an autonomous process of selection, which recruits the outcomes of localised interactions between its components to select a subset of those components for replication and enhancement [20]. Holland [21] describes four key features of CAS: they consist of large numbers of components that interact by sending and receiving signals (i.e., parallelism); the actions of their components depend upon the signals they receive (i.e., conditional action); groups of rules can form subroutines that can be combined to deal with environmental novelties (i.e., modularity); and the components of the system change over time to adapt to the environment and improve performance (i.e., adaptation). Applications of the CAS framework have proliferated across the physical, human and computer sciences, but there is not the scope to survey this literature here. However, to pre-empt the papers to follow, we would note that this framework has already been applied to precisely the same systems that are the foci of our contributors—ranging from metabolic and cellular processes, e.g., [22–26]; to the brain and social processes, e.g., [26–30]; and to artificial intelligence and robotics, e.g., [31–33]. The articles presented in our Special Issue build upon such literature by illustrating how the FEP can afford fresh insights into the dynamics of CAS.

Three of the contributions to our Special Issue leverage the FEP to cast new light on processes intrinsic to biological agents. In Cancer Niches and their Kikuchi Free Energy, Sajid, Convertino et al. [34] examine cancer morphogenetic fields as undesirable stable attractors in the complex dynamics of homeostasis, self-renewal and differentiation, which contributes to their deviation from regular autopoietic homeostasis (the internal molecular dynamics that regulate the production and regeneration of a system’s components). Sajid, Convertino et al. offer a computational model in silico to study communication and information processing at a population level of cancer cell networks within their environment in vivo. By deploying the Kikuchi free energy approximation, which is a generalisation of the Bethe free energy for computing beliefs over large sets of cell clusters, they
account for higher-order interactions and phase transitions between clusters of healthy and oncogenic cells. Here, cancer niche construction can be construed as a Bayes-optimal process for the transmission of information across different levels of cellular networks due to its tendency to minimize the overall Kikuchi free energy over the whole system. Their findings suggest that three distinct cancer trajectories—namely, proliferation (local growth), metastasis and apoptosis—can emerge from the natural evolution of the state function (i.e., free energy) in biological systems. These findings have important implications for our understanding and study of cancer cell growth and apoptosis.

Next, Parr describes how biochemical networks in adaptive biological systems can be recast in terms of an inferential message passing scheme that involves the gradient descent on variational free energy towards the least surprising states, based on the organism’s implicit (generative) model of these states. In *Message Passing and Metabolism* [35], he points out that the biochemical regulation of metabolic processes relies on sparse interactions (message-passing) between coupled reactions, with enzymes creating conditional dependencies between reactants. He then extrapolates the law of mass action (the rate of chemical reaction and the concentrations of reactants involved in this process) and the Michaelis–Menten kinetics (which approximates the dynamics of irreversible enzymatic reactions) from the FEP. Assuming the existence of a causal structure in biochemical (metabolic) networks, one can build the sparse message passing scheme to capture the independence of substrates and products, conditioned upon the enzyme and enzyme-substrate complex within such networks. The temporal evolution of the categorical probability of each state within this system can be described by a chemical master equation that takes into account sparse network interactions. Parr describes how the steady state distribution of these dynamics can be recast as a generative model, which suggests that the biochemistry that underlies metabolism follows an inferential message-passing scheme that seeks to minimize free energy. An important extension of Parr’s model is that metabolic disorder can emerge when an enzymatic disconnection by thiamine depletion interrupts message passing and incites aberrant prior beliefs, which gives rise to false (biochemical) inference.

The third contribution follows more traditional applications of the FEP by accounting for conscious, first-person experience. In *The Radically Embodied Conscious Cybernetic Bayesian Brain*, Safron [36] proposes models of embodied conscious agency based on the FEP, extending the Integrated World Modelling Theory of consciousness proposed elsewhere [37] to explicitly account for aspects of intentional actions and agentic experiences. According to the radically embodied account on offer, what we call attention and imagination emerge from the (sometimes liminal) activity of multimodal, action-oriented body maps and representations, realized as neural attractors in the form of ‘embodied self-models’ (ESMs), which conform to the FEP as cybernetic controllers. When functioning online, ESMs allow for overt interactions with affordances, or structured possibilities for environmental interactions. However, Safron suggests subthreshold activations of such ‘quasi-homuncular’ ESMs also underwrite our (affordance-structured) covert abilities to imagine and pursue courses of action, as well as our ability to intentionally deploy attentional resources. Thus, even seemingly abstract representational capacities may be grounded in twin capacities for embodied action and counterfactual explorations of the world. Safron then applies this radically embodied perspective to core aspects of conscious experience. He attempts to chart a middle way between perspectives in the representation wars in cognitive science, describing brains as hybrid machine learning architectures capable of supporting both symbolic and sub-symbolic processes for 4E agents (where cognition is thought to be embodied, embedded, enacted and extended). Safron’s perspective is ecumenical, deploying information-theoretic constructs and representationalist concepts that would be rejected by hard-line proponents of both 4E cognition and more Cartesian (representationalist) approaches. For example, to account for information flow in mammalian brains, Safron deploys constructs that are typically rejected by 4E theorists, such as Cartesian theatres and quasi-homunculi. However, he does so from a radically embodied
perspective, suggesting that such a “strange inversion of reasoning” follows from principles of cognitive development and computational neuroscience.

Unlike the authors above, Goekoop and de Kleijn look beyond the phenotype to consider how the FEP might apply to groups. In Permutation Entropy as a Universal Disorder Criterion [38], they argue that living systems can be described as hierarchical problem-solving machines that embody predictive models of their econiches, called a goal hierarchy, which incorporates a set of lower-order econiches (goals) and corresponding subniches (subgoals) that the system needs to pursue in order to achieve the global econiche (goal) represented at the top of the hierarchy. Using this scheme to frame the rest of their argument, they concentrate on stress responses in organisms, dyads and collectives. Equating stress with free energy or ‘prediction error’, and stress responses with ‘action’, they suggest that as free energy increases, there is a progressive collapse of (allostatic) hierarchical control, eventually resulting in disordered states characterised by behavioural shifts from long-term goal-directed behaviour (e.g., reproductive success) towards short-term goals and habitual behaviours concentrated on self-preservation (e.g., survival). After introducing permutation entropy as a universal measure of disordered states across such systems, they briefly describe how their model can be used to explain disorder at an individual level, before progressing to the transmission of disorder through interpersonal interactions, and concluding with a brief discussion of population-level dynamics.

The idea that the FEP can be extended to social systems is also taken up by Kaufmann and colleagues. In An Active Inference Model of Collective Intelligence [39], the authors propose an active inference model of alignment, describing the manner in which within-scale local interactions (e.g., individual agents’ behaviors) can align with cross-scale global phenomena (e.g., collective behavior) in multi-scale systems. In so doing, they offer a principled, agent-based model that has the potential to function as a workbench to simulate collective intelligence as an emergent phenomenon, across many scales. Although one obvious target for this modelling approach would be human behavior as an emergent phenomenon that ties physiological, cognitive and cultural dynamics, nothing, in principle, limits the application of Kaufmann and colleagues’ model to human phenomena.

In another paper that illustrates the broad scope of multiscale thinking under active inference, Jesse Hoey, in Equality and Freedom as Uncertainty in Groups [40], shows how agents attempting to align with other group members leads to a quasi-equilibrium, or “sweet spot”, at which the group free energy is minimal and the agent’s predictive capacity of higher order parameters, such as those attributed to the group, matches the group’s capacity to predict an agent’s behavior. Hoey further discusses two intriguing trade-offs. Higher agent model complexity leads to lower individual learning capacity with respect to the complexity of the group, resulting in agents who are hobbled in the pursuit of their own ends, but in a group that is more diverse, innovative, and open to change. On the other hand, lower agent model complexity allows the expression of individual preference towards the group, but the group becomes more homogeneous, secure, and closed, as otherwise the pro-social behaviours of individual agents would be hampered. Hoey suggests that such emergent social dynamics provides insights into concepts such as freedom and equality in society, which correspond to changes in model uncertainties and complexities. Oscillation between radical freedom, where no cooperation is possible, and radical equality, where no discovery is possible, is an emergent phenomenon characteristic of Western society; akin to what Karl Marx called historical materialism, which according to many is the main driver of history itself. Could Hoey’s findings initiate research on an active inference model of history as an emergent phenomenon of human societies?

So far, we have considered a range of applications of the FEP to living systems. However, active inference—the process theory derived from the FEP—is increasingly being applied to machine intelligence in practical settings. Use cases in robotics provide an exciting opportunity to test the applicability of active inference to implement sensory processes and motor control in real time. In their review for our Special Issue, How Active Inference Could Help Revolutionise Robotics, Da Costa and colleagues [41] examine the
usefulness of active inference for several core problems in robotics, such as state estimation in artificial perception, motor control, learning, safety and explainability. They argue that active inference may help advance robotics due to several of its core features: it enables explainable artificial intelligence in a manner that operates in situations involving uncertainty, volatility, and noise. This is especially relevant for human-centric applications, such as human–robot interaction and healthcare.

In closing, it is worth recognising that the majority of submissions presented herein focus chiefly on human systems, despite our call for more wide-ranging applications. Nevertheless, it should be clear that the authors’ proposals can be readily extended to other complex adaptive systems, including biological dynamics intrinsic to other lifeforms [34–36], collective, group-level behaviour [39,40], and even non-living systems [41]. Taken together, we hope that the collection of papers presented in our Special Issue motivate others to consider how the FEP might be gainfully applied to their own systems of interest, living or otherwise.

Funding: ZS is supported by The Long-Term Future Fund, provided by the Centre for Effective Altruism. AC is supported by the Australian Laureate Fellowship project, A Philosophy of Medicine for the 21st Century (Ref: FL170100160), and by a Social Sciences and Humanities Research Council doctoral fellowship (Ref: 752–2019-0065).

Acknowledgments: We would like to thank Lancelot Da Costa, Rutger Goekoop, Jesse Hoey, Rafael Kaufmann, Thomas Parr, Adam Safron, and Noor Sajid for their comments on an earlier draft of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Friston, K. A theory of cortical responses. Philos. Trans. R. Soc. B Biol. Sci. 2005, 360, 815–836. [CrossRef] [PubMed]
2. Friston, K. The free-energy principle: A unified brain theory? Nat. Rev. Neurosci. 2010, 11, 127–138. [PubMed]
3. Kiebel, S.J.; Friston, K.J. Free energy and dendritic self-organization. Front. Syst. Neurosci. 2011, 5, 80. [CrossRef] [PubMed]
4. Calvo, P.; Friston, K. Predicting green: Really radical (plant) predictive processing. J. R. Soc. Interface 2017, 14, 20170096. [CrossRef]
5. Friston, K. Life as we know it. Nat. Rev. Neurosci. 2013, 14, 809. [CrossRef] [PubMed]
6. Campbell, J.O. Universal Darwinism as a process of Bayesian inference. Front. Syst. Neurosci. 2016, 10, 49. [CrossRef]
7. Ramstead, M.J.; Badcock, P.B.; Friston, K. Future climates: Markov blankets and active inference in the biosphere. J. R. Soc. Interface 2020, 17, 20200503. [CrossRef]
8. Badcock, P.B.; Friston, K.J.; Ramstead, M.J. Answering Schrödinger’s question: A free-energy formulation. Phys. Life Rev. 2018, 24, 1–16. [CrossRef]
9. Badcock, P.B.; Friston, K.J.; Ramstead, M.J. The hierarchically mechanistic mind: A free-energy formulation of the human psyche. Phys. Life Rev. 2019, 31, 104–121. [CrossRef]
10. Badcock, P.B.; Friston, K.J.; Ramstead, M.J.; Ploeger, A.; Hohwy, J. The hierarchically mechanistic mind: An evolutionary systems theory of the human brain, cognition, and behavior. Cogn. Affect. Behav. Neurosci. 2019, 19, 1319–1351. [CrossRef]
11. Badcock, P.B. Evolutionary systems theory: A unifying meta-theory of psychological science. Rev. Gen. Psychol. 2012, 16, 10–23. [CrossRef]
12. Ramstead, M.J.; Veissière, S.P.; Kirmayer, L.J. Cultural affordances: Scaffolding local worlds through shared intentionality and regimes of attention. Front. Psychol. 2016, 7, 1090. [CrossRef] [PubMed]
13. Veissière, S.P.; Constant, A.; Ramstead, M.J.; Friston, K.J.; Kirmayer, L.J. Thinking through other minds: A variational approach to cognition and culture. Behav. Brain Sci. 2020, 43, E90. [CrossRef] [PubMed]
14. Constant, A.; Clark, A.; Kirchhoff, M.; Friston, K.J. Extended active inference: Constructing predictive cognition beyond skulls. Mind Lang. 2020, in press. [CrossRef]
15. Constant, A.; Ramstead, M.J.; Veissière, S.P.; Campbell, J.O.; Friston, K.J. A variational approach to niche construction. J. R. Soc. Interface 2018, 15, 20170685. [CrossRef]
16. Bastos, A.M.; Usrey, W.M.; Adams, R.A.; Mangun, G.R.; Fries, P.; Friston, K.J. Canonical microcircuits for predictive coding. Neuron 2012, 76, 695–711. [CrossRef]
17. Friston, K.J.; Daunizeau, J.; Kiebel, S.J. Reinforcement learning or active inference? PLoS ONE 2009, 4, e6421. [CrossRef]
18. Friston, K.; FitzGerald, T.; Rigoli, F.; Schwartenbeck, P.; Pezzulo, G. Active inference: A process theory. Neural Comput. 2017, 29, 1–49. [CrossRef]
19. Friston, K.J.; Rosch, R.; Parr, T.; Price, C.; Bowman, H. Deep temporal models and active inference. Neurosci. Biobehav. Rev. 2018, 90, 486–501. [CrossRef]
20. Levin, S. Complex adaptive systems: Exploring the known, the unknown and the unknowable. *Bull. Am. Math. Soc.* 2003, 40, 3–19. [CrossRef]
21. Holland, J.H. Studying complex adaptive systems. *J. Syst. Sci. Complex.* 2006, 19, 1–8. [CrossRef]
22. Schwab, E.; Pienta, K.J. Cancer as a complex adaptive system. *Med. Hypotheses* 1996, 47, 235–241. [CrossRef]
23. Theise, N.D.; d’Inverno, M. Understanding cell lineages as complex adaptive systems. *Blood Cells Mol. Dis.* 2004, 32, 17–20. [CrossRef] [PubMed]
24. Mallick, P. *Complexity and Information: Cancer as a Multi-Scale Complex Adaptive System*, in *Physical Sciences and Engineering Advances in Life Sciences and Oncology*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 5–29.
25. Rea, T.J.; Brown, C.M.; Sing, C.F. Complex adaptive system models and the genetic analysis of plasma HDL–cholesterol concentration. *Pertsp. Biol. Med.* 2006, 49, 490. [CrossRef] [PubMed]
26. Barandiarian, X.; Moreno, A. Adaptivity: From metabolism to behavior. *Adapt. Behav.* 2008, 16, 325–344. [CrossRef]
27. Lansing, J.S. Complex adaptive systems. *Annu. Rev. Anthropol.* 2003, 32, 183–204. [CrossRef]
28. Shine, J.M. The thalamus integrates the macrosystems of the brain to facilitate complex, adaptive brain network dynamics. *Prog. Neurobiol.* 2021, 199, 101951. [CrossRef] [PubMed]
29. Morowitz, H.J.; Singer, J.L.E. *The Mind, the Brain and Complex Adaptive Systems*; Routledge: London, UK, 2018.
30. Page, S.; Miller, J.H. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*; Princeton University Press: Princeton, NJ, USA, 2009.
31. Neace, K.S.; Chipkevich, M.B.A. Designed complex adaptive systems exhibiting weak emergence. In Proceedings of the NAECON 2018–IEEE National Aerospace and Electronics Conference, Dayton, OH, USA, 23–26 July 2018.
32. Steels, L. Evolving grounded communication for robots. *Trends Cogn. Sci.* 2003, 7, 308–312. [CrossRef]
33. Stanescu, A.M.; Nita, A.; Moisescu, M.A.; Sacala, I.S. From industrial robotics towards intelligent robotic systems. In Proceedings of the 2008 4th International IEEE Conference Intelligent Systems, Varna, Bulgaria, 6–8 September 2008.
34. Sajid, N.; Convertino, L.; Friston, K. Cancer niches and their kikuchi free energy. *Entropy* 2021, 23, 609. [CrossRef]
35. Parr, T. Message Passing and Metabolism. *Entropy* 2021, 23, 606. [CrossRef] [PubMed]
36. Safron, A. The radically embodied conscious cybernetic Bayesian brain: From free energy to free will and back again. *Entropy* 2021, 23, 783. [CrossRef] [PubMed]
37. Safron, A. An Integrated World Modeling Theory (IWMT) of consciousness: Combining integrated information and global neuronal workspace theories with the free energy principle and active inference framework; Toward solving the hard problem and characterizing agentic causation. *Front. Artif. Intell.* 2020, 3, 30. [PubMed]
38. Goekoop, R.; de Kleijn, R. Permutation Entropy as a Universal Disorder Criterion: How Disorders at Different Scale Levels Are Manifestations of the Same Underlying Principle. *Entropy* 2021, 23, 1701. [CrossRef] [PubMed]
39. Kaufmann, R.; Gupta, P.; Taylor, J. An active inference model of collective intelligence. *Entropy* 2021, 23, 830. [CrossRef]
40. Hoey, J. Equality and Freedom as Uncertainty in Groups. *Entropy* 2021, 23, 1384. [CrossRef]
41. Da Costa, L.; Lanillos, P.; Sajid, N.; Friston, K.; Khan, S. How Active Inference Could Help Revolutionise Robotics. *Entropy* 2022, 24, 361. [CrossRef]