Giulio Tononi
Phi – A Voyage from the Brain to the Soul
Pantheon Books 2012

xi Preface
What is consciousness, and what does it mean? How is it related to the world around us? What is it made of, and how is it generated inside the brain? Can science shed some light on it? Perhaps, but consciousness cannot just just inside the shroud of science. Because consciousness is more than an object of science: it is it subject to.
What follows is a story where an old scientist, Galileo, goes through a journey in search of consciousness. In his time, Galileo removed the Observer from nature and opened the way for the objectivity of science. Perhaps this is why Galileo is engaged to return the Observer to nature, to make subjectivity are part of science.
During his journey, Galileo meets people from here's another times, learned many lessons, thinks many thoughts, and sometimes wonders, too, whether he is awake or dreaming. In the first part of the book, he learns the facts of consciousness and the brain - why certain part of the brain are important but not others, or my consciousness fades with sleep. In the second part, he sees how these facts can be unified and understood through a scientific theory of consciousness - a theory that links consciousness to **Φ** (Phi), the symbol of integrated information that gives the book its title.
And finally, in the third part of the book, he realises some of the theories implications, and sees that they concern us all, because consciousness is everything we have, and everything we are. Each experience, Galileo realises, is a unique shape made of integrated information - a shape that is maximally irreducible - the shape of understanding. And it is the only shape that's really real - the most real thing there is.

147 Galileo and the Camera
Consciousness we take for granted, Galileo thought, because we always had it, and it requires no effort. We see dark, we see light, we see a woman, we see any other trillion things - they are just there, she is just there, immediately there, with no need for us to seek, compare, or calculate. And yet that immediacy may be illusary, because our brain can pick and choose from an inexhaustible repertoire - the repertoire of 1000 lifetimes. If we did not, if we had the insignificant repertoire of a photodiode, maybe we would not see her, we would not even see the dark - perhaps we would see nothing at all.

148 To be conscious, Galileo had concluded, the system must be able to distinguish among a large repertoire of possible states. Then a photodiode, with a repertoire vanishingly small - just one state corresponding to dark, and one corresponding to light - could only be minimally conscious, indeed just one bit conscious.

149 Nobody has ever counted the number of possible experiences that are available to me, set Galileo.

153 That being is…/ Indeed it is the same to think and to be…/ Now that all has been named light and dark…/ Everything is full at once of light and of dark night…
The One, if it has being, is One and Many. Parmenides

157 **Integrated Information: The Many and the One**

When is an entity one entity? How can multiple elements be a single thing? The question is simple enough - but one, thought Galileo, that had not yet been answered. Or perhaps, it had not been asked.
The sensor of the digital camera certainly had a large repertoire of states - it could take any possible picture. But was it a single entity? You use the camera as a single entity, you grasp it with your hands as one. You watch the photograph as a single entity. That is within your own consciousness. If it were not for you, the Observer, would it still be a single entity? And

what exactly would that mean?

158 An image came to Galileo. An astronomer was watching the sky and bad work, during an eclipse, and precisely at the same moment, another astronomer was watching the night sky at the anti-poets. Would there be a single consciousness contemplating, in one great image, the entire dome of the sky? That is absurd, thought Galileo whether the two were separated by the diameter of the Earth, or by a fraction of an inch, like two photodiodes on the camera sensor, made no difference. Because in both cases the two parties could not interact. And if they could not interact, they could not form a single entity, and they could not have a single, unified conscious experience.

159 If one measured information the way Shannon did, a camera was better than a brain: the larger the repertoire of states available to a system, the greater the reduction of uncertainty - the greater the information generated by the particular state the system visiting. But is this the right way of measuring information? It should make a difference if the information is generated by a system that is one, rather than just a collection of parts.

164 The information generated by the whole above and beyond its parts - call it **integrated information** - is what distinguishes (consciousness) from a camera.

(James said:) **Integrated information** is the information generated by a system above its parts, where the parts are those that, taken independently, generates the most information. Now that we have a definition, we need a symbol for it.

If you need a symbol, it should be **Φ**, said (Alan Turing), that is the symbol of the golden ratio - the right way of dividing something into parts. And the minimum cut, which reveals how much information is integrated information, is the right way of dividing a system into parts, is it not? You should call it **Φ.**
That would be interesting, said Galileo. After all, the golden ratio was studied by a fellow Pisan, the good old **Φ**ibonacci.
It is better than that, said James, **Φ** is like **Φ**enomenology, like experience, which is what consciousness is.
Better than that, said Galileo. **Φ** has an I, for information, and an O, a circle, for integration. Letters call it **Φ** then.

Scientific American 2009
http://www.scientificamerican.com/article.cfm?id=a-theory-of-consciousness

A "Complex" Theory of Consciousness
Is complexity the secret to sentience, to a panpsychic view of consciousness?

By **Christof Koch**
Do you think that your newest acquisition, a Roomba robotic vacuum cleaner that traces out its unpredictable paths on your living room floor, is conscious? What about that bee that hovers above your marmalade-covered breakfast toast? Or the newborn who finally fell asleep after being suckled? Nobody except a dyed-in-the-wool nerd would think of the first as being sentient; adherents of Jainism, India's oldest religion, believe that bees—and indeed all living creatures, small and large—are aware; whereas most everyone would accord the magical gift of consciousness to the baby.
The truth is that we really do not know which of these organisms is or is not conscious. We have strong feelings about the matter, molded by tradition, religion and law. But **we have no objective, rational method, no step-by-step procedure, to determine whether a given organism has subjective states, has feelings**.

The reason is that **we lack a coherent framework for consciousness**. Although consciousness is the only way we know about the world within and around us—shades of the famous Cartesian deduction cogito, ergo sum—there is no agreement about what it is, how it relates to highly organized matter or what its role in life is. **This situation is scandalous! We have a detailed and very successful framework for matter and for energy but not for the mind-body problem. This dismal state of affairs might be about to change, however**.

The universal lingua franca of our age is **information**. We are used to the idea that stock and bond prices, books, photographs, movies, music and our genetic makeup can all be turned into data streams of zeros and ones. These bits are the elemental atoms of information that are transmitted over an Ethernet cable or via wireless, that are stored, replayed, copied and assembled into gigantic repositories of knowledge. Information does not depend on the substrate. The same information can be represented as lines on paper, as electrical charges inside a PC's memory banks or as the strength of the synaptic connections among nerve cells.
Since the early days of computers, scholars have argued that **the subjective, phenomenal states that make up the life of the mind are intimately linked to the information expressed at that time by the brain**. Yet they have lacked the tools to turn this hunch into a concrete and predictive theory. Enter psychiatrist and neuroscientist **Giulio Tononi** of the University of Wisconsin–Madison. Tononi has developed and refined what he calls the **integrated information theory (IIT) of consciousness**.

An Integrated Theory
IIT is based on two axiomatic pillars. First, **conscious states are highly differentiated; they are informationally very rich**. You can be conscious of an uncountable number of things: you can watch your son's piano recital, for instance; you can see the flowers in the garden outside or the Gauguin painting on the wall. Think of all the frames from all the movies you have ever seen or that have ever been filmed or that will be filmed! Each frame, **each view, is a specific conscious percept**.
Second, this information is highly integrated. No matter how hard you try, you cannot force yourself to see the world in black-and-white, nor can you see only the left half of your field of view and not the right. When you're looking at your friend's face, you can't fail to also notice if she is crying.

Whatever information you are conscious of is wholly and completely presented to your mind; it cannot be subdivided. Underlying this unity of consciousness is a multitude of causal interactions among the relevant parts of your brain.
If areas of the brain start to disconnect or become fragmented and balkanized, as occurs in deep sleep or in anesthesia, consciousness fades and might cease altogether. Consider split-brain patients, whose corpus callosum—the 200 million wires linking the two cortical hemispheres—has been cut to alleviate severe epileptic seizures. The surgery literally splits the person's consciousness in two, with one conscious mind associated with the left hemisphere and seeing the right half of the visual field and the other mind arising from the right hemisphere and seeing the left half of the visual field. **To be conscious, then, you need to be a single, integrated entity with a large repertoire of highly differentiated states**. Although the 60-gigabyte hard disk on my MacBook exceeds in capacity my lifetime of memories, that information is not integrated. For example, the family photographs on my Macintosh are not linked to one another. The computer does not know that the girl in those pictures is my daughter as she matures from a toddler to a lanky teenager and then a graceful adult. To my Mac, all information is equally meaningless, just a vast, random tapestry of zeros and ones. Yet I derive meaning from these images because my memories are heavily cross-linked. And the more interconnected, the more meaningful they become. Indeed, **Tononi's IIT postulates that the amount of integrated information that an entity possesses corresponds to its level of consciousness.**

These ideas can be precisely expressed in the language of mathematics using notions from information theory such as entropy [see box on next page]. **Given a particular brain, with its neurons and axons, dendrites and synapses, one can, in principle, accurately compute the extent to which this brain is integrated**.

From this calculation, the theory derives a single number, Φ (pronounced "fi"). Measured in bits, Φ denotes the size of the conscious repertoire associated with any network of causally interacting parts. Think of Φ as the synergy of the system. The more integrated the system is, the more synergy it has, the more conscious it is. If individual brain regions are too isolated from one another or are interconnected at random, Φ will be low. If the organism has many neurons and is richly endowed with specific connections, Φ will be high—capturing the quantity of consciousness but not the quality of any one conscious experience. (That value is generated by the informational geometry that is associated with Φ but won't be discussed here.)

Explaining Brain Facts

The theory can account for a number of puzzling observations. The **cerebellum**, the "little brain" at the back of the brain that contains more neurons than the convoluted cerebral cortex that crowns the organ, has a regular, crystallinelike wiring arrangement. Thus, its circuit complexity as measured by Φ is low as compared with that of the cerebral cortex. Indeed, if you lose your cerebellum you will never be a rock climber, pianist or ballet dancer, but your consciousness will not be impaired. The **cortex** and its gateway, the thalamus—the quail egg–shaped structure in the center of the brain—on the other hand, are essential for consciousness, providing it with its elaborate content. **Its circuitry conjoins functional specialization with functional integration thanks to extensive reciprocal connections linking distinct cortical regions and the cortex with the thalamus**. This **corticothalamic complex** is well suited to behave as a single dynamic entity endowed with a large number of discriminable states. Lose one chunk of a particular cortical area, and you might be unable to perceive motion. If a different area were lesioned, you would be blind to faces (yet could see the eyes, hair, mouth and ears). When people are woken from deep sleep, they typically recall experiencing nothing or, at best, only some vague bodily feeling; this experience contrasts with the highly emotional narratives our brains weave during rapid-eye-movement (REM) sleep. What is paradoxical is that the average firing activity of individual nerve cells does not differ that much in deep sleep and quiet wakefulness. At the whole system level, though, electroencephalographic electrodes on the skull pick up slow, large and highly synchronized waves during deep sleep. Because these waves are quite regular, they will disrupt the transfer of specific information among brain cells.

Every day, in tens of thousands of surgical operations, patients' consciousness is quickly, safely and transiently turned off and on again with the help of various anesthetic agents. There is no single mechanism common to all. The most consistent regional finding is that anesthetics reduce thalamic activity and deactivate mesial (middle) and parietal cortical regions. Twenty years of electrical recording in anesthetized laboratory animals provided ample evidence that many cortical cells, particularly in primary sensory cortical regions, continue to respond selectively during anesthesia. What appears to be disrupted is **large-scale functional integration in the corticothalamic complex**.

IIT explains why consciousness requires neither sensory input nor behavioral output, as happens every night during REM sleep, in which a central paralysis prevents the sleeper from acting out her dreams. **All that matters for consciousness is the functional relation among the nerve cells that make up the corticothalamic complex**. Within this integrated dynamic entity can be found the dream of the lotus eater, the mindfulness of the meditating monk, the agony of the cancer patient and the Arcadian visions of your lost childhood home. Paraphrasing Oscar Wilde, I would say it is the causal interactions within the dynamic core that make the poppy red, the apple odorous and the skylark sing.

Consciousness Is Universal

One unavoidable consequence of IIT is that all systems that are sufficiently integrated and differentiated will have some minimal consciousness associated with them: not only our beloved dogs and cats but also mice, squid, bees and worms.

Indeed, the theory is blind to synapses and to all-or-none pulses of nervous systems. At least in principle, the incredibly complex molecular interactions within a single cell have nonzero Φ. In the limit, a single hydrogen ion, a proton made up of three quarks, will have a tiny amount of synergy, of Φ. **In this sense, IIT is a scientific version of panpsychism**, the ancient and widespread belief that all matter, all things, animate or not, are conscious to some extent. Of course, IIT does not downplay the vast gulf that separates the Φ of the common roundworm *Caenorhabditis elegans* with its 302 nerve cells and the Φ associated with the 20 billion cortical neurons in a human brain.

The theory does not discriminate between squishy brains inside skulls and silicon circuits encased in titanium. Provided that the causal relations among the transistors and memory elements are complex enough, computers or the billions of personal computers on the Internet will have nonzero Φ. The size of Φ could even end up being a yardstick for the intelligence of a machine.

Future Challenges
IIT is in its infancy and lacks the graces of a fully developed theory. A major question that it so far leaves unanswered is, Why should natural selection evolve creatures with high Φ? What benefit for the survival of the organism flows from consciousness? One answer that I hope for is that intelligence, the ability to assess situations never previously encountered and to rapidly come to an appropriate response, requires integrated information. Another possible answer, though, could be that high-Φ circuits do not have any special status in terms of their survival. Just as electrical charge is a fundamental feature of the universe without a function, consciousness might also lack any specific evolutionary role. It just is.

A second stumbling block with IIT is that Φ is exceedingly difficult to compute even for very small systems. To accurately evaluate Φ for the roundworm is utterly unfeasible, even if using all of Google's more than 100,000 computers. Can we find other algorithms to more easily compute Φ?

A third issue to understand is why so much brain processing and so many of our daily behaviors are unconscious. Do the neural networks that mediate these unconscious, zombielike behaviors have lower Φ than the ones that give rise to consciousness?

Tononi's integrated information theory of consciousness could be completely wrong. But it challenges us to think deeply about the mind-body problem in a novel, rigorous, and mathematically and empirically minded manner. And that is a great boon to this endeavor. If Tononi's equation for Φ proves to plumb the hitherto ineffable—consciousness itself—it would validate the ancient Pythagorean belief that "number is the ruler of forms and ideas and the cause of gods and demons."

---

http://en.wikipedia.org/wiki/Integrated_Information_Theory

The **Integrated Information Theory** is a recently formulated theory which attempts to quantitatively measure consciousness. It was developed by psychiatrist and neuroscientist Giulio Tononi of the University of Wisconsin–Madison.[1]

The theory is based on two key observations. The first is that every observable conscious state contains a massive amount of information. A common example of this is every frame in a movie. Upon seeing a single frame of a movie you have watched you instantly associate it with a "specific conscious percept."[2] That is to say you can discriminate a single frame from a film with any other single frame, including a blank, black screen. **The mind, therefore, can discriminate amongst a massive number of possible visual states**. This is a tremendous amount of information being represented. Compare our visual awareness to a simple photodiode which only can discriminate the presence of light from dark. It doesn't matter if the light is a lightbulb, a scene from *Ben Hur* or the bright light of noon on a summer day, the photodiode represents only minimal information. **The hypothesis then is that the amount of consciousness an entity has is equal to the amount of information processing it contains. This brings us to the second key observation of the theory**.

All of the information you have gleaned from conscious states is highly, and innately, integrated into your mind. **It is impossible for you to see the world apart from all of the information that you are conscious of**. When you are looking at an orange, for example, you cannot separate the color of the fruit (orange) from its shape (round). **Consciousness is "integrated"**; even though color processing and spatial processing are separately localized in the brain (a stroke victim can lose color perception yet maintain perfect spatial awareness, for example) **conscious experiences cannot be atomized into distinct parts**.

Giulio Tononi's initial ideas were further developed by Adam Barrett, who created similar measures of integrated information [3] such as "phi empirical".

Definition of Consciousness
In this theory, consciousness arises as a property of a physical system, its 'integrated information'. Integrated information is an exact quantity that can be measured using the following equations:

Information

Given: a system (including current probability distribution) and Mechanism (which specifies the possible next state probability distribution, if the current state is perturbed with all possible inputs).
You can determine: Actual Distribution - Possible system states at time t = -1 Thus: System and Mechanism constitute information (about the system's previous state), in the classic sense of 'reduction of uncertainty.'

Relative Entropy/Effective Information

Effective Information = relative entropy H between the actual and potential repertoires
= Kullback-Leibler divergence

It is implicitly specified by mechanism and state, so it is an 'intrinsic' property of the system. One can calculate the actual repertoire of states by perturbing the system in all possible ways to obtain the forward repertoire of output states. After that, one applies Bayes' Rule.

Example

System of two Binary elements - Four possible states (00, 01, 10, 11)

The first binary element operates randomly. The second binary element will be whatever the first element was in the previous state. Initially: (0, 0). maximum entropy: p = (1/4, 1/4, 1/4, 1/4) Given, at time t, state is 11 Previous state must have been 11 or 10, p = (0, 0, 1/2, 1/2) Generated one bit of information

since $ei(X(mech, x_1)) = H[p(X_0(mech, x_1)) \parallel p(X_0(maxH))]$ where X is our system, mech is that system's mechanism, x1 is a state of the system, and p(X0(maxH)) is the uniform or potential distribution.

Integration (Φ)

$\Phi(X(mech, x_1)) = H[p(X_0(mech, x_1)) \parallel \Pi p(^kM_0(mech, \mu_1))]$ for $^kM_0 \in MIP$

where X is our system, mech is that system's mechanism, x_1 is a state of the system, $\Pi(p(^kM_0(mech, \mu_1)))$ is the product of all the probability distributions of each part of the system in the minimal information partition.

It's clear then that Φ will be high when there is a lot of information generated among the parts of a system as opposed to within them.

Complexes

A complex is a set of elements that generate integrated information that is not fully contained in a larger set of higher Φ.

This then leads naturally to the notion of a main complex, which is the complex in a system that generates the largest amount of Φ. Note that a main complex can partially contain complexes of lower Φ within it.

Interpretations of different aspects of consciousness

Quality of consciousness

We begin by defining a multi-dimensional space called qualia space, or Q-space. This space has an axis for every state of the system. A point in this space, then, has a component for every state; if we restrict the components to be numbers from 0 to 1, then we can view the components as probabilities that the system is in that state. Thus a point in Q-space represents a probability distribution. Now again using relative entropy we can measure the amount of information generated by a single connection c within the system with the following equation:

$$\Phi_c = H[p(X(mech, x)) \parallel p(Y(mech, y))]$$

where Y is the system with that connection removed. Thus there are points Y and X in Q-space that correspond to the probability distributions of the system with and without the connection c, respectively. We can then draw a vector from Y to X that has length Φ_c. This vector is associated with the connection c and is called a q-arrow. So a q-arrow is a representation of the informational relationship specified by a connection.

Properties of q-arrows

Context dependency

Q-folds

Entanglement

References

---

1. ^ Tononi, Giulio (2008-12). "Consciousness as integrated information: a provisional manifesto". *The Biological Bulletin* **215** (3): 216–242. ISSN 0006-3185. PMID 19098144.
2. ^ Koch, Christof. "A "Complex" Theory of Consciousness". Scientific American. Retrieved 2012-04-18.
3. ^ Barrett, A.B., & Seth, A.K. (2011). Practical measures of integrated information for time-series data. PLoS Comput. Biol., 7(1): e1001052

External links

- Integrated Information Theory: A Provisional Manifesto
- Sizing Up Consciousness by Its Bits
- A Bit of Theory: Consciousness as Integrated Information Theory
- Scientific American Article
- Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework

http://www.biolbull.org/content/215/3/216.full

Consciousness as Integrated Information: a Provisional Manifesto
Giulio Tononi⇓
Department of Psychiatry, University of Wisconsin, Madison, Wisconsin
To whom correspondence should be addressed. E-mail: gtononi@wisc.edu

Abstract
The **integrated information theory (IIT)** starts from phenomenology and makes use of thought experiments to **claim that consciousness is integrated information**. Specifically:
(i) the quantity of consciousness corresponds to the amount of integrated information generated by a complex of elements;
(ii) **the quality of experience is specified by the set of informational relationships generated within that complex**.

Integrated information (Φ) is defined as the amount of information generated by a complex of elements, above and beyond the information generated by its parts.
Qualia space (Q) is a space where each axis represents a possible state of the complex, each point is a probability distribution of its states, and arrows between points represent the informational relationships among its elements generated by causal mechanisms (connections). Together, the set of informational relationships within a complex constitute a shape in Q that completely and univocally specifies a particular experience.
Several observations concerning the neural substrate of consciousness fall naturally into place within the IIT framework. Among them are the association of consciousness with certain neural systems rather than with others; the fact that neural processes underlying consciousness can influence or be influenced by neural processes that remain unconscious; the reduction of consciousness during dreamless sleep and generalized seizures; and the distinct role of different cortical architectures in affecting the quality of experience. Equating consciousness with integrated information carries several implications for our view of nature.

- **Φ, integrated information**

- **IIT, integrated information theory**
- **MIP, minimum information partition**