Information Integration, Data Integration and Machine Consciousness

David Gamez[1]

Abstract. Information integration is a property of systems of connected elements that expresses the extent to which they are capable of entering a large number of states that result from causal interactions among their elements. In recent years a number of people have claimed that there is a link between information integration and consciousness and a number of algorithms for measuring information integration have been put forward. This paper gives an overview of the conceptual and experimental issues surrounding information integration and explores some of the links between information integration and machine consciousness.

1 INTRODUCTION

Neuroscience research often uses information measures, such as mutual information or transfer entropy, to identify the anatomical, functional and effective connections between different parts of the brain or a neural simulation [1, 2]. Many authors have observed that integration is a key feature of conscious states [3, 4], and it has been a natural progression to suggest that measures of functional and effective connectivity could be used to identify the parts of the brain that are highly integrated and thus correlated with conscious states. Tononi [5, 6] has gone beyond this correlation approach to claim that consciousness actually *is* integrated information, and proposed algorithms for identifying the areas of maximum information integration in a system [7, 8].

One of the key attractions of information integration theories of consciousness is that they are precise enough to be experimentally tested. For example, Tononi's most recent theory [5] can predict the areas of a system that are associated with consciousness, the amount of consciousness that is present and the qualitative character of this consciousness for each state of a system. This precision of information integration theories points the way towards a more scientific approach to consciousness, in which falsifiable predictions made by different mathematically formulated theories of consciousness can be systematically compared (see Section 4). A second advantage of information integration theories is that they can be applied to both artificial and natural systems: if a link could be established between information integration and consciousness in humans, then it would be possible to make convincing predictions about the consciousness of artificial systems as well.

While information integration theories are promising, there are many issues that need to be addressed. Some conceptual difficulties surrounding the nature of information and integration are covered in Section 3, and there are a number of practical problems with measuring information integration, including the performance of current algorithms, how accuracy can be evaluated, and the selection of a particular level of the system for analysis. These practical issues are examined in Section 4, which gives an overview of how information integration theories could be experimentally tested. Some of the potential applications of information integration algorithms are covered in Section 5.

2 BACKGROUND

Research on information integration and consciousness is closely linked to work on the identification of functional and effective relationships between neurons and neuron groups using neural complexity [9, 10], transfer entropy [11] and other measures. There has been some research comparing neural complexity measures and graph theory [12], and these measures have been used by a number of people to examine the anatomical, functional and effective connectivity of biological networks, either using scanning or electrode data, or large-scale models of the brain. One example of this type of work is Honey et al. [13], who used transfer entropy to study the relationship between anatomical and functional connections on a large-scale model of the macaque cortex, and demonstrated that the functional and anatomical connectivity of their model coincided on long time scales. Another example is Brovelli et al. [14], who used Granger causality to identify the functional relationships between recordings made from different sites in two monkeys as they pressed a hand lever during the wait discrimination task. Information-based analyses have also been used to guide and study the evolution of artificial neural networks connected to simulated robots [15, 16]. An overview of this type of research can be found in [1, 2].

A number of people have suggested that there is a link between information integration and consciousness [3, 4], or that information integration actually *is* consciousness [5, 6], and several algorithms for calculating information integration have been put forward. These include neural complexity [10], stateless Φ [8], state-based Φ [7],[3] causal density [17], liveliness [18], and an information integration measure that can be applied to time series data [19]. Seth et al. [20] gives a review of earlier work, identifies a number of weaknesses in Tononi and Sporns' [8] method and criticizes the link between information integration and consciousness.

There has been a limited amount of experimental work on the link between information integration and consciousness. For example, Lee et al. [21], made multi-channel EEG recordings from eight sites in conscious and unconscious subjects and constructed a covariance matrix of the recordings on each frequency band that was used to identify the complexes within

[1] Dept. of Computing, Imperial College, London SW7 2BT, UK. Email: dgamez@imperial.ac.uk.

[3] The information integration measure put forward by Tononi and Sporns [8] will be referred to as "stateless Φ" to distinguish it from the related state-based measure of Φ of Balduzzi and Tononi [7].

the 8 node network using Tononi and Sporns' [8] algorithm. This experiment found that the information integration capacity of the network in the gamma band was significantly higher when subjects were conscious. Massimini et al. [22, 23] have carried out experiments in which a TMS pulse was applied to the subject's brain and the resulting activity was recorded using EEG. Massimini et al. found that the activity resulting from the TMS pulse was more localized and less differentiated when the subjects were unconscious, suggesting that a combination of integration and differentiation is linked to conscious states. In the simulation work on information integration detailed predictions have been made about the amount and distribution of consciousness in an 18,000 neuron network [24], and some preliminary comparisons have been carried out between state-based Φ and liveliness [18].

3 INFORMATION OR DATA INTEGRATION?

3.1 The Nature of Information

> Information is notorious for coming in many forms and having many meanings. It can be associated with several explanations, depending on the perspective adopted and the requirements and desiderata one has in mind.
>
> Floridi [25], p.1

In the work using information theory to measure brain connectivity, the exact nature of information is not important because Shannon's information equations are used as mathematical tools to identify the functional and effective connections between groups of neurons. However, the nature of information does become important when a property of the system called information integration is linked to consciousness. In this shift it is no longer the *relationships* between biological neurons that are important, but the *presence* of information integration in the system. It then becomes necessary to say what information is and how it can be identified in an arbitrary system.

To understand the problem of identifying information, imagine that an aubergine is lying on the table in front of you. Simply through its existence on the table, this aubergine contains the information that there is an aubergine on the table in front of you. Cut the aubergine open and a pattern of seeds are revealed, which can be interpreted as letters in a particular language. The genetic code of the aubergine can also be read and the aubergine contains an enormous amount of information about the location of each of its atoms relative to a particular reference point. Each of these types of information at different levels of the aubergine can be transformed into other types of information. For example, the sequence of nucleotides in the aubergine's DNA could be remapped into a sequence of numbers representing an image or sound.

This example of the aubergine suggests that physical objects can be interpreted as containing a virtually infinite quantity of information, which is relative to an observer and level of abstraction that define the information states in a system. A theory of information that can handle this apparent relativity is the General Definition of Information (GDI). According to Floridi's [26] formulation of the GDI, σ is an instance of information, understood as semantic content, if and only if:

GDI.1) σ consists of n data, for $n >= 1$;
GDI.2) the data are well-formed;
GDI.3) the well formed data are meaningful.

This GDI is based on the notion of *data*, which Floridi describes as a lack of uniformity in the world. The data that is accessible to us and can be read depends on pre-theoretical differences in the physical world, called *dedomena*, which cannot be experienced without an interpretation that is applied to the world. These *dedomena* are the conditions of possibility for experienced data - something like Kant's *noumena* or Locke's *substance* - that make the differences that we can measure and manipulate possible. For example, *dedomena* might make the measureable difference between higher and lower charge in a battery possible, and this type of measureable difference between physical states can in turn be used to create higher levels of data, such as symbols. The interface that defines the scope and type of data in a system is called a level of abstraction by Floridi.

From the point of view of the information integration theory of consciousness, this distinction between *dedomena* and data that we understand and manipulate is important because only *dedomena* can be considered to be an objective property of the system. The data or sets of differences that we actually extract will always be the result of a particular interpretation. This is not a problem if we are interested in correlations between information integration and consciousness, but it is an issue for Tononi's claim that information integration *is* consciousness because consciousness is typically thought to be an objective feature of the world, not a subjective interpretation of a system by an observer. One potential way around this problem would be to say that consciousness is the integration between differences in a physical attribute of the system, such as its electric field, which is thought to be more than just a subjective interpretation.

Once a method for identifying data in a system has been defined, the next stage is to specify a syntax that will enable *well-formed* data to be extracted in a systematic way. For example, the sequences of nucleotides that constitute the aubergine's genetic code can only be read when we can distinguish between sequences coding proteins and junk DNA. If we want to interpret the aubergine's seeds as Arabic writing, the seed pattern will have to conform to the shapes of the letters in Arabic, the order of the letters will have to conform to the orthography of Arabic and the order of the words will have to conform to the grammar of Arabic.

While it is relatively easy to extract well-formed data from a system, this data might may not be *meaningful* in any way. Suppose that the aubergine is in half, covered with a grid of millimetre squares and a 1 is read off if the square contains an even number of seeds and a 0 is read off if the square contains an odd number of seeds. This sequence of 1s and 0s is well formed data because it conforms to a specified syntax, but since it lacks meaning, it is not information according to the GDI. The question of what makes data meaningful is much more difficult than the identification of data in a system, and Floridi's approach is to describe semantic content as a combination of data and queries. So, for example, the proposition "The earth only has one moon" can be interpreted as a piece of meaningful data in which the semantic content is the question "Does the Earth only have one moon?" and the answer "yes" is a single bit of data.

Starting with the work of Shannon [27], there has been an extensive amount of work on the communication of information, describing the entropy of an information source, the mutual information between two devices and the maximum rate of communication over a channel. However, as Floridi points out, Shannon's mathematical theory of communication (MTC) is a theory about data transmission, not about information transmission, because it does not take the meaning of the messages into account: "since MTC is a theory of information without meaning (not in the sense of meaningless, but in the sense of not yet meaningful), and since we have seen that [information – meaning = data], 'mathematical theory of data communication' is a far more appropriate description of this branch of probability theory than 'information theory'."([26], p. 33). This suggests that "information integration" algorithms based on Shannon's work, such as [8], are actually measures of *data* integration, unless it can be shown that the integrated data carries semantic content. This distinction between data and information poses a separate question about whether there is a link between information integration and consciousness, where information is understood as meaningful data. This is particularly relevant when considering embodied theories of consciousness, since meaningful data could be data that co-varies with the world.

3.2 Integration of Information

There are many ways of interpreting the notion of integrated information, including data fusion, meta data about information and statistical and causal relationships between items of information. The type of information integration that is claimed to be linked to consciousness has a very specific meaning because it is not just the *integration* that is important, but the *differentiation* of the information states as well. Tononi [5, 6] illustrates this idea of differentiated integration using the example of a digital camera sensor with a million photodiodes. This sensor is highly differentiated because it can enter $2^{1,000,000}$ different states but in each of these states the photodiodes are acting independently and there is no integration between them. In contrast, consider a million Christmas lights connected to a single switch: when the switch is on, the lights are on; when the switch is off, the lights are off. In this system there is a high level of integration between the switch and the lights, but almost no differentiation because the system can only enter two possible states: all lights on or all lights off. In between the camera photodiode and the Christmas lights are systems that are both differentiated and integrated: they can enter a large number of different states and these states are the result of causal interactions between the elements. According to Tononi [5, 6], a key example of differentiated and integrated systems are the areas associated with consciousness in the human brain.

The algorithms that have been put forward for measuring information integration - for example [7, 8] - are intended to quantify the balance between differentiation and integration in a system of connected elements. While algorithms for measuring the integration between items of information are relatively straightforward – for example, statistical or causal measures – and the differentiation of a system can be quantified using information entropy, it is a challenging task to find an algorithm that can quantify the combination of differentiation and integration. Some of the performance and accuracy issues raised by the current algorithms are covered in Section 4.3.

4 TESTING INFORMATION INTEGRATION

4.1 Introduction

Information integration is an empirical theory about a link between a measured feature of the physical world and phenomenal experience. Its great strength is that it makes strong claims about the world that can be shown to be false. Many other theories of consciousness, such as higher order thought [28], might be thought to be intuitively plausible, but they are not scientific if they cannot be experimentally tested.

Experiments on the link between information integration and consciousness are likely to involve the following steps:

1. Select a system that is known to be conscious or commonly agreed to be conscious.
2. Measure the information integration of the system.
3. Measure the consciousness of the system.
4. Identify correlations between information integration and consciousness.
5. Test predictions made by information integration about the consciousness of the system.

The following sections cover each of these stages in more detail. Although this discussion is framed in terms of the information integration theory of consciousness, a similar approach could be applied to any theory of consciousness that is expressed in a precise mathematical or algorithmic form.

4.2 The Platinum Standard System

To establish whether information integration in a physical system is linked to conscious states it is necessary to start with a physical system that is known or commonly agreed to be associated with consciousness. Although we typically assume that infants and higher mammals are conscious, the only system that is confidently associated with consciousness is the awake normal adult human brain. By 'normal' it is meant that the brain is undamaged and its functions and measurements fall within two standard deviations for the human species. 'Awake' is intended in a non-technical sense to indicate that the brain is functioning in a way that is typically considered 'conscious'. This type of wakefulness is distinct from the medical definition, since apparently wakeful states can be exhibited by people in a vegetative state who are unlikely to be conscious [29]. While there will be times when the awake normal adult human brain is not conscious – for example, epileptic automatism [30] - a science of consciousness has to start somewhere, and the awake normal adult human brain is the physical system that we are most certain is typically associated with conscious states.

The awake normal adult human brain will be referred to as the platinum standard system. Just as a platinum-iridium standard bar in Paris was used to define the length of a metre, the awake normal adult human brain is our platinum standard for a conscious system. If this physical system is not associated with conscious states most of the time, then nothing is. A further assumption, that the consciousness associated with different platinum standard systems is roughly the same, may become necessary for detailed predictions about the contents of consciousness.

It is important to note that artificial systems cannot be used to test the link between information integration and consciousness because it is not clear whether they are associated with conscious states. If a link between information integration and consciousness could be established, then it would become possible to make predictions about the consciousness of artificial systems using information integration, but this link has to be demonstrated on a platinum standard system first.

4.3 Measuring Information Integration in the Physical Platinum Standard System

To investigate the link between information integration and consciousness it is necessary to measure the amount of information integration in the platinum standard system. The first stage is the definition of the level of abstraction that will be used in the experiments. Data in the brain can be defined at many different levels – for example, sub-atomic, atomic, molecular, neural or neuron group - and it is far from clear whether different levels of abstraction will lead to different amounts of information integration in the system, or whether the levels will coincide. As an example, consider the problem of measuring colour in a sack of oranges. If colour is measured at the level of individual oranges, then the sack of oranges will be pronounced orange. Likewise, an analysis at the level of segments will result in an orange colour, but analyses at the sub-atomic level or at the level of pips will result in zero orange colour. Within information integration analyses, the key experimental challenge is to identify whether the levels coincide or contradict – for example, whether information integration analyses at the level of ions match analyses based on areas of the brain. There is also a challenging question about whether the interpretation of the neural code will affect the amount of information integration – for example, rate-based analyses could give very different results from analyses based on polychromous groups [31].

The presence of multiple information integration algorithms that apparently measure the same objective property of the physical world raises the question about which is the most *accurate* algorithm and how this accuracy can be measured. The accuracy of information integration algorithms could be evaluated by making the (problematic) assumption that information integration is correlated with consciousness, and carrying out experiments – for example, using fMRI or EEG – that measure the correlation between the output of the information integration algorithms and the reports of conscious states from the platinum standard system. If information integration is correlated with consciousness, then the algorithms that most accurately predict consciousness would be the most accurate measures of information integration. The main problem with this approach is that some or all of the algorithms might be measuring a property of the brain that is correlated with consciousness, but which has nothing to do with information integration. There is also the issue that our spatial and temporal access to the brain is severely limited, which makes it very difficult to measure information integration in humans.

Another way of measuring the accuracy of information integration algorithms is to create simulated networks with regions that we expect to have high information integration – for example, a neural network with several highly intra-connected modules would be expected to have higher information integration within the modules. Different information integration algorithms could be run on these networks and their output compared with the areas of expected maximum information integration. This approach has the problem that our intuitions about the areas of maximum information integration might not be correct, but there does not appear to be a way of measuring the information integration of a network that does not depend on a particular algorithm. While the simulated networks approach is problematic, until our access to the brain improves it appears to be the only method available for the comparison of different measures of information integration.

A second problem with measuring information integration is the performance of some of the current algorithms - for example, it has been predicted to take 10^{9000} years to fully analyze an 18,000 neuron network using Tononi and Sporns' algorithm [32] and it could take 100 million years to analyze a network of 30 elements using Balduzzi and Tononi's algorithm [18]. Some work on addressing this issue has been carried out by Aleksander and Gamez [18], who developed an algorithm for measuring information integration based on liveliness that scales linearly with the number of neurons and connections. However, even with these improvements, current supercomputers are likely to struggle with a full analysis of the platinum standard system, which has around 10^{10} neurons and 10^{15} synapses. Although artificial systems cannot be used as platinum standard systems, they are an ideal test environment for benchmarking the performance and accuracy of different ways of measuring information integration.

A final problem with some of the current algorithms, such as state-based Φ and liveliness, is that they rely on knowledge about the underlying causal structure of the system. This is not a problem with artificial systems, where the causal structure is usually known, but typical measurements of the platinum standard system, such as fMRI, EEG or electrode data, are a sequence of states whose causal relationship is unknown. This is not an issue for Seth's Granger causality measure [33], and it could be partially addressed for the other measures by inferring the causal structure from the sequence of states using Granger causality, transfer entropy or mutual information.

4.4 Measuring Consciousness in the Platinum Standard System

In experiments on the link between information integration and consciousness, the information integration of the physical system is compared with measurements of conscious states. In the platinum standard system, consciousness is typically measured through first person reports, although other behaviours can be used to infer the presence of consciousness and its contents. It is also possible to use cognitive abilities that are systematically linked to consciousness to reliably infer the presence of consciousness in a platinum standard system [34]. There are numerous problems with the measurement of consciousness, such as the dependence on potentially fallible memory and the limited bandwidth and accuracy of human language – see [34] for a more detailed discussion.

4.5 Correlations between Information Integration and Consciousness in the Platinum Standard System

While falsifiable predictions are the gold standard for scientific theories (see Section 4.6), initial work on the link between information integration and consciousness is likely to focus on the identification of correlations between information integration and consciousness. Early experiments are likely to study whether the presence of consciousness is correlated with higher information integration; eventually research will move on to examine whether different degrees of consciousness are linked to different amounts of information integration and whether the contents of consciousness vary in the way suggested by information integration theories.

4.6 Predictions about Consciousness in the Platinum Standard System

> ...the real test of a scientific theory of consciousness is its ability to make falsifiable predictions: I shall certainly admit a system as empirical or scientific only if it is capable of being tested by experience. These considerations suggest that not the verifiability but the falsifiability of a system is to be taken as a criterion of demarcation ... I shall require that its logical form shall be such that it can be singled out, by means of empirical tests, in a negative sense: it must be possible for an empirical system to be refuted by experience.
>
> Popper [35], p.18

A substantial amount of the current work on consciousness is based on theories that are felt to be more or less intuitively plausible. To become truly scientific, the study of consciousness has to move towards a situation in which predictions made about the consciousness of a platinum standard system are compared with the platinum standard system's behavioural reports about its consciousness.

The information integration theory that is most capable of making predictions [5] can predict the areas of a physical system that are associated with consciousness, the amount of consciousness in these areas and the qualitative character of this consciousness. When our ability to measure the human brain has increased its spatial and temporal resolution (and if the performance of current algorithms can be improved), it should become possible to make predictions about the consciousness of a platinum standard system and compare these predictions with first person reports. The information integration theory of consciousness would become widely accepted if it could make large numbers of accurate predictions about the contents of consciousness of human subjects using only physical information about the system.

4.7 The Path Ahead

The problems identified in this paper currently make it impractical to systematically test information integration algorithms on platinum standard systems. Instead, initial work in this area is likely to use artificial neural networks, possibly embodied in robots, to develop more efficient algorithms and investigate novel ways of analyzing networks for information integration. This work on artificial systems will feed into experimental work using scanning data with low spatial and/or temporal resolution, such as the experiments discussed in Section 2. During this research, the predictions that are made about the consciousness of artificial systems will *not* be believed because a link will not have been established between information integration and consciousness on the platinum standard system. Eventually it is hoped that improvements in the speed of computers, increased spatial and temporal accuracy of brain scanning, and greater efficiency of information integration algorithms will make it possible to establish whether there are systematic correlations between information integration and consciousness in the platinum standard system.

If empirical evidence could establish that information integration is systematically linked with consciousness in the platinum standard system, then information integration could be used to make predictions about the consciousness of other systems. For example, we could use information integration to make believable predictions about the consciousness of artificial systems, infants and animals. We could also look back retrospectively at the systems that were analyzed for information integration in the past, and believe that these systems were conscious to the degree predicted because the information integration theory will have been rigorously proved on the platinum standard system.

It is possible that empirical research will demonstrate that information integration is *not* correlated with consciousness in the platinum standard system, or that it fails to make accurate predictions about consciousness. In this case, information integration theories should be abandoned and better approaches sought.

5 APPLICATIONS

Accurate predictions about consciousness in humans have many applications, such as measuring the degree of consciousness in coma patients, identifying whether a person is unconscious during an operation, and it might be possible for paraplegic patients to use the predicted contents of their consciousness to control artificial limbs. Accurate predictions about consciousness would also raise serious privacy issues - for example military and police interrogators could read a suspect's mind, and people's intentions are often used to identify their degree of criminal guilt: the difference between murder and manslaughter is largely a matter of intention. Predictions about animal consciousness could be used to minimize animal suffering – it might even become possible to genetically engineer food animals with little or no consciousness. Predictions about the consciousness of early embryos would have applications in abortion legislation.

Within work on machine consciousness, an accurate algorithmic measure of consciousness could be used to measure the degree to which an artificially conscious system has been constructed. It could also determine whether artificially conscious systems are suffering – one of the objections to machine consciousness raised by Metzinger [4] is that it amounts to the creation of a race of retarded infants for experimentation. This worry about the suffering and confusion of artificial systems is becoming more pressing because scanning technologies are developing to the point at which it may soon become possible to get exact data about the location and connections of every neuron in a mouse brain [36, 37]. This would enable the real-time simulation of a particular mouse's brain, which might be capable of experiencing the same pain as the original mouse.

There has been a substantial amount of discussion of the possibility that people could scan their brains into a computer and achieve a form of digital immortality [38, 39]. After a person dies their brain would be preserved and a succession of very thin slices would be scanned and integrated together to build up a complete picture of their neurons and connections.[4] This information would then used to build a neural network with the same neurons and connections, and in theory this could be accurate enough to produce the same global behaviour. Since the person's brain will have developed in a close relationship with their body, it might be necessary to connect the simulated brain to their original body, perhaps using electrodes attached to nerves in the spinal column.

People who pay for this procedure might only be interested in perpetuating the *external behaviour* of their brain after their death, which might be possible if the neural simulation works in real time, is accurate enough, and is connected to a suitable body. However, a key question will remain as to whether this simulation of a person's brain will be as conscious as the person was before their death, or whether the simulation will be a zombie that just replicates the person's external behaviour. If information integration or a similar theory could be shown to make accurate predictions about consciousness, then it could be used to predict the extent to which a simulation of a person's brain is as conscious as the brain was before the person's death. It is possible that some simulations of a person's brain will not be conscious at all - perhaps because they are running on a time-sliced computer - so predictions about the consciousness of digitized brains might make people think very carefully about the way in which they want to be 'uploaded' after their death.

6 CONCLUSIONS

This paper has given an overview of some of the conceptual and experimental issues surrounding work on the possible link between information integration and consciousness. Information integration theories of consciousness are interesting because they open up the possibility of making predictions about consciousness, which could open up a new chapter in the scientific study of natural and artificial minds.

While artificial systems cannot be used to test the link between information integration and consciousness, they can play a key role in improving and understanding the current algorithms and addressing questions about their performance and accuracy. If a link between information integration and consciousness could be proved, then it would become possible to use information integration to make believable predictions about the consciousness of animals, artificial systems and simulated scanned brains. Although information integration is a promising approach, we are only just beginning to explore mathematical and algorithmic theories of consciousness, and experimental work may show that information integration is poorly correlated with consciousness. In this nascent field many different approaches may have to be tried before we discover a high performance high accuracy scientific theory of consciousness.

[4] It might eventually become possible to make a detailed scan of a person's brain before their death, but this is well beyond the reach of current technology.

ACKNOWLEDGEMENTS

This work was supported by a grant from the EPSRC (EP/F033516/1).

REFERENCES

[1] O. Sporns, D. R. Chialvo, M. Kaiser, and C. C. Hilgetag. Organization, Development and Function of Complex Brain Networks. *Trends Cogn Sci,* 8: 418-25 (2004).
[2] O. Sporns. (2011). *Scholarpedia Article on Brain Connectivity*. Available: http://www.scholarpedia.org/article/Brain_connectivity
[3] B. J. Baars, *A Cognitive Theory of Consciousness*. Cambridge University Press, Cambridge England ; New York (1988).
[4] T. Metzinger, *Being No One : The Self-Model Theory of Subjectivity*. MIT Press, Cambridge, Mass. (2003).
[5] G. Tononi. Consciousness as Integrated Information: A Provisional Manifesto. *Biol Bull,* 215: 216-42 (2008).
[6] G. Tononi. An Information Integration Theory of Consciousness. *BMC Neurosci,* 5: 42 (2004).
[7] D. Balduzzi and G. Tononi. Integrated Information in Discrete Dynamical Systems: Motivation and Theoretical Framework. *PLoS Comput Biol,* 4: e1000091 (2008).
[8] G. Tononi and O. Sporns. Measuring Information Integration. *BMC Neurosci,* 4: 31 (2003).
[9] G. Tononi, G. M. Edelman, and O. Sporns. Complexity and Coherency: Integrating Information in the Brain. *Trends Cogn Sci,* 2: 474-84 (1998).
[10] G. Tononi, O. Sporns, and G. M. Edelman. A Measure for Brain Complexity: Relating Functional Segregation and Integration in the Nervous System. *Proc Natl Acad Sci U S A,* 91: 5033-7 (1994).
[11] T. Schreiber. Measuring Information Transfer. *Physical Review Letters,* 85: 461-464 (2000).
[12] M. Shanahan. Dynamical Complexity in Small-World Networks of Spiking Neurons. *Physical Review E,* 78, (2008).
[13] C. J. Honey, R. Kotter, M. Breakspear, and O. Sporns. Network Structure of Cerebral Cortex Shapes Functional Connectivity on Multiple Time Scales. *Proc Natl Acad Sci U S A,* 104: 10240-5 (2007).
[14] A. Brovelli, M. Ding, A. Ledberg, Y. Chen, R. Nakamura, and S. L. Bressler. Beta Oscillations in a Large-Scale Sensorimotor Cortical Network: Directional Influences Revealed by Granger Causality. *Proc Natl Acad Sci U S A,* 101: 9849-54 (2004).
[15] O. Sporns and M. Lungarella. Evolving Coordinated Behavior by Maximizing Information Structure. In: *Artificial Life X: Proceedings of the 10th International Conference on the Simulation and Synthesis of Living Systems*, pp. 322-329 (2006).
[16] A. K. Seth and G. M. Edelman. Environment and Behavior Influence the Complexity of Evolved Neural Networks. *Adaptive Behavior,* 12: 5-20 (2004).
[17] A. K. Seth. Cognitive Networks in Simulated Neural Systems. *Cognitive Neurodynamics* 2: 49-64 (2008).
[18] I. Aleksander and D. Gamez. Informational Theories of Consciousness: A Review and Extension. In: *Proceedings of BICS 2010*, Madrid (2010).
[19] A. B. Barrett and A. K. Seth. Practical Measures of Integrated Information for Time-Series Data. *PLoS Comput Biol,* 7: e1001052 (2011).
[20] A. K. Seth, E. Izhikevich, G. N. Reeke, and G. M. Edelman. Theories and Measures of Consciousness: An Extended Framework. *Proc Natl Acad Sci U S A,* 103: 10799-804 (2006).
[21] U. Lee, G. A. Mashour, S. Kim, G. J. Noh, and B. M. Choi. Propofol Induction Reduces the Capacity for Neural Information

Integration: Implications for the Mechanism of Consciousness and General Anesthesia. *Conscious Cogn,* 18: 56-64 (2009).

[22] F. Ferrarelli, M. Massimini, S. Sarasso, A. Casali, B. A. Riedner, G. Angelini, G. Tononi, and R. A. Pearce. Breakdown in Cortical Effective Connectivity During Midazolam-Induced Loss of Consciousness. *Proc Natl Acad Sci U S A,* 107: 2681-6 (2010).

[23] M. Massimini, M. Boly, A. Casali, M. Rosanova, and G. Tononi. A Perturbational Approach for Evaluating the Brain's Capacity for Consciousness. *Prog Brain Res,* 177: 201-14 (2009).

[24] D. Gamez. Information Integration Based Predictions About the Conscious States of a Spiking Neural Network. *Consciousness and Cognition,* 19: 294-310 (2010).

[25] L. Floridi, *Information : A Very Short Introduction*. Oxford University Press, Oxford (2010).

[26] L. Floridi. Philosophical Conceptions of Information. *Lecture Notes in Computer Science,* 5363: 13-53 (2009).

[27] C. E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal,* 27: 379–423, 623–656 (1948).

[28] D. M. Rosenthal. Two Concepts of Consciousness. *Philosophical Studies* 49: 329-59 (1986).

[29] S. Laureys, S. Antoine, M. Boly, S. Elincx, M. E. Faymonville, J. Berre, B. Sadzot, M. Ferring, X. De Tiege, P. van Bogaert, I. Hansen, P. Damas, N. Mavroudakis, B. Lambermont, G. Del Fiore, J. Aerts, C. Degueldre, C. Phillips, G. Franck, J. L. Vincent, M. Lamy, A. Luxen, G. Moonen, S. Goldman, and P. Maquet. Brain Function in the Vegetative State. *Acta Neurol Belg,* 102: 177-85 (2002).

[30] V. S. Ramachandran and S. Blakeslee, *Phantoms in the Brain : Probing the Mysteries of the Human Mind*, 1st ed. William Morrow, New York (1998).

[31] E. M. Izhikevich. Polychronization: Computation with Spikes. *Neural Comput,* 18: 245-82 (2006).

[32] D. Gamez. *The Development and Analysis of Conscious Machines*. PhD, Department of Computer Science, University of Essex, 2008.

[33] A. K. Seth. Causal Networks in Simulated Neural Systems. *Cogn Neurodyn,* 2: 49-64 (2008).

[34] M. Shanahan, *Embodiment and the Inner Life : Cognition and Consciousness in the Space of Possible Minds*. Oxford University Press, Oxford (2010).

[35] K. R. Popper, *The Logic of Scientific Discovery*. Routledge, London (2002).

[36] M. Helmstaedter, K. L. Briggman, and W. Denk. 3d Structural Imaging of the Brain with Photons and Electrons. *Curr Opin Neurobiol,* 18: 633-41 (2008).

[37] A. Li, H. Gong, B. Zhang, Q. Wang, C. Yan, J. Wu, Q. Liu, S. Zeng, and Q. Luo. Micro-Optical Sectioning Tomography to Obtain a High-Resolution Atlas of the Mouse Brain. *Science,* 330: 1404-8 (2010).

[38] R. Kurzweil, *The Age of Spiritual Machines : How We Will Live, Work and Think in the New Age of Intelligent Machines*. Phoenix, London (1999).

[39] D. Chalmers. The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies,* 17: 7-65 (2010).