이학석사학위논문

Relationship Between Integrated Information and Network Property of Functional Connectome

기능적 커넥톰의 통합 정보와 네트워크 특성 간의 관련성 연구

2019년 2월

서울대학교 대학원

협동과정 인지과학전공

배 홍 주

Relationship Between Integrated Information and Network Property of Functional Connectome

기능적 커넥톰의 통합 정보와 네트워크 특성 간의 관련성 연구

지도교수 이 경 민

이 논문을 이학석사 학위논문으로 제출함
2018년 12월

서울대학교 대학원
협동과정 인지과학전공
배 홍 주

배홍주의 이학석사 학위논문을 인준함
2018년 12월

위 원 장	김 홍 기	(인)
부위원장	김 청 택	(인)
위 원	이 경 민	(인)

Abstract

Many complex phenomena can be interpreted as a dynamic function of a network system. Approaches using the network theory to describe complex systems, including the connectomics, have now become classic. The Integrated Information Theory of consciousness (IIT) proposes Φ as a measure of the system's integrated information. Initially, Φ has been developed in order to construct the mathematical model of subjective consciousness, but it can be also used as an index on capturing intrinsic and integrative properties of complex dynamic systems, such as causality. Analysis of the integrated information generated on these complex dynamic systems may reveal the nature of the intrinsic structure of the system. In this paper, the correlation between Φ and general network properties used in the conventional network theory will be examined to discover how integrated information indicates a distinct property and helps to interpret information in a specific way.

Keywords: Integrated information theory, Network theory, Connectomics, Functional connectivity, Consciousness, Causality

Student Number: 2016-20095

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Explaining the subjective consciousness has been a long-standing mystery that has not been solved by any scientific method until now. From this point, Integrated Information Theory of consciousness (IIT), proposed by Giulio Tononi and his colleagues, seeks to give the answer about consciousness using scientific approaches. IIT explains consciousness as a feature stemming from the intrinsic property of a causal system. According to IIT, if a complex dynamic network containing a causal mechanism, such as the brain, generates more information than by just the sum of its parts, then the system should be considered conscious. This information above and beyond the system's parts is defined as integrated information(Tononi 2004, 2008; Oizumi, Albantakis, and Tononi 2014).[1]

However, the explanatory power of integrated information is not limited to only describe the consciousness. For instance, integrated information can successfully work as a metric for group interaction(Engel and Malone 2018).

1. IIT possesses certain pitfalls in both theoretical and technical aspects and has been the subject of controversy upon the science of consciousness(Moon and Pae 2018). However, to support or criticize IIT is not the purpose of this paper. Therefore, discussing whether IIT is a convincing theory of consciousness will be excluded from this paper.

Having been initially proposed as a measure for the causal system's integration, it can be applied to define the system's genuine causality(Albantakis et al. 2017). Such versatile use of integrated information is enabled by the fact that it measures the integration only from the system's intrinsic perspective, which is completely irrelevant from the external environment of the system. Said differently, integrated information is unique since it measures causal information from the intrinsic point of view. Therefore, in order to give plentiful explanations of various causal dynamic systems, the integrated information and other metrics used for measuring the property of the network system should be compared.

Another aspect to mention is that IIT is basically a theory inspired by the connectivity of the brain structure(Tononi and Sporns 2003). The earliest form of IIT was an approach to explain the extraordinary complex connectivity among the thalamocortical region. This aligns IIT with connectomics, a field of neuroscience that analyzes brain connectivity from the viewpoint of the network. Hence, investigating the relationships between integrated information and general metrics used in connectomics might be plausible and even crucial. Therefore, in the present study, the correlation between integrated information and several network properties widely used in connectomics will be investigated on artificially configured networks.

1.1 Core concepts of integrated information

To identify unique features of integrated information, it is necessary to look through the theoretical composition of IIT in detail.[2] Since IIT attempts to

2. The original text in Sections 1.1 - 1.1.3 has been published as (Moon and Pae 2018). The content was written by the present author, so it is directly quoted here.

explain how conscious experience arises from physical substrates, there are several explanatory concepts describing this bottom-up process. While IIT has kept updating its version from 1.0 to 3.0(Tononi 2001, 2004, 2008, 2012; Balduzzi and Tononi 2008, 2009; Oizumi, Albantakis, and Tononi 2014), those core concepts remain to be fundamentals of the theory throughout all versions.

1.1.1 Mechanisms, states, connections, and repertoires

The central focus of IIT is on the physical substrates of experience and their causal structures. IIT analyzes candidate physical substrates of experience in a bottom-up manner; physical *elements*, which can causally interact with each other, are under consideration. Any set of elements can be considered as a *mechanism*. Furthermore, any set of mechanisms can be thought of as a higher-order mechanism or a *system of mechanisms* (in short, *system*). The system is composed of elements so that the system itself also can be a mechanism or a set of elements. On the other hand, causal structures of physical substrates are analyzed by two central notions of IIT; mechanisms, or systems, can be in a *state*, which corresponds to outputs of their elements. For instance, if three elements—A, B, and C—with the binary output 1 or 0 compose a mechanism, and these element's outputs are respectively 1, 0, and 0, the state of the mechanism ABC is represented as 100(see Oizumi, Albantakis, and Tononi 2014, Figure 1A). Further, such mechanism in a state can have a *connection*, which corresponds to a set of causal connections among elements of the mechanism(Balduzzi and Tononi 2009). For example, if causal connections c^1, c^2, c^3, and c^4 are given, there might be a set of connections, such as as $\{c^1, c^2\}$, $\{c^1, c^3\}$, $\{c^1, c^2, c^3\}$ or $\{c^1, c^2, c^3, c^4\}$, etc. Any causal relationships could be characterized as a connection, such as synapses between neurons, which could be ideally represented as logic gates with simple computational functions.

From states and connections of the mechanism, one can have *repertoires*. A repertoire is defined as a *probability distribution* to possible states of the mechanism. In IIT, the causal structure of the mechanism must be known *a priori*. When the state and connection of a mechanism are given at time t, one can infer which past or future states of which mechanism—including the mechanism *itself*—could be causes or effects of the given state of the mechanism, and how much probabilities would be distributed to each possible cause or future effect states. Therefore, these probability distributions are probabilistic expressions of how the mechanism's particular state could cause or be caused by a certain mechanism's past or future states. In this sense, the mechanism in the state *specifies* repertoires, or its possible causes and effects.

The notions of mechanisms, states, connections, and repertoires are the very fundamentals in IIT. Without these concepts, calculating information from a mechanism's causal structure is not possible. As explained above, repertoires are derived from states and connections of the mechanism. Furthermore, the very concept of information is formally defined by repertoires and related notions. The concepts of mechanisms, states, connections, and repertoires tie causation and information together and enable us to calculate how much information is generated from the causal structure of the mechanism. In part, this is the reason why they survived several updates so far.

These notions also provide IIT with a quite liberal view about possible physical substrates of consciousness. None of these notions tells about what kind of materials should be considered as a candidate for the physical base of experience. Therefore, when something has its state and connection and specifies repertories, it can be at least considered regarding if it produces experience. Given that mechanisms or systems in a state are not limited to biological substrates, chemical structures such as silicon chips can be legitimate candidates

for the physical base of consciousness. Thus, under the framework of IIT, the question "Is this cellular phone conscious?" is not a category-mistaken question that should be *a priori* rejected. As far as the cellular phone can be considered as a "system of mechanisms in a state", we can at least consider the possibility of its consciousness. In principle, anything that has its states and connections can be a mechanism, and any mechanism can be a possible candidate for a conscious mechanism(Tononi and Koch 2015).

1.1.2 Intrinsic and causal information

According to IIT, an amount of information generated by a mechanism is calculated from repertoires. This calculation is performed by measuring the distance between the unconstrained and constrained repertories. For the past or future state, IIT supposes the unconstrained repertoire as a probabilistic base. Given the system's causal structure, the repertoire is unconstrained in that such uncertainty is not constrained yet by the given state of the mechanism. Using Bayes' theorem, one can infer the constrained repertoire from the given state of the mechanism. It is this distance between unconstrained and constrained repertoires that is defined as information throughout all versions of IIT.

The crucial point here is that those repertoires involved in information should be inferred from mechanisms *within* a considered system. To calculate repertoires specified by the mechanism in the state, one must consider past or future states of mechanisms within the system under consideration. No mechanism *outside* of the considered system should be taken into account. For example, to calculate the amount of information generated by the mechanism mentioned in Section 1.1.1, ABC in 100, one should consider mechanisms only within a considered system; suppose that with the mechanism of ABC constitutes a certain system under consideration. Other elements, such as D

and E, are out of the considered system. Then, according to IIT, ABC in 100 cannot specify repertoires of mechanisms such as D, DE, or even AD, AE, ABD, ABE, $ABCD$, $ABCE$. It only specifies repertories of mechanisms A, B, C, AB, AC, BC, ABC. Those repertoires would represent possible causes or effects of ABC's being in 100 that are in the considered system with their probabilities. In short, mechanisms in a certain system under consideration only can specify repertoires of mechanisms within that system. In this specific sense, in IIT, repertoires specified by the mechanism in a state express *intrinsic causal power* of the mechanism. As repertoires represent the intrinsic causal power of the mechanism, information in IIT is essentially *intrinsic* and *causal*. Information generated by the mechanism is measured as the distance between repertoires. Of note, these repertoires involve nothing external to the system. They solely depend on possible causes or effects within the system. Therefore, information is intrinsic to the system in that it does not require anything external to the system. In addition, information has nothing to do with input/output signals that can be detected only by the external observer. Rather, it is about causes and effects that can be detected only from *the system's own intrinsic perspective*(Tononi 2008, 2012; Oizumi, Albantakis, and Tononi 2014). Moreover, given that repertoires specify possible causes or effects and their probabilities, information produced by the mechanism is causal. This is why IIT repeatedly emphasizes the notion of information as *"differences that make a difference"* (Bateson 1972). In IIT, for instance, the mechanism in a state specifies which past states of a certain mechanism ("differences") would likely to cause the mechanism's being in that state ("a difference"). This further implies that only something that can be *selectively* caused or cause can produce information. This intrinsic and causal notion of information is the hallmark of IIT, which distinguishes IIT from other information theories:

anything informative has an intrinsic causal power, and anything intrinsically causal has information. This intrinsic and causal nature of information is directly inherited by the most central concept in IIT, integrated information. Although integrated information is defined in a sophisticated manner, in so far as it is information, it also should be intrinsic and causal. The intrinsic and causal information is fundamental to IIT in that it determines what kind of information the theory deals with.

1.1.3 Integrated information

The notion of *integration* first stems from the phenomenological aspects of experience: "[p]henomenologically, every experience is an integrated whole, one that means what it means by virtue of being one, and which is experienced form a single point of view"(Tononi 2012, p.295). To be a physical underpinning of such integrated, unified experience, what should a mechanism be like? Here, IIT suggests one of its thought experiments: let's compare a highly informative, but unconscious mechanism and a conscious mechanism. For example, what is the difference between a conscious brain and an unconscious digital camera that consists of thousands of photodiodes? According to the IIT, the most significant difference is that while the former is causally integrated, the latter is not(Tononi 2012). Causal interactions within the brain are so highly integrated with each other that, once they are fragmented, the whole brain's performance might break down. This thought experiment on the camera model suggests that producing information is not sufficient for a mechanism to generate consciousness. Even if the mechanism is equipped with complicated connections and distinguishes vast repertoires, if its elements do not specify a maximum of integrated information, the mechanism cannot give rise to experience.

As a mechanism with a causal structure produces intrinsic information,

one with integrated causal structure generates integrated information. The integrated information is integrated in the sense that, as a whole, the mechanism generates more information than the sum of its parts. Said differently, it is information produced only from the mechanism as a whole. By definition, the integrated information of the system is irreducible to its parts. Therefore, according to IIT, the amount of integrated information generated by the mechanism is calculated by partitioning the system by disconnecting the connections between the mechanisms. That is, if the information disappears by partitioning, it would be the information generated by the mechanism as a whole, not by individual parts. The informational difference between the mechanism as a whole and the system's partitions' mechanism is defined as integrated information. Nonetheless, considering that there are many possible ways of how the mechanism is partitioned, it becomes crucial to decide which partition should be used in calculating integrated information. IIT chooses the partition which causes the least loss of information, which is called *minimum information partition*(MIP). Finally, depending on the level of calculation, the calculated values of integrated information are represented as Φ.

1.1.4 Mathematical notations of integrated information

Thus far, numerous methods of measuring or assuming integrated information have been proposed(Barrett and Seth 2011; Ay 2015; Tegmark 2016; Oizumi et al. 2016; Oizumi, Tsuchiya, and Amari 2016; Marshall, Gomez-Ramirez, and Tononi 2016). However, not every proposed measure is available for practical uses. In the present study, integrated information was calculated using Phi Toolbox (Kitazono and Oizumi 2018). Phi Toolbox offers total 4 different metrics that are available for applying in practical data using the Gaussian assumption. These metrics are Φ_G, Φ^*, Φ_{MI} and Φ_{SI}. All of these metrics cal-

culates Φ based on the definition from IIT version 2.0 as it is much elaborated than from the version 1.0, also being less complicated and thus relatively easy to simplify the formulas than from the version 3.0.[3] In this section, how those metrics were designed will be shortly explained.

As mentioned in Sections 1.1.1 - 1.1.3, *integrated information* Φ is basically computed as the minimum of informational difference of *intrinsic information* between the *fully-connected system*, or the non-partitioned system, and the *disconnected system*, or the partitioned system. In IIT version 2.0, the probability distribution indicating the state change of the system between two time points (past and present) is defined as the intrinsic information. In general, the state change between time points is represented as the joint probability distribution(Tononi 2008). However, the exact method of defining intrinsic information into probability distribution varies through the metrics of Φ, as it is the core part where all the metrics become unique and different. Similarly, while the Kullback-Leibler divergence is generally used to calculate the informational difference between fully-connected and disconnected system(Tononi 2008), it may differ along with the metrics of Φ. The general calculation of the integrated information Φ based on joint probability distribution and the Kullback-Leibler divergence is illustrated in Figure 1.1.

3. IIT gets much more refined and sophisticated as the version gets updated to 3.0. For instance, intrinsic information is defined using three time points (past, present, and future); furthermore, large Φ and small ϕ are distinguished, which enables a more detailed modeling of subjective consciousness. However, since computational complexity burgeons as there are more factors to consider, no measures for practical use of Φ based on version 3.0 have been proposed yet (see (Moon and Pae 2018) for further detail on the changes and updates of IIT in version 3.0).

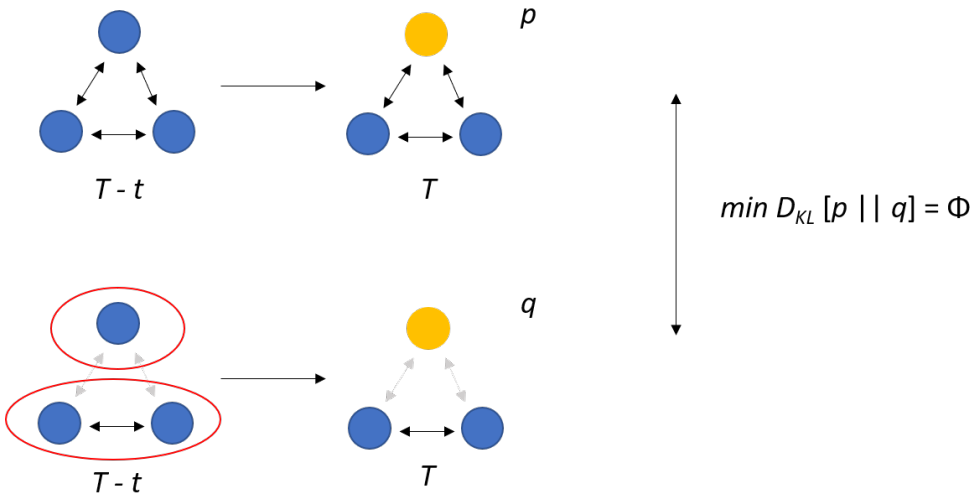


Figure 1.1: General calculation of Φ from IIT version 2.0. Each circle indicates *elements* that constitute a causal *system*. As time flows from past $(T-t)$ to present (T), the *state* of the elements changes, as they depend on the elements' causal mechanism. The blue circles indicate the non-activated state, while the yellow circle indicates the activated state. The causal connections between the elements are shown as black arrows. In the *fully-connected system* p, all elements are connected; while some connections are disconnected in the *disconnected system* q, which are indicated as gray arrows. In this figure, disconnected system q is *partitioned* into two parts. Partitioning of the system is visualized as the red outline grouping the system into two. Finally, the integrated information Φ is calculated as the minimum of the Kullback-Leibler divergence between the probability distribution p and q.

Note again that integrated information is defined as the smallest informational loss between the fully-connected system and the disconnected system. Therefore, to compute integrated information, it is necessary to find the partition that generates the smallest informational loss (MIP, see Section 1.1.3). However, while many recent studies have proposed various efficient methods of searching the MIP(Hidaka and Oizumi 2018; Kitazono, Kanai, and Oizumi 2018; Toker and Sommer 2018), searching the exact MIP still requires excessive amount of computational power. Therefore, in the present study, inte-

grated information among *atomic partition* was computed as Φ, as suggested in (Albantakis et al. 2014) and (Oizumi, Tsuchiya, and Amari 2016) for practical approximation of Φ. That is, all integrated information calculated in the present study is based on the number of partition $i = 30$, which equals to the number of nodes in the networks. Φ calculated using the atomic partition naturally becomes the upper-bound, or the biggest value of the system's possible integrated information, as atomic partition refers to the partition that destroys all connections of the target network. Disconnected system q for the system's atomic partition is shown in Figure 1.2.

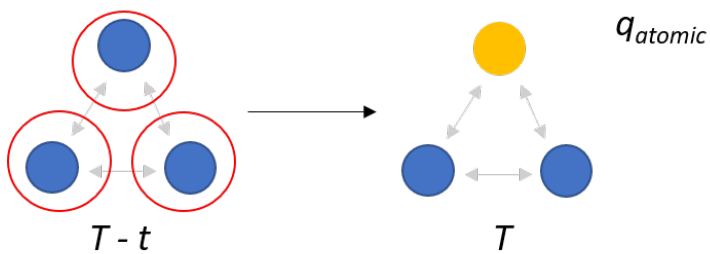


Figure 1.2: Atomic partition of the system. Note that the number of partitioning, grouped by the red outline, is equal to the number of elements.

Using the joint probability distribution as the definition of intrinsic information and using Kullback-Leibler divergence to measure the informational difference, geometric integrated information Φ_G can be calculated. Φ_G is a metric that, based on information geometry, indicates information integration from both equal-time interactions and time-lagged interactions(Oizumi, Tsuchiya, and Amari 2016). Here, the intrinsic information of fully-connected system X for given present state T and past state $T - t$ is defined as joint

probability distribution p :

$$p(X_T, X_{T-t}) \tag{1.1}$$

For Φ_G, intrinsic information of disconnected system X^i can also be expressed as a joint probability distribution. The partitioned probability distribution q of the system for i-th subsystem is defined as follows :

$$q(X^i_{T-t}|X_T) = q(X^i_{T-t}|X^i_T) \quad (\forall i) \tag{1.2}$$

as the transition of the elements in the partitioned systems should be mutually independent(Oizumi, Tsuchiya, and Amari 2016, Figure 1).
Finally, geometric integrated information Φ_G for probability distribution p of a fully-connected system X and partitioned probability distribution q of the system X^i which is disconnected into i parts can be defined as follows :

$$\Phi_G = min D_{KL}[p(X_T, X_{T-t})||q(X^i_T, X^i_{T-t})] \tag{1.3}$$

where $D_{KL}[p|q]$ stands for the Kullback-Leibler divergence between distributions p and q.
Here, if the time-series data from the system X follow normal distribution, one can apply multivariate linear regression model on X for the transition over time T to $T-t$. The regression equation for X can be expressed as follows :

$$X_T = AX_{T-t} + E \tag{1.4}$$

where A is the regression coefficient matrix and E is residual matrix from the regression model that estimates the fully-connected system. Since X follows normal distribution and the random variables of X are all Gaussian, it is

possible to express regression model of disconnected system X^i using X, as proven by (Oizumi, Tsuchiya, and Amari 2016). Therefore, regression equation on disconnected system X^i can be expressed as follows :

$$X_T = A'X_{T-t} + E' \tag{1.5}$$

where A' is the regression coefficient matrix and E' is the residual matrix from the regression model that estimates the disconnected system. Finally, as previously demonstrated by (Oizumi, Tsuchiya, and Amari 2016) and (Toker and Sommer 2018), geometric integrated information Φ_G can be obtained by calculating the Kullback-Leibler divergence, using the covariance of E and E' :

$$\Phi_G = \frac{1}{2} \log \frac{|Cov(E')|}{|Cov(E)|} \tag{1.6}$$

where $|Cov(E)|$ is determinant of covariance of E, and $|Cov(E')|$ is determinant of covariance of E'. According to the definition above, geometric integrated information Φ_G quantifies how much the partitioned system X^i deviates from the fully-connected system X, or how much informational loss the partition causes, as Φ_G increases with an increase of the difference of each covariance.

The next metric to be explained is the integrated information explained by the decoding perspective, Φ^* (called as *phi-star*)(Oizumi et al. 2016). Φ^* is defined by using the mutual intrinsic information. The mutual intrinsic information I for a fully-connected system X is defined as follows :

$$I(X_{T-t}, X_T) = H(X_{T-t}) - H(X_{T-t}|X_T) = p(X_T|X_{T-t}) \tag{1.7}$$

where $H(\ )$ is the entropy and $H(\ |\ )$ is the conditional entropy. This is also

noted as *matched decoding*, by interpreting the mutual information as the information of past state of the system X_{T-t} decoded by the present state of the given system X_T. Therefore, I is equal to the probability distribution of the present state of a fully-connected system given past states, which is noted as $p(X_T|X_{T-t})$ from Equation 1.7. Said differently, matched decoding is to predict the system's past states with "true" conditional distribution p given. Likewise, the mutual intrinsic information I^* for the disconnected system X^i is defined as follows :

$$I^*(X_{T-t}, X_T) = q(X_T|X_{T-t}) \tag{1.8}$$

where q indicates the partitioned probability distribution. Since Φ^* defines partitioned probability distribution q as the product of the conditional probability distribution in each partition of i, q can be expressed as follows :

$$q(X_T|X_{T-t}) = \prod_i p(X_T^i|X_{T-t}^i) \tag{1.9}$$

By this way, I^* indicates *mismatched decoding*, as using "false" conditional distribution q on predicting the system's past states. Here, Φ^* is formulated by subtracting the partitioned system's mutual information(mismatched decoding) from the fully-connected system's mutual information(matched decoding) :

$$\Phi^* = min\ I(X_{T-t}, X_T) - I^*(X_{T-t}, X_T) \tag{1.10}$$

As like as Φ_G, Φ^* can be analytically computed using the Gaussian assumption if probability distribution p of fully-connected system X follows normal distribution. The MATLAB function for computing Φ^* while assuming Gaussian is offered in the Phi Toolbox, following the statistical methods described in

(Oizumi et al. 2016).

The third metric, multi information, or as called mutual information(note that this notion is totally different from I of above), Φ_{MI}, is originally proposed by (Barrett and Seth 2011) and calculated using the joint entropy $H(\ ,\ )$. Φ_{MI} is noted as Φ_I in (Oizumi et al. 2016). Φ_{MI} assumes partitioned probability distribution q as follows :

$$q(X_{T-t}, X_T) = \prod_i q(X^i_{T-t}, X^i_T) \tag{1.11}$$

Using the definition mentioned above for q, multi information Φ_{MI} is defined as follows :

$$\Phi_{MI} = \sum_i H(X^i_{T-t}, X^i_T) - H(X_{T-t}, X_T) \tag{1.12}$$

Note that Φ_{MI} disconnects all causal transitions between elements of past and present, by defining q as a product of joint entropy(Oizumi, Tsuchiya, and Amari 2016, Figure 1). If the distribution of X is normal, Φ_{MI} can be transformed into as follows :

$$\Phi_{MI} = \sum_i \log |Cov(X^i_{T-t}, X^i_T)| - |Cov(X_{T-t}, X_T)| \tag{1.13}$$

where Cov indicates covariance. Finally, it is possible to calculate Φ_{MI} by substituting determinant of covariance from fully-connected and partitioned systems, as done from Φ_G.

The last metric, stochastic information, Φ_{SI}, is originally proposed by (Ay 2015). It is calculated using the conditional entropy $H(\ |\ )$. Φ_{SI} is noted as Φ_H in (Oizumi et al. 2016). Φ_{SI} assumes partitioned probability distribution q as

follows :

$$q(X_{T-t}|X_T) = \prod_i q(X^i_{T-t}|X^i_T) \tag{1.14}$$

Using the definition mentioned above for q, stochastic information Φ_{SI} is defined as follows :

$$\Phi_{SI} = \sum_i H(X^i_{T-t}|X^i_T) - H(X_{T-t}|X_T) \tag{1.15}$$

Note that Φ_{SI} disconnects causal transitions between elements of present, by defining q as a product of conditional entropy(Oizumi, Tsuchiya, and Amari 2016, Figure 1). If the distribution of X is normal, Φ_{SI} can be transformed into as follows :

$$\Phi_{SI} = \sum_i \log |Cov(X^i_{T-t}|X^i_T)| - |Cov(X_{T-t}|X_T)| \tag{1.16}$$

where Cov indicates covariance. Finally, it is possible to calculate Φ_{SI} by substituting determinant of covariance from fully-connected and partitioned systems, as done from Φ_G and Φ_{MI}.

1.2 Network-based analysis of brain connectivity

The central principle of *connectomics* is to regard the brain as a network. To explain the complex mechanism of the brain, connectomics suggests network-based analysis. Thus, in connectomics, brain connectivity is considered to be the core factor that determines the function of the brain. The connectivity pattern of the brain, which is illustrated as a network, becomes the target of the analysis, and is called *connectome*. All connectome should be illustrated as a network without self-loop and can be either directed or undirected, weighted

or binary. This makes possible to apply graph theory in order to analyze the properties of the connectome. The network is then analyzed among various topological and spatial organizations by calculating network properties, such as modularity, density and efficiency(Sporns 2011; Kaiser 2011). By classifying the connectome according to its own network structure, it is possible to make a correlation between various functional traits and the network property(Heuvel and Sporns 2011; Avena-Koenigsberger, Misic, and Sporns 2018). Therefore, connectome is frequently used as, for instance, a biomarker for diagnosing neurological disorders(Stam 2014).

Depending on the substance the network is composed of, the network of a brain can be defined in three different ways: anatomical connectivity, functional connectivity, and effective connectivity(Sporns 2007, 2010, 2011). First, *anatomical connectivity*, or structural connectivity, configures the brain network by the anatomical connection. It is specified by the physical connections between neuronal groups, widely ranging from micro-level such as synapse, to macro-level such as lobes of the brain. Second, *functional connectivity* configures the brain by statistical correlations measured from neuronal data. The network indicating functional connectivity can be considered to be functional interactions between regions; however, it is solely based on statistics. Indeed, one of the most common ways to define a functional connectome is to calculate the correlation between each region of interest(ROI) in neuronal time-series data(Varoquaux and Craddock 2013). For example, the correlation between channels in time-series EEG data or ROI in time-series fMRI data can be calculated and expressed into a correlation (or covariance) matrix. After removing noise through thresholding, the correlation matrix becomes identical to the adjacency matrix for functional connectome. Third, *effective connectivity* illustrates causal interactions between neuronal units. It can be understood as

a combined network of functional connectivity and anatomical connectivity. In effective connectivity, the network represents a genuine causal structure itself from the brain. As the connection is defined based on the causality, the algorithm for calculating the causality becomes very important(Sporns 2007).

Among three different kinds of connectivity, in the current study, functional connectivity is used for the correlation analysis. Since functional connectome is purely a statistical relationship between brain regions, it is sufficient to use practical methods for calculating the integrated information involving the Gaussian assumption. As noted in Section 1.1.4, a time-series data that follows normal distribution is necessary for efficient computing of the integrated information. Since the adjacency matrix of the functional connectome is considered to be the same as the correlation matrix, it is possible to generate normally distributed time-series data that follow the correlation directly derived from the network's adjacency matrix. This method makes it possible to compute integrated information of the target network without any loss of statistical significance. In short, analyzing the relationship between the integrated information and network properties could be conveniently performed when the target network is assumed to be the functional connectome. Therefore, all networks generated in the present study were considered to be an artificial functional connectomes.[4]

4. Since integrated information is defined based on causal relationships, comparing it with properties of effective connectome would seem plausible in a glance. However, computing integrated information based on the effective connectome becomes harder than computing based on the functional connectome; there is no choice to go around for practical purpose. For instance, one has to know *complete* causal mechanisms of the connectome to compute the system's joint probability distribution, which is difficult to be done practically. Moreover, one also has to arbitrarily set the state of the system, since effective connectome does not provide any information about the network's state. This makes integrated information rather dependent upon how one set the system's state and elements' causal mechanisms upon the experiment, which may drive the correlation slightly arbitrary and *ad-hoc*.

1.2.1 Mathematical notations of network property

As connectome itself is a network, network property can be directly measured by its adjacency matrix. In the present study, 5 different measures on network property—graph density, global efficiency, local efficiency, average clustering coefficient, and transitivity—were calculated and analyzed. To calculate these properties, the adjacency matrix, or a connectivity matrix, of a target network was needed. The target network can be either directed or undirected, either weighted or unweighted, but must not include self-loops. To control the variables efficiently in purpose, the weighted undirected network was generated and analyzed. In what follows, mathematical formula of such properties will be explained.

First, to ensure the consistency over all networks, *graph density* is calculated. Graph density D for network G is defined as the fraction of present connections to possible connections, and can be formulated as follows(Rubinov and Sporns 2010) :

$$D = \frac{2K}{N^2 - N} \tag{1.17}$$

where N is the number of all nodes and K is the number of all edges in the network. Weights of edges were ignored when calculating graph density. Since degree distribution of a node is strictly regulated as almost equal in the present study, graph density should not deviate largely for all 66 networks.

To measure the integration of network, *global efficiency* was also calculated. In general, efficiency of a path between two edges is defined as the inverse of their distance. It is classified global efficiency and local efficiency by its scope. Global efficiency is defined among the entire network, and is calculated by the

average of node efficiency, which is the average of inverse shortest path length. Global efficiency E_G of weighted network G is defined as follows(Latora and Marchiori 2001) :

$$E_G = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j \in N, j \neq i} \frac{1}{d_{ij}^w}}{n-1} \tag{1.18}$$

where n is the number of nodes of certain group, w is the weight of the edge, and d_{ij} is the connection status between node i and j. High global efficiency means that the network is effectively connected each other, thus making the average path between nodes short.

Furthermore, to measure the segregation of network, three different property measures were calculated. The first measure is *local efficiency*. Local efficiency is the global efficiency computed among the subgraph, or the neighborhood of each nodes, and is related to the clustering coefficient. Local efficiency E_L of weighted network G is defined as follows(Latora and Marchiori 2001) :

$$E_L = \frac{1}{n} \sum_{i \in N} \frac{\sum_{j,h \in N, j \neq i} (w_{ij} w_{ih} [d_{jh}^w (N_i)]^{-1})^{1/3}}{k_i (k_i - 1)} \tag{1.19}$$

where w is weight, k is degree of node, and ij, ih and jh are the edges between nodes i, j and h, respectively. High local efficiency means that the network among a certain community is effectively interconnected. Note that a high local efficiency is compatible with a high global efficiency, as global efficiency is not inversely proportional to the number of community.

Average clustering coefficient was also used to measure segregation of the network. Clustering coefficient is the average intensity, or geometric mean, of all triangles associated with each node. Clustering coefficient CC of weighted

network G is defined as follows(Onnela et al. 2005) :

$$CC = \frac{1}{n} \sum_{i \in N} \frac{2t_i^w}{k_i(k_i - 1)} \tag{1.20}$$

where t is the number of triangles around the node. Nodes with high clustering coefficient indicate that the network tends to cluster together in that node. Since clustering coefficient is a metric applied to each node, in the present study, the average of them was calculated to indicate average clustering coefficient of the networks' nodes.

Finally, for segregation of the network, *transitivity* is measured. Transitivity is the ratio of 'triangles to triplets' in the network; therefore, it is a measure similar to clustering coefficient. Transitivity T of weighted network G is computed as follows(Newman 2003) :

$$T = \frac{\sum_{i \in N} 2t_i^w}{\sum_{i \in N} k_i(k_i - 1)} \tag{1.21}$$

Chapter 2

Methods

To analyze the correlation between integrated information and network properties, the research started from generating networks representing the functional connectivity of a system. All networks of the experimental group were designed to have community-and-hub structure. Depending on the distribution of weight over edges, they were considered either as integrated or as segregated. Using the adjacency matrix of the networks, network properties were then measured. The next step was to generate the normally distributed time-series data that follows the covariance derived from the designed functional connectivity network. These data corresponded to the neuronal time-series data directly measured from the system. By performing multivariate auto-regression analysis on the time-series data, the regression model of the system was obtained. Finally, using the regression coefficient matrix A and the covariance matrix of the residual $Cov(E)$ from the regression model, integrated information Φ of the system was calculated.

2.1 Designing the model networks

The sample networks were generated using Graph Online, a website-based graph visualization tool(Unicksoft 2015). In the present study, a total of 66 weighted undirected networks were generated, including 60 structure-intended networks for the experimental group and 6 random networks for the control group. All networks had 30 nodes with the average degree distribution around 4.

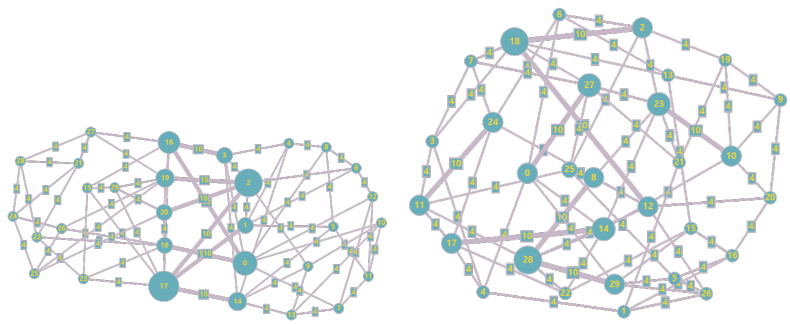
To control the properties of the network, the structure of all networks was intended. Firstly, *communities*, or *modules*, of the network were constructed. Each of 20 networks with 2, 3, and 5 communities were generated. For the networks with 2 communities, both communities consisted of 15 nodes each, and the average degree distribution of all nodes in each community was controlled by 4.13. Likewise, the networks with 3 communities had 10 nodes each and had the average degree distribution of 4 over all community nodes; finally, the networks with 5 communities had 6 nodes each and had the average degree distribution of 3.34 over all community nodes. In this way, connections over community nodes were designed to be equally distributed , consisting of almost equal graph density among all sample networks.

In the next step of generating specific network structure, communities of all networks were connected to each other among a fixed number of *hubs*. Connections between each community were strictly regulated, as they were only allowed in hub nodes. The networks with 2 communities had 5 hub nodes in each community, having a total of 8 hub-hub connections. Networks with 3 communities had 3 hub nodes in each community, having a total of 7 hub-hub connections. Finally, networks with 5 communities had 1 hub node in each community and a total of 7 hub-hub connections. Hub-hub connections were

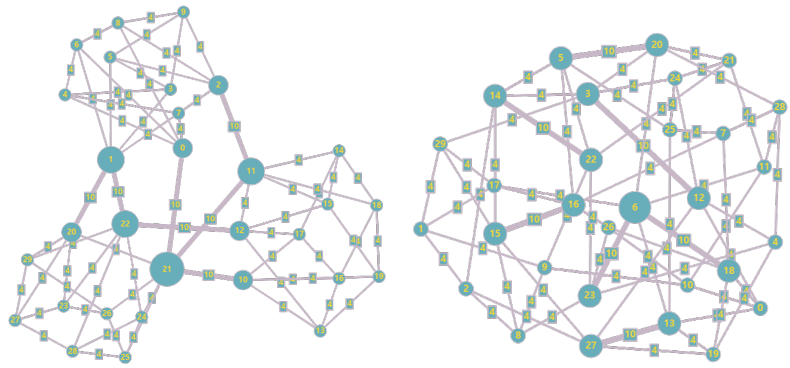
randomly distributed in every network.

In the final step, all connections were *weighted*. The distribution of weight depended on whether the network was designed as an *integrated network* or as a *segregated network*. In the present study, the edges with the weight of 10 were intended as a strong connection, while the edges with the weight of 4 were intended as weak connection. Applying such conditions, hub-hub connection was weighted as 10, and the other edges were weighted as 4 for an integrated network. For a segregated network, hub-hub connection was weighted as 4, and the other edges were weighted as 10. For each different community type, a total of 20 networks consisting of 10 integrated networks and 10 segregated networks were generated. Note that each pair of integrated and segregated network shared the same network topology, as their difference arose only from the distribution of weight over the edges.

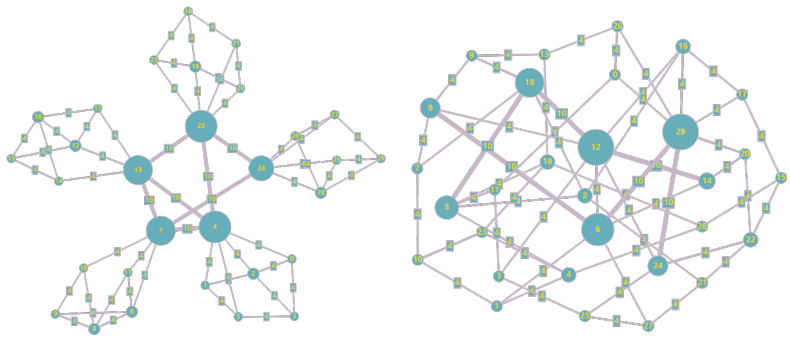
Random networks as control group were also generated for each different community type: one for an integrated network and another for a segregated network. They were generated by randomly rewiring the edges, while preserving weight and degree distribution of all nodes. By doing so, community structures of network were destroyed. The topological difference between an intended network and a random network along with community type is visualized in Figure 2.1.



(a) Integrated with 2 communities (b) Random network of (a)



(c) Integrated with 3 communities (d) Random network of (c)



(e) Integrated with 5 communities (f) Random network of (e)

Figure 2.1: Topological difference of an intended network and a random network shown along with difference of community type. Size of the nodes(blue circle) and thickness of the edges are visualized differently based on weight.

2.2 Calculation of network properties

It is important to note that the networks generated as described in Section 2.1 did not guarantee the precise value of property measures, even if they were designed to obtain some properties related with integration and segregation. That is, networks with a high weight on hub-hub connections were intended to be more integrated than those having less weight on hub-hub connections; however, this did not guarantee that the former's global efficiency would be higher than the latter's. Therefore, calculating the exact value of each network properties was important for quantitative analysis. In the present study, graph density, global efficiency, local efficiency, average clustering coefficient, and transitivity were calculated with Brain Connectivity Toolbox(Rubinov and Sporns 2010) in MATLAB version 9.5.0.944444 (R2018b). Mathematical details of all measures are explained in Section 1.2.1.

2.3 Linear regression of time-series data from model networks

The adjacency matrix of the networks generated as described in Section 2.1 was considered to be identical to the functional connectivity matrix, as noted in Section 1.2.1. In general, since functional connectivity is defined as a statistical dependence among the region of interests on a neuronal time-series data, the connectivity matrix can be considered as equivalent to the correlation (or covariance) matrix of such time-series data(Varoquaux and Craddock 2013). Hence, if the functional connectivity matrix is known, the time-series data that satisfy such a functional relationship can be obtained.

To use the adjacency matrix as a covariance matrix, the normalizing pro-

cess was needed. This was achieved by putting a large number into the diagonal of the adjacency matrix without changing any other weighted connections. In this way, we got a positive definite covariance matrix from all 66 adjacency matrices, exactly preserving the intended correlation from the adjacency matrix. In this study, the adjacency matrix with number 30 in diagonal was considered as a covariance matrix. Applying the Cholesky decomposition to the covariance matrix, we generated correlated time-series data that followed a normal distribution. Time-series data from one node of the integrated network with 5 communities are visualized in Figure 2.2.

Note that the time-series data must follow the normal distribution in order to apply efficient calculation methods of integrated information. As explained in Section 1.1.4, to calculate the integrated information of a multivariate normal data, regression coefficient A and covariance of residual $Cov(E)$ are needed from the linear regression model(Oizumi, Tsuchiya, and Amari 2016; Toker and Sommer 2018). Therefore, multivariate linear regression was applied to the generated time-series data. In this study, regression model and its coefficients were obtained by using function `VAR()` from R package `vars`(Pfaff 2008). Using function `VAR()`, multivariate autoregression towards 2000 time points of data with time lag $t = 1$ was performed. It is important to note that the covariance matrix derived from the adjacency matrix and the covariance of the regression model should be almost identical.

All statistical analyses reported in this section, including generating time-series data, were performed using R version 3.2.4 Revised(R Core Team 2016).

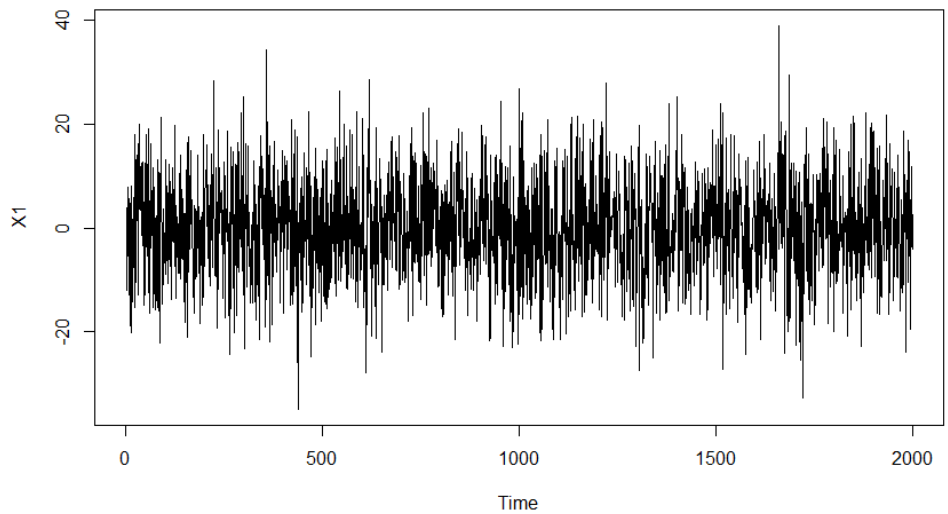


Figure 2.2: Time-series data of a single node, which follow the covariance matrix of integrated network with 5 communities. Data are generated for 2000 time points.

2.4 Calculation of integrated information

In the present study, all integrated information was calculated by using Phi Toolbox(Kitazono and Oizumi 2018) in MATLAB version 9.5.0.944444 (R2018b). Phi Toolbox offers a total of 4 different metrics—Φ_G, Φ^*, Φ_{MI} and Φ_{SI}—for the integrated information, which are all available for applying in practical data using the Gaussian assumption. Using regression coefficient A and covariance of residual $Cov(E)$ from the regression model, the integrated information of the system was calculated according to the formulas explained in Section 1.1.4. Next, time point T was set at 2000, and time lag t was set at 1. As noted in Section 1.1.4, partition of the distribution q was set by atomic partition $i = 30$ instead of MIP.

Chapter 3

Results

Based on the computed value of each measure, the relationship between integrated information and network property, as well as the discrepancy of integrated information and network property among differently structured networks was closely examined.

In this chapter, the results are presented in tables and plots on each page. Each table shows the detailed value of total 9 measures for each network having a different community structure. The correlation plot shows the statistically calculated correlation among 9 measures. The bar graph shows the distinction between network properties among differently designed networks, along with integrated information.

network	Φ_{MI}	Φ_{SI}	Φ^*	Φ_G	D	E_G	E_L	CC	T
c2i1	0.11	0.34	0.22	0.22	0.16	2.35	1.25	0.10	0.94
c2i2	0.11	0.32	0.21	0.21	0.16	2.48	1.26	0.11	0.87
c2i3	0.13	0.36	0.23	0.23	0.16	2.44	1.23	0.10	0.99
c2i4	0.10	0.33	0.22	0.22	0.16	2.44	1.05	0.08	0.76
c2i5	0.13	0.37	0.24	0.24	0.16	2.44	1.05	0.09	0.79
c2i6	0.15	0.38	0.22	0.22	0.16	2.40	1.12	0.09	0.87
c2i7	0.16	0.41	0.25	0.25	0.16	2.42	1.08	0.09	0.75
c2i8	0.11	0.35	0.23	0.23	0.16	2.35	1.22	0.10	1.00
c2i9	0.12	0.35	0.23	0.23	0.16	2.37	1.17	0.10	1.05
c2i10	0.13	0.35	0.22	0.22	0.16	2.38	1.02	0.08	0.82
c2irand	0.11	0.32	0.21	0.21	0.16	2.38	0.60	0.06	0.56
c2s1	0.11	0.32	0.21	0.21	0.16	1.53	1.11	0.23	0.78
c2s2	0.13	0.35	0.23	0.23	0.16	1.55	1.26	0.27	0.87
c2s3	0.13	0.35	0.23	0.23	0.16	1.54	1.02	0.21	0.75
c2s4	0.11	0.34	0.23	0.23	0.16	1.54	1.00	0.20	0.73
c2s5	0.11	0.33	0.22	0.22	0.16	1.55	1.01	0.21	0.73
c2s6	0.11	0.34	0.23	0.23	0.16	1.54	1.02	0.21	0.75
c2s7	0.11	0.32	0.21	0.21	0.16	1.54	1.08	0.23	0.75
c2s8	0.10	0.33	0.23	0.23	0.16	1.53	1.05	0.22	0.81
c2s9	0.10	0.32	0.21	0.21	0.16	1.55	0.90	0.19	0.67
c2s10	0.10	0.33	0.22	0.22	0.16	1.55	0.86	0.18	0.63
c2srand	0.11	0.33	0.22	0.22	0.13	1.70	0.48	0.12	0.36

Table 3.1: Frequency table illustrating a comparison between integrated information and network properties on the network with 2 communities. i refers to an integration-designed network, s refers to a segregation-designed network, and *rand* refers to a randomly rewired network.

network	Φ_{MI}	Φ_{SI}	Φ^*	Φ_G	D	E_G	E_L	CC	T
c3i1	0.11	0.35	0.25	0.24	0.15	2.36	1.35	0.11	0.94
c3i2	0.11	0.34	0.23	0.23	0.15	2.25	1.05	0.09	0.89
c3i3	0.12	0.35	0.24	0.24	0.15	2.31	1.16	0.10	0.88
c3i4	0.13	0.37	0.24	0.24	0.15	2.21	1.42	0.11	1.07
c3i5	0.11	0.33	0.22	0.22	0.15	2.30	1.50	0.12	1.03
c3i6	0.13	0.36	0.23	0.23	0.15	2.29	1.42	0.11	0.94
c3i7	0.12	0.34	0.22	0.22	0.15	2.31	1.52	0.12	1.03
c3i8	0.10	0.33	0.23	0.23	0.15	2.30	1.39	0.11	0.94
c3i9	0.11	0.34	0.24	0.24	0.15	2.31	1.49	0.11	1.03
c3i10	0.10	0.32	0.23	0.23	0.15	2.25	1.44	0.11	1.03
c3irand	0.11	0.34	0.23	0.23	0.15	2.35	0.56	0.05	0.50
c3s1	0.10	0.33	0.23	0.23	0.15	1.34	1.35	0.27	0.94
c3s2	0.12	0.35	0.23	0.23	0.15	1.35	1.02	0.22	0.82
c3s3	0.12	0.33	0.21	0.21	0.15	1.34	1.09	0.23	0.81
c3s4	0.12	0.33	0.22	0.22	0.15	1.29	1.37	0.27	0.99
c3s5	0.12	0.33	0.21	0.21	0.15	1.33	1.44	0.28	0.96
c3s6	0.12	0.34	0.22	0.22	0.15	1.30	1.42	0.28	0.94
c3s7	0.11	0.34	0.23	0.23	0.15	1.30	1.45	0.28	0.95
c3s8	0.10	0.33	0.23	0.23	0.15	1.30	1.39	0.27	0.94
c3s9	0.12	0.33	0.21	0.21	0.15	1.35	1.43	0.28	0.96
c3s10	0.11	0.34	0.23	0.23	0.15	1.33	1.40	0.27	0.96
c3srand	0.12	0.34	0.22	0.22	0.15	1.90	0.44	0.10	0.39

Table 3.2: Frequency table illustrating a comparison between integrated information and network properties on the network with 3 communities. i refers to an integration-designed network, s refers to a segregation-designed network, and *rand* refers to a randomly rewired network.

network	Φ_{MI}	Φ_{SI}	Φ^*	Φ_G	D	E_G	E_L	CC	T
c5i1	0.10	0.35	0.24	0.24	0.13	2.09	2.39	0.19	1.64
c5i2	0.12	0.34	0.22	0.22	0.13	2.15	2.37	0.19	1.74
c5i3	0.11	0.33	0.22	0.22	0.13	2.01	2.35	0.19	1.71
c5i4	0.11	0.33	0.22	0.22	0.13	2.12	2.40	0.19	1.76
c5i5	0.11	0.35	0.24	0.24	0.13	2.12	2.32	0.19	1.60
c5i6	0.11	0.34	0.23	0.23	0.13	2.15	2.38	0.20	1.88
c5i7	0.11	0.34	0.23	0.23	0.13	2.01	2.39	0.19	1.71
c5i8	0.12	0.36	0.24	0.24	0.13	2.04	2.46	0.19	1.85
c5i9	0.11	0.33	0.22	0.22	0.13	2.06	2.38	0.19	1.67
c5i10	0.11	0.33	0.21	0.21	0.13	2.12	2.42	0.19	1.77
c5irand	0.11	0.35	0.24	0.24	0.13	2.28	0.62	0.06	0.45
c5s1	0.11	0.34	0.22	0.22	0.13	1.06	2.27	0.45	1.34
c5s2	0.11	0.34	0.23	0.23	0.13	1.07	2.21	0.45	1.31
c5s3	0.12	0.35	0.23	0.23	0.13	1.05	2.21	0.43	1.41
c5s4	0.11	0.33	0.22	0.22	0.13	1.06	2.24	0.45	1.32
c5s5	0.11	0.34	0.23	0.23	0.13	1.06	2.22	0.45	1.32
c5s6	0.12	0.34	0.22	0.22	0.13	1.07	2.22	0.45	1.32
c5s7	0.12	0.33	0.21	0.21	0.13	1.05	2.25	0.44	1.41
c5s8	0.11	0.33	0.22	0.22	0.13	1.05	2.24	0.44	1.40
c5s9	0.10	0.33	0.23	0.23	0.13	1.06	2.25	0.45	1.37
c5s10	0.12	0.34	0.23	0.23	0.13	1.06	2.25	0.45	1.34
c5srand	0.10	0.33	0.23	0.23	0.13	1.70	0.45	0.11	0.31

Table 3.3: Frequency table illustrating a comparison between integrated information and network properties on the network with 5 communities. i refers to an integration-designed network, s refers to a segregation-designed network, and *rand* refers to a randomly rewired network.

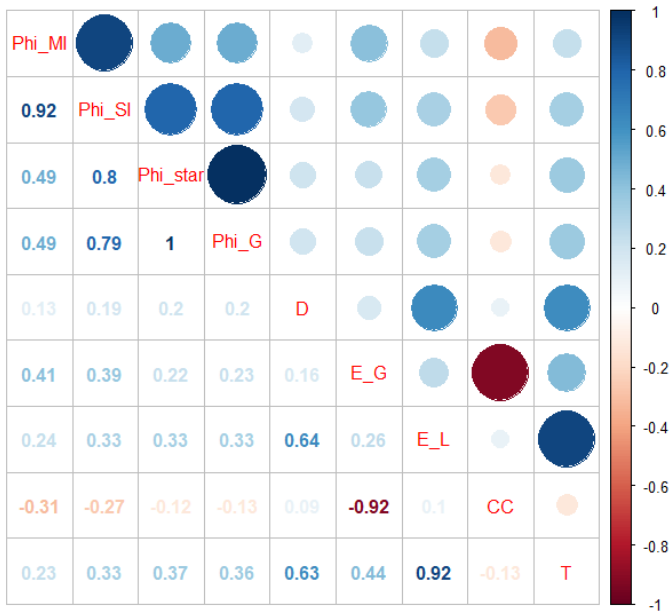


Figure 3.1: Correlation plot of intrinsic information and network properties on network with 2 communities. The randomly rewired network was excluded.

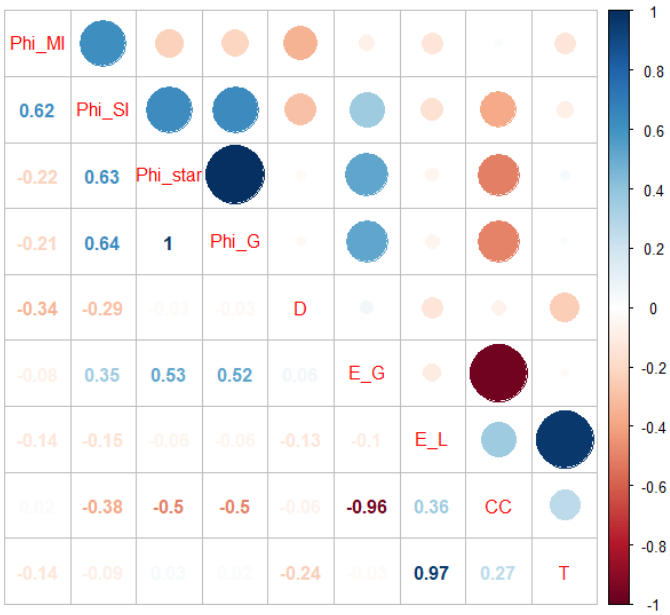


Figure 3.2: Correlation plot of intrinsic information and network properties on network with 3 communities. The randomly rewired network was excluded.

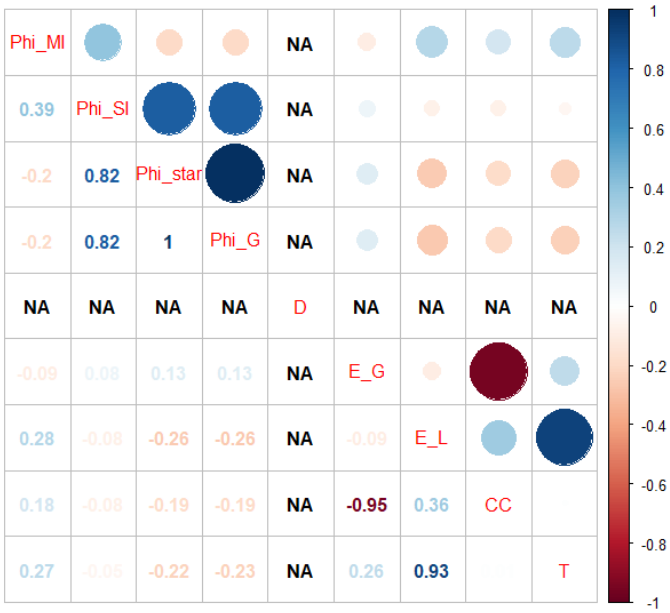


Figure 3.3: Correlation plot of intrinsic information and network properties on network with 5 communities. Correlation of graph density could not be calculated, as the graph density was equal among all 20 networks. The randomly rewired network was excluded.

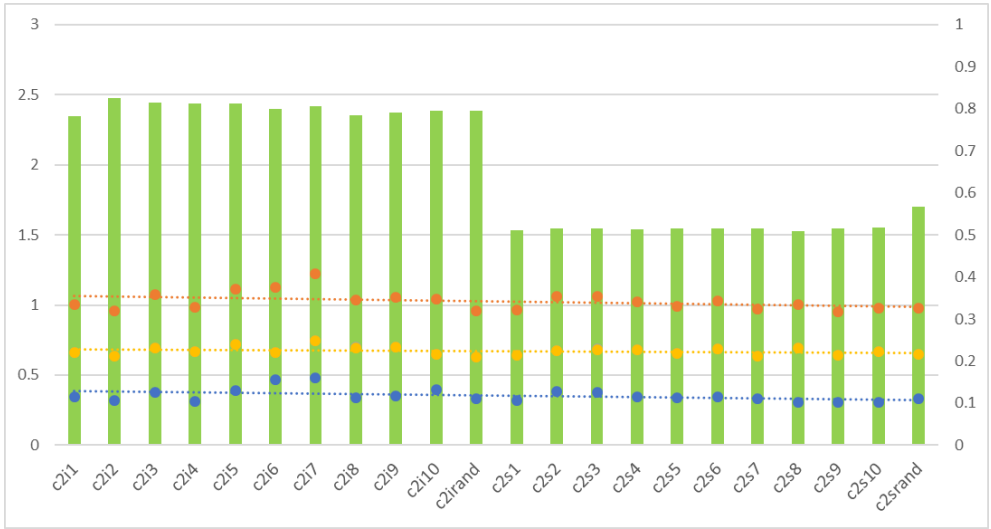


Figure 3.4: Comparison of global efficiency and Φ among the network structure, on the network with 2 communities. Green bars indicate global efficiency, units on the left. Red dotted line indicates Φ_{SI}, yellow dotted line indicates Φ^* and Φ_G, blue dotted line indicates Φ_{MI}, units on the right.

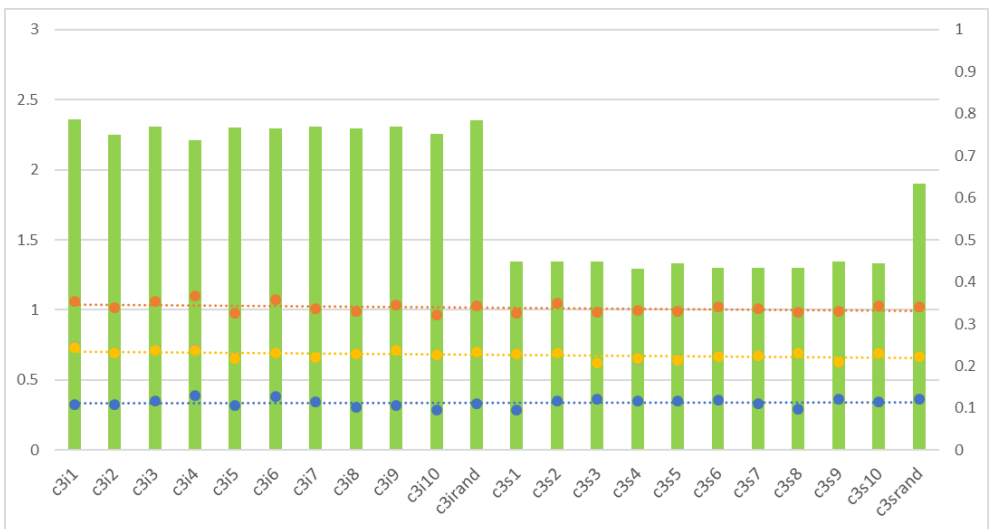


Figure 3.5: Comparison of global efficiency and Φ among the network structure, on the network with 3 communities. Green bars indicate global efficiency, units on the left. Red dotted line indicates Φ_{SI}, yellow dotted line indicates Φ^* and Φ_G, blue dotted line indicates Φ_{MI}, units on the right.

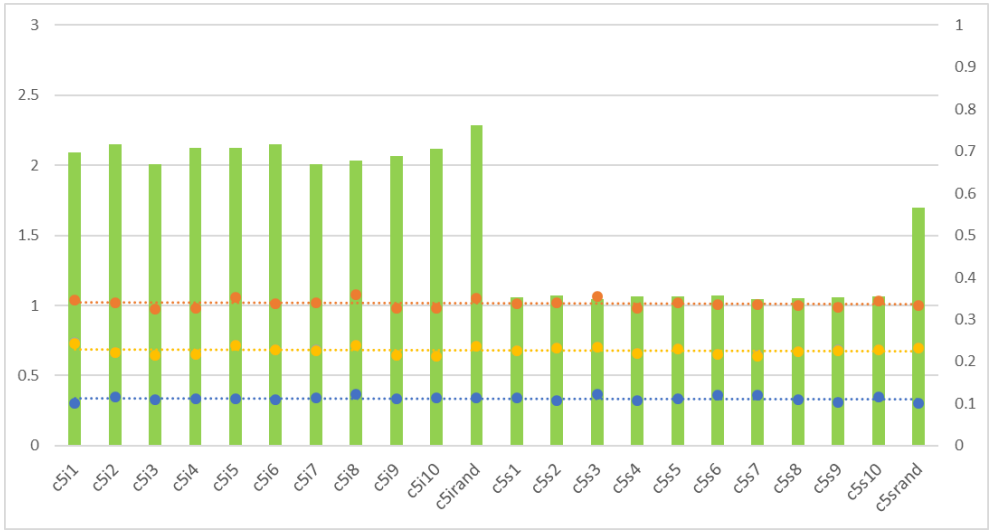


Figure 3.6: Comparison of global efficiency and Φ among the network structure, on the network with 5 communities. Green bars indicate global efficiency, units on the left. Red dotted line indicates Φ_{SI}, yellow dotted line indicates Φ^* and Φ_G, blue dotted line indicates Φ_{MI}, units on the right.

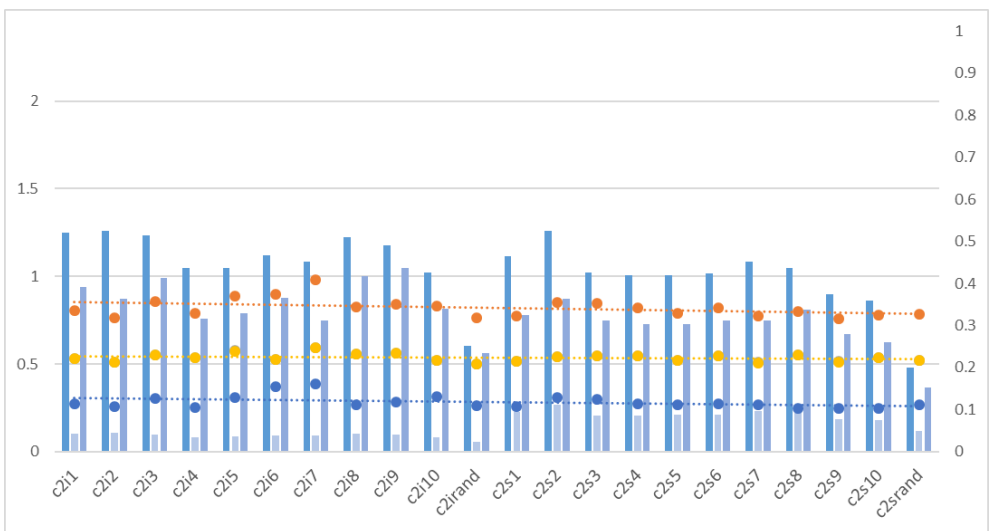


Figure 3.7: Comparison of network segregation on 2 communities. Blue bars indicate local efficiency, average clustering coefficient and transitivity in sequential order, units on the left. Red dotted line indicates Φ_{SI}, yellow dotted line indicates Φ^* and Φ_G, blue dotted line indicates Φ_{MI}, units on the right.

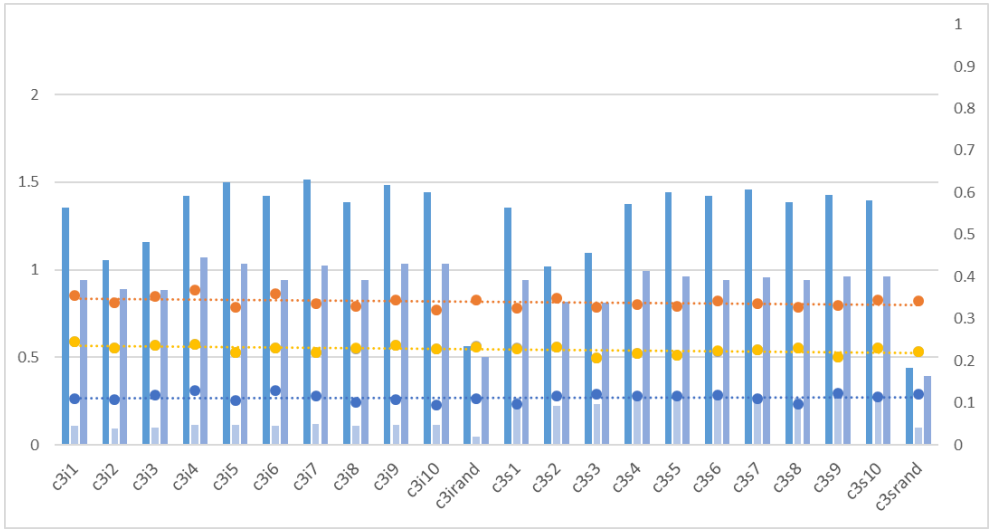


Figure 3.8: Comparison of network segregation on 3 communities. Blue bars indicate local efficiency, average clustering coefficient and transitivity in sequential order, units on the left. Red dotted line indicates Φ_{SI}, yellow dotted line indicates Φ^* and Φ_G, blue dotted line indicates Φ_{MI}, units on the right.

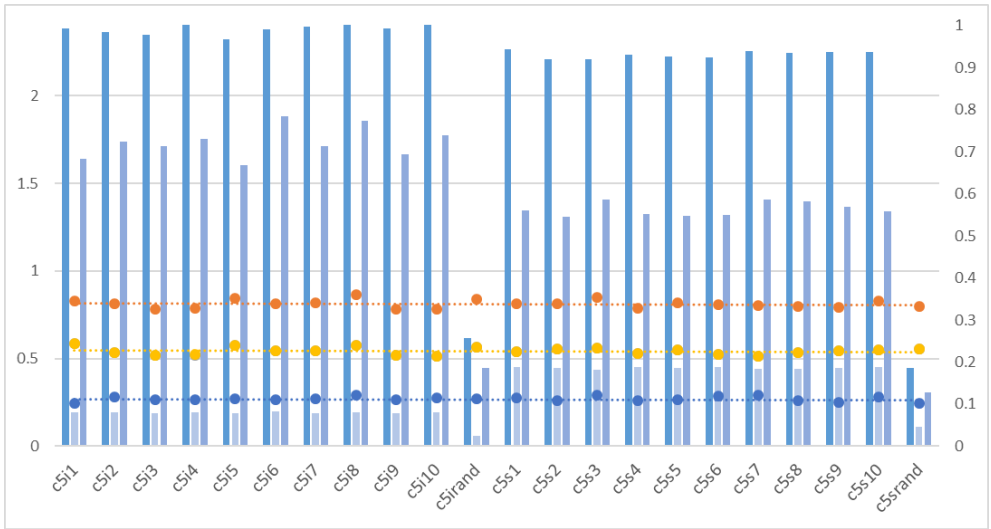


Figure 3.9: Comparison of network segregation on 5 communities. Blue bars indicate local efficiency, average clustering coefficient and transitivity in sequential order, units on the left. Red dotted line indicates Φ_{SI}, yellow dotted line indicates Φ^* and Φ_G, blue dotted line indicates Φ_{MI}, units on the right.

3.1 Relationship between integrated information and network property

Amazingly, integrated information and network properties did not show any relevant correlation upon every network structure. The correlation plot did not show any significant correlation between the measures of Φ and the network properties, as the determinants of correlations were around or below 0.3. The results shown in bar graphs in this chapter suggest that, regardless of network properties, the values of each Φ were consistently 0.1, 0.2 and 0.3 respectively. All measures of Φ turned out to be almost consistent for each randomly rewired network, as well as for the integrated network and the segregated network within the same community structure. The values of Φ between different numbers of communities were found to be consistent. The results reported in tables in this chapter indicate that the graph density was consistently equal among the same community structure; however, the correlation plots showed there was no statistically significant correlation between graph density and Φ, as all determinants of correlations remained below 0.3. Therefore, it appears that upper-bound of integrated information has no correlation with the topological structure of the correlation matrix.

3.2 Network property among differently structured networks

As shown in the tables in this chapter, it turns out that all networks with the same number of community structures have equal graph density. This implies there was less variation on hub nodes' degree distribution, which means the intended design was well-controlled and consistent.

Furthermore, randomly rewired networks turned out to have much less local efficiency, average clustering coefficient, and transitivity than the experimental group. However, for global efficiency, rewired networks only from the segregated structure had a significant difference as compared to the experimental group. In conclusion, network properties of segregation were found to be much more affected than global efficiency by community structures being destroyed.

Another aspect to note here is that the network properties of segregation increased with an increase of the number of community structure. The trend can be clearly seen through the bar graph in Figures 3.7 - 3.9. Meanwhile, the segregation properties of rewired networks hardly changed. Therefore, it is plausible that the topological segregation of the network is clearly defined through local efficiency, average clustering coefficient, and transitivity.

Moreover, there was a noticeable correlation between certain network properties. As shown in the correlation plot in Figures 3.1 - 3.3, the global efficiency and clustering coefficient had a strong negative correlation, and local efficiency and transitivity had a strong positive correlation among all numbers of the community structures. Although their relationship was not warranted as in a formula, since global efficiency indicates the integration of network, and local efficiency, clustering coefficient, and transitivity indicate the segregation of network, all these results were sufficiently predictable.

3.3 Integrated information among differently structured networks

As shown by the results reported above, all 4 different measures of Φ had consistent values among all 66 networks, regardless of the network structure. Their range can be shown as $\Phi_{MI} < \Phi_G = \Phi^* < \Phi_{SI}$, corresponding to 0.1, 0.2

and 0.3, respectively. This suggests that, as also noted in 1.1.4, the definition of the amount of integrated information is different for each measure. Of note, however, Φ_G was almost equal to Φ^*, as can be seen from the tables and the bar graphs shown above. The statistical correlation between Φ_G and Φ^* was 1 in all three different community structures.

The relationship between different Φ was also notable. Φ_{MI} and Φ_{SI} had a strong correlation of 0.92 in the networks with 2 communities. This correlation as decreased with an increase of the number of the community, as it decreases into 0.39 on the networks with 5 communities. The correlation between Φ_{SI} and Φ^* (or Φ_G, as those two values were almost identical) was fairly positive but not as consistent, meanwhile, the correlation between Φ_{MI} and Φ^* was completely inconsistent.

Chapter 4

Discussion

The results of the present study demonstrated that the upper-bound of integrated information does not statistically correlate with the network property of functional connectome. On the other hand, the network property well explains the topological structure and connectivity of the functional connectome. Therefore, it can be concluded that the topological structure of functional connectome, which indicates statistical dependencies between the region of interests, does not affect the upper-bound value of Φ.

Of note, in the present study, the atomic partition was used instead of MIP, and thus calculated the largest possible Φ of the system, which is the upper-bound of Φ. The Φ value is determined by the informational difference between the joint (or conditional) probability of the fully-connected system p and the probability of the disconnected system q. By definition, both factors have a decisive influence on the calculation of integrated information. Since atomic partition was used in the present study, the probability of the disconnected system q was applied equally among all networks. This means all discrepancies among integrated information should have been caused by the difference from the distribution p. However, in the present study, the consistent value of Φ was

calculated regardless of the network property. This suggests that the **probability distribution of the fully-connected system p is determined irrespective of the distribution of system-wide correlation**. Said differently, the change upon functional connectivity did not change distribution p. This means that the probability of the fully-connected system p, or the intrinsic causal information of the system, was determined in the first place as irrelevant from the distribution of system-wide correlation. Moreover, since all upper bounds of each value were nearly identical among all 66 networks, the maximum value of the Φ a system could have no relation to the distribution of system-wide correlation.

The results also suggest that the maximum value of each Φ forms a relative range when one causal system is determined. This range should be closely related with the distribution p defined by each Φ. In particular, the maximum value of Φ_G and Φ^* can have are identical. That is, **the difference between the value of Φ_G and Φ^* is determined by the calculation of q, not by p.**

Let's take a deeper look at the result that Φ_G and Φ^* were identical. This result indicates that Equation 1.3 and 1.10 have the same distribution p from the right side of the equation. That is, maximally decomposing the temporal correlation of each element and maximally mismatched decomposition are identical to each other. As shown in (Oizumi, Tsuchiya, and Amari 2016), Φ_{MI}, Φ_{SI}, and Φ_G were compared and analyzed at the level of information geometry, where each relationship was illustrated as (Oizumi, Tsuchiya, and Amari 2016, Figure 1). However, Φ^*, which was developed from the viewpoint of decoding perspective has not yet been fully compared from the perspective of information geometry. Considering the results of the present study, it can be expected that **each distribution p from Φ_G and Φ^* may have a geometrically**

equivalent, or at least symmetrical, "shape" in the information space.

The results of the present study demonstrate that functional connectivity does not affect the upper-bound of Φ. Said differently, functional connectivity does not affect the probability distribution p. Discovering the exact MIP of a system is necessary to figure out how does the functional connectivity affects the distribution q. Therefore, it is crucial to determine the MIP of a system, for further study to configure the general relationship between Φ and functional connectivity.

Bibliography

Albantakis, Larissa, Arend Hintze, Christof Koch, Christoph Adami, and Giulio Tononi. 2014. Evolution of integrated causal structures in animats exposed to environments of increasing complexity. *PLOS Computational Biology* 10, no. 12 (December): 1–19. doi:10.1371/journal.pcbi.1003966.

Albantakis, Larissa, William Marshall, Erik Hoel, and Giulio Tononi. 2017. What caused what? an irreducible account of actual causation. https://arxiv.org/abs/1708.06716.

Avena-Koenigsberger, Andrea, Bratislav Misic, and Olaf Sporns. 2018. Communication dynamics in complex brain networks. *Nature Reviews Neuroscience* 19:17–33. doi:10.1038/nrn.2017.149.

Ay, Nihat. 2015. Information geometry on complexity and stochastic interaction. *Entropy* 17 (4): 2432–2458. doi:10.3390/e17042432.

Balduzzi, David, and Giulio Tononi. 2008. Integrated information in discrete dynamical systems: motivation and theoretical framework. *PLOS Computational Biology* 4 (6): e1000091. doi:10.1371/journal.pcbi.1000091.

Balduzzi, David, and Giulio Tononi. 2009. Qualia: the geometry of integrated information. *PLOS Computational Biology* 5 (8): e1000462. doi:10.1371/journal.pcbi.1000462.

Barrett, Adam B., and Anil K. Seth. 2011. Practical measures of integrated information for time-series data. *PLOS Computational Biology* 7 (1): e1001052. doi:10.1371/journal.pcbi.1001052.

Bateson, Gregory. 1972. *Step to ecology of mind.* University of Chicago Press.

Engel, Davia, and Thomas W. Malone. 2018. Integrated information as a metric for group interaction. *PLOS ONE* 13, no. 10 (October). doi:10.1371/journal.pone.0205335.

Heuvel, Martijn P. van den, and Olaf Sporns. 2011. Rich-club organization of the human connectome. *Journal of Neuroscience* 31 (44): 15775–15786. doi:10.1523/JNEUROSCI.3539-11.2011.

Hidaka, Shohei, and Masafumi Oizumi. 2018. Fast and exact search for the partition with minimal information loss. *PLOS ONE* 13 (9): e0201126. doi:10.1371/journal.pone.0201126.

Kaiser, Marcus. 2011. A tutorial in connectome analysis: topological and spatial features of brain networks. Special Issue: Educational Neuroscience, *NeuroImage* 57 (3): 892–907. ISSN: 1053-8119. doi:10.1016/j.neuroimage.2011.05.025.

Kitazono, Jun, Ryota Kanai, and Masafumi Oizumi. 2018. Efficient algorithms for searching the minimum information partition in integrated information theory. *Entropy* 20 (3): 173. doi:10.3390/e20030173.

Kitazono, Jun, and Masafumi Oizumi. 2018. Practical PHI Toolbox (September). doi:10.6084/m9.figshare.3203326.v10. https://figshare.com/articles/phi_toolbox_zip/3203326.

Latora, Vito, and Massimo Marchiori. 2001. Efficient behavior of small-world networks. *Phys. Rev. Lett.* 87 (19): 198701. doi:10.1103/PhysRevLett.87.198701.

Marshall, William, Jaime Gomez-Ramirez, and Giulio Tononi. 2016. Integrated information and state differentiation. *Frontiers in Psychology* 7:926. doi:10.3389/fpsyg.2016.00926.

Moon, Kyumin, and Hongju Pae. 2018. Making sense of consciousness as integrated information: evolution and issues of iit. https://arxiv.org/abs/1807.02103v2.

Newman, M. 2003. The structure and function of complex networks. *SIAM Review* 45 (2): 167–256. doi:10.1137/S003614450342480.

Oizumi, Masafumi, Larissa Albantakis, and Giulio Tononi. 2014. From the phenomenology to the mechanisms of consciousness: integrated information theory 3.0. *PLOS Computational Biology* 10 (5): e1003588. doi:10.1371/journal.pcbi.1003588.

Oizumi, Masafumi, Shun-ichi Amari, Toru Yanagawa, Naotaka Fujii, and Naotsugu Tsuchiya. 2016. Measuring integrated information from the decoding perspective. *PLOS Computational Biology* 12 (1): e1004654. doi:10.1371/journal.pcbi.1004654.

Oizumi, Masafumi, Naotsugu Tsuchiya, and Shun-ichi Amari. 2016. Unified framework for information integration based on information geometry. *Proceedings of the National Academy of Sciences* 113 (51): 14817–14822. doi:10.1073/pnas.1603583113.

Onnela, Jukka-Pekka, Jari Saramäki, János Kertész, and Kimmo Kaski. 2005. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E* 71 (6): 065103. doi:10.1103/PhysRevE.71.065103.

Pfaff, Bernhard. 2008. Var, svar and svec models: implementation within R package vars. *Journal of Statistical Software* 27 (4). http://www.jstatsoft.org/v27/i04/.

R Core Team. 2016. *R: a language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rubinov, Mikail, and Olaf Sporns. 2010. Complex network measures of brain connectivity: uses and interpretations. Computational Models of the Brain, *NeuroImage* 52 (3): 1059–1069. ISSN: 1053-8119. doi:10.1016/j.neuroimage.2009.10.003.

Sporns, Olaf. 2007. Brain connectivity. *Scholarpedia* 2 (10): 4695. doi:10.4249/scholarpedia.4695.

———. 2010. *Networks of the brain.* The MIT Press. ISBN: 9780262014694. https://mitpress.mit.edu/books/networks-brain.

———. 2011. The human connectome: a complex network. *Annals of the New York Academy of Sciences* 1224 (1): 109–125. doi:10.1111/j.1749-6632.2010.05888.x.

Stam, Cornelis J. 2014. Modern network science of neurological disorders. *Nature Reviews Neuroscience* 15:683–695. doi:10.1038/nrn3801.

Tegmark, Max. 2016. Improved measures of integrated information. *PLOS Computational Biology* 12 (11): e1005123. doi:10.1371/journal.pcbi.1005123.

Toker, Daniel, and Friedrich T. Sommer. 2018. Information integration in large brain networks. https://arxiv.org/abs/1708.02967v2.

Tononi, Giulio. 2001. Information measures for conscious experience. *Archives Italiennes de Biologie* 139 (4): 367–371.

———. 2004. An information integration theory of consciousness. *BMC Neuroscience* 5:42. doi:10.1186/1471-2202-5-42.

———. 2008. Consciousness as integrated information: a provisional manifesto. *The Biological Bulletin* 215 (3): 216–242. doi:10.2307/25470707.

———. 2012. Integrated information theory of consciousness: an updated account. *Archives Italiennes de Biologie* 150 (2-3): 290–326.

Tononi, Giulio, and Christoph Koch. 2015. Consciousness: here, there, and everywhere? *The Royal Society* 370 (1668). doi:10.1098/rstb.2014.0167.

Tononi, Giulio, and Olaf Sporns. 2003. Measuring information integration. *BMC Neuroscience* 4, no. 1 (December): 31. doi:10.1186/1471-2202-4-31.

Unicksoft. 2015. *Graph online: creation and easy visualization of graph and shortest path searching.* https://www.http://graphonline.ru/en/.

Varoquaux, Gaël, and R. Cameron Craddock. 2013. Learning and comparing functional connectomes across subjects. Mapping the Connectome, *NeuroImage* 80:405–415. doi:10.1016/j.neuroimage.2013.04.007.

초 록

네트워크 시스템의 관점을 택하면 복잡계의 수많은 현상들을 설명할 수 있다. 커넥토믹스를 비롯하여, 복잡계를 설명하고자 네트워크 중심적으로 현상에 접근하는 것은 이제는 하나의 고전이 되었다. 한편 통합정보이론(IIT) 에서는 시스템의 정보통합량에 대한 척도로서 Φ를 제시한다. Φ는 본래 의식에 대한 수리적 모델링을 위해 고안되었지만, Φ가 뜻하는 내재적 통합정보의 의미를 응용하면 인과성과 같은 복잡계의 속성을 포착하는 지표로서 사용될 수 있다. 따라서 복잡계 네트워크에서 생성된 정보통합량을 분석하면 시스템의 내재적이며 본질적인 구조를 파악할 수 있을 것으로 보인다. 본 논문에서는 기존의 네트워크 이론에서 이용되는 네트워크 속성에 대한 척도들과 Φ 간의 상관관계를 분석하여 통합정보의 고유한 속성을 확인하고자 한다.

주요어: 통합정보이론, 네트워크이론, 커넥토믹스, 기능적 커넥톰, 의식, 인과성
학 번: 2016-20095

Acknowledgements

Special thanks to my colleague Kyumin Moon for sharing the extraordinary journey towards understanding Integrated Information Theory of consciousness, suggesting fruitful comments upon the research.