

Title: Being vs. Happening: information from the intrinsic perspective of the system itself

Date: Apr 12, 2018 03:30 PM

URL: <http://pirsa.org/18040120>

Abstract: When applied to a physical system, the two main, established notions of information, Shannon Information and Algorithmic Information, explicitly neglect the mechanistic structure of the system under evaluation. Shannon information treats the system as a channel and quantifies correlations between the system's inputs and outputs, or between its past and future states. Algorithmic information quantifies the length of the shortest program capable of reproducing the system's outputs or dynamics. The goal in both cases is to predict the system's behavior from the perspective of an extrinsic investigator. From the intrinsic perspective of the system itself, however, information must be physically instantiated to be causally relevant. For every bit, there must be some mechanism that is in one of two (or several) possible states, and which state it is in must matter to other mechanisms. In other words, the state must be a difference that makes a difference and implementation matters. By examining the informational and causal properties of artificial organisms (animats) controlled by small, adaptive neural networks (Markov Brains), I will discuss necessary requirements for intrinsic information, autonomy, and agency.



Being vs. Happening:

Information from the intrinsic perspective of the system itself

04/12/18 – Larissa Albantakis

Motivation: Consciousness

To be conscious is to have an experience



Integrated information theory (IIT) - *Giulio Tononi*

*Consciousness is what goes away
when one falls into dreamless sleep*



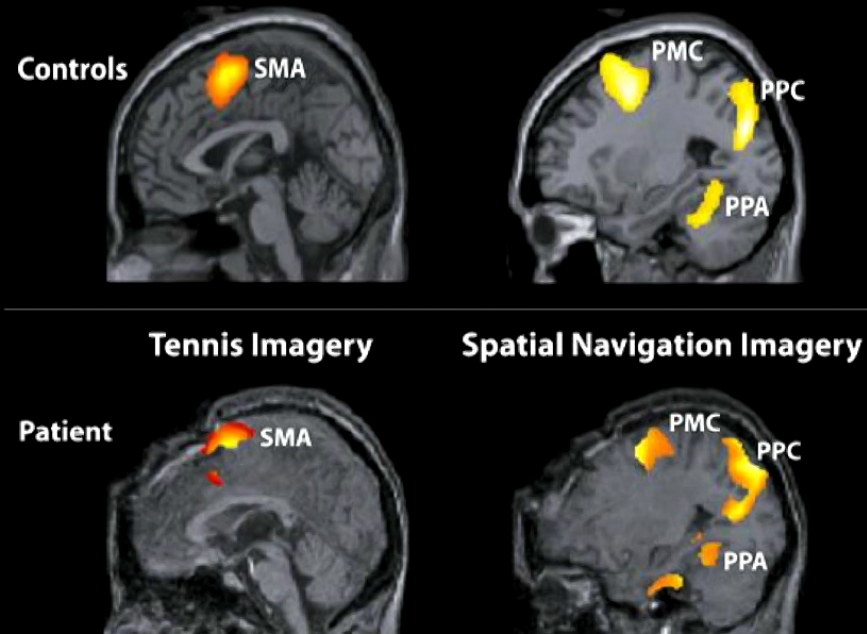
Seurat: Le dormeur

Consciousness is not just self-reflection



Rodin: Le penseur

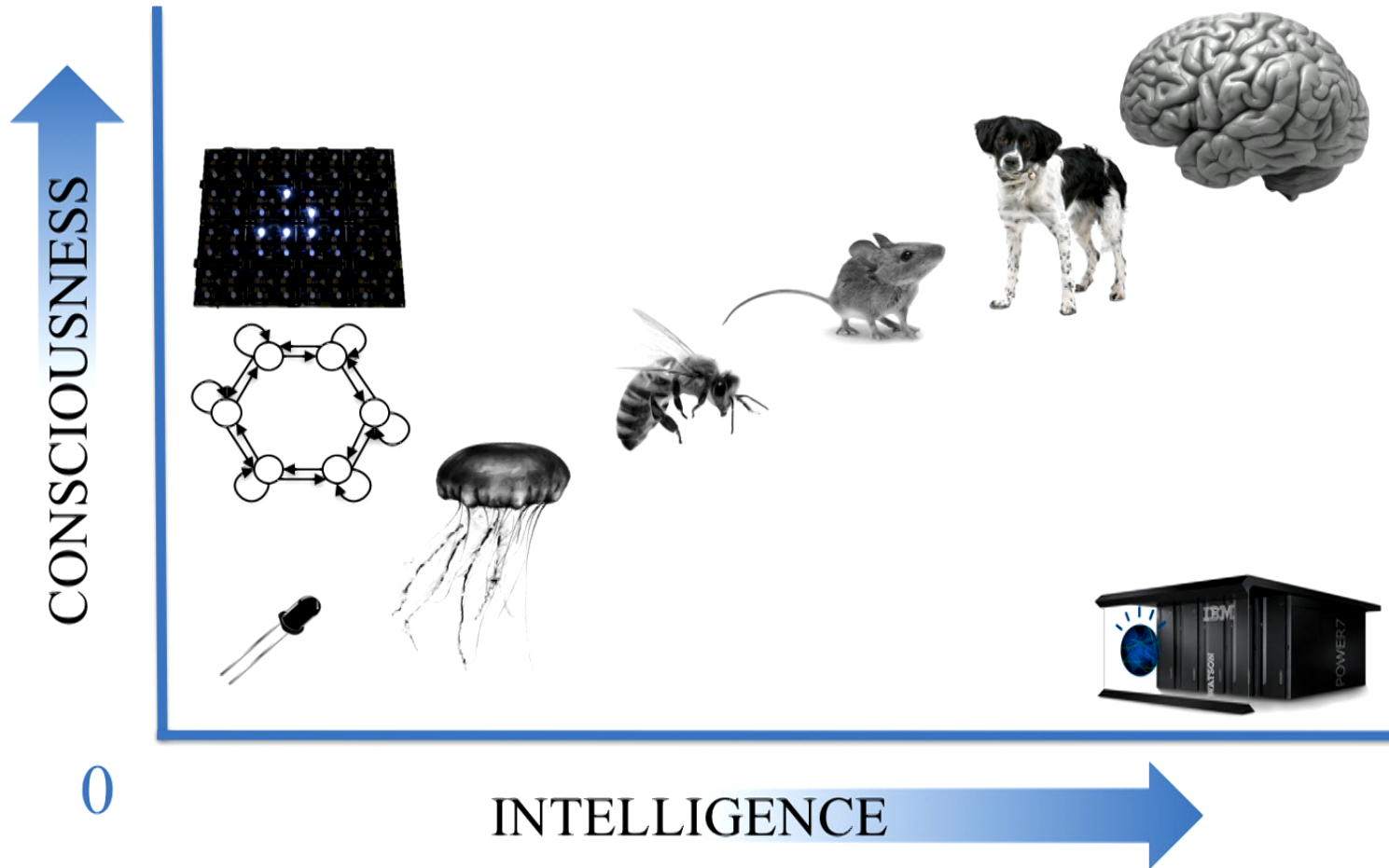
Consciousness ≠ Responsiveness



Imagine playing tennis in the 'vegetative state'

Owen et al, Science 2006; Boly et al. Neuroimage 2007, Boly, Current Opinion Neurology 2011, Sanders et al., Anesthesiology 2014

Intelligence vs. Consciousness



Consciousness is not just of the environment

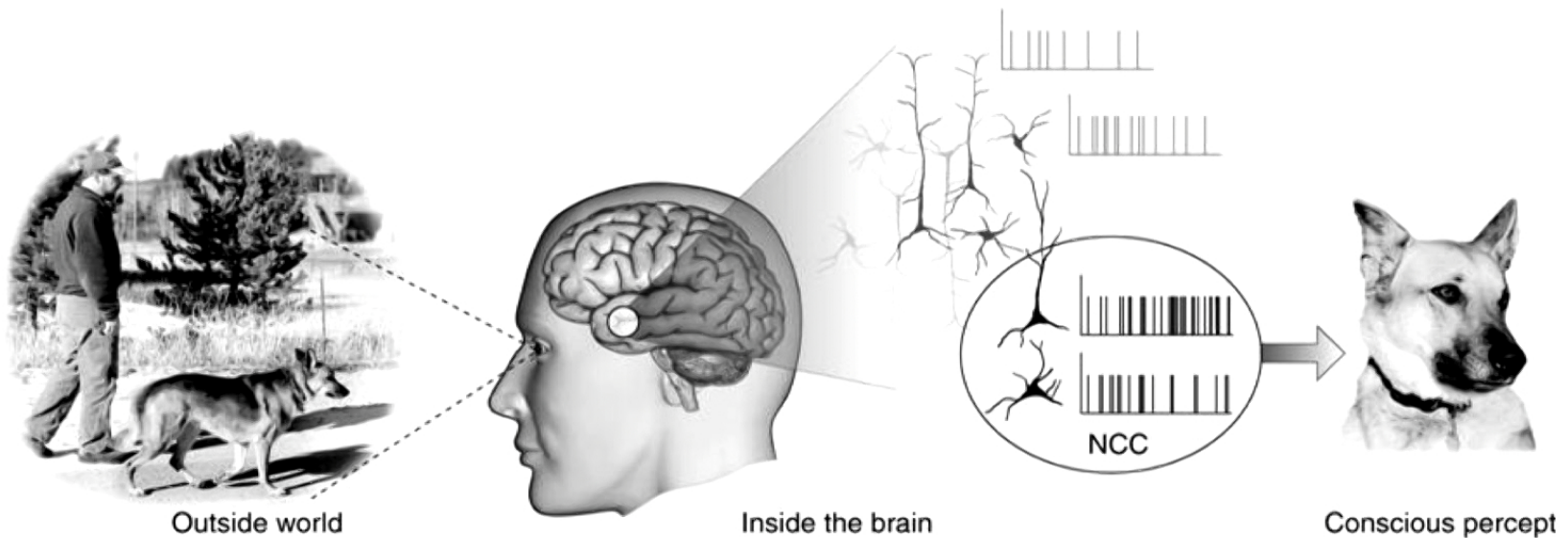


Magritte: Le dormeur temeraire

Perception: “Information processing”?

Extrinsic Information
(deviations from independence)

Intrinsic Information
(differences that make a difference)

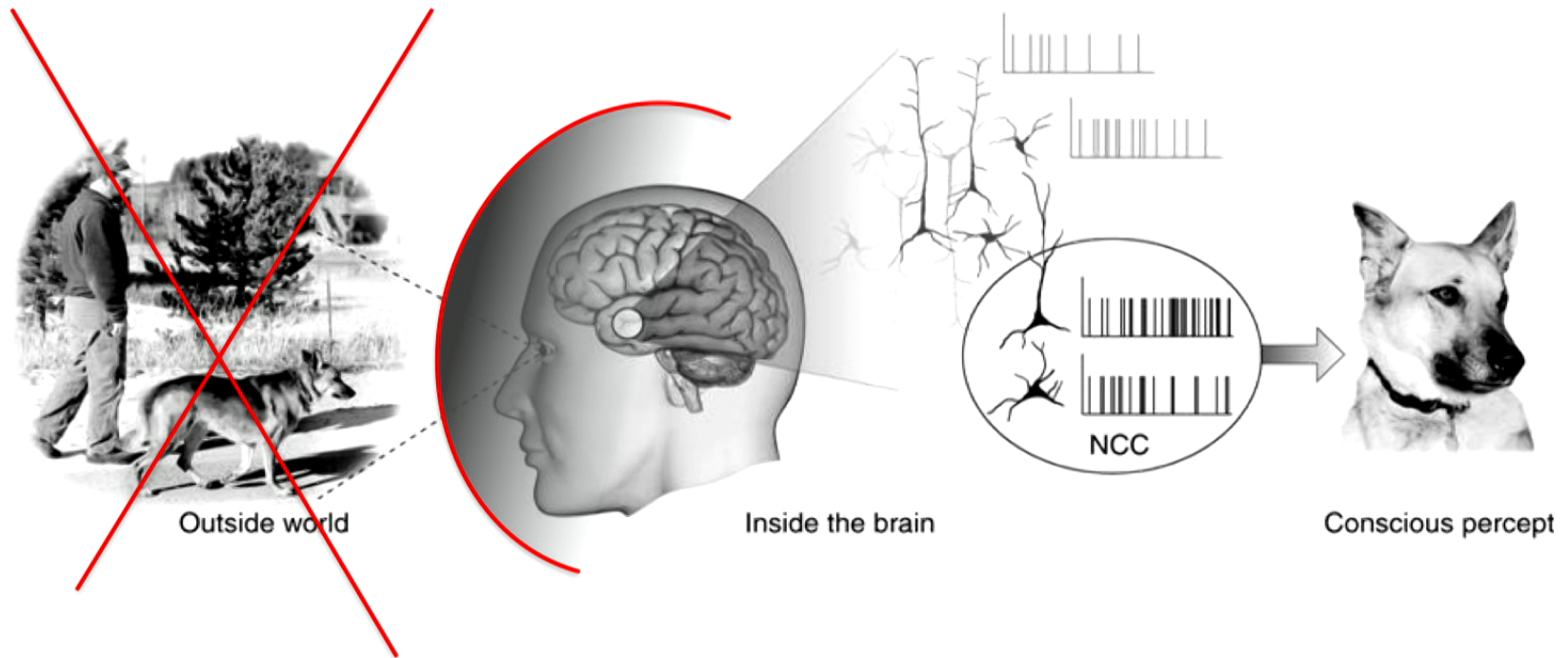


Koch C (2004) The Quest for Consciousness: A Neurobiological Approach. Roberts, Denver, CO.

Experience is intrinsic

Extrinsic Information
(deviations from independence)

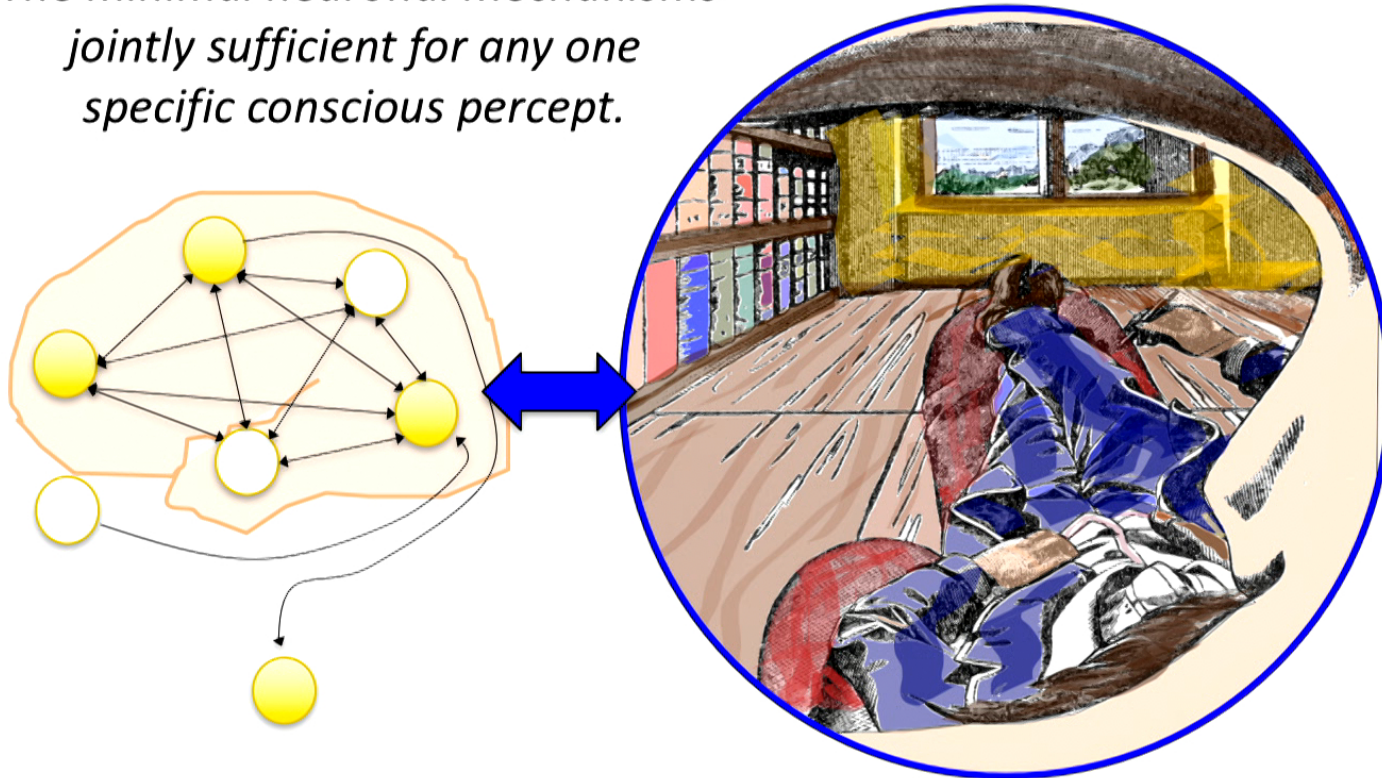
Intrinsic Information
(differences that make a difference)



Koch C (2004) The Quest for Consciousness: A Neurobiological Approach. Roberts, Denver, CO.

Neural correlates of consciousness (NCC)

The minimal neuronal mechanisms jointly sufficient for any one specific conscious percept.

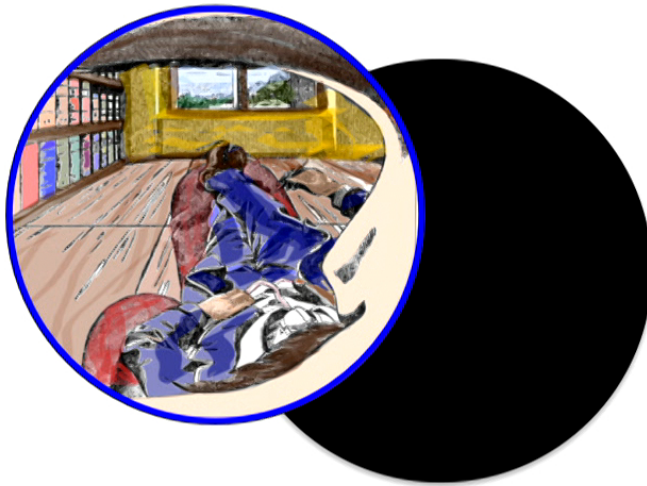


Crick and Koch, Nat Neurosci 2003

Level and content of consciousness

Level of consciousness:

- irrespective of content
- Wake vs dreamless sleep, anesthesia, or coma (brain damaged patients)

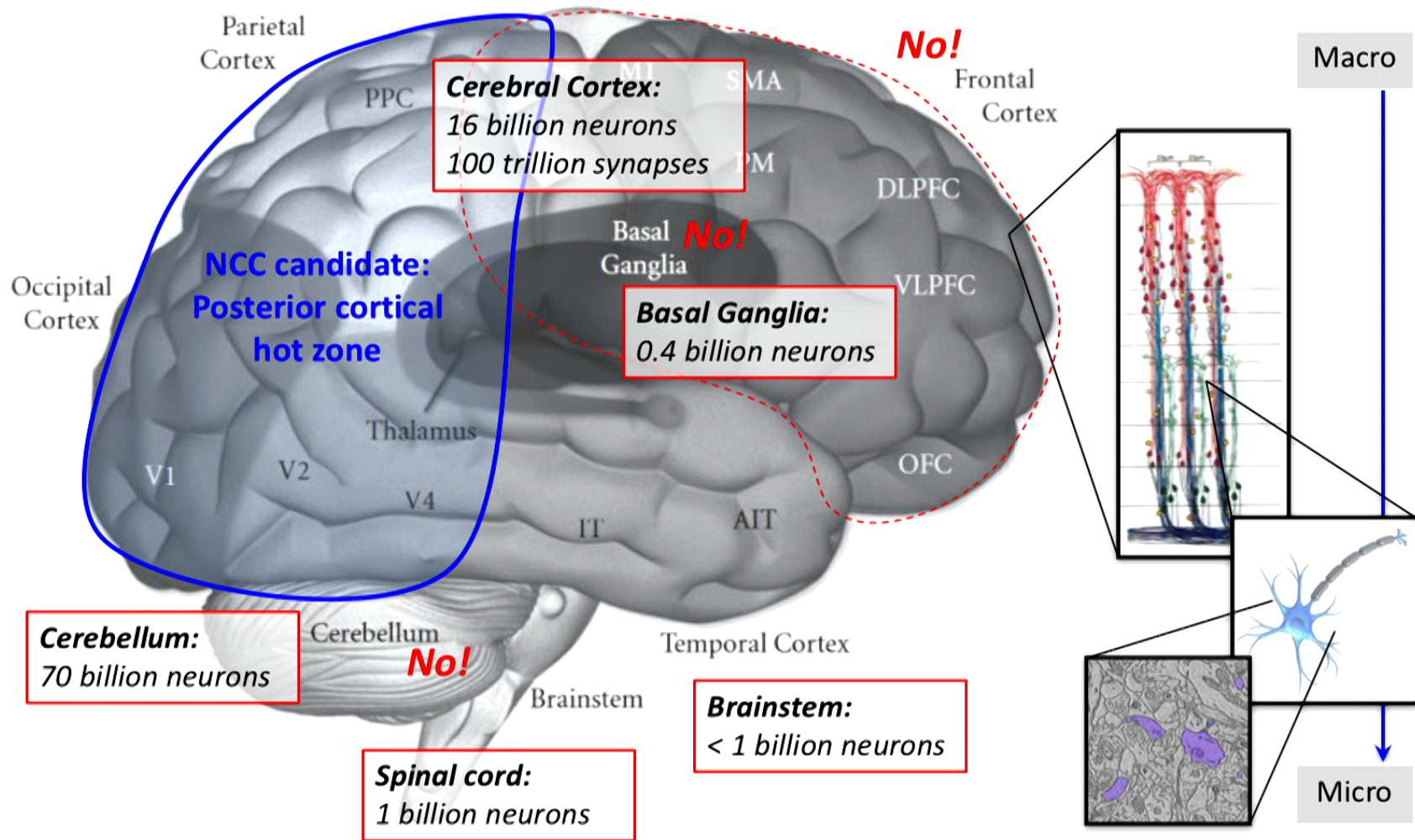


Content of consciousness:

- experience is structured, it has many aspects
- Subjects report perceiving stimulus or not (Masking, binocular rivalry, inattentional blindness)



The neural substrate of consciousness?



From BCC to NCC, and back ... to the hard problem

500 ms fixation, 500 ms mask, 50 ms prime, 33 ms mask, 700 ms target

Kouider and Dehaene, 2007

I don't see it

I see it

BCC

invisible

visible

Visibility rating

P3 size

Dehaene and Changeux, 2011

P3b : 576 ms

NCC

Chalmers 1995

?

*One cannot start from matter
and “squeeze” consciousness out of it*

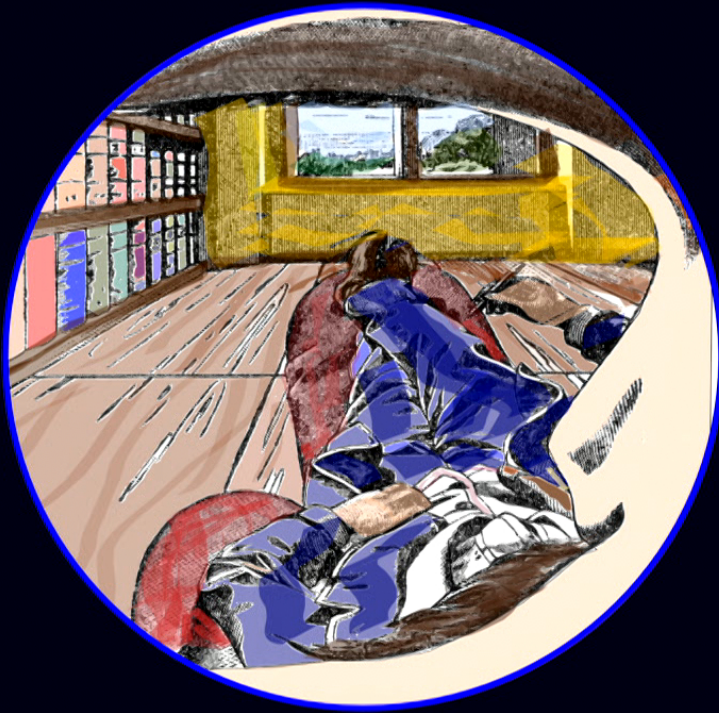


PHYSICS



PHENOMENOLOGY

*But one can start from consciousness itself
and ask what physical system could account for its
properties*



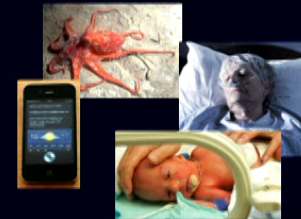
PHENOMENOLOGY



PHYSICS

Integrated Information Theory (IIT)

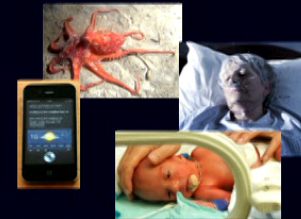
- Starts from *phenomenology*, not from behavioral or neural correlates
- Identifies the *essential* properties of every experience (**axioms**)
- Derives the requirements that physical systems must satisfy to account for them (**postulates**)
- Has *explanatory*, *predictive* and *inferential* power
- Leads to *measures* of both quantity and quality of consciousness
- Proposes a *formal framework* to characterize the intrinsic cause-effect structure of discrete dynamical systems



Oizumi*, Albantakis* and Tononi, 2014; Tononi, Scholarpedia 2015; Tononi et al. 2016 in Nat Rev Neurosci

Integrated Information Theory (IIT)

- Starts from *phenomenology*, not from behavioral or neural correlates
- Identifies the *essential* properties of every experience (**axioms**)
- Derives the requirements that physical systems must satisfy to account for them (**postulates**)
- Has *explanatory*, *predictive* and *inferential* power
- Leads to *measures* of both quantity and quality of consciousness
- Proposes a *formal framework* to characterize the intrinsic cause-effect structure of discrete dynamical systems



Oizumi*, Albantakis* and Tononi, 2014; Tononi, Scholarpedia 2015; Tononi et al. 2016 in Nat Rev Neurosci

Integrated Information Theory (IIT)

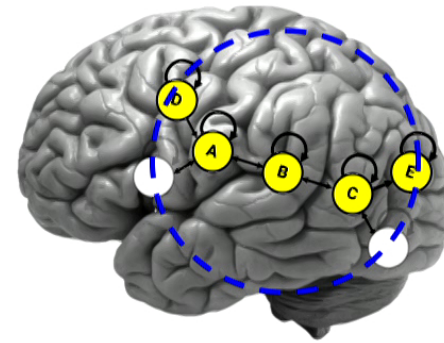
*Phenomenal
existence*



Experience



*Physical
existence*



Cause-effect power



Essential properties of phenomenal existence

(immediately and indubitably true of every experience)

Intrinsicity



Composition



Information



Integration

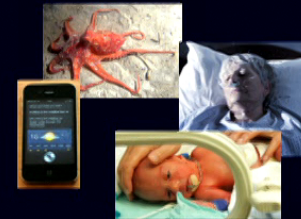


Exclusion



Integrated Information Theory (IIT)

- Starts from *phenomenology*, not from behavioral or neural correlates
- Identifies the *essential* properties of every experience (**axioms**)
- Derives the requirements that physical systems must satisfy to account for them (**postulates**)
- Has *explanatory*, *predictive* and *inferential* power
- Leads to *measures* of both quantity and quality of consciousness
- Proposes a *formal framework* to characterize the intrinsic cause-effect structure of discrete dynamical systems

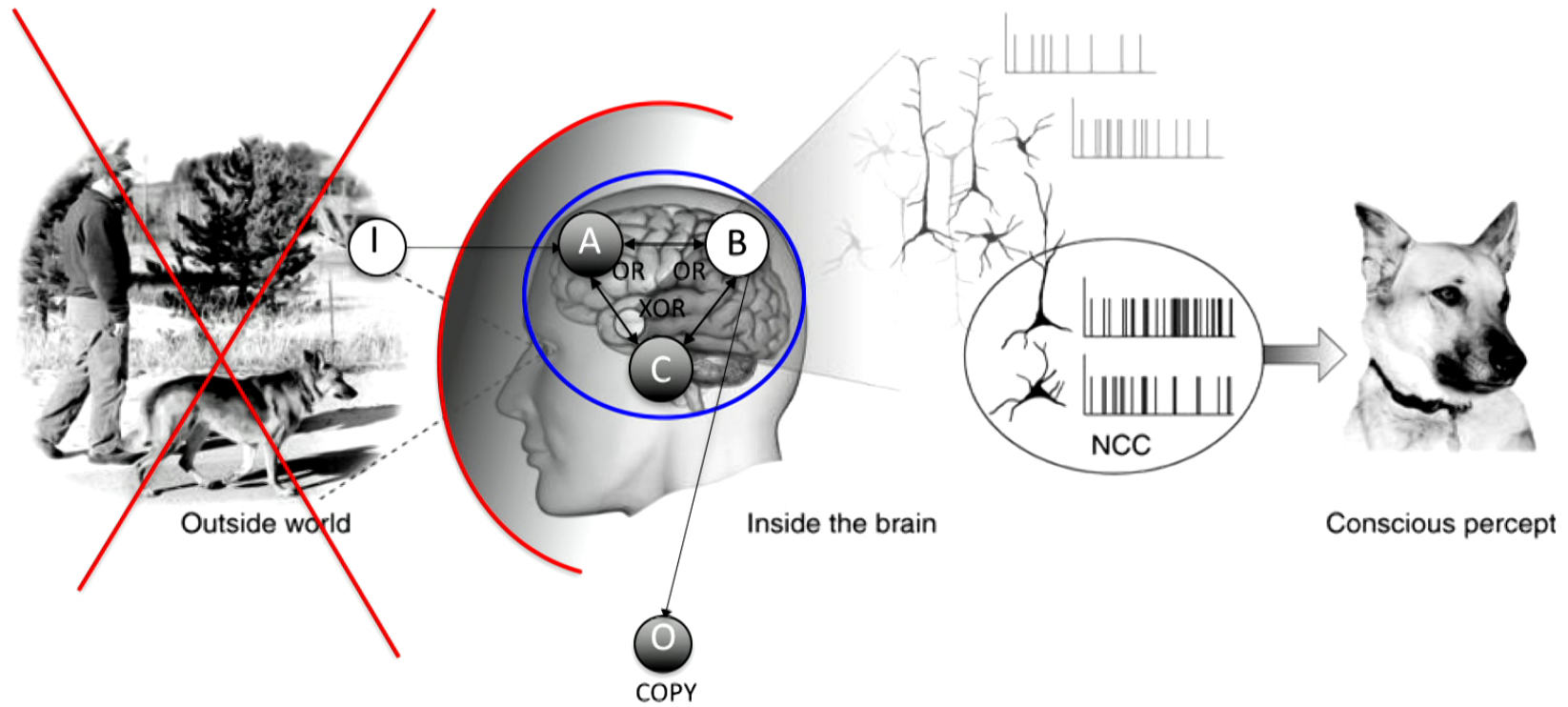


Oizumi*, Albantakis* and Tononi, 2014; Tononi, Scholarpedia 2015; Tononi et al. 2016 in Nat Rev Neurosci

A system of mechanisms in a state

Extrinsic Information
(deviations from independence)

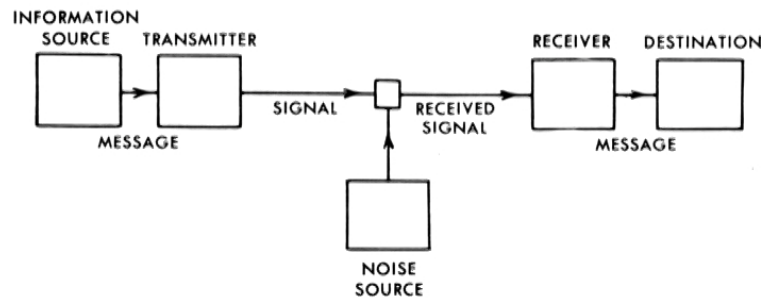
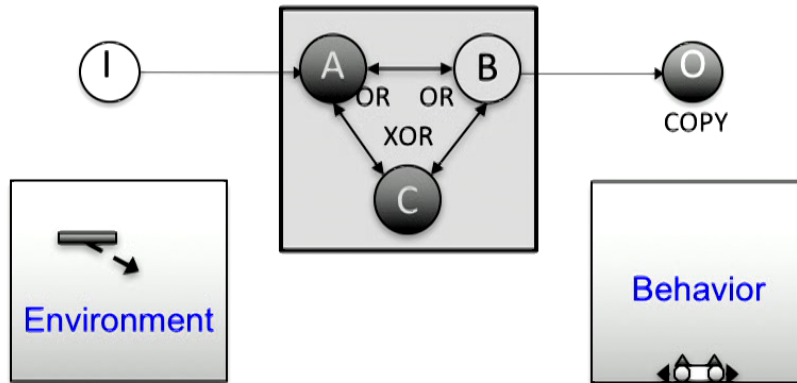
Intrinsic Information
(differences that make a difference)



Koch C (2004) The Quest for Consciousness: A Neurobiological Approach. Roberts, Denver, CO.

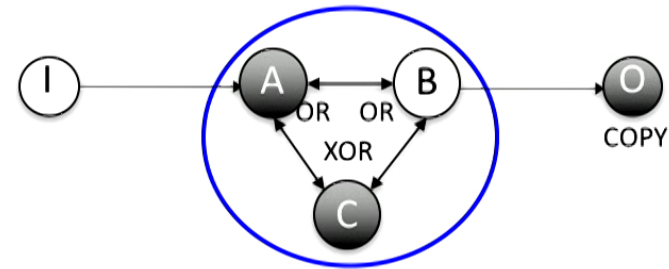
Information: Extrinsic perspective and intrinsic perspective

Extrinsic: signal transmission across a channel (I/O correlations) from the extrinsic perspective of an **observer**

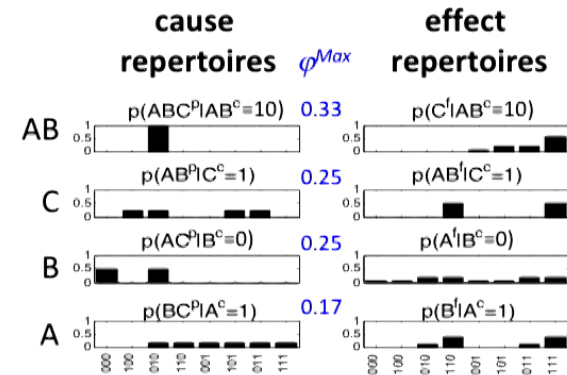


Correlational measures, such as mutual information between **I** and **O**

Intrinsic: differences that make a difference (causes/effects) from the intrinsic perspective of a system

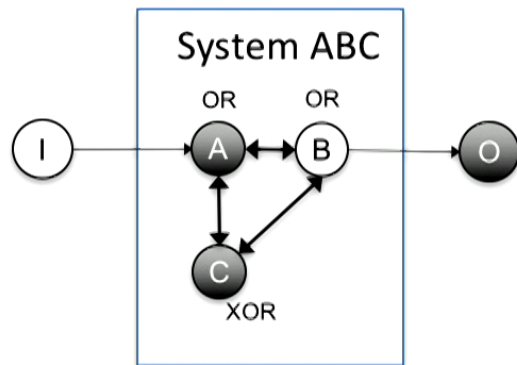


set of concepts: Cause-effect structure (**C**)



within the system **ABC** state-dependent

Formalism to assess intrinsic cause-effect structure



Transition Probability Matrix (TPM)

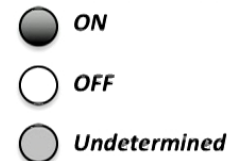
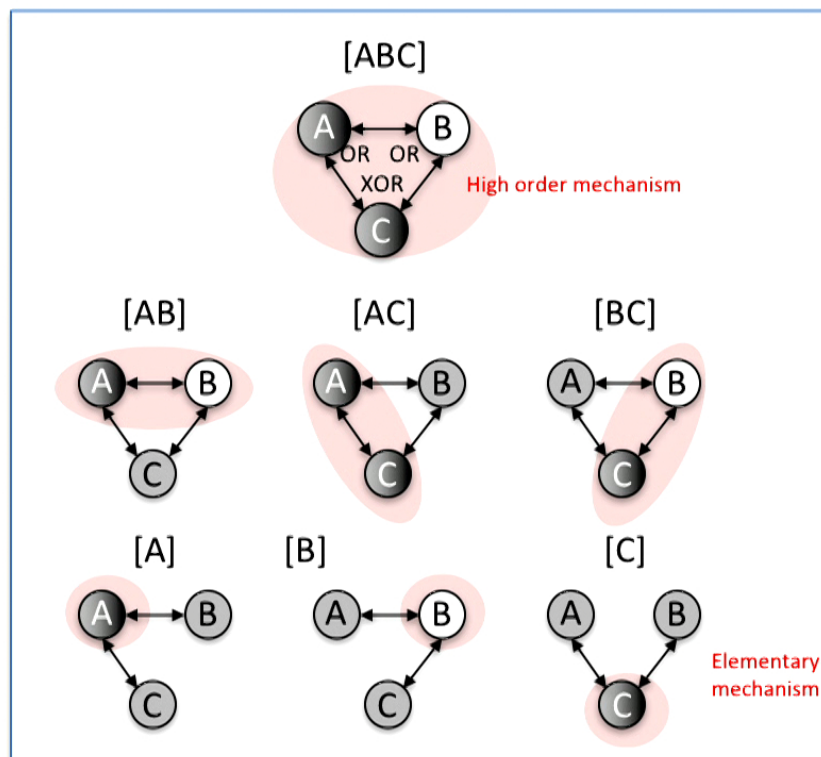
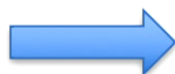
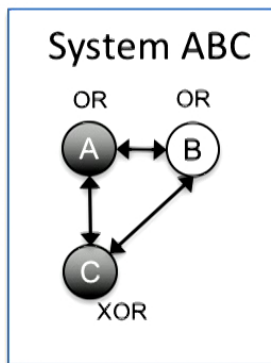
		t_1								
		ABC	000	100	010	110	001	101	011	111
t_0	ABC									
	000	1	0	0	0	0	0	0	0	0
	100	0	0	0	0	0	0	0	1	0
	010	0	0	0	0	0	1	0	0	0
	110	0	0	0	1	0	0	0	0	0
	001	0	0	0	1	0	0	0	0	0
	101	0	0	0	0	0	0	0	0	1
	011	0	0	0	0	0	0	0	0	1
111	0	0	0	1	0	0	0	0	0	

INTRINSICALITY

From the intrinsic perspective of a system, only “differences that make a difference” within the system matter.

Therefore, the system’s mechanisms must constrain the past and future of the system.

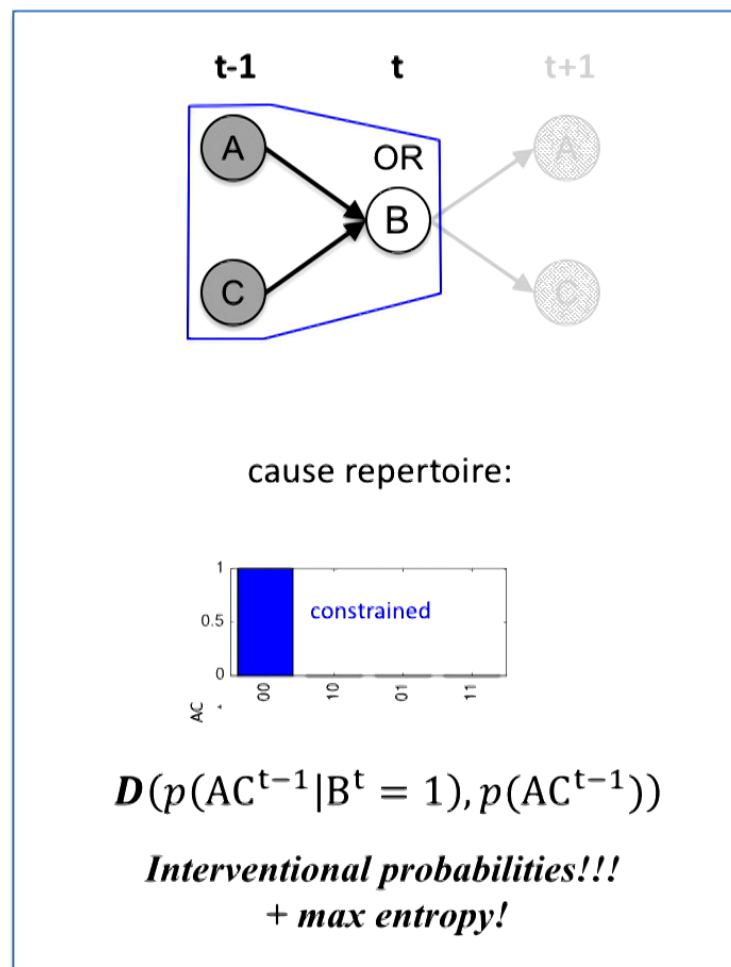
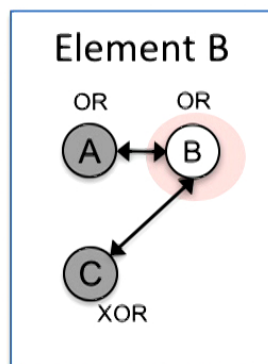
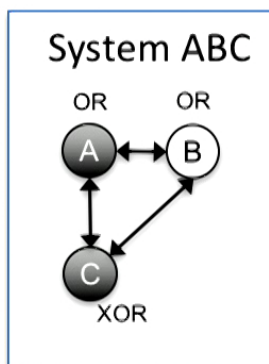
Formalism to assess intrinsic cause-effect structure



IIT principles:

- **Composition**
- **Information**
(Differences that make a difference)
- **Integration**
(Irreducible to parts)
- **Exclusion**
(Only maxima of integrated information count)

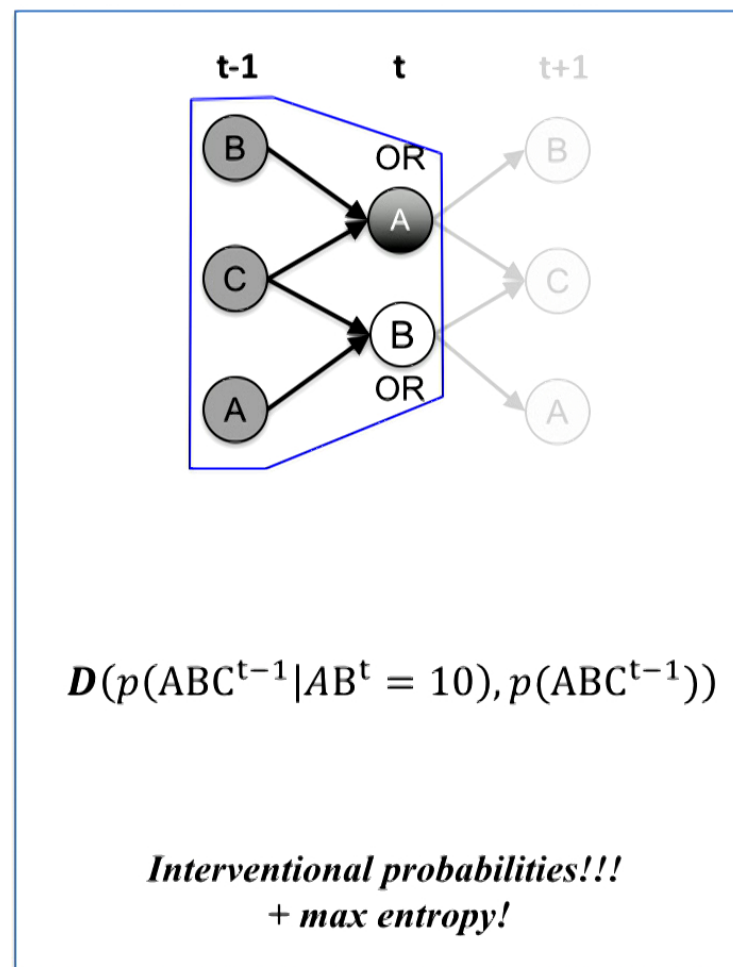
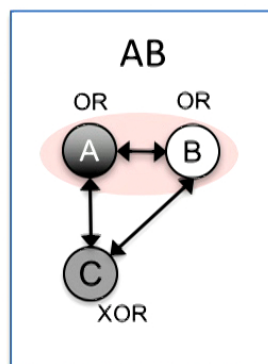
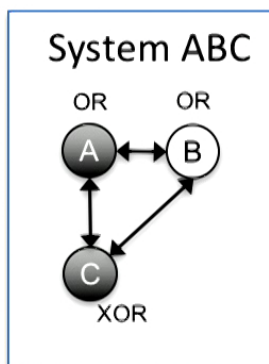
Formalism to assess intrinsic cause-effect structure



IIT principles:

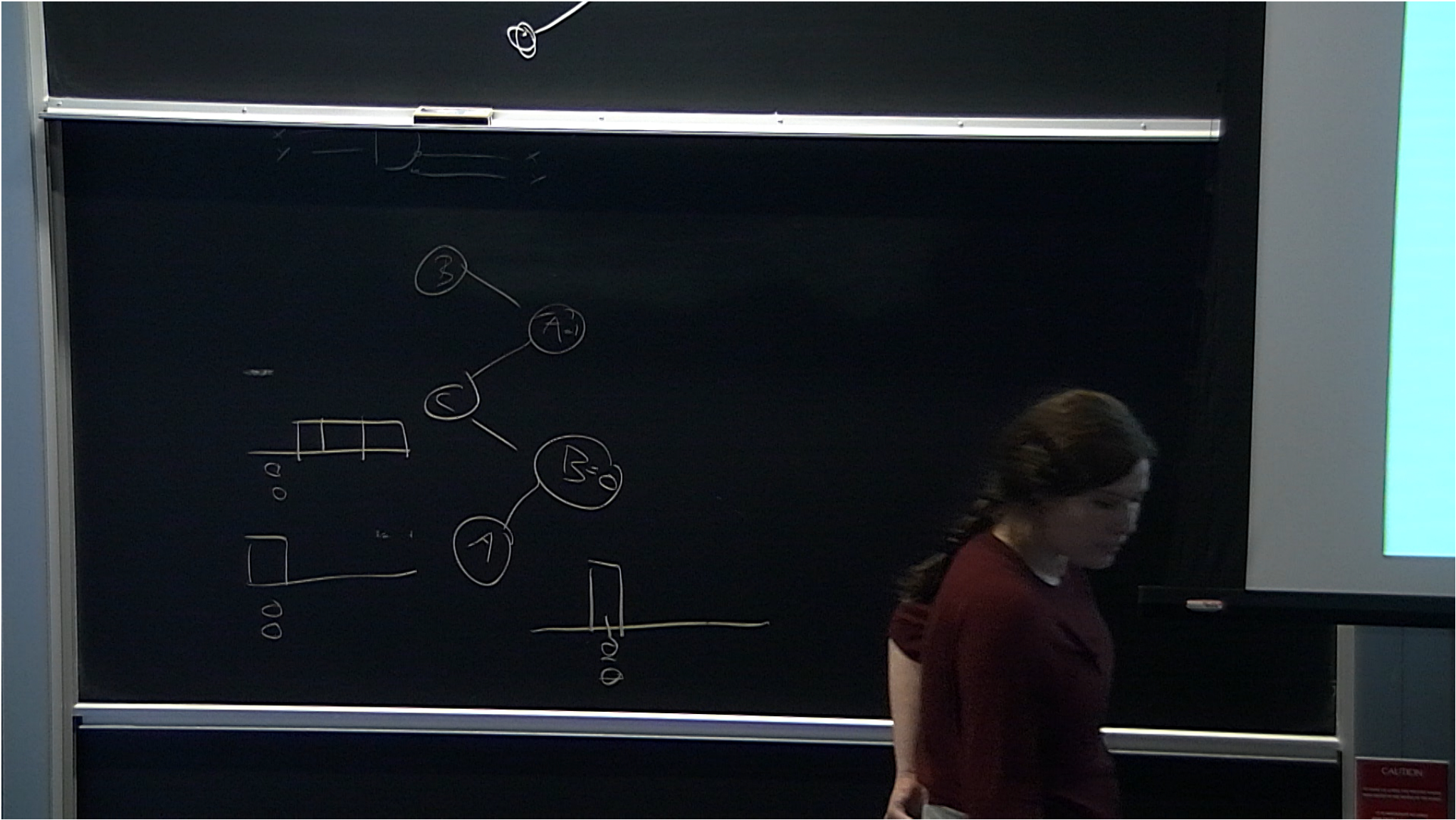
- Composition
- **Information**
(Differences that make a difference)
- Integration
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

Formalism to assess intrinsic cause-effect structure

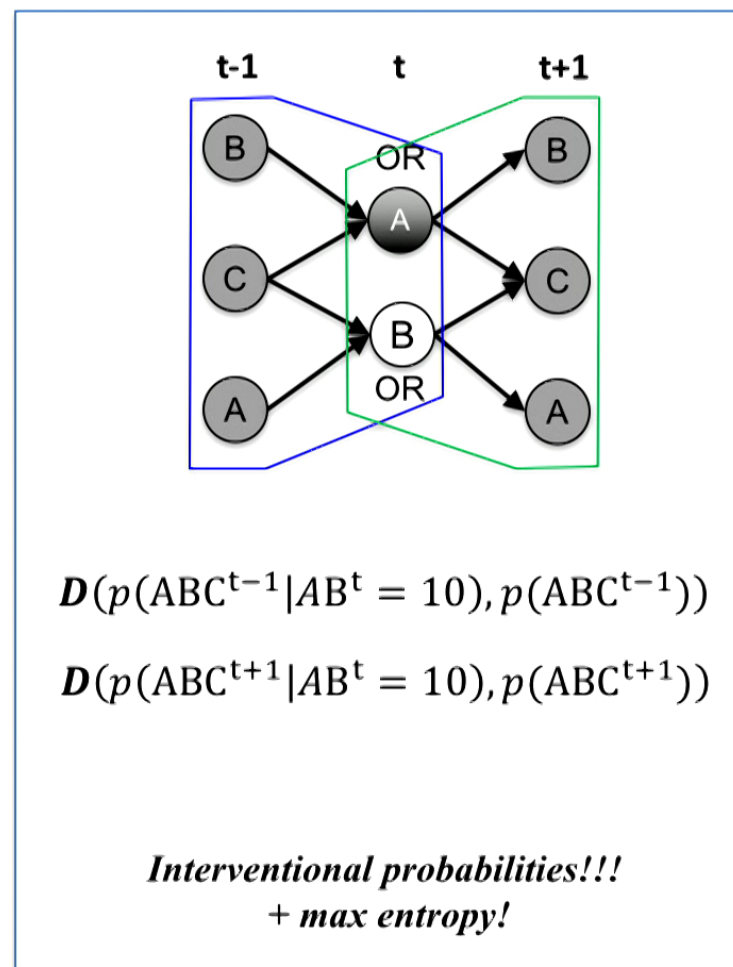
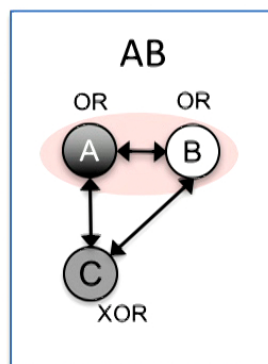
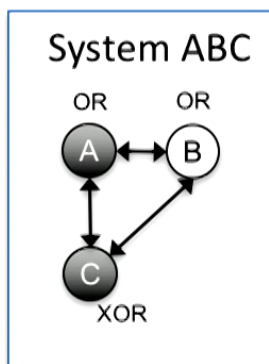


IIT principles:

- Composition
- Information
(Differences that make a difference)
- Integration
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)



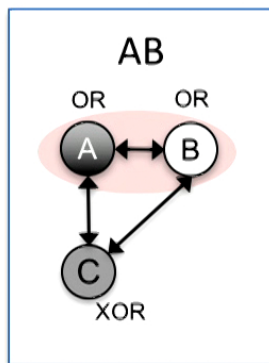
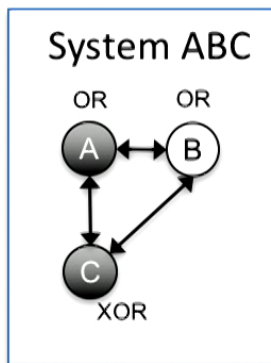
Formalism to assess intrinsic cause-effect structure



IIT principles:

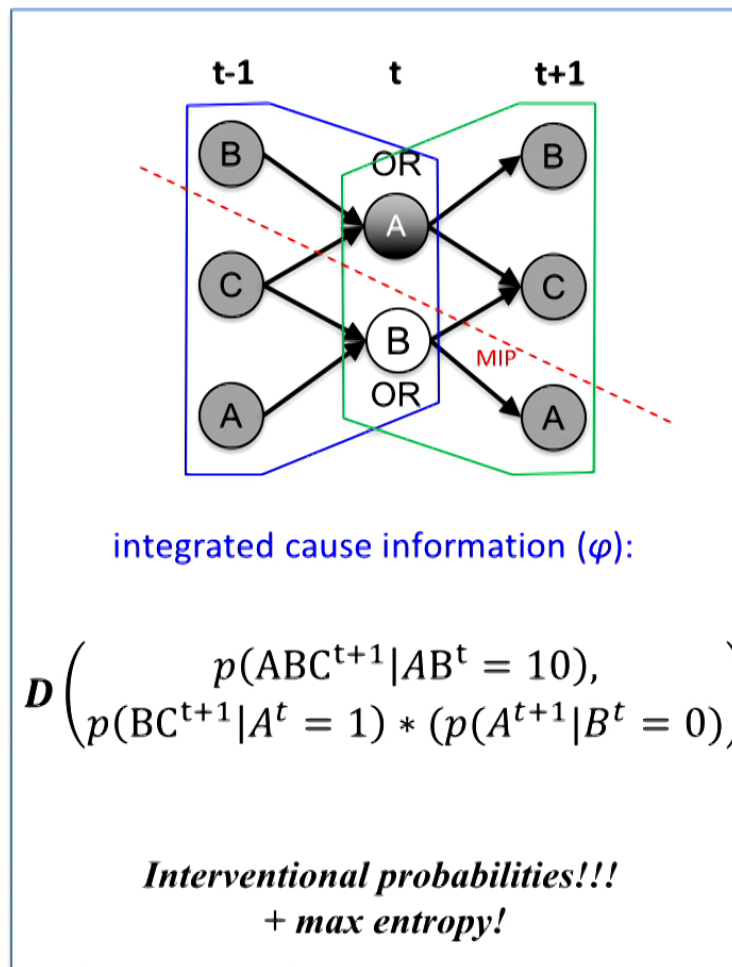
- **Composition**
- **Information**
(Differences that make a difference)
- **Integration**
(Irreducible to parts)
- **Exclusion**
(Only maxima of integrated information count)

Formalism to assess intrinsic cause-effect structure

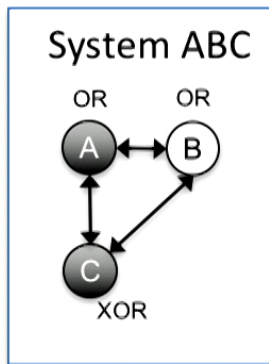


IIT principles:

- Composition
- Information
(Differences that make a difference)
- Integration
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

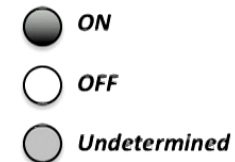
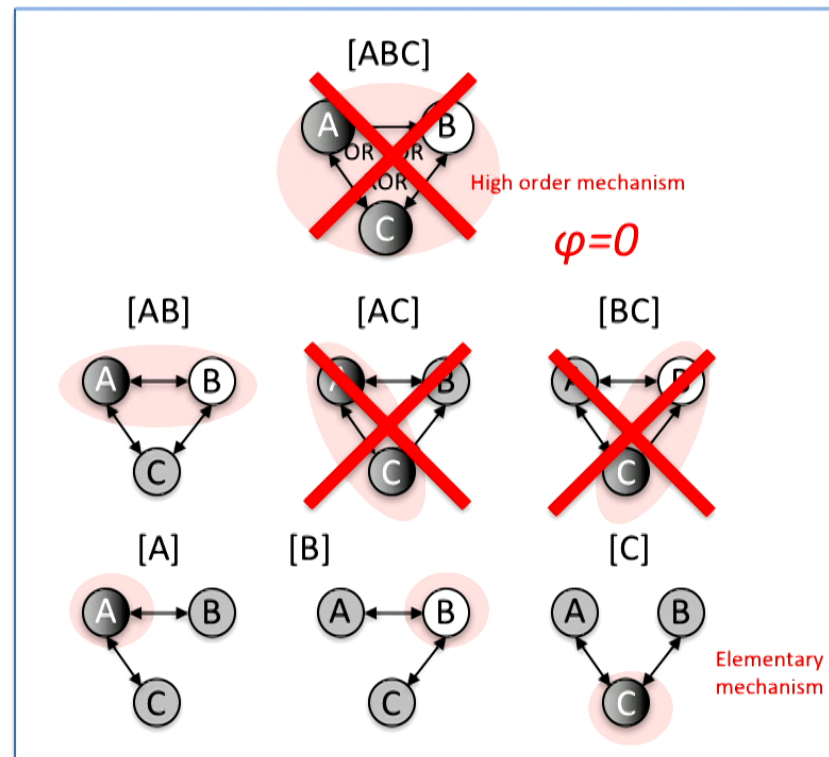


Formalism to assess intrinsic cause-effect structure

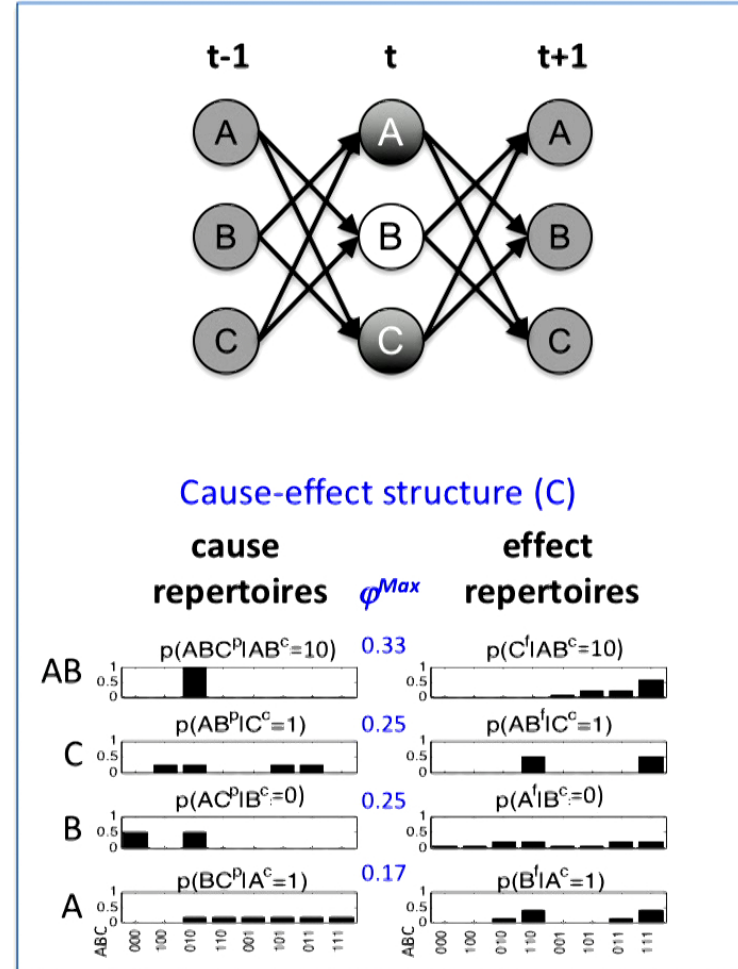
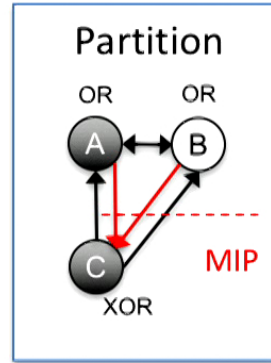
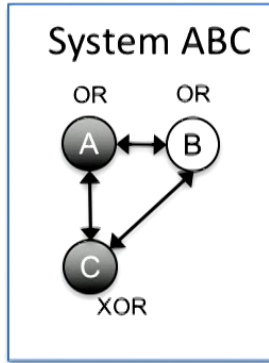


IIT principles:

- Composition
(Differences that make a difference)
- Information
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

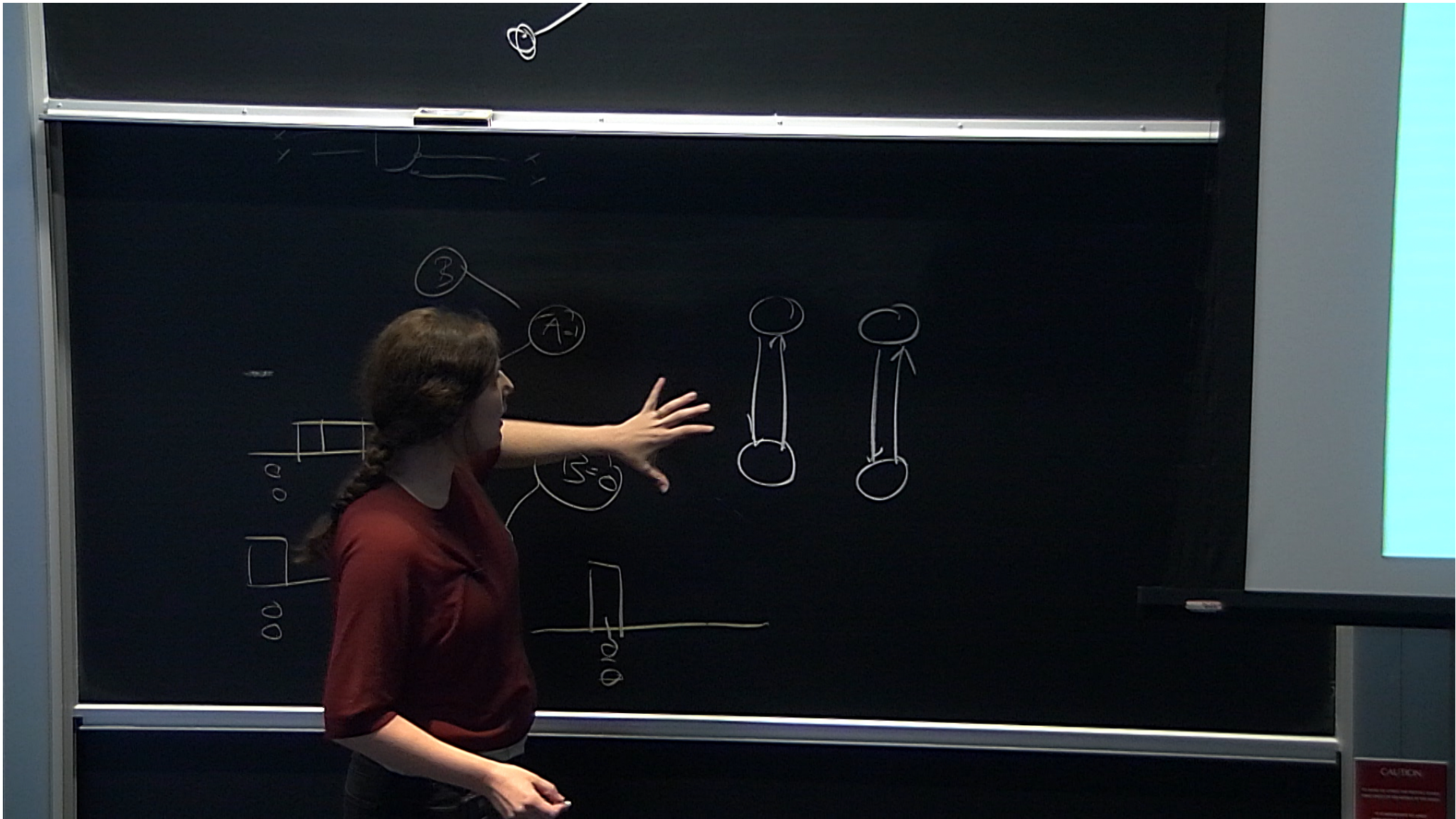


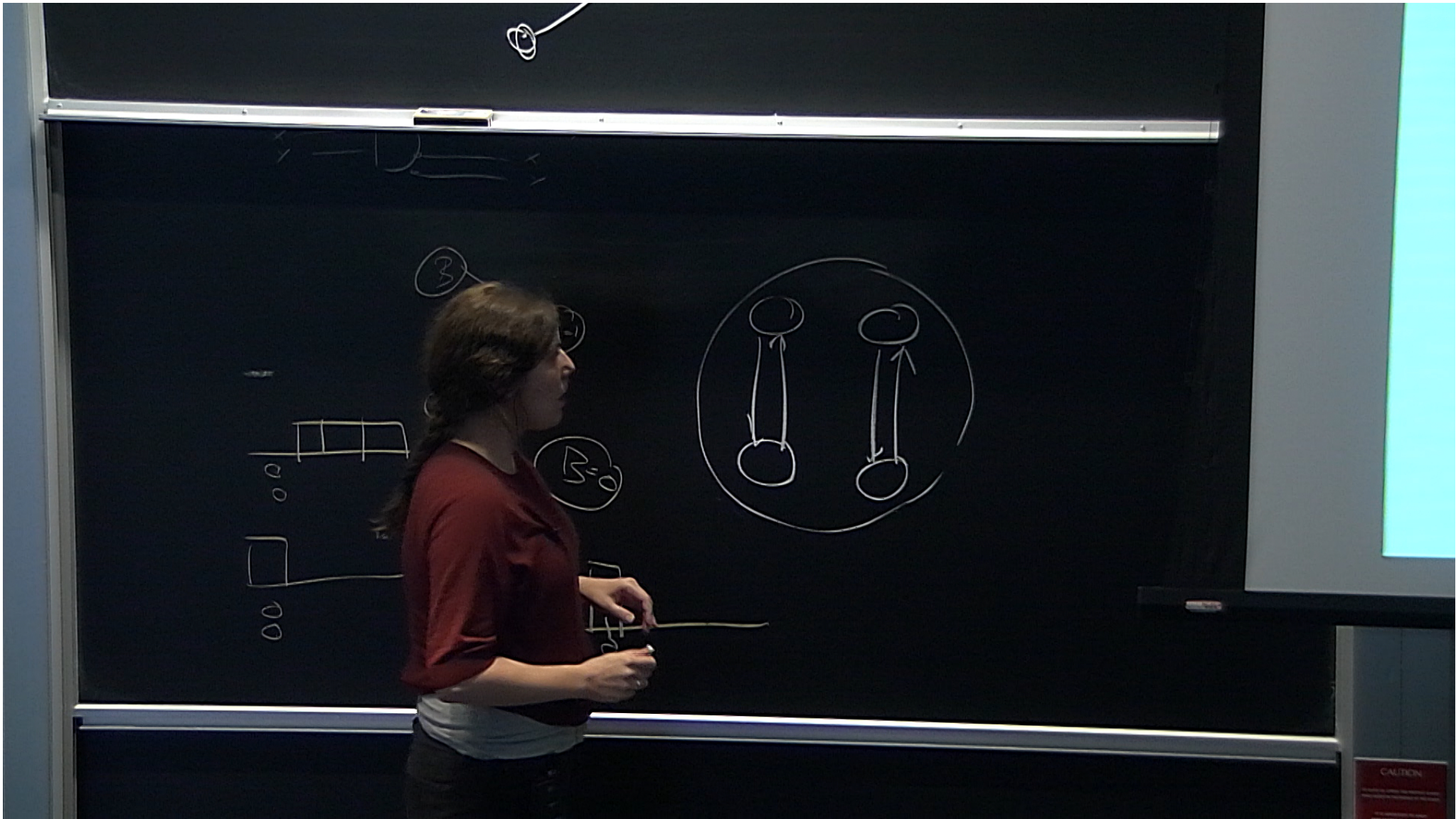
Formalism to assess intrinsic cause-effect structure



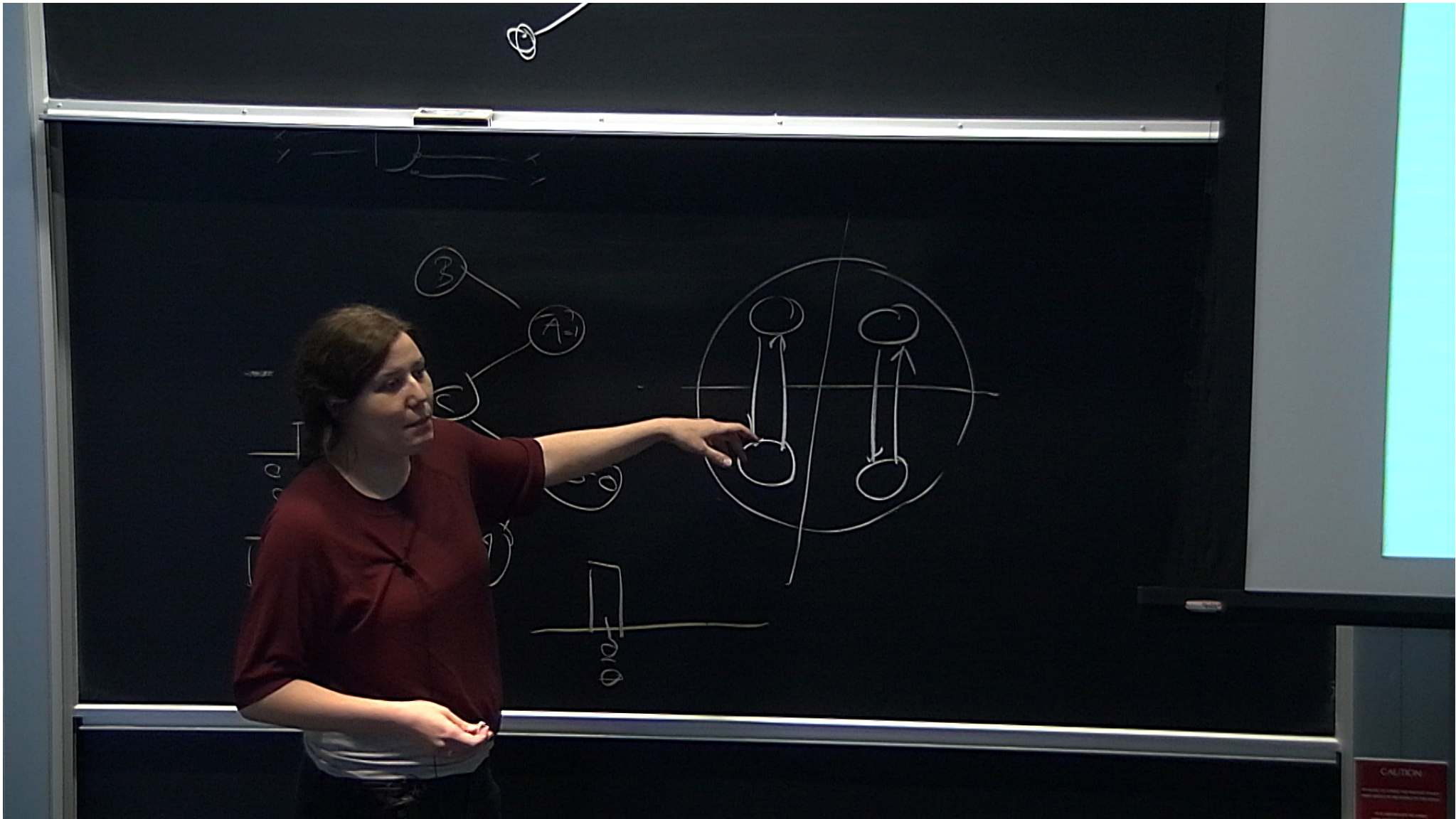
IIT principles:

- Composition
- Information
(Differences that make a difference)
- Integration
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

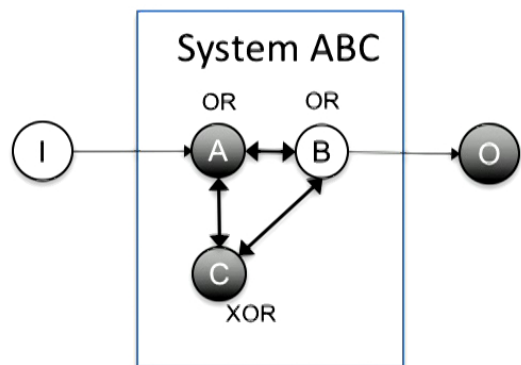






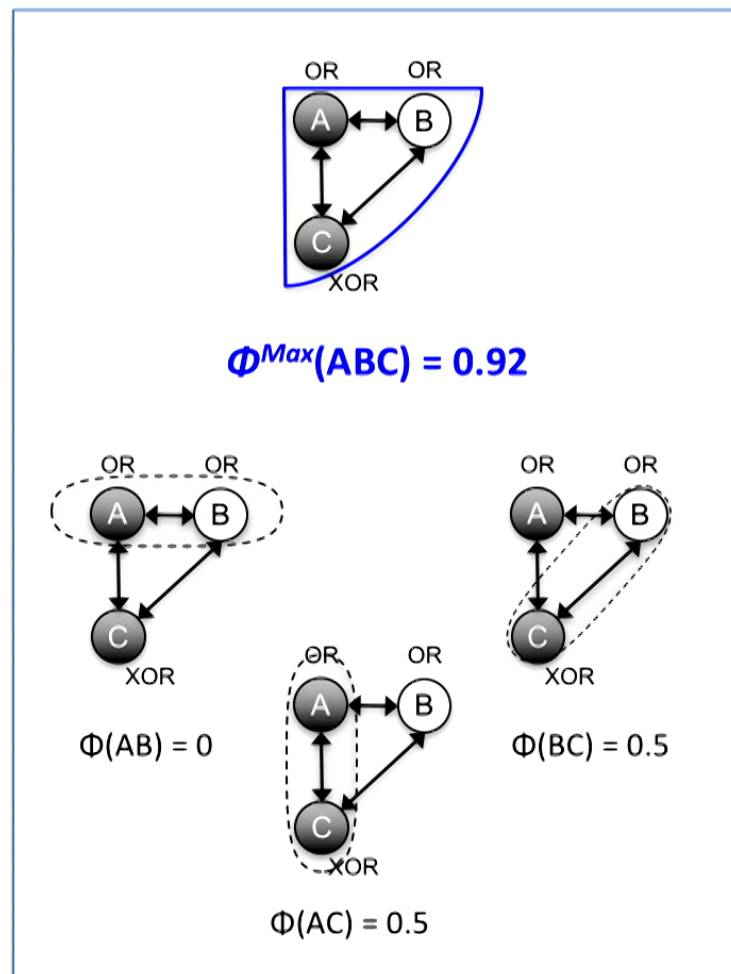


Formalism to assess intrinsic cause-effect structure

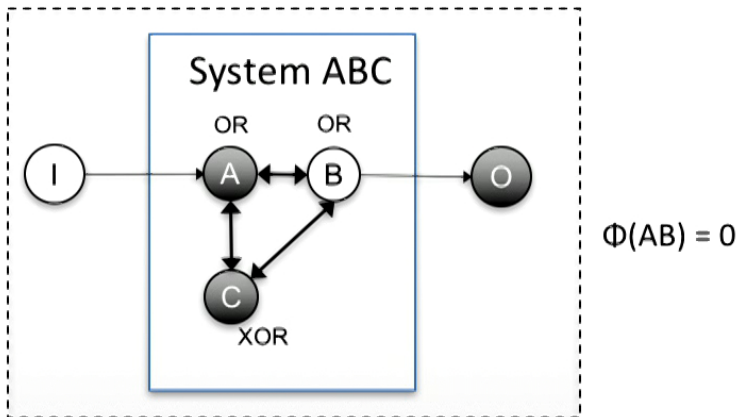


IIT principles:

- Composition
- Information
(Differences that make a difference)
- Integration
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

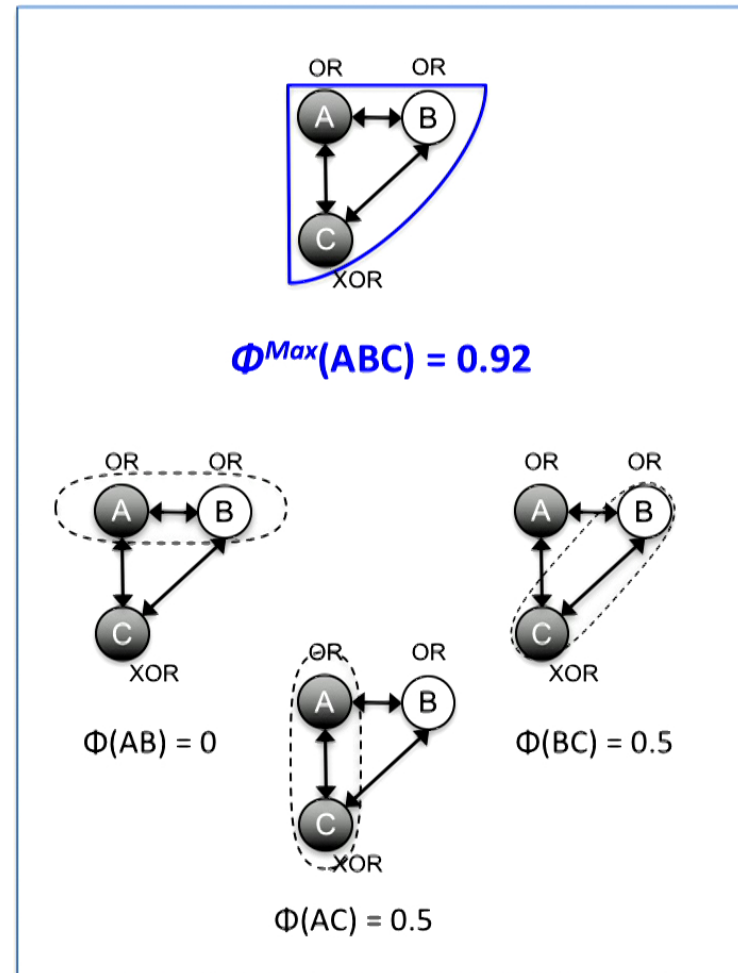


Formalism to assess intrinsic cause-effect structure

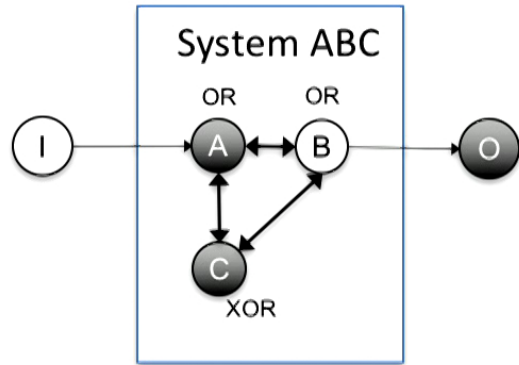


IIT principles:

- Composition
- Information
(Differences that make a difference)
- Integration
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

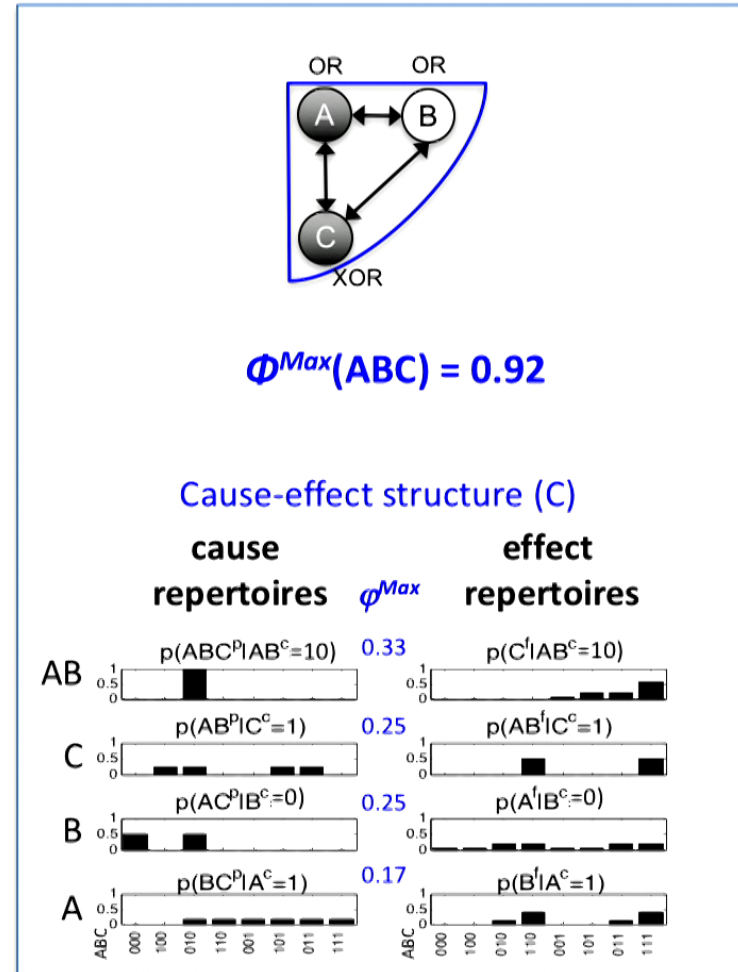


Formalism to assess intrinsic cause-effect structure



IIT principles:

- Composition
(Differences that make a difference)
- Information
(Irreducible to parts)
- Exclusion
(Only maxima of integrated information count)

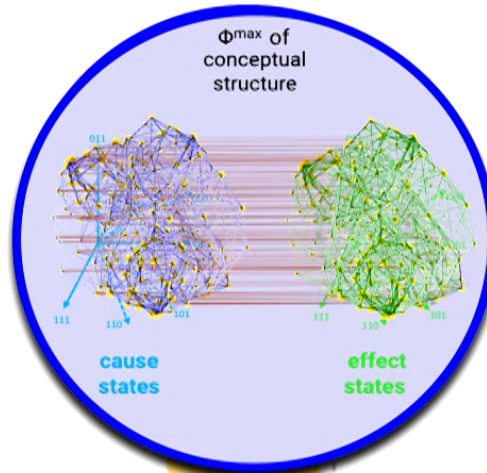


*An experience is a global maximum (Φ) of cause-effect power
(a cause-effect structure)*

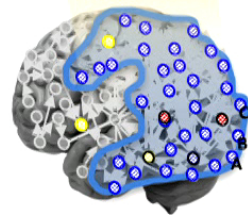


Experience

=



Cause-effect structure



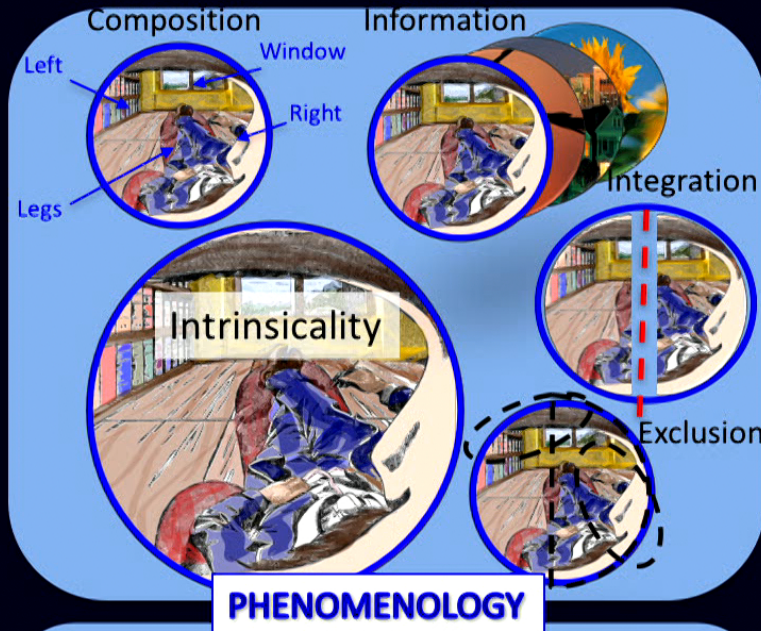
Physical substrate

Quantity:
*irreducibility Φ
of the cause-effect structure*

Quality:
*'form'
of the cause-effect structure*

*Tononi et al
Nature Rev Neurosci 2016*

From phenomenology to mechanisms, and back



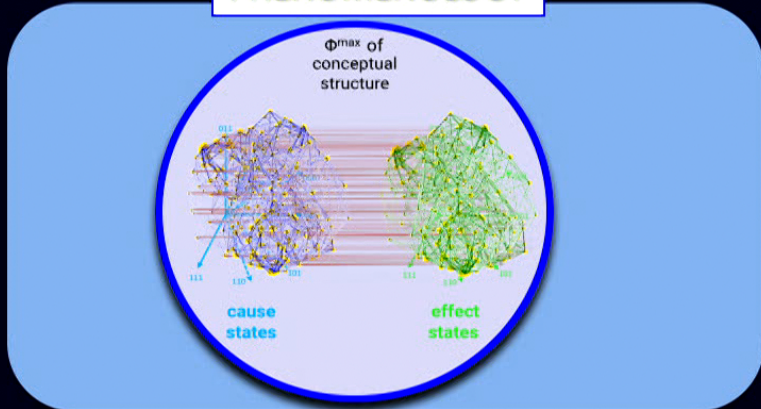
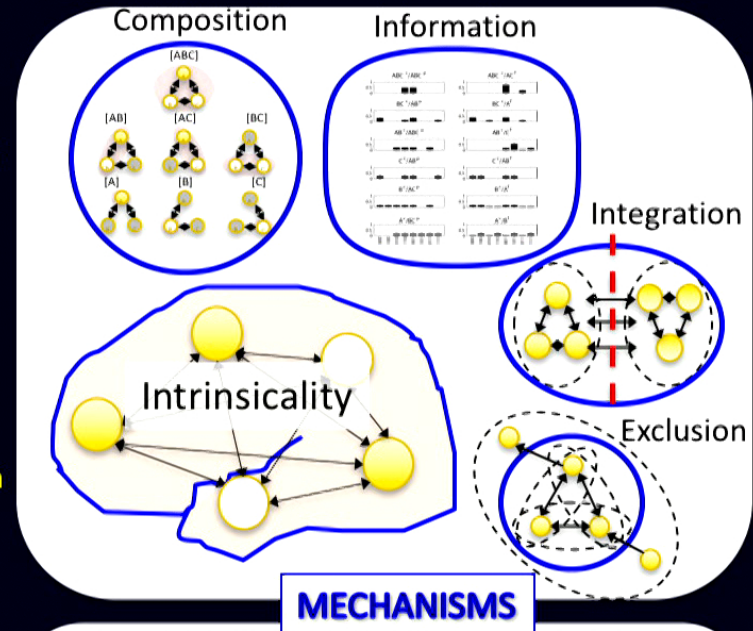
empirical



prediction



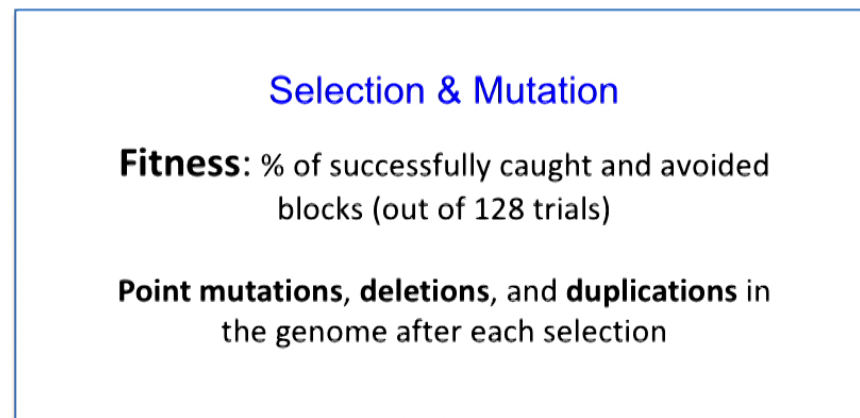
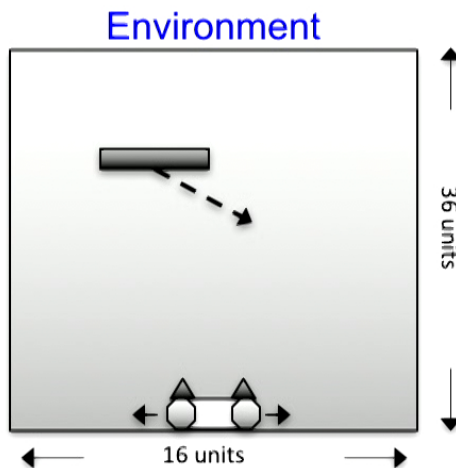
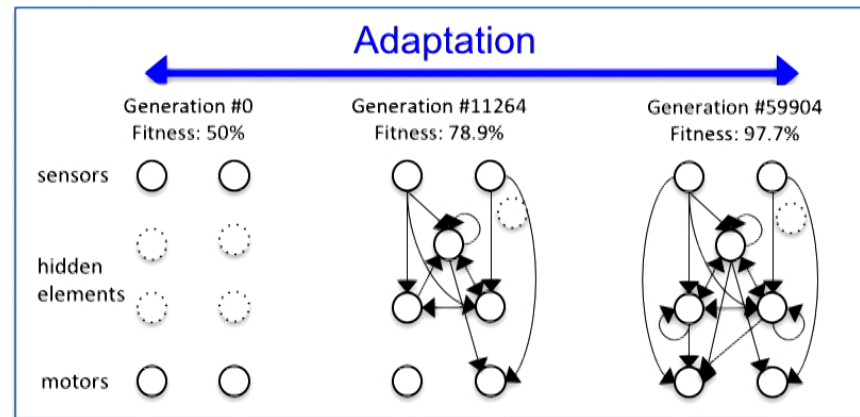
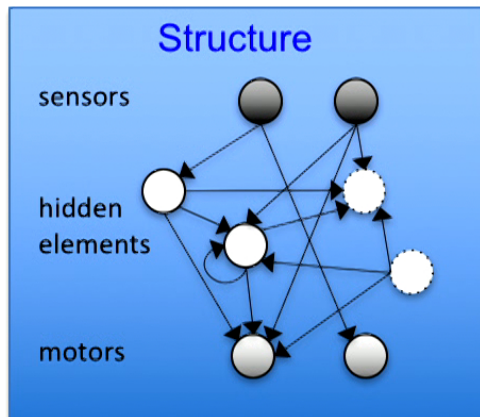
prediction



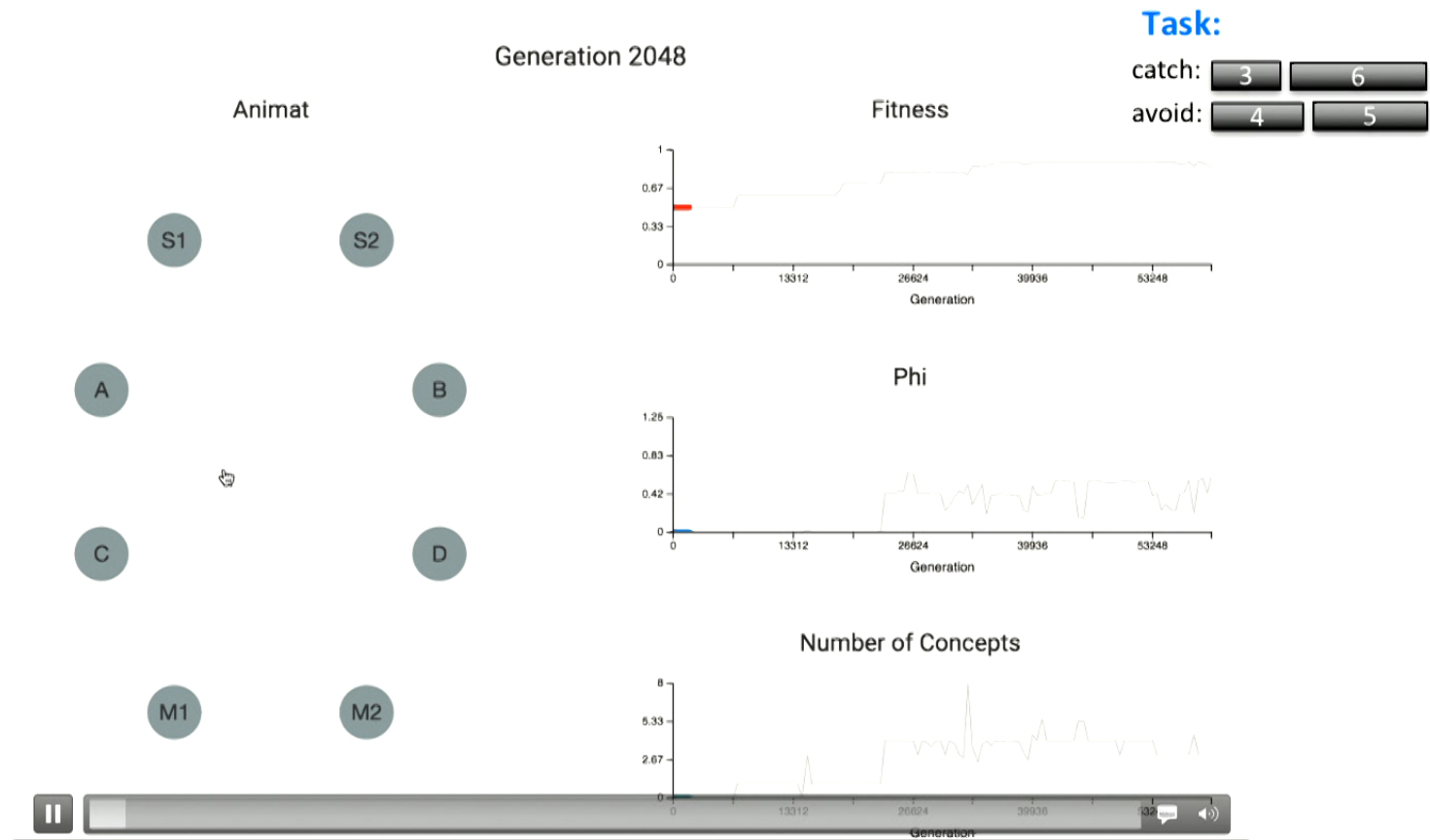
inference



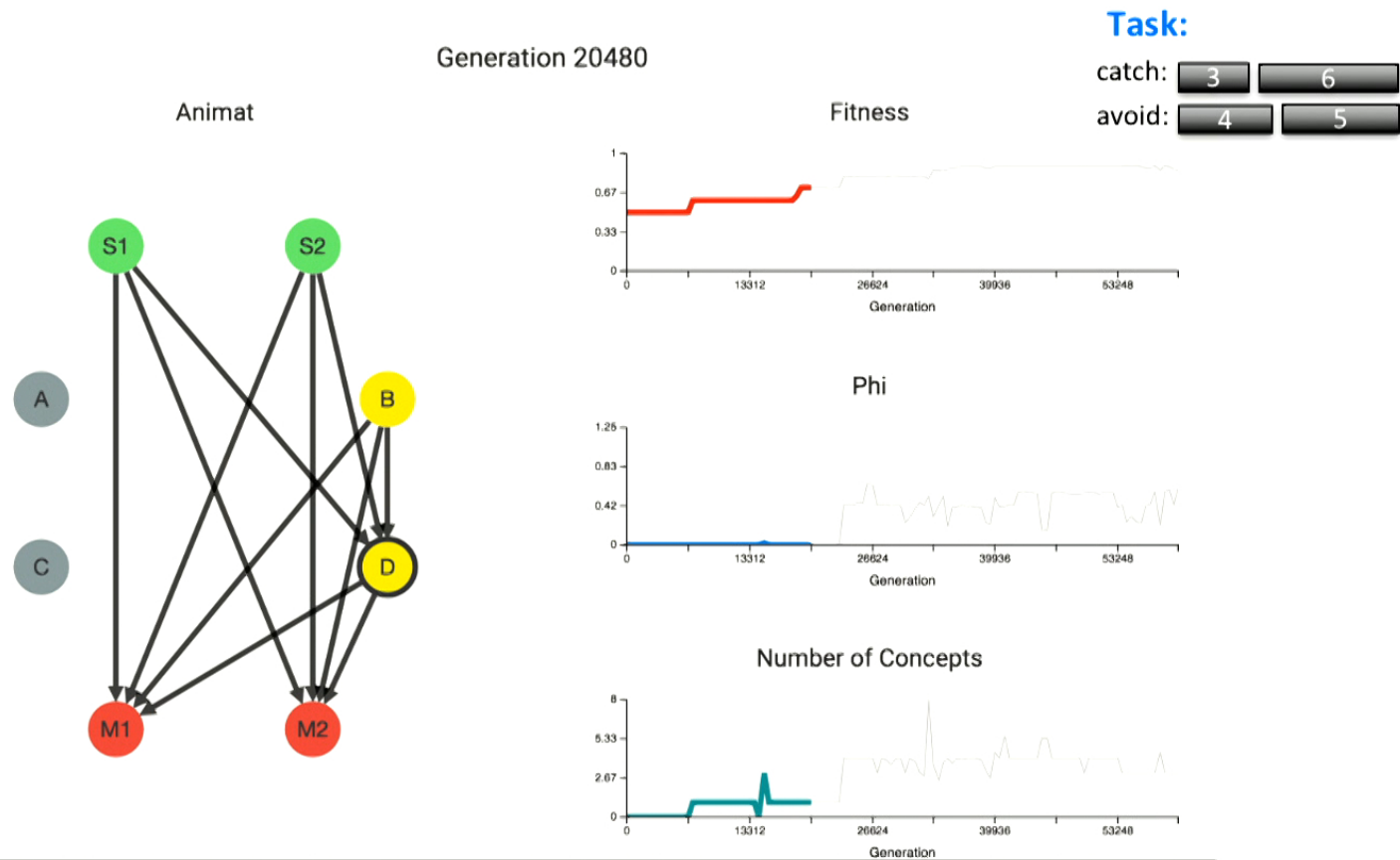
Animats – adaptive, behaving systems



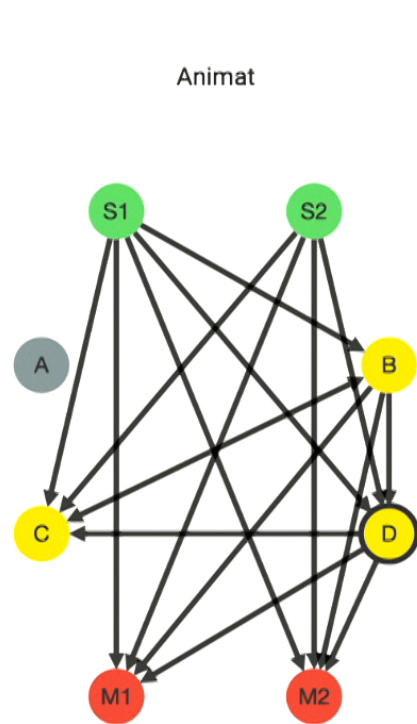
Integrated information Φ increases during evolution



Integrated information Φ increases during evolution



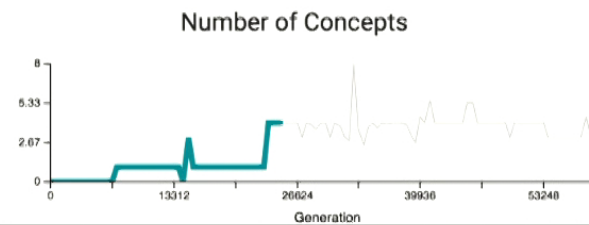
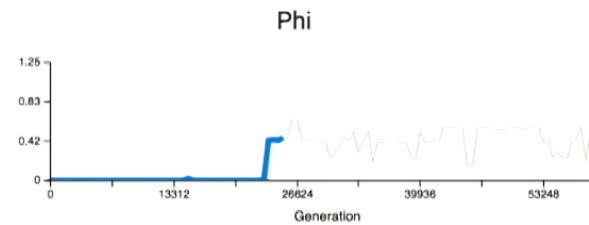
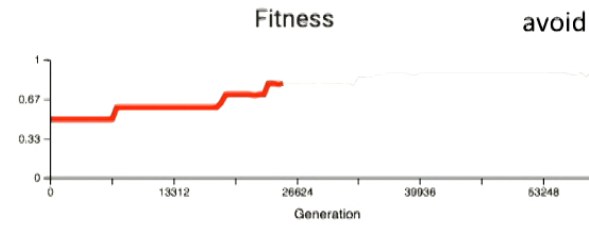
Integrated information Φ increases during evolution



Generation 25600

Task:

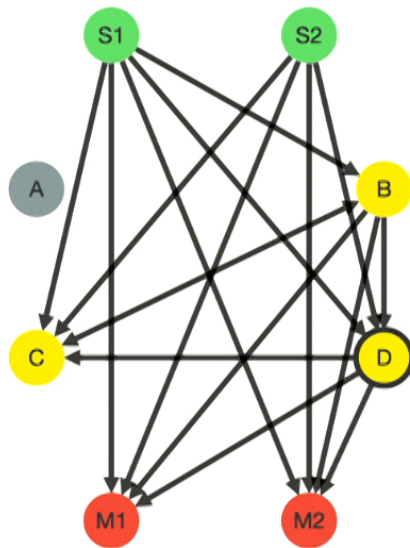
catch:
avoid:



Integrated information Φ increases during evolution

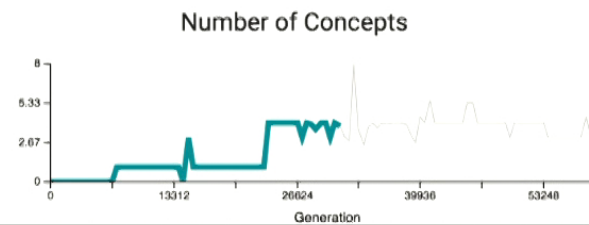
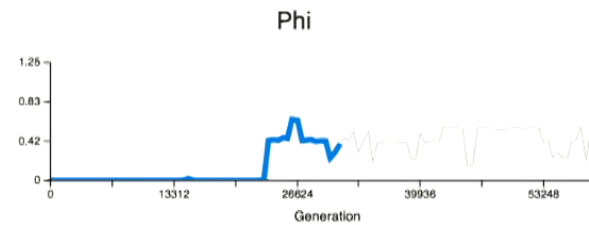
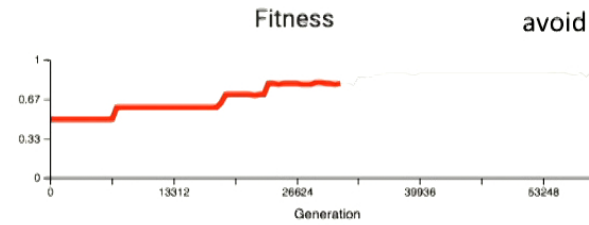
Generation 31744

Animat



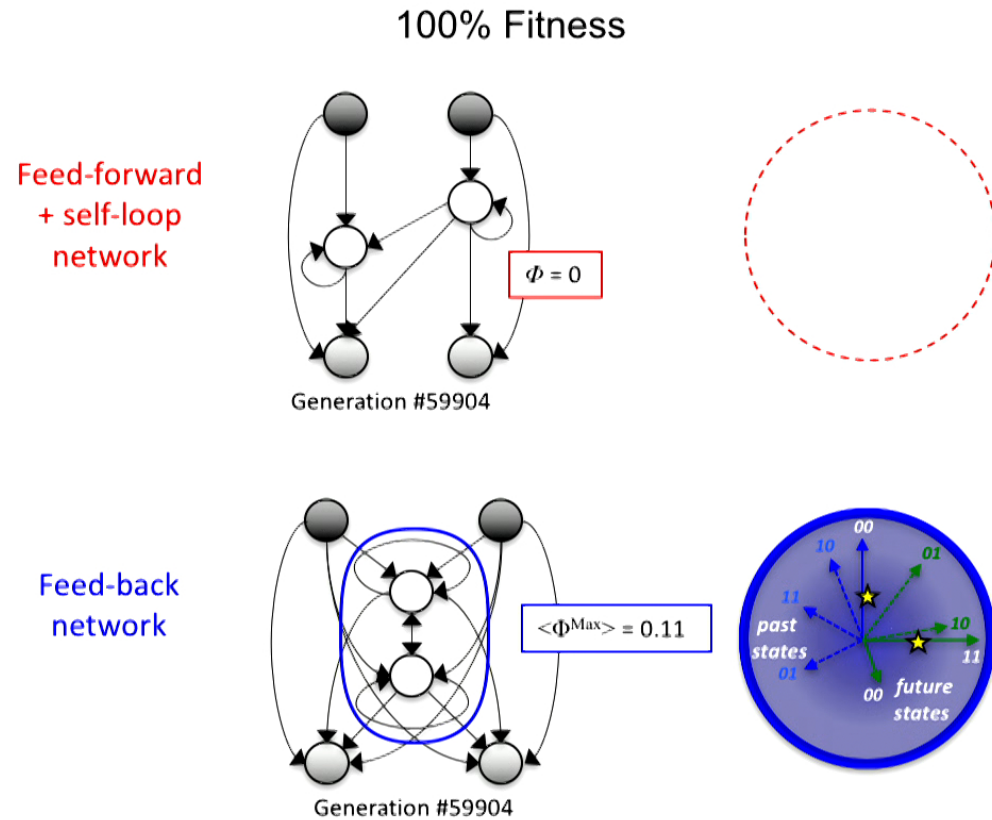
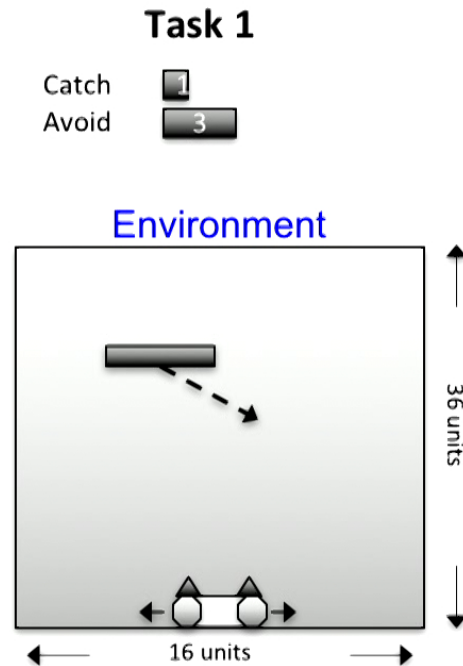
Task:

catch:
avoid:



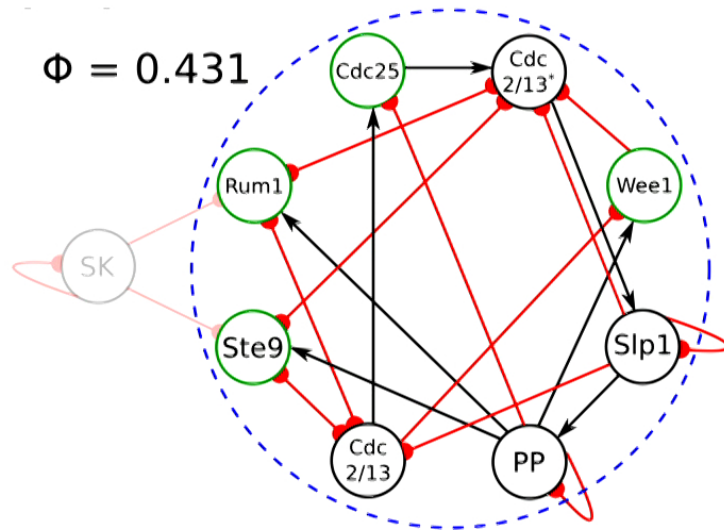
Evolved network structures

Simple task can be solved in a modular and integrated manner



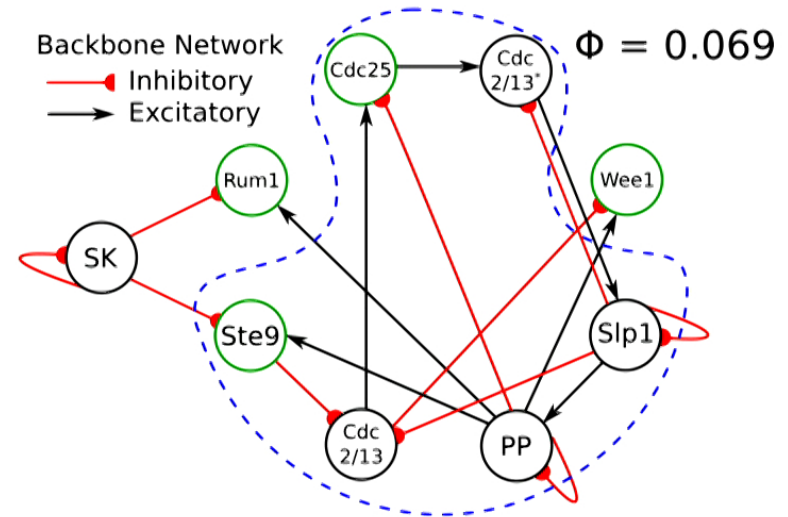
Robustness and Autonomy

fission yeast cell cycle model



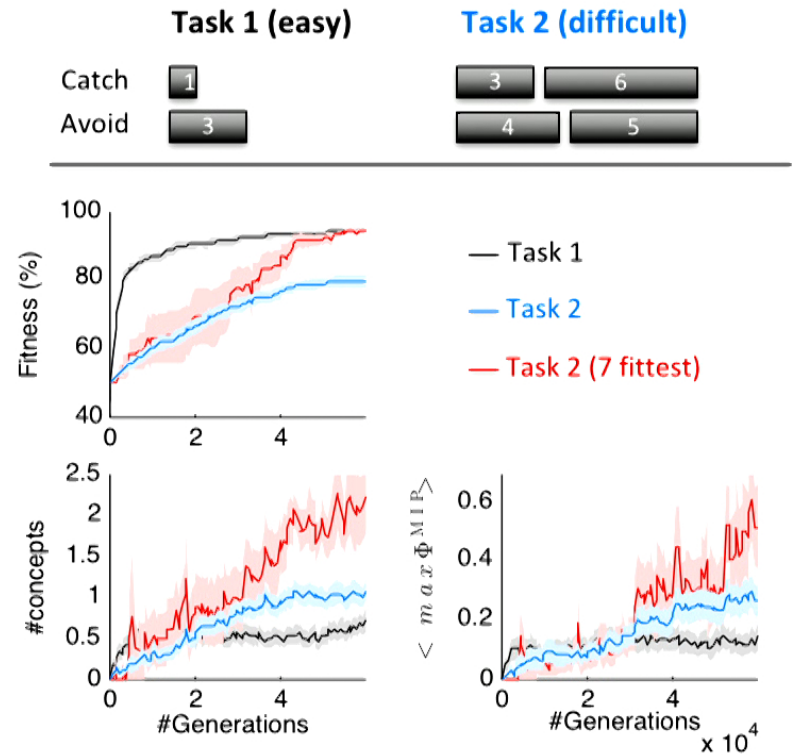
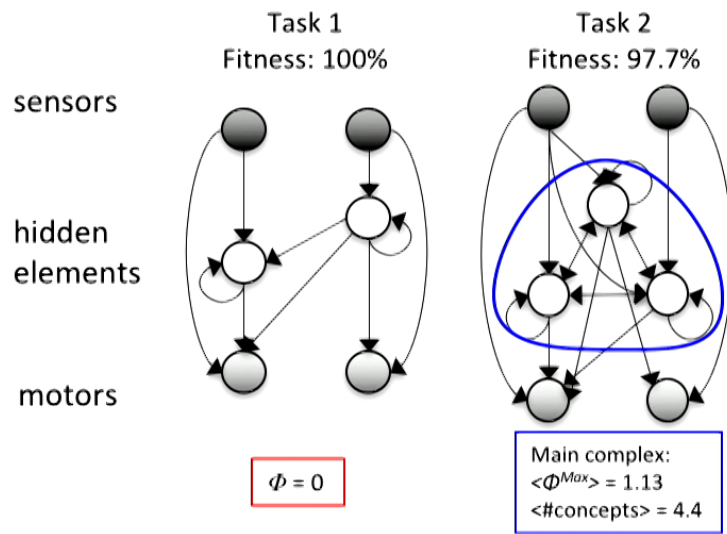
maximum of Φ in every state in the cycle

functionally identical
'backbone' network



Marshall W, Kim H, Walker SI, Tononi G, Albantakis L (in press) How causal analysis can reveal autonomy in models of biological systems.

The more difficult the task, the higher the integrated information
 ($\langle \Phi^{max} \rangle$ and # of mechanisms) in the fittest animats

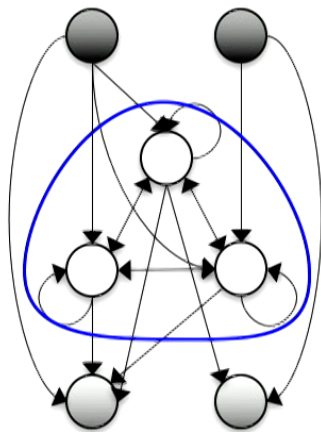


Albantakis et al., 2014

Autonomous Agents with context dependent behavior

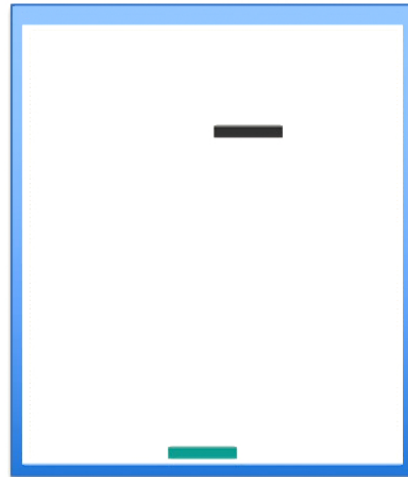
Task 2

catch: 3 6
 avoid: 4 5



Main complex:
 $\langle \phi^{Max} \rangle = 1.13$
 $\langle \#concepts \rangle = 4.4$

Generation #59904
 Fitness: 97.7%

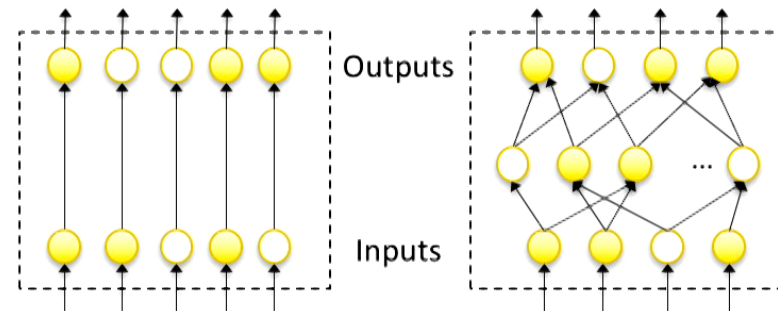


Animat is partially autonomous due to memory.



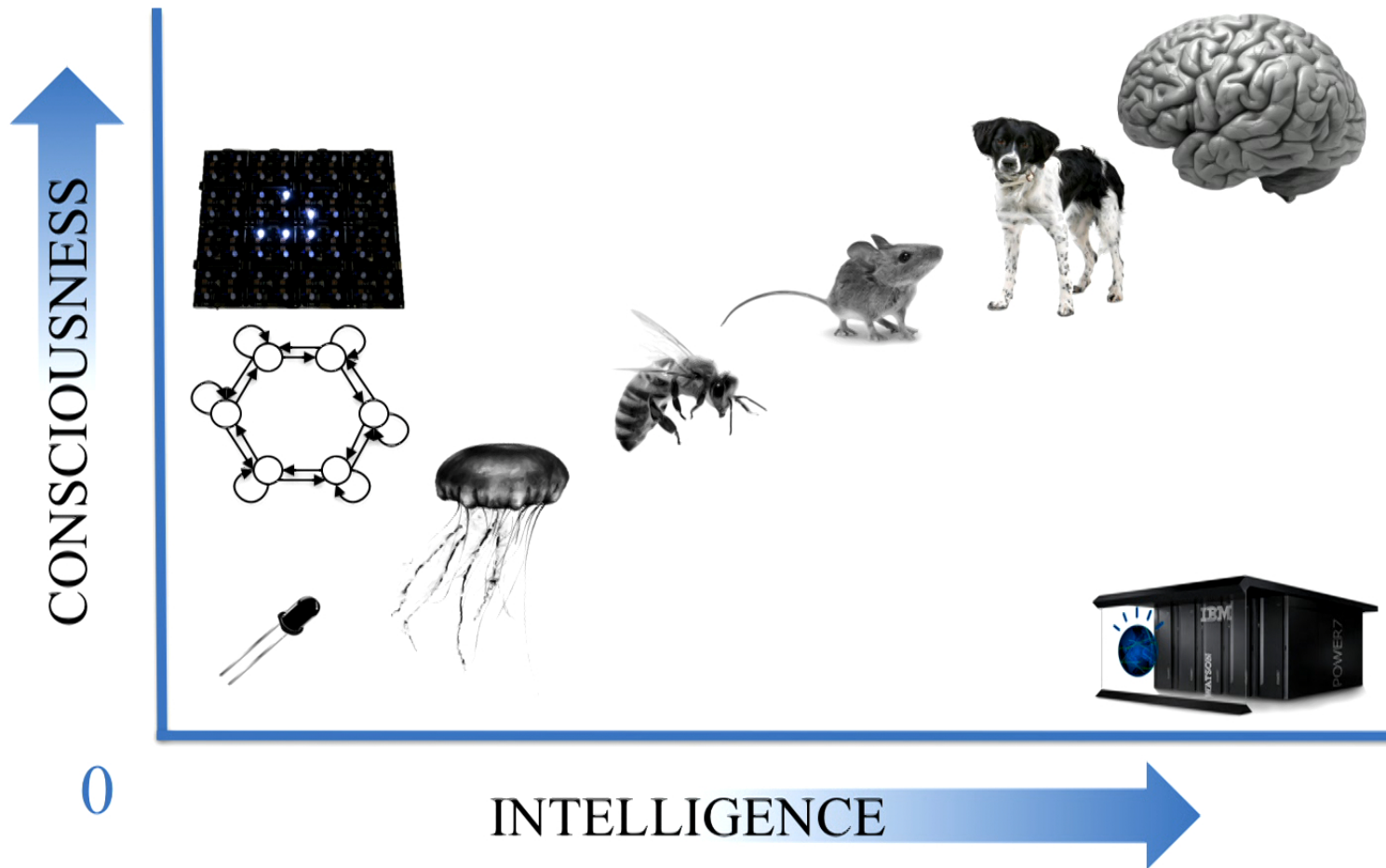
Modular/ Feed-forward system

$$\Phi = 0$$



Behavior entirely driven by the environment.

Intelligence vs. Consciousness



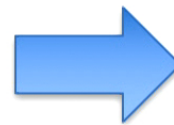
Why did consciousness evolve?

In adaptive systems, evolution to an environment with a rich causal structure provides a link between the system's behavior and its intrinsic cause-effect structures, leading to behaviorally fit systems with high Φ .

Axioms (phenomenal existence)



Postulates (physical existence)
→ IIT formalism

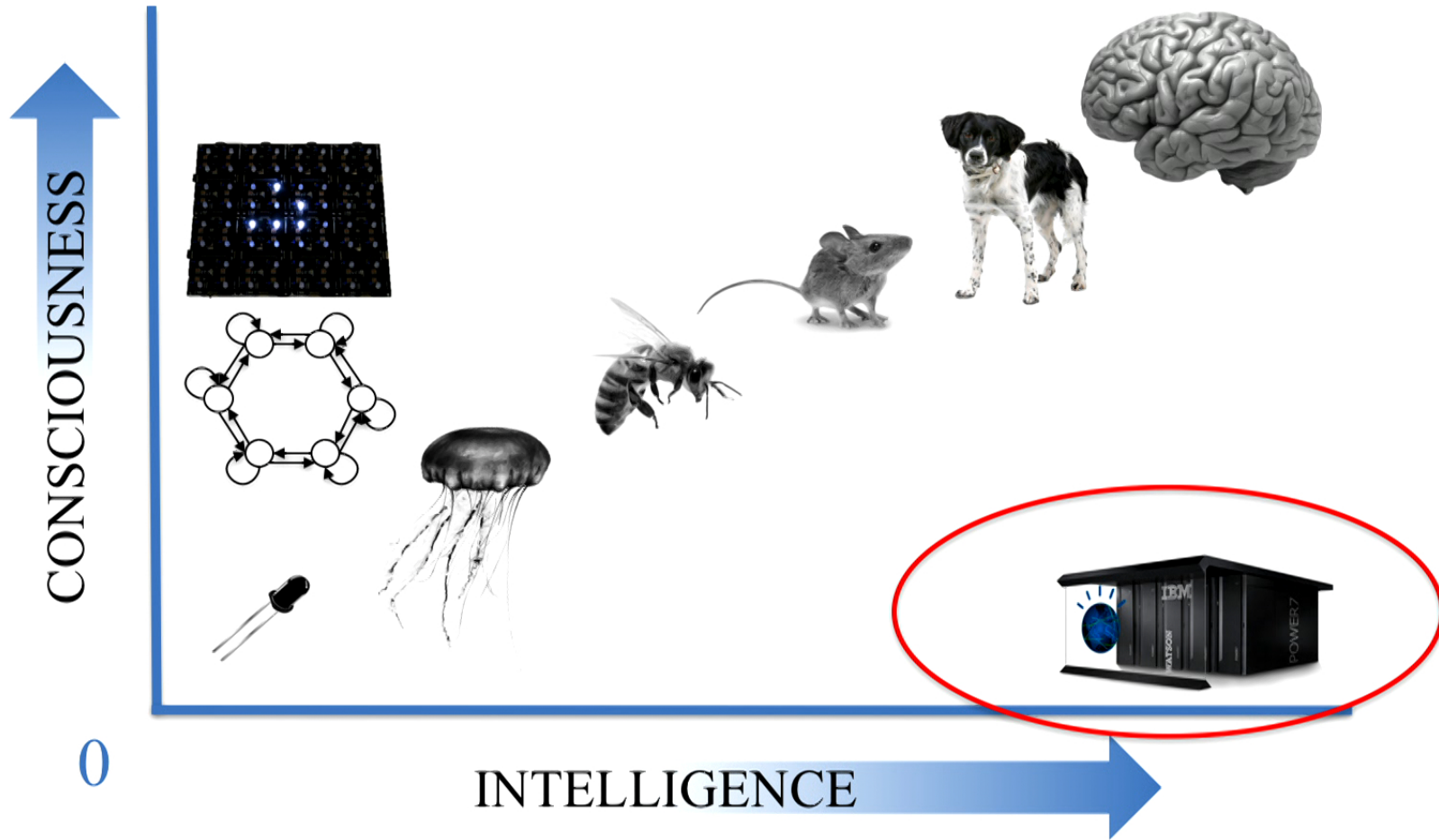


Why did consciousness evolve?

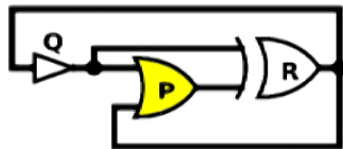
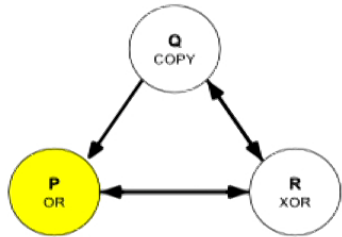


Why did integrated information evolve?

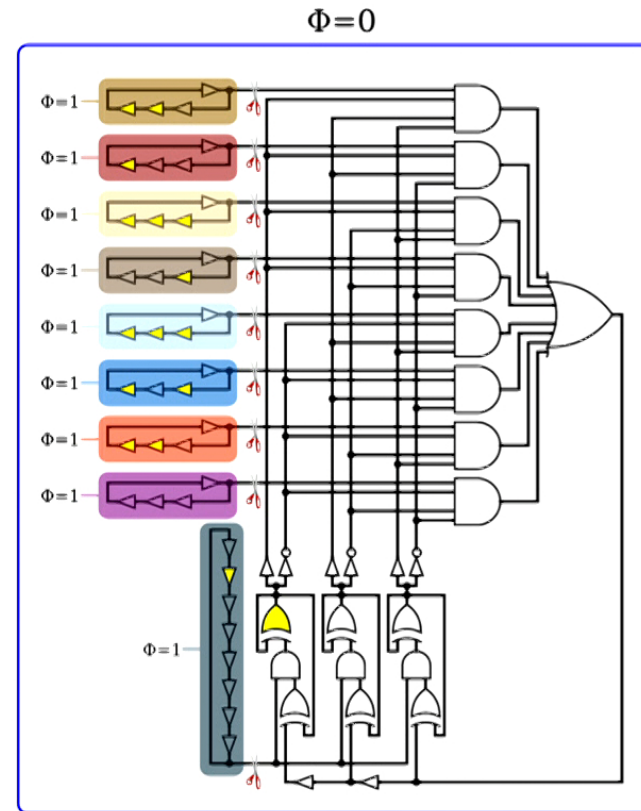
Intelligence vs. Consciousness



Functional equivalence does not imply phenomenal equivalence



$\Phi = 2.31$

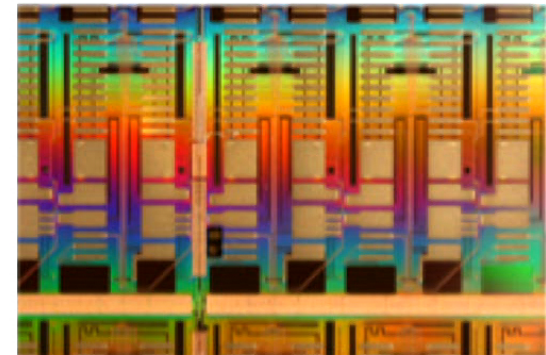
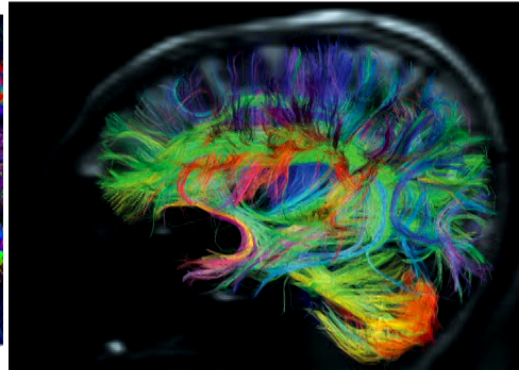
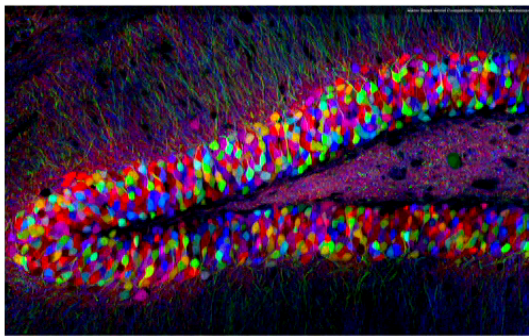


Findlay et al., in preparation

Computers simulating our behavior (BCC) are not conscious



Computers simulating our neural organization (NCC) are not conscious

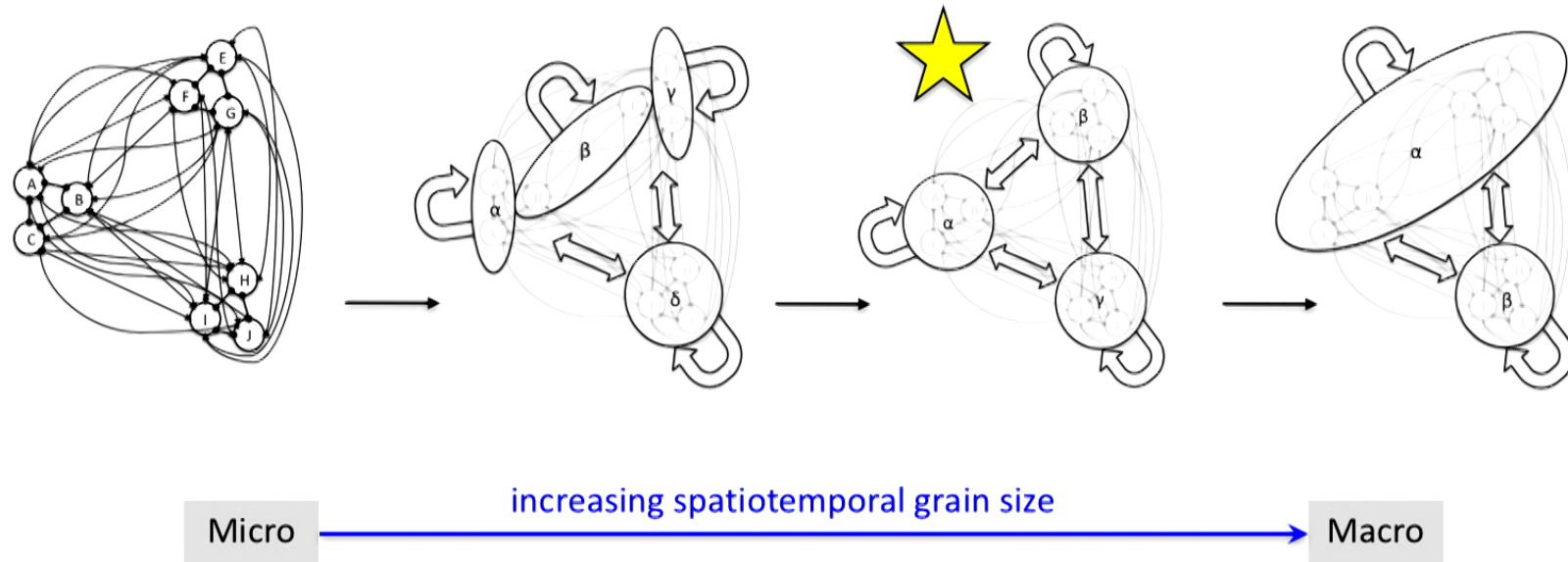


Question

At what spatiotemporal grain should we study the brain if our goal is to understand conscious experience?

Macro elements: Coarse-grains of micro elements

Proof-of-concept that Φ can peak at a macro level in small networks of logic-gates



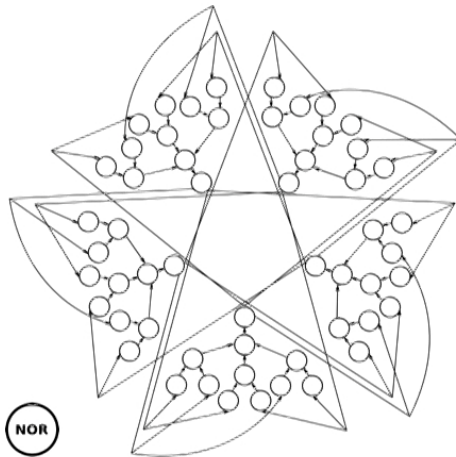
Hoel EP, Albantakis L, Tononi G (2013) Quantifying causal emergence shows that macro can beat micro. PNAS 110: 19790–19795.

Hoel EP, Albantakis L, Marshall W, Tononi G (2016) Can the macro beat the micro? Integrated information across spatiotemporal scales. Neurosci Conscious 2016.

Macro elements: Black boxes of micro elements

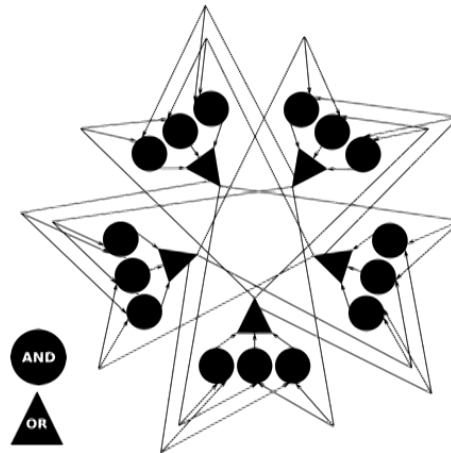
Proof-of-concept that Φ can peak at a macro level in small networks of logic-gates

$\Phi = 0.453$



- 55 micro elements
- 55 first order concepts
- 0 high order concepts

$\Phi = 0.080$



- 20 macro elements
- 20 first order concepts
- 0 high order concepts



Micro

increasing spatiotemporal grain size

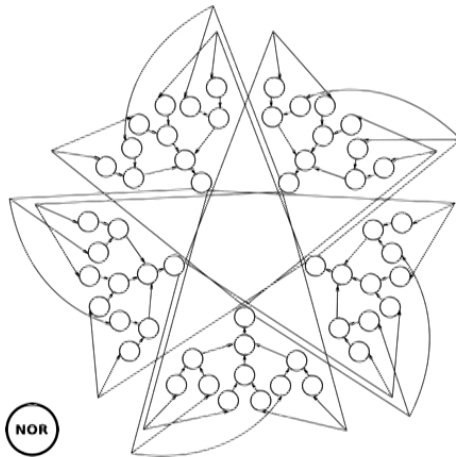
Macro

Marshall W, Albantakis L, Tononi G (2016) Black boxing and cause-effect power. ArXiv:1608.03461.

Macro elements: Black boxes of micro elements

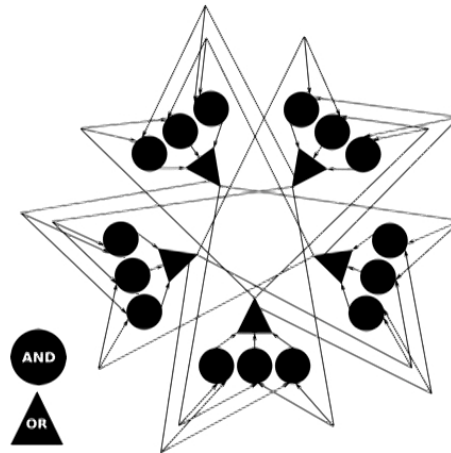
Proof-of-concept that Φ can peak at a macro level in small networks of logic-gates

$\Phi = 0.453$



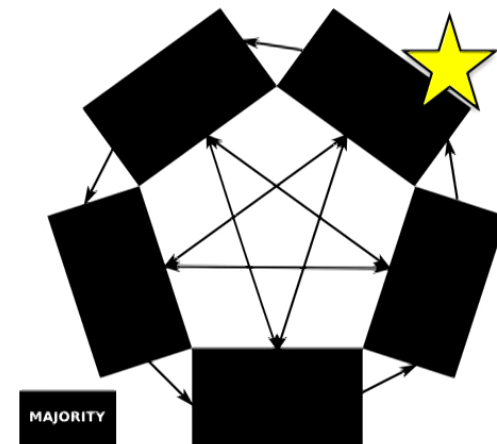
- 55 micro elements
- 55 first order concepts
- 0 high order concepts

$\Phi = 0.080$



- 20 macro elements
- 20 first order concepts
- 0 high order concepts

$\Phi = 2.33$



- 5 macro elements
- 5 first order concepts
- 25 high order concepts



Marshall W, Albantakis L, Tononi G (2016) Black boxing and cause-effect power. ArXiv:1608.03461.

Remarks on intrinsic information

- Intrinsic information is not about prediction. It's about (causal) structure. "Being vs. happening".
- Structural equivalence between phenomenology and cause-effect structure is testable. Not everything that can be decoded is part of phenomenology.
- Redundancy is meaningful.
- Implementation matters.

?

- Is IIT's causal analysis compatible with physics or not? QM?
- Is there convergence or real noise in the physical world? (E.g, Do real AND gates exist?)

Albantakis L (2017) A Tale of Two Animats: What does it take to have goals? ArXiv 1705.10854:1–6.

Integrated Information Theory

- physical information
- causal composition and higher order interactions
- complexity
- identifying causal/informational boundaries
- informational/causal measures of autonomy
- causal exclusion and emergence
- practical approximations of integrated information
- applications

Submission Deadline: **30 November 2018**

Guest Editors

Dr. Larissa Albantakis

University of Wisconsin–Madison

Prof. Giulio Tononi

University of Wisconsin–Madison

Author Benefits

- ❖ High visibility
- ❖ Rapid publication
- ❖ Unlimited and free access for readers
- ❖ No Copyright Constraints
- ❖ Thorough Peer-Review
- ❖ No Space Constraints, No Extra Space or Color Charges