Supplementary information S1 (box): The entropy of sensory states and their causes

This box shows that the entropy of hidden states in the environment is bounded by the entropy of sensory states. This means that if the entropy of sensory signals is minimised, so is the entropy of the environmental states that caused them. For any agent or model m the entropy of generalised sensory states $\tilde{s}(t) = [s, s', s'', \ldots]^T$ is simply their average surprise $-\ln p(\tilde{s} \mid m)$ (with a sight abuse of notion)

$$H(\tilde{s} \mid m) := \int -p(\tilde{s} \mid m) \ln p(\tilde{s} \mid m) d\tilde{s} = \lim_{T \to \bullet} \int_0^T -\ln p(\tilde{s}(t) \mid m) dt \qquad \text{S1.1}$$

Under ergodic assumptions, this is just the long-term time or path-integral of surprise. We will assume sensory states are an analytic function of hidden environmental states plus some generalised random fluctuations

$$\begin{aligned} \tilde{s} &= g(\tilde{x}, \theta) + \tilde{z} \\ \dot{\tilde{x}} &= f(\tilde{x}, \theta) + \tilde{w} \end{aligned} \qquad \text{S1.2}$$

Here, hidden states change according to the stochastic differential equations of motion (with parameters θ) in S1.2. Because $\tilde{x}$ and $\tilde{z}$ are statistically independent, we have (see Eq. 6.4.6 in Jones 1979, p149)

$$I(\tilde{s}, \tilde{z}) = H(\tilde{s} \mid m) - H(\tilde{x} \mid m) - \int p(\tilde{x} \mid m) \ln |\partial_{\tilde{x}} g| d\tilde{x} \qquad \text{S1.3}$$

Here, $I(\tilde{s}, \tilde{z}) \geq 0$ is the mutual information between the sensory states and noise. By Gibb's inequality this cross-entropy or Kullback-Leibler divergence is non-negative (Theorem 6.5; Jones 1979, p151). This means the entropy of the sensory states is greater than the entropy of the sensory mapping. Here. $\partial_{\tilde{x}} g$ is the sensitivity or gradient of the sensory mapping with respect to the hidden states. The integral in S1.3 reflects the fact that entropy is not invariant to a change of variables and assumes that the sensory mapping $g:\tilde{x}\rightarrow\tilde{s}$ is diffeomorphic (i.e., bijective and smooth). This requires the hidden and sensory state-spaces to have the same dimension, which can be assured by truncating generalised states at an appropriately high order. For example, if we had n hidden states in m generalised coordinates of motion, we would consider m sensory states in n generalised coordinates; so that $\dim(\tilde{x})=\dim(\tilde{s})=n\times m$. Finally, rearranging S1.3 gives

$$H(\tilde{x}\,|\,m)\le H(\tilde{s}\,|\,m)-\int p(\tilde{x}\,|\,m)\ln|\partial_{\tilde{x}}g|d\tilde{x} \qquad \text{S1.4}$$

In conclusion, the entropy of hidden states is upper-bounded by the entropy of sensations, assuming their sensitivity to hidden states is constant, over the range of states encountered.

Clearly, the ergodic assumption in S1.1 only holds over certain temporal scales for real organisms that are on a trajectory from birth to death. This scale can be somatic (e.g., over days or months, where development is locally stationary) or evolutionary (e.g., over generations, where evolution is locally stationary).

Reference

Jones, DS. (1979). *Elementary information theory*. Publisher: Oxford: Clarendon Press; New York: Oxford University Press