Free Energy Principle as a model of biological and cognitive self-organization

Tomasz Korbak

November 12, 2017

Institute of Philosophy and Sociology,
Polish Academy of Sciences

FEP and mind-life continuity

Free energy principle (FEP) is a normative theory of brain function derived from the properties required by active agents to survive in dynamic, changing environments.

A wide range of neurobiological and psychological phenomena can be accounted for within FEP (think predictive coding), but its aims are even more ambitious – to provide a general theory of life and mind as processes of inference-based self-organization in an uncertain world.

The goal of this talk is review the prospects of generalizing FEP to subsume enactive theory of mind.

Basic assumptions behind FEP

Let us consider an arbitrary biological agent in an arbitrary environment.

- **A1** The agent prefers certain environmental states over others.
- **A2** The agent believes in its own survival.

Basic assumptions behind FEP

A1 The agent prefers certain environmental states over others.

- Thermodynamically, this is equivalent of being a far-from-equilibrium system.
- From a cognitive science point of view, these preferences can be seen as constrains imposed due to embodiment of the agent.
- Mathematically, preference is formulated as a (utility) function over states.

A2 The agent believes in its own survival.

- *Belief* here is to be understood in Bayesian terms (as a probability distribution)
- *Survival* means occupying (on average) a limited set of preferred states.

The generative model

Agent's optimistic belief about fulfillment of its preferences is to be understood as having a generative model of its $p(s, o)$, which breaks down into

1. $p(o)$, a prior on expected/preferred sensory stimuli
2. $p(s|o)$, a model of hidden causes behind these stimuli

where $s \in S$ denotes hidden states and $o \in O$ denotes sensory states.

Survival means occupying (on average) a limited set of preferred states

The expectation of experiencing a limited set of stimuli means that the entropy of $p(o)$ is low and should be minimized:

$$H(O) = -\int_{o \in O} p(o) \log p(o) \; do \tag{1}$$

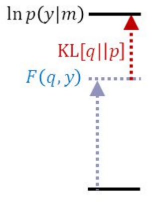
Under ergodic assumptions (equality of time-average and ensemble-average), minimizing entropy of beliefs about sensory states boils down to minimizing surprise ($-\log p(o_t)$) over time.

$$H(O) = \lim_{T \to \infty} -\frac{1}{T} \int_0^T \log p(o_t) \; dt \tag{2}$$

Surprise minimization as variational inference

However, true statistical properties of $p(o)$ are not known for the agent. Inferring them would entail computing the marginal distribution:

$$p(o) = \int_{s \in S} p(o|s)p(s) \; ds \tag{3}$$



but

integrating over all possible states $s \in S$ renders exact inference *intractable*.

Instead, we employ an approximate method known as *variational inference*: we optimize a surrogate model $q(s)$ to minimize Kullback-Leibler divergence between $q(s)$ and $p(s|o)$, which is known.

Minimizing this divergence places an upper bound on surprisal known as the *free energy* (F):

$$F(o, u) = -\log p(o) + D_{KL}(q_u(s)||p(s|o)) \tag{4}$$

FEP aspires to be a general theory of life and mind

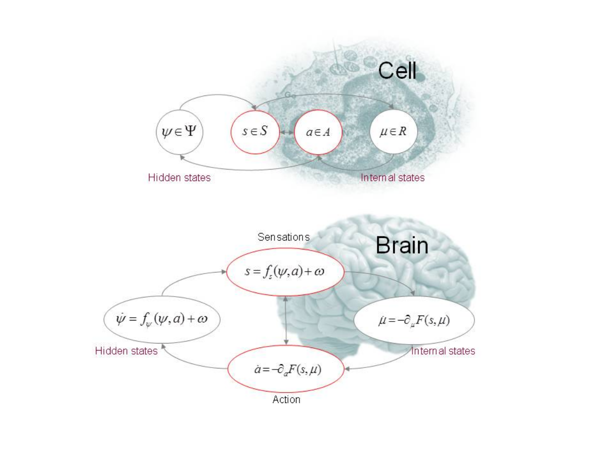


Figure 1: As free energy is a functional of sensory and internal states, its minimization is with respect to both. This corresponds to active inference (changing the state of the world such that sensory inputs it generates are less surprising) or perceptual inference (changing the model such that current sensory input is less surprising), respectively.

Autopoiesis as a theory of life and mind

An autopoietic system, argued to be a minimal model of a living creature, is an organized system characterized by two properties:
(a) it continuously realizes the networks of processes that maintain its existence,
(b) it has a concrete spatial boundary (specified by the topological domain of its realization as a network).

FEP subsumes autopoiesis

1. continuously realizes the networks of processes that maintain its existence
   The mathematical formulation of FEP explicitly captures the distinctive circular causality (known as "action-perception loop") between sensory, action, internal and hidden states, which enactive cognitive scientists find to be the hallmark of mind and life
2. it has a concrete spatial boundary (specified by the topological domain of its realization as a network).
   Maintaining an upper bound on the entropy of sensory states entails existence of a *Markov blanket* consisting of action states (which are dependent on internal states but not external states) and sensory states (which are dependent on external states but not internal states)

FEP enhances autopoiesis

Mere autopoiesis seems hardly enough for a viable account of the continuity between life and cognition

1. Constructive requirements: autopoiesis is too *static* and *abstract* to distinguish the living from the non-living, it is lacking all consideration of *material* and *energetic* conditions for life as a far-from-equilibrium process.
2. Interactive requirements: autopoiesis is too *isolated*; it thus ignores a possible incorporation of environment and a dependence on other living systems.
3. Normative requirements: autopoesis cannot account for normative activity such as *adaptivity* and *goal-directed action*.
4. Historical requirements: autopoiesis is lacking a capacity for *memory*, *learning*, *evolution*, and *development*.
5. Phenomenological requirements: autopoiesis is insufficient for grounding a *lived perspective of concern*

[Froese and Stewart, 2012]

FEP enhances autopoiesis

1. Constructive requirements: autopoiesis is too *static* and *abstract* to distinguish the living from the non-living, it is lacking all consideration of *material* and *energetic* conditions for life as a far-from-equilibrium process.
2. Interactive requirements: autopoiesis is too *isolated*; it thus ignores a possible incorporation of environment and a dependence on other living systems.
   FEP explicitly assumes living systems to be materially and thermodynamically open systems. Additionally, while the leap from a self-sufficient cell to multicellular organism and multi-organism ecosystems poses a challenge to autopoiesis, the concept of Markov blanket can be easily cast in terms of a nested hierarchy of blankets-within-blankets [Clark, 2017].

FEP enhances autopoiesis

1. Normative requirements: autopoesis cannot account for normative activity such as *adaptivity* and *goal-directed action*.
   FEP cast in Bayesian terms in essentially a normative theory of what should an organism to maintain its integrity in a changing environment (the answer is, obviously, that it should act to minimize its free energy)
2. Historical requirements: autopoiesis is lacking a capacity for *memory*, *learning*, *evolution*, and *development*.
   Positing tunable internal states FEP straightforwardly enables learning.
3. Phenomenological requirements: autopoiesis is insufficient for grounding a *lived perspective of concern*
   The notion of consciousness as a process of inference seems to be a fruitful framework for explaining a number of psychiatric diseases (or, altered states of consciousness) [Friston, 2017].

Conclusions

1. According to Free Energy Principle (FEP), cognition boils down to free energy minimization.
2. Minimizing free energy is equivalent to approximate Bayesian inference.
3. According to autopoietic enactivism (AE), cognition is a process of self-construction under a functional boundary (autopoisesis)
4. Minimizing free energy entails a process of self-construction under a functional boundary.
5. For a theory of mind to be interesting and plausible, it should not only demonstrate the coherence of enactive and computational perspectives, but also impose new constraints.
6. FEP imposes additional constraints on autopoiesis (i.e. thermodynamic, interactive).
7. FEP imposes additional constraints on computational approaches (i.e. algorithmic, implementational, ecological)
8. Therefore, FEP implies autopoiesis and is a plausible and interesting theory of mind.

Thanks for your attention! Questions?

This work was done within a project “Formal and computational models of self-organization in cognitive science”, which is supported by Polish Ministry of Science and Higher Education grant DI2015010945.

Alternative formulations of free energy

There are two alternative formulations of the free energy equation.

$$F = D_{\mathrm{KL}}(q(s) \,||\, p(s) - E_q[\log p(o|s)] \tag{5}$$

The complexity-accuracy tradeoff formulation, which posed FEP as two adversarial objectives: of maximizing accuracy of the model and minimizing its complexity. This is intuitive in the light of machine learning, when we add a regularization term to prevent overfitting of the model to its training data.

$$F = E_q[-\log p(o, s)] - H[q(s)] \tag{6}$$

This formulation shows a peculiar analogy with Helmholtz thermodynamic free energy, i.e. internal energy of the system minus its entropy (at a constant temperature and volume). It also relates to maximum entropy principle.

This work was done within a project "Formal and computational models of self-organization in cognitive science", which is supported by Polish Ministry of Science and Higher Education grant DI2015010945.

Clark, A. (2017).
How to Knit Your Own Markov Blanket:.
Technical report, Theoretical Philosophy/MIND Group JGU Mainz.
DOI: 10.15502/9783958573031.

Friston, K. (2017).
The mathematics of mind-time.
Aeon.

Froese, T. and Stewart, J. (2012).
Life After Ashby: Ultrastability and the Autopoietic Foundations of Biological Autonomy.
Cybernetics and Human Knowing, 17(4).

This work was done within a project "Formal and computational models of self-organization in cognitive science", which is supported by Polish Ministry of Science and Higher Education grant DI2015010945.