



Review

A simple theory of every ‘thing’

Ines Hipolito

University of Wollongong, Northfields Ave, Wollongong, 2500, Australia

Available online 25 October 2019

Communicated by Felix Schoeller

Abstract

One of the criteria to a strong principle in natural sciences is simplicity. This paper claims that the Free Energy Principle (FEP), by virtue of unifying particles with mind, is the simplest. Motivated by Hilbert’s 24th problem of simplicity, the argument is made that the FEP takes a seemingly mathematical complex domain and reduces it to something simple. More specifically, it is attempted to show that every ‘thing’, from particles to mind, can be partitioned into systemic states by virtue of self-organising symmetry break, i.e. self-entropy in terms of the balance between risk and ambiguity to achieve epistemic gain. By virtue of its explanatory reach, the FEP becomes the simplest principle under quantum, statistical and classical mechanics conditions. Crown Copyright © 2019 Published by Elsevier B.V. All rights reserved.

Keywords: Hilbert’s 24th problem; Simple proof; Free energy principle

Under a given set of conditions there can be but one simplest proof.
[David Hilbert, *Mathematische Notizbücher*]

1. Introduction

Other things being equal, simpler proofs are better. We can identify them *a posteriori* because they survive the proof of time. Through the sieve of simplicity monumental conjectures became obsolete. Today’s demonstrations, theories and principles, covering hundreds of pages of mathematical journals, will get discarded a few generations from now. Simple proofs survive the test of time and together constitute today the landscape of science.

The difficulty however resides in how to identify, in our time, simplicity in mathematics. This endeavour is so pertinent in mathematics that David Hilbert added it in his selection of mathematical problems – the 24th problem.¹ He described it as,

Criteria of simplicity, or proof of the greatest simplicity of certain proofs. Develop a theory of the method of proof in mathematics in general. Under a given set of conditions there can be but one simplest proof. Quite generally,

E-mail address: inesh@uow.edu.au.

¹ For detailed accounts of Hilbert 24th problem see the Special Issue by Hipólito and Kahle [8].

if there are two proofs for a theorem, you must keep going until you have derived each from the other, or until it becomes quite evident what variant conditions (and aids) have been used in the two proofs. Given two routes, it is not right to take either of these two or to look for a third; it is necessary to investigate the area lying between the two routes.²

The simplified claim ‘there can be but one simplest proof’ will surely be rejected by a majority of scholars but Hilbert was careful enough to qualify it by adding ‘under a given set of conditions’. It might still be doubtful whether ‘under a given set of conditions’, indeed, ‘there can be but one simplest proof’. Nevertheless, Hilbert at least made clear, that we have to identify this ‘given set of conditions’ for different proofs [8].

In the published version of his contribution to the International Congress of Mathematician,³ Hilbert addresses simplicity in relation to rigour,

Besides, it is an error to believe that rigour in the proof is the enemy of simplicity. On the contrary, we find it confirmed by numerous examples that the rigorous method is at the same time, the simpler and the more easily comprehended. The very effort for rigour forces us to find out simpler methods of proof [6].

Hilbert also relates simplicity with consistency. In his talk at the Swiss Mathematical Society with the title *Axiomatisches Denken* (Axiomatic Thinking),⁴ in which Hilbert addresses the question of consistency as not isolated but “belonging to a wide range of most difficult epistemological questions of specific mathematical colouring” which includes the *criterion for the simplicity* of mathematical proofs”.

Looking today at Hilbert’s 24th problem, the fundamental question, in the era of computational models, is the relation between simplicity and complexity. Particularly how to define the natural world – from matter to mind – in terms of the simplest theories or principles. Crucially, this means addressing the challenge of developing theories and principles that take seemingly complex domains and reduce them to something simple. If the explanatory power of theories, on the one hand, is supposed to be assessed by its resistance to falsification, principles, on the other hand, cannot be assessed the same way. More specifically, as self-consistent statements in a formal mathematical system, principles cannot be tested by the means of falsification. The difficulty then is how to determine the explanatory power of principles constituting the landscape of natural science. Given they obey a formal mathematical system, the explanatory power of principles can be determined, following Hilbert’s 24th problem, by their simplicity.

This paper attempts to show that the Free Energy Principle qualifies to a simple principle in natural science. This is attempted to show via Markov blanket partition and, then, their integration within a large-scale Bayesian network via self-organising. It is concluded that the FEP advances an explanation that unifies particles to organic life under quantum, statistical and classical mechanics conditions.

2. Simplifying: partition of the systemic states of things

Free energy, in thermodynamics, is a measure of the amount of work that can be extracted from a system. This can be probabilistically measured via a probability distribution over the exchanges among systems [3]. Within these exchanges, the Free Energy Principle (FEP) aims to explain how systems maintain their order (non-equilibrium steady-state) by restricting themselves to a limited number of states. As a principle for adaptive systems, the FEP is becoming popular to explain cognition ([4,5,17]). This paper attempts to show a stronger claim. That, beyond a unitary principle to cognition, the FEP moreover unifies all things in the natural world, from particles to organismistic life.⁵ This is specifically shown in this section by the partition of systemic states.

The FEP is the philosophical conception of mathematical formalisation of the boundaries between distinguishable things that causally influence each other. Such a principle is highly ambitious because it necessarily entails spelling out a theory of every ‘thing’ – from metaphysics to physics – and their relations, considering that things do not ex-

² The English translations were given by Thiele [18]. Find in full length the originals of Hilbert’s notebooks in Hipólito and Kahle [8].

³ See Hilbert (1902) [6].

⁴ See Hilbert (1918) [7].

⁵ See how Linson et al. [11] make a similar argument for the explanatory power of active inference for natural and artificial embodied cognition.

ist in isolation. Fundamentally, any ‘thing’ is something that can be distinguishable from any other ‘thing’ [2]. Put otherwise, something that exists must possess states that can be distinguished from states that do not constitute a thing.

This is where generative models of probability distributions come handy. We can model the dependencies between things by drawing different models about what is the most likely cause of the coupling. Each model corresponds to an alternative hypothesis about, explicitly, why a specific coupling was caused. The question that follows is how to determine what is relevant to a generative model. Though it depends on the level of enquiry as we will see later, this is the question of how to differentiate each component and their active couplings from the complex system in which they are embedded.

One way to do it is to look into states. Any ‘thing’ is differentiable from another ‘thing’ by its states [2]. States comprise the minimal set of conditional independencies that allow identifying the self-organisation of a complex system, i.e. by separating its internal from its external states. This separation implies the existence of a Markov blanket.⁶ An organism, as a ‘thing’, is composed of its Markov blankets of smaller things. Importantly, this does not mean that each system demarcated, or made salient by Markov blankets has its one specific set of rules by which systems become insulated or dissociable from each other – for example, the brain would become insulated from the rest of the organism, or the organism from its environment. On the contrary, it is because all the systems obey a common principle – as we will see, the FEP – that it is possible to, on the one hand, make them statistically salient, and, on the other, integrate them within a large-scale Bayesian network (under Markov blankets).

There is no privileged level of description. Depending on the level of enquiry, any level constitutes an appropriate level, with many internal microscopic states lying below. For example, on a descending scale: a star, a planet, an environment, an organism, a brain, a neural network, a neuron, neuronal processes, exchanges among intracellular organelles, and so on to atomic and subatomic levels. Crucially, as we will see in detail, what unifies these systems is that all of them self-organise, and thereby can be made salient and be explained in terms of Markov blankets. Inside any Markov blanket we find intrinsic or internal states that themselves are constituted by (mixtures of) ensembles of blanket states – nested Markov blankets [2,12,13]. Each system is what it is by its nested Markov blankets.

The Markovian demarcation is not a metaphysical insulation of systems or internal states. It is a statistical way of highlighting states: internal and external states, which can be further partitioned into sensory states and active states, respectively: $b = \{s, a\} \in B$. Put simply, the existence of a Markov blanket implies a partition of states into external, sensory, active and internal states: $x = \{\eta, s, a, \mu\} \in X$. External states cause sensory states that influence – but are not influenced by – internal states, while internal states cause active states that influence – but are not influenced by – external states. This is by virtue of conditional independencies among the partition of states that result from precluding an influence of external states on active and internal states – and an influence of internal states on sensory states. This dynamical structure is summarised in terms of a marginal flow lemma and its corollaries.

Though the argument has been made that Markov blankets speak to ontological segregation [9], it is important to note that the dependencies induced by Markov blankets create a circular causality: external states cause changes in internal states, via sensory states, while the internal states couple back to the external states through active states, such that internal and external states influence each other in a circular way. Given that there is no privileged level of description, interactive couplings can be found at any physical scale.

An illustration of this is presented in Fig. 1 from Friston [2]. The upper panel shows this partition as it would be applied to action and perception in a brain. In this setting, self-organisation of internal states then corresponds to perception, while active states couple brain states back to external states. Organisms, as actuators, explore these exchanges by virtue of active inference. However, importantly, the lower panel shows the same dependencies but rearranged so that the internal states are associated with the intracellular states of a bacteria *Bacillus*, where the sensory states become the surface states or cell membrane overlying active states (e.g., the actin filaments of the cytoskeleton). Note that the only missing influences are between internal and external states – and directed influences from external (respectively internal) to active (respectively sensory) states. Still, bacteria display essential features for self-organisation: particular states comprised of autonomous and sensory states (See Box 1).

Importantly, we can also find external states influencing internal states in coupled ways in the world of particles. For example, in thermodynamics microstates and microstates (e.g. of the causal relations between pressure-volume or entropy-temperature). Also in quantum phase transitions, normally in Hermitian operators in a Hilbert space, recent

⁶ The term Markov blanket was introduced in the context of Bayesian networks or graphs and refers to the children of a set (the set of states that are influenced), its parents (the set of states that influence it) and the parents of its children [2].

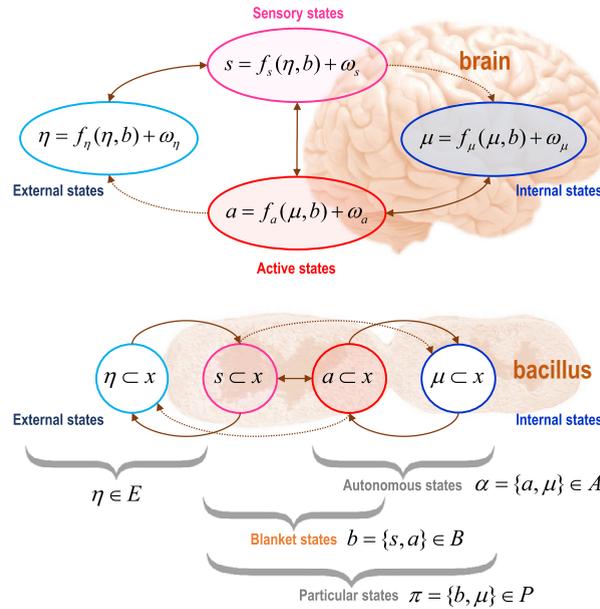


Fig. 1. *Markov blankets*. This probabilistic graphical model illustrates the partition of states into *internal* states (blue) and hidden or *external* states (cyan) that are separated by a Markov blanket – comprising *sensory* (magenta) and *active* states (red). [2].

years have seen an outburst of experimental activities studying quantum correlated systems, such as intermediate-coupling phenomena and extreme-interaction regimes ([19]; see also [14,10]). Very recently, the theoretical idea of ultrastrong coupling between light and matter has been made an experimental reality [1].

Active matter is another example. Each particle corresponds to the Lorentz systems⁷ ‘dressed’ with blanket states to create an internal state and enable interactions among particles [2]. Equations of motion can be used to simulate coupled (random) dynamic systems (i.e. particles) to illustrate self-organisation, i.e. the conditional dependencies among particles, where each particle comprises its Markov blanket and internal states (see Fig. 2). These couplings can be shown by the emergence in terms of Markov blankets and internal states. Importantly, these systems – from particles to brains – maintain continuous, strong interactions with the environment, so to avoid dissipation.

The last question is, how is this implemented? Any system can be interpreted as a network system under the FEP. A shared goal unifies all elements of a system: to minimise its free energy. Each system, as a self-organising system, possesses a phase space of all possible states it can occupy within which a subspace of its most likely states. Put differently, by virtue of it being a random dynamic system, it possesses an itinerant orbit in phase space that keeps revisiting the same states. These states are termed as a pullback or random global attractor.

The key point is that any system that possesses an attractor, and thereby minimises the free energy, must maintain some sort of synchrony with its environment. Openness to the environment is a property of self-organising systems. Brains, organisms, cells, particles, are never isolated by virtue of being continuously in an exchange of the internal and external states of a system. Depending on the level of description, this exchange comprises the existence of a set of internal states and a set of external states; and, finally, a set of boundary states that separates them, i.e. a Markov blanket [15]. By virtue of a non-privileged scale, exchanges and delimitations between things can be elegantly formalised by Markov blankets, all the way up to the brain, all the way down to the world of particles.

In conclusion, by the light of the Markov blanket, it is possible to explain the exchanges and dependencies within a random dynamical system. The set of state that separates internal and external states, i.e. a Markov blanket, has been expressed in terms of how the (marginal) flow of certain states depends on other states [2]. The crucial point for the Markov blanket, at any scale, is that its boundaries are dictated by flows that depend upon certain states. It is by

⁷ This implies that, in any physical system, in the absence of perfect knowledge of the initial conditions, our ability to predict its future course will always fail. Which, notably, underscores that physical systems can be completely deterministic and yet still be inherently unpredictable even in the absence of quantum effects.

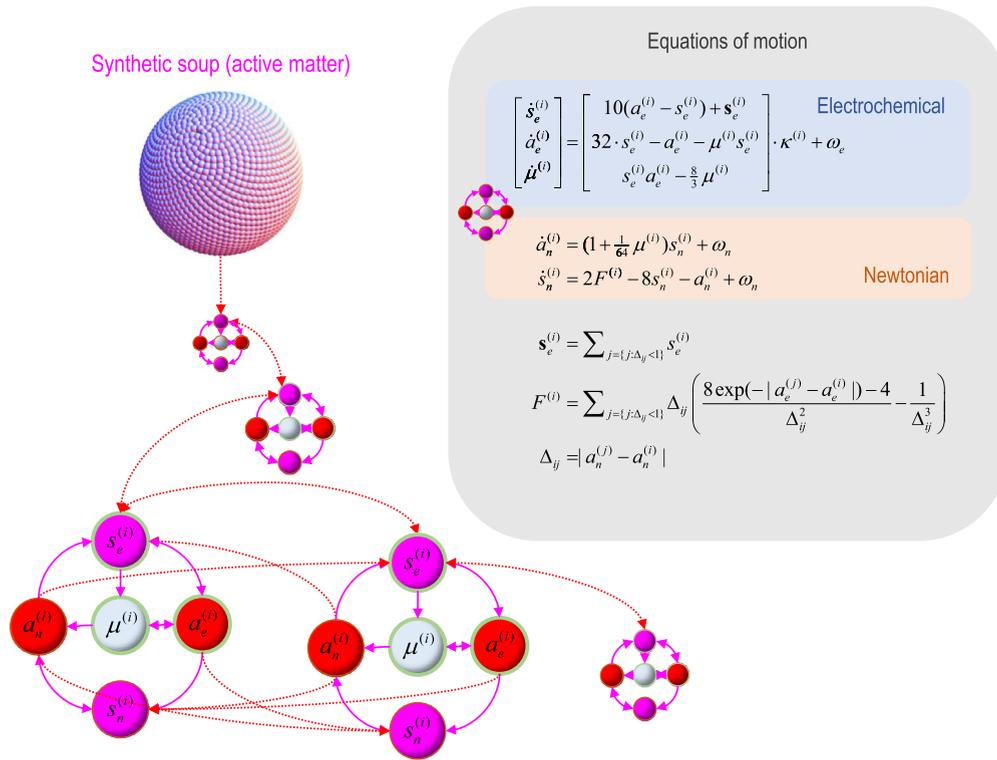


Fig. 2. Synthetic soup and active matter. The states with orange outlines are electrochemical states and the remaining pair constitute Newtonian states. Active states (red circles) play the role of the position, while sensory states (magenta circles) become velocity that depends on active states [2].

their flexibility that Markov blankets allow explaining itinerant couplings while still drawing statistical boundaries. Which is possible because any random dynamical system possesses information-length to time-dependent densities with Markov blankets. Therefore, hierarchical levels, in any system, are a necessary consequence of any random dynamical system that possesses a Markov blanket. Markov blankets demarcate itinerant boundaries of couplings and contextual functions in *ad infinitum* levels of inquiry.

3. Self-organising systems and symmetry breaking

Having set the partition of systemic states of things in the previous section, we now turn to how to integrate them within a large-scale Bayesian network via self-organising in terms of symmetry breaking (by diverging to different regimes of phase-space). This is central because it elucidates how all systems share the formal explanation of intricate self-organised coupling between particular dynamics and external states.

Self-organising systems reduce their self-entropy to the extent allowed by coupling to external states (or other Markov blankets) and random fluctuations. Following Friston [2], using information theory, it is possible to interpret and explain this kind of behaviour in terms of the avoidance of complexity cost (i.e. risk) and inaccuracy (i.e. ambiguity). Self-organising systems are fundamentally characterised by the possibility of non-equilibrium dynamics. Which means that they can diverge from near trajectories to different regimes of phase space – this is known as symmetry breaking. Things, such as particles, cells or organisms, can be described in terms of exploring their state-space while having a well-defined attracting manifold with low measure (i.e. low entropy). In colloquial terms, things are ‘curious’ enough to explore their range of possibilities as a game between reducing ambiguity and risk control (otherwise, the particle dissipates). Because of this important feature, much of life sciences is concerned with the problem of how to formalise multiscale, itinerant and chaotic dynamics. Without the partitions allowed by Markov blankets, as seen above, we can only talk about entropy of a density and how it changes with time. The introduction of Markov blankets

changes the game as it allows us to establish a crucial distinction between external and internal states of a particle in terms of ensemble densities [2].

BOX 1

PARTICULAR STATES: STATES THAT CONSTITUTE A PARTICLE, NAMELY AUTONOMOUS AND SENSORY STATES (ALSO KNOWN AS BLANKET OR INTERNAL STATES)

ENTROPY: LACK OF ORDER OR PREDICTABILITY; GRADUAL DECLINE INTO DISORDER

PARTICULAR SELF-ENTROPY: ENTROPY OF PARTICULAR STATES.

AUTONOMOUS STATES: STATES THAT ARE NOT INFLUENCED BY EXTERNAL STATES

MUTUAL INFORMATION (OR RELATIVE ENTROPY): THE UNCERTAINTY ABOUT PARTICULAR STATES MINUS THE UNCERTAINTY, GIVEN THE EXTERNAL STATES.

COMPLEXITY: THE DIVERGENCE BETWEEN A POSTERIOR AND PRIOR DISTRIBUTION OVER EXTERNAL (HIDDEN) STATES.

ACCURACY: EXPECTED LOG PROBABILITY OF PARTICULAR STATES, UNDER THE POSTERIOR.

AMBIGUITY: THE MARGINAL FLOW OF AUTONOMOUS STATES WILL APPEAR TO MINIMISE THE UNCERTAINTY ABOUT SENSORY STATES, GIVEN EXTERNAL STATES.

CONDITIONAL ENTROPY: EXPECTED INACCURACY (I.E. AMBIGUITY), WHILE MUTUAL INFORMATION, BECOMES THE EXPECTED COMPLEXITY COST (I.E. RISK).

INFORMATION GAIN: OR EPISTEMIC VALUE CORRESPONDS TO THE PROBABILITY DENSITY OVER EXTERNAL STATES AFFORDED BY SENSORY STATES, I.E. THE KULLBACK-LEIBER (KL) DIVERGENCE BETWEEN THE POSTERIOR DENSITY WITH AND WITHOUT SENSORY STATES, CONDITIONED UPON AUTONOMOUS STATES.

One way to capture this non-equilibrium dynamics of steady-state densities is with a separation between autonomous $\alpha = \{a, \mu \in A$ and sensory states $s \in S$. Particles are constituted by their particular states, namely autonomous and sensory states (see Box 1). Autonomous states are states that cannot be influenced by external states. Ultimately what a particle tends to avoid is its particular self-entropy (dissipation). To do so, a system must self-organise in terms of an autonomous suppression of self-entropy (i.e. surprisal) towards relative entropy. Relative entropy, also known as mutual information, refers to the uncertainty about particular states minus the uncertainty, given the external states. For example, when there is no reduction in uncertainty afforded by knowing the external states, mutual information is zero. In short, suppressing entropy, on the one hand, means minimisation of mutual information, where it plays the role of risk, and on the other minimisation of ambiguity (towards accuracy), where it requires a maximising of mutual information, i.e. information gain.

In turn, self-entropy can be decomposed into risk (resp. complexity) and ambiguity (resp. inaccuracy) resolving components that look as if the flow of autonomous states mediates them. Where complexity is the divergence between a posterior and prior distribution over external (hidden) states⁸; and ambiguity refers to the minimising of the uncertainty about sensory states, given external states. In more detail, self-organisation will appear to seek out regimes of phase-space in which external states cause unambiguous sensory-states. These dynamics are self-organising in the sense that (on average), autonomous states will appear to reduce the entropy of particular states. That particular entropy is the mutual information between blanket and external states plus their conditional uncertainty, conditioned on the external states. In other words, autonomous states will appear to minimise the statistical coupling (mutual information) with external states while, at the same time, resisting their dispersion, under any given hidden states.

The conditional entropy becomes expected inaccuracy (i.e. ambiguity), while mutual information becomes the expected complexity cost (i.e. risk). Its goal is information gain or epistemic value, that is probability density over external states afforded by sensory states, i.e. the Kullback-Leiber (KL) divergence between the posterior density with and without sensory states, conditioned upon autonomous states. This is also the mutual information between the sensory and hidden states, afforded by autonomous activity – this measure is also called, in the life sciences,

⁸ Notably, reducing complexity cost underwrites Occam's principle, i.e., the best explanation provides an accurate account with the smallest change in posterior beliefs relative to prior beliefs (Penny et al. [16]).

Bayesian surprise or salience. In short, information gain brings us back to mutual information as the key against entropy or dissipation, as it comprises the information between external and sensory states minus the third-order mutual information (among external, sensory and autonomous states).

Ultimately, the goal of this information gain, in terms of mutual information, is not to represent the truth out there with its veridical details, but to deal with the consequences of exploratory actions to improve future interactions, on atomic or subatomic levels.

4. Conclusion

The FEP constitutes a singular and simple explanation for all the systems, from particles to the brain: all are explained by minimization of expected free energy, expected surprise or uncertainty. This was shown by demonstrating that every ‘thing’ can be distinguished from another ‘thing’ and further partitioned into particular states, all the way down, up and out (via Markov blankets). These ‘things’ have also been illustrated as self-organising systems that reduce their self-entropy to the extent allowed by coupling to external states (or other Markov blankets) and random fluctuations. Expressing self-organisation in terms of self-entropy, risk and ambiguity crucially allows talking about – and quantifying – self-organisation in terms of (inter)active behaviour instead of representational systems.

Declaration of competing interest

No conflict of interest exists.

Funding

This work was funded by the University of Wollongong (Australia).

Intellectual property

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

References

- [1] Di Stefano O, Settinieri A, Macri V, Garziano L, Stassi R, Savasta S, et al. Resolution of gauge ambiguities in ultrastrong-coupling cavity quantum electrodynamics. *Nat Phys* 2019;1.
- [2] Friston K. A free-energy principle for a particular physics. arXiv preprint arXiv:1906.10184, 2019.
- [3] Friston K, Kilner J, Harrison L. A free energy principle for the brain. *J Physiol* 2006;100(1–3):70–87.
- [4] Friston K. The free-energy principle: a unified brain theory?. *Nat Rev Neurosci* 2010;11(2):127.
- [5] Gershman SJ. What does the free energy principle tell us about the brain?. arXiv preprint arXiv:1901.07945, 2019.
- [6] Hilbert D. Mathematical problems. *Bull Am Math Soc* 1902;8:437–79.
- [7] Hilbert D. Axiomatisches Denken. *Math Ann* 1918;78:405–15. English translation.
- [8] Hipólito I, Kahle R. Discussing Hilbert’s 24th problem. *Philos Trans R Soc Lond, Ser A* 2019;377:20180040.
- [9] Hohwy J. The self-evidencing brain. *Noûs* 2016;50(2):259–85.
- [10] Hwang MJ, Rabl P, Plenio MB. Dissipative phase transition in the open quantum Rabi model. *Phys Rev A* 2018;97(1):013825.
- [11] Linson A, Clark A, Ramamoorthy S, Friston K. The active inference approach to ecological perception: general information dynamics for natural and artificial embodied cognition. *Front Robot. AI* 2018;5:21.
- [12] Kirchhoff MD, Kiverstein J. How to determine the boundaries of the mind: a Markov blanket proposal. *Synthese* 2019;1–20.
- [13] Kirchhoff M, Parr T, Palacios E, Friston K, Kiverstein J. The Markov blankets of life: autonomy, active inference and the free energy principle. *J R Soc Interface* 2018;15(138):20170792.
- [14] Morales A, Zupancic P, Léonard J, Esslinger T, Donner T. Coupling two order parameters in a quantum gas. *Nat Mater* 2018;17(8):686.
- [15] Palacios ER, Isomura T, Parr T, Friston K. The emergence of synchrony in networks of mutually inferring neurons. *Sci Rep* 2019;9(1):6412.
- [16] Penny WD, Stephan KE, Mechelli A, Friston KJ. Comparing dynamic causal models. *Neuroimage* 2004;22:1157–72.
- [17] Ramstead MJ, Constant A, Badcock PB, Friston KJ. Variational ecology and the physics of sentient systems. *Phys Life Rev* 2019;31:188–205. <https://doi.org/10.1016/j.plrev.2018.12.002> [in this issue].
- [18] Thiele R. Hilbert’s twenty-fourth problem. *Am Math Mon* 2003;110:1–24. <https://doi.org/10.1080/00029890.2003.11919933>.
- [19] Vojta M. Quantum phase transitions. *Rep Prog Phys* 2003;66(12):2069.