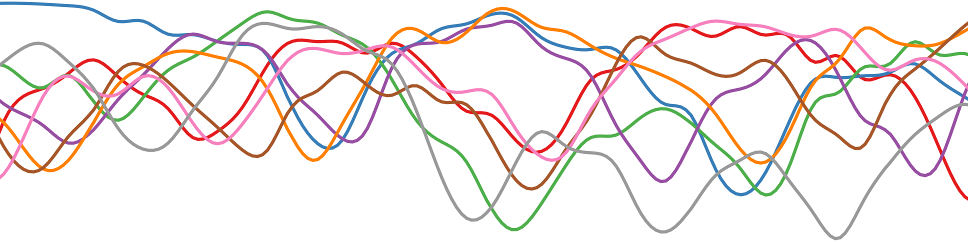
Reconciling Emergences

An information-theoretic approach to identify causal emergence in multivariate data



Pedro A.M. Mediano

plogp@pm.me

UNIVERSITY OF CAMBRIDGE

Imperial College London

ACKNOWLEDGEMENTS

Fernando Rosas

Daniel Bor

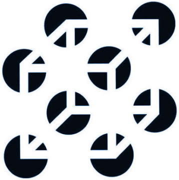
Adam Barrett

Thanks to:

- Anil Seth
- Henrik Jensen
- Robin Carhart-Harris

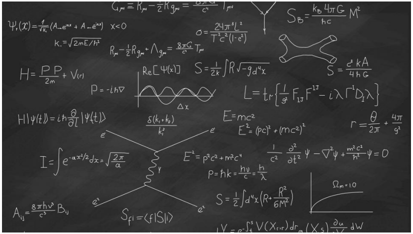
EMERGENCE: WHAT IT IS, AND WHY IT MATTERS

Informally: “the whole is more than the sum of its parts.”

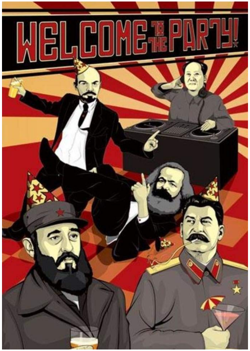


THEORIES OF EMERGENCE

1. **Reductionism**: There's only physics. Laplace's demon rocks.
2. **Emergentism**: Some things can't be explained by "microstates."
   - 2A *Strong*: true emergence possible in principle.
   - 2B *Weak*: true emergence only apparent in practice.



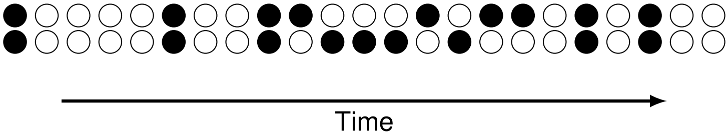
$\not\rightarrow$



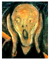
CAUSAL EMERGENCE

MINIMAL EXAMPLE

- Example system with two binary variables:
  - With probability γ, the future has the same parity as the past.
  - Otherwise, they have the opposite parity.



- The dynamics are not visible in *any subset* of the system.

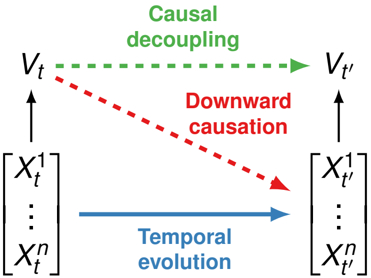


Many measures can't pick this up: $\mathrm{TE} = \mathrm{MI} = \Phi = 0$.

This is a minimal example of **causal emergence**.

CAUSAL EMERGENCE

OUTLINE



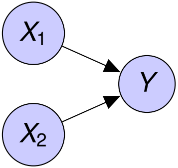
1. Provide a formal definition of causal emergence.
2. Distinguish between two different *kinds* of emergence.
3. Propose a practical criterion and show it in action.

INFORMATION DECOMPOSITION

PARTIAL INFORMATION DECOMPOSITION

Two predictors X_1, X_2 and one target Y.

- Joint predictability: $I(X_1 X_2; Y)$
- Marginal predictability: $I(X_1; Y),\ I(X_2; Y)$



However, sometimes:

$$\underbrace{I(X_1 X_2; Y)}_{\text{“the whole“}} > \underbrace{I(X_1; Y) + I(X_2; Y)}_{\text{“the parts“}}$$

The *Partial Information Decomposition* (PID) postulate:

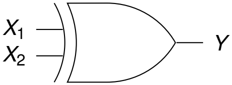
$$I(X_1 X_2; Y) = \underbrace{I_\partial^{\{1\}\{2\}}}_{\text{redundancy}} + \underbrace{I_\partial^{\{1\}} + I_\partial^{\{2\}}}_{\text{unique info}} + \underbrace{I_\partial^{\{12\}}}_{\text{synergy}}$$

(Williams & Beer, 2010)

EXAMPLE: XOR LOGIC GATE

Perfect example of synergy: XOR.

X_1	X_2	Y
0	0	0
0	1	1
1	0	1
1	1	0



- Knowing one input tells you nothing:

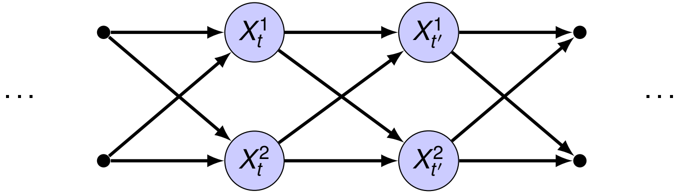
$$I(X_1; Y) = I(X_2; Y) = 0$$

- Knowing both inputs tells you everything:

$$I(X_1 X_2; Y) = 1$$

INFORMATION AND DYNAMICAL SYSTEMS

- PID decomposes information many sources have about one target.
- **BUT** we care about multivariate systems evolving *jointly* over time.




Problem: PID cannot deal with multiple targets!

INFORMATION DECOMPOSITION

Can we extend PID to multivariate time series?

Yes! With *Integrated Infomation Decomposition*, ΦID.

$$I(\boldsymbol{X}_t; \boldsymbol{X}_{t'}) = \sum_{\boldsymbol{\alpha},\boldsymbol{\beta}\in\mathcal{A}} I_\partial^{\boldsymbol{\alpha}\to\boldsymbol{\beta}}$$

Beyond integrated information: A taxonomy of information dynamics phenomena

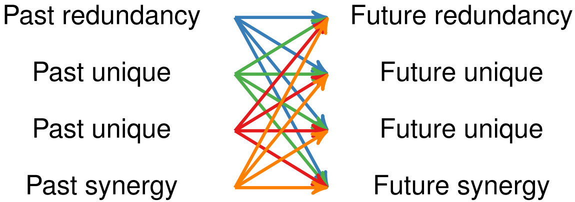
Pedro A.M. Mediano,[1] Fernando Rosas,[2,3,4] Robin L. Carhart-Harris,[2] Anil K. Seth,[5] and Adam B. Barrett[5,6]
[1] *Department of Computing, Imperial College London, London SW7 2RH*[*]
[2] *Center for Psychedelic Research, Department of Medicine, Imperial College London, London SW7 2DD*
[3] *Data Science Institute, Imperial College London, London SW7 2AZ*
[4] *Center for Complexity Science and Department of Mathematics, Imperial College London, London SW7 2AZ*[*]
[5] *Sackler Center for Consciousness Science, Department of Informatics, University of Sussex, Brighton BN1 9RH*
[6] *The Data Intensive Science Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*

INTEGRATED INFORMATION DECOMPOSITION

- In PID there are 4 terms: redundancy, unique (2x), and synergy:

$$\underbrace{I_\partial^{\{1\}\{2\}}}_{\text{redundancy}},\ \underbrace{I_\partial^{\{1\}},\ I_\partial^{\{2\}}}_{\text{unique info}},\ \underbrace{I_\partial^{\{12\}}}_{\text{synergy}}$$

- In ΦID, we have all combinations of past and future PID:



- In total, $4 \times 4 = 16$ terms.

Integrated information decomposition

Examples:

- $I_\partial^{\{1\}\{2\}\to\{1\}\{2\}}$: redundant stored information
- $I_\partial^{\{1\}\to\{2\}}$: unique transferred information
- $I_\partial^{\{1\}\to\{1\}\{2\}}$: “duplicated” information
- ...

DEFINING EMERGENCE

PID NOTATION

- We need to define a *coarse-grained* PID:
  - $\mathtt{Un}(X \to Y|\boldsymbol{Z})$: unique information that X has about Y that no individual variable Z^i has.
  - $\mathtt{Syn}(\boldsymbol{X} \to Y)$: information about Y that no individual X^i has (but $\boldsymbol{X}$ as a whole does).

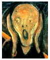
DEFINING EMERGENCE

- Setting:
  - System with n components $\boldsymbol{X}_t = (X_t^1, X_t^2, \ldots, X_t^n)$
  - Candidate emergent feature $V_t = F(\boldsymbol{X}_t)$
- Informal definition: A feature V_t that says something about the future that individual micro elements don't.
- Formal definition:

Definition (causal emergence):

A supervenient feature $V_t = F(X_t)$ exhibits causal emergence if $\mathtt{Un}(V_t \to \boldsymbol{X}_{t'} | \boldsymbol{X}_t) > 0$.

DEFINING EMERGENCE



To compute $\mathtt{Un}(V_t \to \boldsymbol{X}_{t'} | \boldsymbol{X}_t)$ we need to know V_t in advance.

Solution: use more PID!

Result:

A system has causally emergent features if and only if $\mathtt{Syn}(\boldsymbol{X}_t \to \boldsymbol{X}_{t'}) > 0$.

- Synergy quantifies the *emergence capacity* of a system.

A TAXONOMY OF EMERGENCE

- Previous definition tells us *whether* there is emergence, but not *what kind* of emergence it is.
- We introduce two types of emergence:
  - **Downward causation**: macroscopic features affect individual elements.
  - **Causal decoupling**: macroscopic features affect other macroscopic features.

A TAXONOMY OF EMERGENCE

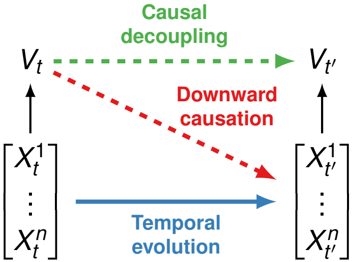
- Formal definitions:

Downward causation:

$$\texttt{Un}(V_t \rightarrow X^i_{t'} | \boldsymbol{X}_t) > 0$$

Causal decoupling:

$$\texttt{Un}(V_t \rightarrow V_{t'} | \boldsymbol{X}_t, \boldsymbol{X}_{t'}) > 0$$



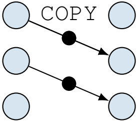
A TAXONOMY OF EMERGENCE

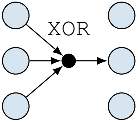
We can decompose causal emergence using ΦID:

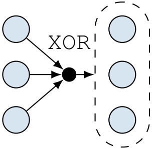
$$\underbrace{\texttt{Syn}(\boldsymbol{X}_t \rightarrow \boldsymbol{X}_{t'})}_{\text{total emergence capacity}} = \underbrace{\mathcal{G}(\boldsymbol{X}_t)}_{\text{causal decoupling}} + \underbrace{\mathcal{D}(\boldsymbol{X}_t)}_{\text{downward causation}}$$

SIMPLE EXAMPLES

- Example: as feature, take V_t as the parity of $\boldsymbol{X}_t$.







$\mathtt{Un}(V_t \to \boldsymbol{X}_{t'} \mid \boldsymbol{X}_t) = 0$	$\mathtt{Un}(V_t \to \boldsymbol{X}_{t'} \mid \boldsymbol{X}_t) = 1$	$\mathtt{Un}(V_t \to \boldsymbol{X}_{t'} \mid \boldsymbol{X}_t) = 1$
$\mathcal{D}(\boldsymbol{X}_t) = \mathcal{G}(\boldsymbol{X}_t) = 0$	$\mathcal{D}(\boldsymbol{X}_t) = 1$	$\mathcal{G}(\boldsymbol{X}_t) = 1$
✗ Not emergent	✓ Emergent	✓ Emergent

PRACTICAL TOOLS

A feature V_t is causally emergent **if** $\Psi > 0$.

$$\Psi_{t,t'}^{(k)}(V) := I(V_t; V_{t'}) - \sum_{j=1}^{n} I(X_t^j; V_{t'})$$

Pros:

- ✓ Uses only standard mutual information.
- ✓ Uses only pairwise marginals (no curse of dimensionality).
- ✓ No false positives.

Cons:

- ✗ Needs a candidate feature V_t.
- ✗ Double-counts redundancy (which reduces sensitivity).
- ✗ Inconclusive if $\Psi \leq 0$.

INTERIM SUMMARY

CAUSAL EMERGENCE

- So far, we have:

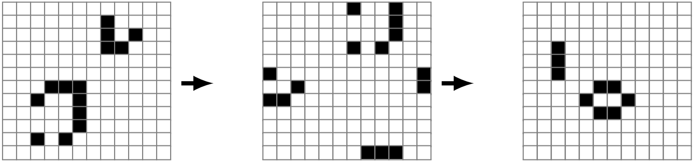
1. Formulated a rigorous definition of causal emergence.
2. Provided an intrinsic criterion of CE based on synergy.
3. Decomposed emergence into $\mathcal{D}$ and $\mathcal{G}$.
4. Provided practical tools to test for emergence in data.

CASE STUDIES

CAUSAL EMERGENCE

RESULTS

Canonical example of emergence: the **Game of Life.**



- **Micro variable**: cell states, $\boldsymbol{X}_t \in \{0, 1\}^n$.
- **Macro variable**: particle type, $V_t \in \{\texttt{blinker}, \texttt{glider}, \dots\}$.

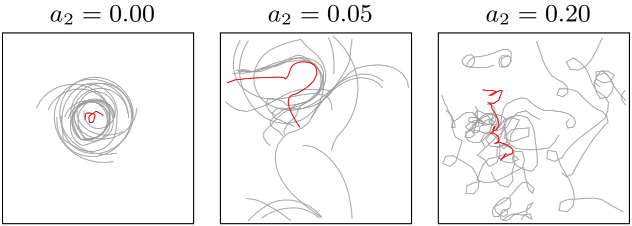
$\longrightarrow$ $\Psi_{t,t'}(V) = 0.58$ bit

CAUSAL EMERGENCE

RESULTS

Example of emergence: **flocking behaviour.**

- **Micro variable**: bird position, $\boldsymbol{X}_t \in \mathbb{R}^{2n}$.
- **Macro variable**: center of flock, $V_t \in \mathbb{R}^2$.

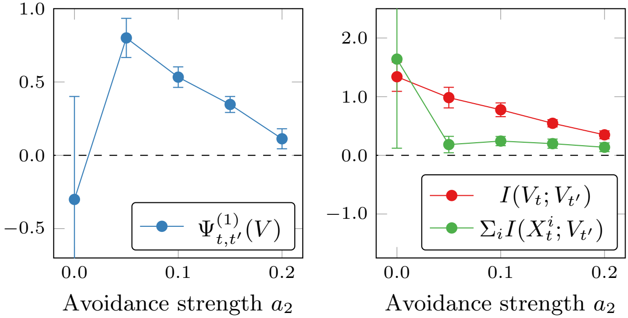


CAUSAL EMERGENCE

RESULTS

Example of emergence: **flocking behaviour.**

- **Micro variable**: bird position, $\boldsymbol{X}_t \in \mathbb{R}^{2n}$.
- **Macro variable**: center of flock, $V_t \in \mathbb{R}^2$.

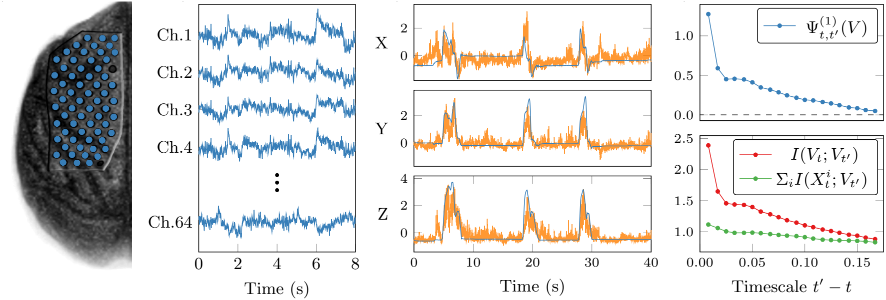


CAUSAL EMERGENCE

RESULTS

Example of emergence: **neural activity** during motor control.

- **Micro variable**: ECoG channels, $\boldsymbol{X}_t \in \mathbb{R}^{64}$.
- **Macro variable**: decoded hand motion, $\boldsymbol{V}_t \in \mathbb{R}^3$.



WRAP-UP

WRAP-UP

- ✓ We proposed a quantitative, rigorous theory of causal emergence.
- ✓ Our theory agrees with intuition in paradigmatic examples of emergence (e.g. Game of Life).
- ✓ New family of information metrics to analyse neural (or other) data.
  → www.github.com/pmediano/ReconcilingEmergences

plogp@pm.me

Thank you for listening!