http://integratedinformationtheory.org/

http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003588

From the Phenomenology to the Mechanisms of Consciousness: Integrated Information Theory 3.0

Masafumi Oizumi, Larissa Albantakis, Giulio Tononi

PLOS Published: May 8, 2014

DOI: 10.1371/journal.pcbi.1003588

Featured in PLOS Collections

Abstract

IIT starts from phenomenological axioms: information says that each experience is specific – it is what it is by how it differs from alternative experiences; integration says that it is unified – irreducible to non-interdependent components; exclusion says that it has unique borders and a particular spatio-temporal grain. These axioms are formalized into postulates that prescribe how physical mechanisms, such as neurons or logic gates, must be configured to generate experience (phenomenology).

By applying it's postulates, IIT concludes the following: an experience is a maximally irreducible conceptual structure (MICS, a constellation of concepts in qualia space), and the set of elements that generates it constitutes a complex. According to IIT, a MICS specifies the quality of an experience and integrated information ΦMax its quantity.

Author Summary

Integrated information theory (IIT) approaches the relationship between consciousness and its physical substrate by first identifying the fundamental properties of experience itself: existence, composition, information, integration, and exclusion. IIT then postulates that the physical substrate of consciousness must satisfy these very properties. We develop a detailed mathematical framework in which composition, information, integration, and exclusion are defined precisely and made operational. This allows us to establish to what extent simple systems of mechanisms, such as logic gates or neuron-like elements, can form complexes that can account for the fundamental properties of consciousness. Based on this principled approach, we show that IIT can explain many known facts about consciousness and the brain, leads to specific predictions, and allows us to infer, at least in principle, both the quantity and quality of consciousness for systems whose causal structure is known. For example, we show that some

simple systems can be minimally conscious, some complicated systems can be unconscious, and two different systems can be functionally equivalent, yet one is conscious and the other one is not.

As will be illustrated below, IIT offers a way to analyze systems of mechanisms to determine if they are properly structured to give rise to consciousness, how much of it, and of which kind. As reviewed previously [4], [5], [12], [13], the fundamental principles of IIT, such as integration and differentiation, can provide a parsimonious explanation for many neuroanatomical, neurophysiological, and neuropsychological findings concerning the neural substrate of consciousness.

IIT leads to experimental predictions, for instance that the loss and recovery of consciousness should be associated with the breakdown and recovery of information integration. This prediction has been confirmed using transcranial magnetic stimulation in combination with high-density electroencephalography.

While the central assumptions of IIT have remained the same, its theoretical apparatus has undergone various developments over the years….

The original formulation, which may be called IIT 1.0, introduced the essential notions including causal measures of the quantity and quality of consciousness. However, to simplify the analysis, IIT 1.0 dealt exclusively with stationary systems.

Let's go to IIT 1.0:

In what follows, we first present the axioms and the postulates of IIT. We then provide the mathematical formalism and motivating examples for each of the postulates.

Axioms, postulates, and identities of IIT are presented, followed by a mathematical formalism.

The axioms and postulates are very general statements about consciousness, such as that it exists, and it is structured.

A concept structure is DEFINED to be identical to its experience.

The existence postulate, the "zeroth" postulate of IIT, claims that mechanisms in a state exist. Within the present framework, "mechanism" simply denotes anything having a causal role within a system, for example, a neuron in the brain, or a logic gate in a computer.

The IIT analysis defines concepts to be identical to experience, and assumes that consciousness is caused by the binary (on/off) connections between neurons; the more the connections, the higher the level of consciousness.