Integrated Information: A Formulation of Dynamical Complexity for Neuronal and Artificial Systems

Xerxes D. Arsiwalla[1], Daniel Pacheco[1], Paul Verschure[1,2]

[1]Institute for Bioengineering of Catalonia (IBEC), [2]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Email: x.d.arsiwalla@gmail.com

CDAC Project: (ERC-2013- ADG 341196)

Motivation:

- A **global complexity** measure that simultaneously takes into account network structure and dynamics
- Useful for calibrating **complexity of brain states** associated to rest and behavior
- A **clinical measure** of consciousness, cognition & volition
- How does one compare **complexity** of **biological agents** such as bacteria, bees, C. elegans, primates, humans, etc. to **AI agents** such as deep neural nets, multi-agent robots, Siri, Watson, etc.?
- How can one generalize wakefulness & awareness for non-biological systems to obtain **scales of consciousness** for artificial systems?

Mathematical Formulation:

The information gain following a state transition equals reduction in uncertainty about initial state $\mathbf{X_0}$. Effective information is defined as the subset of the above that is generated by the system as a whole and is given by the **Kullback-Leibler Divergence** between the multivariate conditional distribution of the system versus the joint independent composite of its parts

$$\Phi_{\mathcal{P}\to r}(\mathbf{X_0}\to\mathbf{X_1}=\mathbf{x}') = D_{KL}\left(\mathbf{P}_{\mathbf{X_0}|\mathbf{X_1}=\mathbf{x}'}\middle\|\prod_{k=1}^{r}\mathbf{P}_{\mathbf{M_0^k}|\mathbf{M_1^k}=\mathbf{m}'}\right)$$

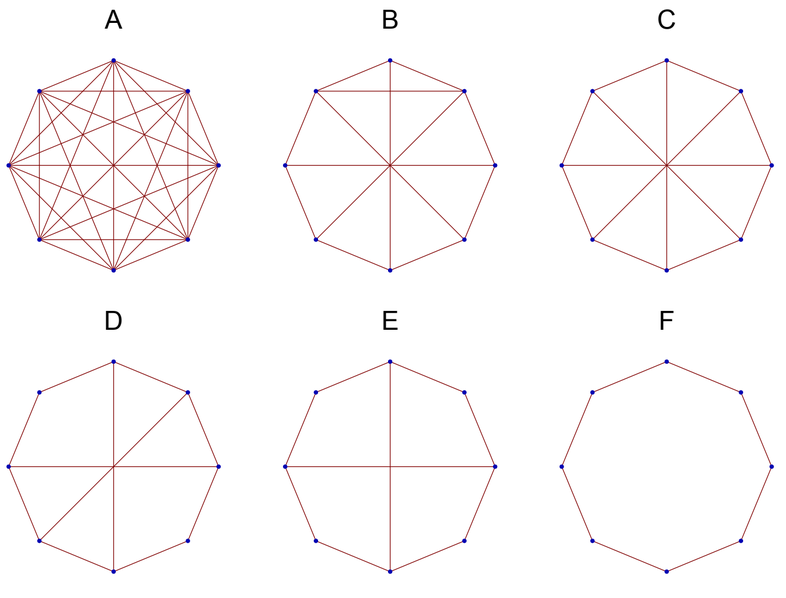
Network states: $\mathbf{X_{t_0}} = \begin{pmatrix} x1_{t_0} \\ \cdot \\ \cdot \\ \cdot \\ x_{n_{t_0}} \end{pmatrix} \longmapsto \mathbf{X_{t_1}} = \begin{pmatrix} x1_{t_1} \\ \cdot \\ \cdot \\ \cdot \\ x_{n_{t_1}} \end{pmatrix}$

Dynamical model: $\mathbf{X_{t_1}} = \mathcal{A}\,\mathbf{X_{t_0}} + \mathcal{E}$ (Linear Dynamics with Noise)

$\mathbf{X_0} \sim \mathcal{N}(\bar{\mathbf{x}}_\mathbf{0}, \mathbf{\Sigma}(\mathbf{X_0}))$ $\quad$ $\mathbf{X_1} \sim \mathcal{N}(\bar{\mathbf{x}}_\mathbf{1}, \mathbf{\Sigma}(\mathbf{X_1}))$ (Gaussian variables)

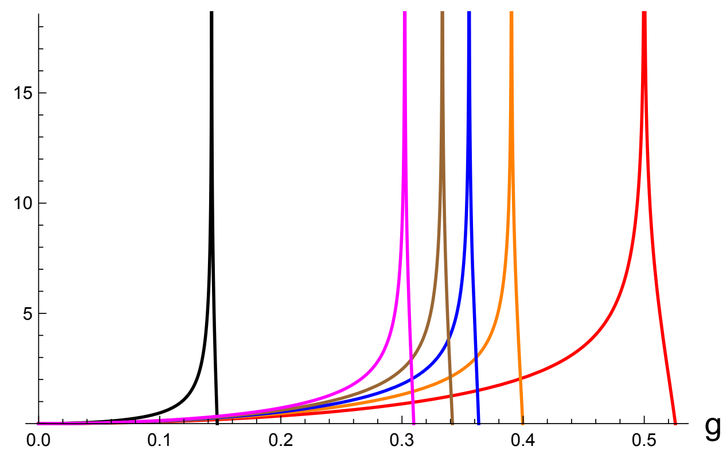
Conditional entropy: $\mathbf{H}(\mathbf{X_0}|\mathbf{X_1}) = \frac{1}{2}n\log(2\pi e) - \frac{1}{2}\log\left[\det\mathbf{\Sigma}(\mathbf{X_0}|\mathbf{X_1})\right]$

$$\langle\Phi\rangle_{\mathcal{P}\to r}(\mathbf{X_0}\to\mathbf{X_1}) = -\mathbf{H}(\mathbf{X_0}|\mathbf{X_1}) + \sum_{k=1}^{r}\mathbf{H}(\mathbf{M_0^k}|\mathbf{M_1^k})$$

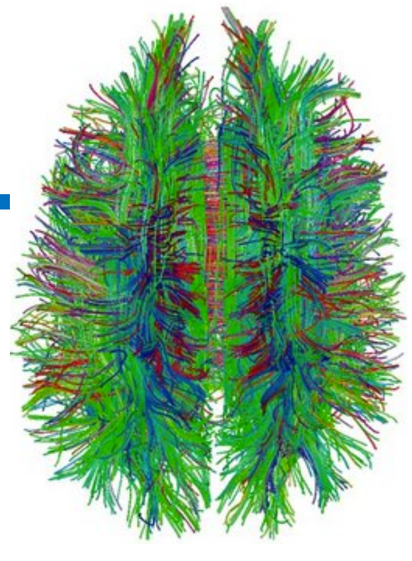


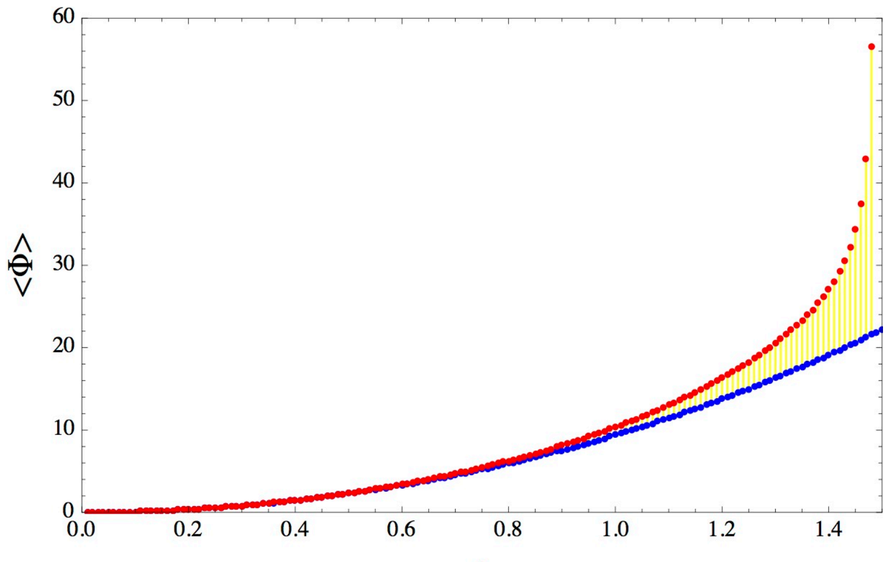
Figures: (Left): Six network topologies upon which we apply the above formulation. **(Right):** The morphospace of consciousness showing a taxonomy of biological and artificial agents using three kinds of complexity scales.

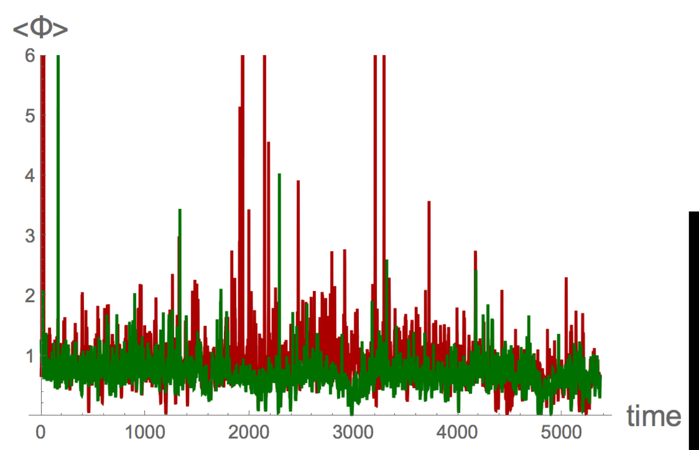
Results:

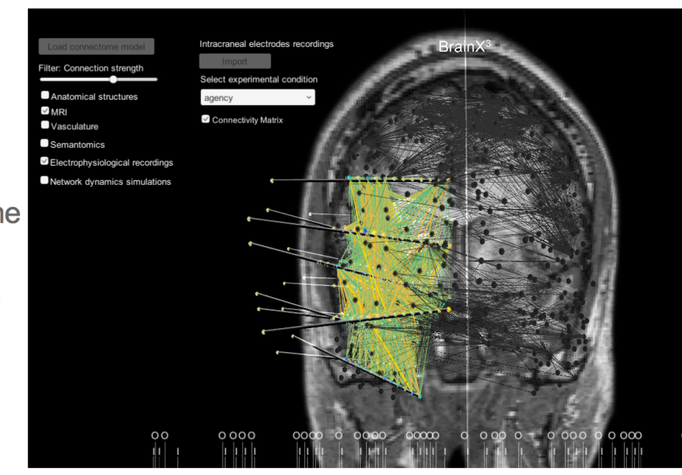


Figures: **(Left)** Phi profiles for the 6 networks shown below. Stable solutions until couplings hit critical point showing high phi near criticality.
(Below): Tractography data from the human connectome (Hagmann 2008) and corresponding phi profile for data vs. randomized network.

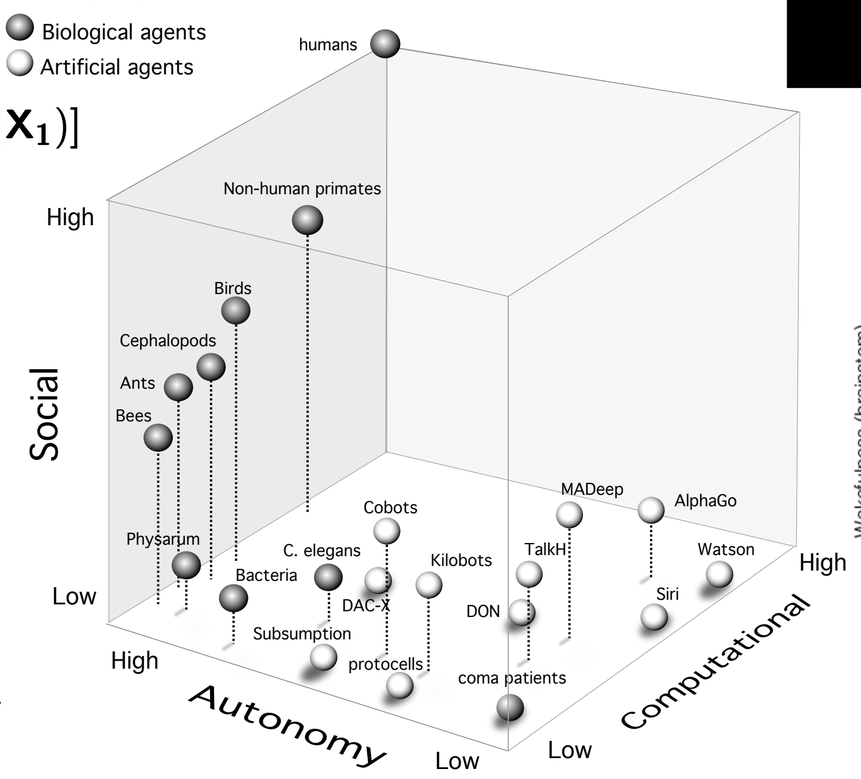


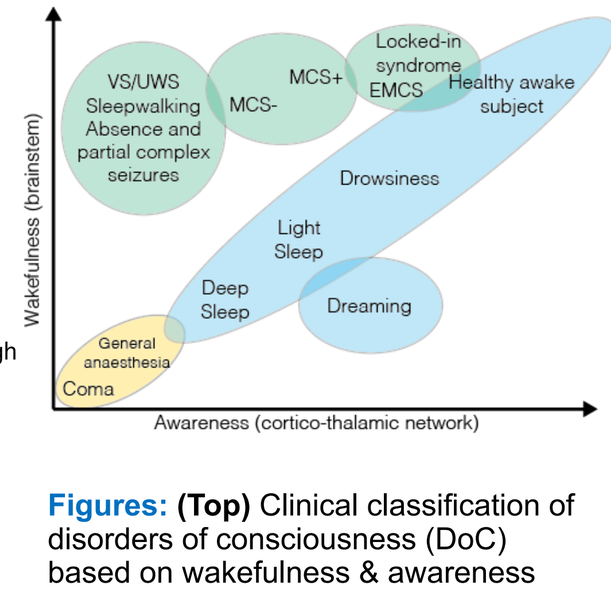






Figures: **(Right)**: FC from epilepsy patients using iEEG recordings. **(Top)**: Temporal phi profiles from FCs obtained in active vs. passive navigation tasks.





Figures: **(Top)** Clinical classification of disorders of consciousness (DoC) based on wakefulness & awareness (Laureys 2005).

Conclusions:

- An analytic framework for computing integrated information for large biological networks
- High phi near the edge of instability
- As a clinical measure of cognitive complexity
- Generalization to complexity of AI systems

References: **[1]** Arsiwalla, X. D., & Verschure, P. F. (2016). The Global Dynamical Complexity of the Human Brain Network. *Applied Network Science*, *1*(1), 16.
[2] Arsiwalla, X. D., Sole, R., Moulin-Frier, C., Herreros, I., Sanchez-Fibla, M., & Verschure, P. (2017). The Morphospace of Consciousness. *arXiv:1705.11190*.
[3] Arsiwalla, X. D., Pacheco, D., Principe, A,. Rocamora, R., & Verschure, P. F. (2018). A Temporal Estimate of Integrated Information for Intracranial Functional Connectivity. ICANN 2018, Springer.