Causality, autonomy, and emergence in neural systems

Anil K Seth

Reader and EPSRC Leadership Fellow
Co-Director, Sackler Centre for Consciousness Science
University of Sussex, UK

Paris, November 2009

www.anilseth.com
What is causality?

Granger (G-)causality

Causal networks in neural systems

G-autonomy

G-emergence

Consciousness
What is causality?
8th century illustration of the Sutra of ‘cause and effect’
Aristotle

- Material cause
- Formal cause
- Efficient cause
- Final cause

Max Born

- Law of cause and effect
- Antecedence
- Contiguity
Connectivity and causality

- **Structural connectivity**
  - The (possibly directed) network of physical connections in a system.

- **Functional (dynamical) connectivity**
  - Statistical dependencies between system elements (variables) over time (correlation, coherence, etc.).

- **Effective connectivity**
  - The directed influence one variable exerts on another.

- **Causal connectivity**
  - *Directed* functional/dynamical connectivity.
Assessing causality

Stimulation

Ablation

Bressler and Seth (in press). Neuroimage.
Time series inference

Causality based on temporal relations among recorded time series.

Data driven (exploratory)
- All variance in data assigned to causal structure

Model driven (confirmatory)
- Variance tested to see whether it fits pre-specified model of data generation.

Stochastic vs deterministic
Granger causality

Norbert Wiener

Clive Granger
Granger (G-)causality

**Causality based on prediction:** If a signal $X_1$ causes a signal $X_2$, then knowledge of the past of both $X_1$ and $X_2$ should improve the predictability of $X_2$, as compared to knowledge of $X_2$ alone.

$$X_1(t) = \sum_{j=1}^{m} A_{11} X_1(t-j) \ldots + \xi_1(t)$$

$$X_2(t) = \sum_{j=1}^{m} A_{22} X_2(t-j) + \xi_2(t)$$

$$\mathcal{F}_{X_1 \rightarrow X_2} = \ln \frac{\text{var}(\varepsilon_t)}{\text{var}(\varepsilon'_t)}.$$
Assumptions

Data must be covariance stationary.
- Unit root test
- Autocorrelation test

AR model ‘order’ must be determined.
- Bayesian/Akaike information criterion

AR model must account for the data.
- Consistency test
- Whiteness test

Spectral G-causality

- G-causality can be assessed in the frequency domain, as well as in the time domain.

\[
\begin{pmatrix}
A_{11}(f) & A_{12}(f) \\
A_{21}(f) & A_{22}(f)
\end{pmatrix}
\begin{pmatrix}
X_1(f) \\
X_2(f)
\end{pmatrix}
= \begin{pmatrix}
E_1(f) \\
E_2(f)
\end{pmatrix},
\]

\[
A_{lm}(f) = \delta_{lm} - \sum_{j=1}^{p} A_{lm}(j)e^{-i2\pi f j},
\]

\[
\begin{align*}
\delta_{lm} &= 0 \quad (l = m) \\
\delta_{lm} &= 1 \quad (l \neq m).
\end{align*}
\]

\[
I_{j\rightarrow i}(f) = -\ln \left( 1 - \frac{\left( \Sigma_{ji} - \frac{\Sigma_{ij}}{\Sigma_{ii}} \right) H_{ij}(f)^2}{S_{ii}(f)} \right)
\]

Partial G-causality

- Partial G-causality controls (partly) for artifacts due to common inputs and latent variables (analogous to partial coherence).

\[ F_1 = \ln \left( \frac{|R_{XX|Z}|^{(1)}}{|R_{XX|Z}|^{(2)}} \right) = \ln \left( \frac{S_{11} - S_{12}S_{22}^{-1}S_{21}}{\Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}} \right) \]

\[ F_2 = \ln \left( \frac{|S_{11}|}{|\Sigma_{11}|} \right) \]

Transfer entropy (TE)

• The TE from $y$ to $x$ is the degree to which $y$ disambiguates the future of $x$ beyond the degree to which $x$ disambiguates its own future.

• Framed in terms of information theory and \textit{resolution of uncertainty}.

• Naturally nonlinear.

• Challenging to accurately estimate conditional entropies.

• For Gaussian variables, GC and TE are exactly equivalent!


Multivariate G-causality

- Causality between sets of elements (perhaps conditioned on other sets).

- For multivariate predictor, no problem.

- For multivariate predictee, can use either trace or determinant of residual covariance matrix.

\[ \mathcal{F}_{Y \rightarrow X \mid Z} \equiv \ln \left( \frac{|\Sigma(\varepsilon_t)|}{|\Sigma(\varepsilon'_t)|} \right) \]

Ladroue et al. (2009). *PLoS One*
Causal networks

• Represent G-causality interactions as a directed graph

• Allows useful summary statistics:
  • Causal flow
  • Causal density

\[ cd(X) = \frac{1}{n(n-1)} \sum_{i \neq j} F_{X_i \rightarrow X_j | X_{[ij]}} \]


Causal density

- Causal density provides a useful measure of complexity.
- Independent elements will have low causal density, as will elements that behave identically.

Multivariate causal density

- Extend causal density using multivariate G-causality.

\[
mcd(X) = \frac{1}{3^n} \sum_{\{M^k\} \in \mathcal{X}_3} \sum_{\sigma \in \mathcal{S}_3} \frac{1}{|M^{\sigma(1)}| \cdot |M^{\sigma(2)}|} \mathcal{F}_{M^{\sigma(2)} \rightarrow M^{\sigma(1)} | M^{\sigma(3)}}
\]

Barrett & Seth. (in preparation)
Causal networks in neural systems
Darwin X

Darwin X

“Morris” water maze

The device learns the task …

... and develops ‘place cells’
Darwin X

Trial 1

Tri-synaptic

Perforant

% of causal pathways

<table>
<thead>
<tr>
<th></th>
<th>Trial 1</th>
<th>Trial 17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tri-synaptic</td>
<td>42.0</td>
<td>29.0</td>
</tr>
<tr>
<td>Perforant</td>
<td>14.8</td>
<td>21.6</td>
</tr>
</tbody>
</table>

Causal cores

a. Pick a ‘Neural Reference’ (NR)

c. G-causality analysis for each synapse.

d. Identify causal core.

e. The causal core ‘in vivo’.

b. Identify the ‘context network’

Darwin X: Causal cores

Causal cores

Multilevel dynamics of learning

How to get from this …

… to this

Darwin X: Causal cores

Lymnaea stagnalis

Passaro et al. (in preparation)
G-autonomy

Seth, A.K. (2009), *Artif. Life*
G-autonomy

• Autonomy as **self-determination**: 

• An autonomous system should not be fully determined by its environment

• A random system should not be autonomous

• A variable is **G-autonomous** if and only if:
  – it is dependent on its own past history, and
  – these dependencies are not accounted for by external factors.

Bertschinger et al. (2007), *Biosys.*
Seth, A.K. (2009), *Artif. Life*
**G-autonomy**

- A process $X_1$ is **G-autonomous** if knowledge of its own past helps predict its future over and above predictions based on the past of a set of (external) processes $X_2 \ldots X_N$

\[
X_1(t) = \sum_{j=1}^{m} A_{12} X_2(t-j) + \epsilon_1(t) \\
X_2(t) = \sum_{j=1}^{m} A_{21} X_1(t-j) + \epsilon_2(t)
\]

\[
\text{gap}_{X_1|X_2} = \log \frac{\text{var}(\xi_{1R(11)})}{\text{var}(\xi_{1U})}
\]

Seth, A.K. (2009), *Artif. Life*
• X1 is non-autonomous because it is white-noise.
• X2 is non-autonomous because it is white noise is influenced by more white noise (X1)
• X3 is autonomous because it has an AR component that is not accounted for by other variables.
• X4 is non-autonomous because, although it has an AR component, this can be accounted for by X3
• X5 is autonomous for the same reason as X3
G-autonomy: example

<table>
<thead>
<tr>
<th></th>
<th>$X_1$</th>
<th>$X_2$</th>
<th>$X_3$</th>
<th>$X_4$</th>
<th>$X_5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X_1$</td>
<td>0</td>
<td>0.03</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
</tr>
<tr>
<td>$X_2$</td>
<td>8.4*</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_3$</td>
<td>0.6</td>
<td>0.4</td>
<td>7.9*</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>$X_4$</td>
<td>0</td>
<td>0</td>
<td>7.7*</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$X_5$</td>
<td>7.3*</td>
<td>0.7</td>
<td>0.3</td>
<td>0</td>
<td>7.0*</td>
</tr>
</tbody>
</table>

- $X_3$ is G-autonomous, G-causes $X_4$ but is not G-caused by anything.
- $X_5$ is G-autonomous, is G-caused by $X_1$ but does not G-cause anything.
G-emergence

“An emergent property is somehow more than the sum of its component parts”
Emergence

- **Nominal emergence**: A property that can be possessed by macro-level objects, but not by their micro-level constituents.

- **Strong emergence**: A macro-level property that is in principle not identifiable from micro-level observations.

- **Weak emergence**: A macro-level property that is derived from the interaction of micro-level components, but in complicated ways such that the macro-level property has no simple micro-level explanation.

Bedau (1997), *Philosophical Perspectives*
Weak emergence

• Weakly emergent properties are *ontologically dependent* on micro-level causal factors, but are *epistemologically irreducible* to these factors (Bedau 1997)

• **Binary weak emergence (Bedau):** Underderivability except by simulation.

• **Continuous weak emergence (Seth):** A macro-property is weakly emergent *to the extent* that it is not identifiable from micro-level observations.

• A weakly emergent property is simultaneously (i) *autonomous from* and (ii) *dependent upon* its underlying causal factors.

Bedau (1997), *Philosophical Perspectives*  
Seth, A.K. (2009), *Artif. Life*
G-emergence

- A macro-variable $M$ is **G-emergent** from a set of micro-variables $m$ to the extent that:
  - $M$ is **G-autonomous** from $m$, and
  - $M$ is **G-caused** by $m$

\[
g_{eM|m} = g_{aM|m} \left( \frac{1}{N} \sum_{i=1}^{N} F_{m_i \rightarrow M} \right)
\]

- G-emergence will be zero either if $M$ is independent of $m$ or if $M$ is fully predicted by $m$.

Seth, A.K. (2009), *Artif. Life*
Example: boids

- Bird flocking is a canonical instance of emergence.
- Three simple rules can simulate flocking behavior: aggregation, avoidance, and velocity matching.

Reynolds (1987), Comp. Graphics
Example: boids

- Objective is to measure the G-emergence of the centre-of-mass of a 10-boid flock with respect to the individuals.

\( R \) (random) \hspace{1cm} \( L \) (low emergence) \hspace{1cm} \( H \) (high emergence)

Seth, A.K. (2009), *Artif. Life*
Results: boids

Seth, A.K. (2009), Artif. Life
Downward causality

- Does emergence involve ‘downward causation’ from macro-properties to micro-properties?

- G-causality offers a metaphysically innocent means of characterizing downward causality.

Seth, A.K. (2009), *Artif. Life*
Consciousness
“Consciousness is the appearance of a world.”

Metzinger (2009)
Conscious *content* vs conscious *level*
not differentiated  not integrated

differentiated and integrated

Measuring complexity

• Neural complexity (Tononi, Sporns, Edelman, 1994)
• Information integration ($\Phi$) (Tononi, 2004, 2008)
• Causal density (Seth, 2005, 2008)

\[
C_N(X) = \sum_k \langle MI(X^k_j; X - X^k_j)\rangle,
\]

\[
\phi(x_1) = H \left[ p(X_0 \rightarrow x_1) \bigg| \prod_{M^k \in P^{MIP}} p(M^k \rightarrow \mu^k_1) \right]
\]

\[
\text{cd}(X) \equiv \frac{1}{n(n-1)} \sum_{i \neq j} \mathcal{F}_{X_i \rightarrow X_j | X_{[i,j]}},
\]
Causal density and conscious level

significantly different channels


Transcranial magnetic stimulation (TMS) applied during sleep and wakefulness.

Massimini et al. (2005)
Impaired consciousness during temporal lobe seizures is related to increased long-distance cortical–subcortical synchronization

Marie Arthuis,1 Luc Valton,2 Jean Régis,1,3 Patrick Chauvel,1,2,3 Fabrice Wendling,4,5 Lionel Naccache,6 Christophe Bernard1,3 and Fabrice Bartolomei1,2,3

Short Communication

Intracranial EEG power spectra and phase synchrony during consciousness and unconsciousness

Susan Pockett a,*, Mark D. Holmes b

a Department of Physics, University of Auckland, Private Bag 92019, Auckland, New Zealand
b Department of Neurology, University of Washington, Seattle, Washington, USA
Strong emergence and consciousness

• **Chalmers**: ‘There is exactly one clear case of strong emergence, and that is consciousness’.
  - Even a complete neuroscience will not provide an understanding of the redness of red
  - Conscious experiences must have causal efficacy (downward causality).

• **Hypothesis**: *experiences of volition correlate with high G-emergence of corresponding global brain states.*

Seth, A.K. (2009), *Cognit. Computation*
Last words

• The ability to measure a phenomenon is fundamental to its effective scientific description.

• **G-causality**, **G-autonomy**, and **G-emergence** reflect new ways of adapting time series analysis techniques to core problems of measurement in 21st century cognitive science and biology.

• www.anilseth.com (MATLAB toolbox)
Last words