Article

A Compositional model of Consciousness based on subjectivity as a fundamental feature of nature

Camilo Miguel Signorelli [1,2,*] 0000-0002-2110-7646, **Quanlong Wang** [1,3,*]

[1] Department of Computer Science, University of Oxford; cam.signorelli@cs.ox.ac.uk
[2] Cognitive Neuroimaging Unit, INSERM U992, NeuroSpin
[3] Cambridge Quantum Computing Ltd.
* Correspondence: cam.signorelli@cs.ox.ac.uk; quaang@cs.ox.ac.uk

Version April 14, 2020 submitted to Preprints

Abstract: The scientific studies of consciousness are mainly based on objective neural mechanism, relying on objects whose existence is independent of any consciousness, but generating epistemic and ontological problems. Alternatively, in this paper consciousness is assumed as fundamental, and the main feature of consciousness characterized as the other-dependent. This approach is mainly inspired by the Buddhism philosophy of the Yogacara school. Therefore, we set up a framework of compact closed category whose morphisms are composed of a set of generators being specified by relations with other generators. The framework naturally subsumes the other-dependent feature. Moreover, it is general enough, i.e. parameters in the morphisms take values in arbitrary commutative semirings, from which any finitely dimensional system can be dealt with, fitting well into a compositional model of consciousness. Finally, as a preliminary application of our framework, we explore a solution to a toy model of the feature biding problem.

Keywords: Consciousness; Conscious Agents; Compositionality; Binding problem; Mathematics of Conciousness; Monoidal Categories.

1. Introduction

The science of consciousness have gained considerable understanding of objective neural mechanisms of consciousness, however, this strategy has also failed in recovering subjective features such as the unity of consciousness from these objective and measurable mechanisms. Thus, we present an alternative approach (Section 2), which takes inspiration from the Yogacara school [1,2], but also to some extent in line with the hypothesis of conscious agents [3], phenomenology [4], as well as other elements from the unified field hypothesis [5]. Following this approach, subjective aspects of reality, rather than physical objects, are here postulated as primitive and fundamental (Section 2.1), without falling into idealism nor dualism (Section 2.2). Meanwhile, the key feature of consciousness is characterised by other-dependent. This allows us to propose a compositional model for consciousness based on process theory, in other words, symmetric monoidal categories (Section 3 and 3.1). Process theory has proved successful at the moment to understand principles and mathematical structures of physical theories [6], such as quantum theory [7,8], causal models [9,10], relativity [11] and interestingly also natural language [12] and cognition [13,14]. At the core of process theory, there lies the principle of compositionality, which describes unity as the composition of basic elements [15,16]. Moreover, process theory is mathematically abstract thus ontologically neutral. All these make process theory suitable to search for structural properties of consciousness [17]. Specifically, in our model, we use generators in terms of diagrams as basic processes which are defined by interdependent relations between them (Section 3.2, 3.2.1 and 3.2.2). This clearly shows the consciousness feature of other-dependent. Our framework comes with a standard interpretation for each diagram (Section 3.2.3 and 3.2.4), making our theory sound, i.e. without contradictions inside. One goal of our framework is giving a mathematical formalism to target important questions about consciousness. For instance, unity of consciousness may naturally arise as result of composition, so here a toy model for the binding problem is described as an application of our framework (Section 4). Eventually, the ultimate goal is recovering objective physical theories (e.g. standard quantum mechanics) from primitive notions of subjectivity that indeed would correspond to each other, avoiding ontological claims and without the need of invoking any physical realization but pure mathematical entities (Section 5).

2. Philosophical considerations

2.1. Consciousness as Fundamental

The science of consciousness has proved elusive. On the one hand, biology and neuroscience have acquired considerable comprehension of objective neural mechanisms of consciousness [18]. On the other hand, the subjective aspects of conscious experience are mainly neglected by these approaches [19,20] or at least postponed for further developments [21]. The basic assumption is that subjective aspects of experience would emerge from the objective physical properties of the brain. In other words, the world, considered as both objective and subjective, might be entirely constructed by measurable physical generators, and subjective features of reality are merely consequences of the objective and measurable properties of the world. In this line, one would expect that taking a physical objective and mathematical theory, the subjective aspects of the experience may naturally emerge from the interaction and combination of these physical and mathematical generators. Nevertheless, scientific approaches to consciousness have failed in recovering subjectivity from the objective and measurable reality [19,20,22].

It is well recognised that objectivity is a basic assumption of science. Objectivity relates to a perceived or unperceived object while subjectivity to a perceiving subject. The object is meant to exist independently of any subject to perceive it, and as such, objectivity is commonly associated with concepts like truth and reliability [23]. Contrary, subjectivity is always interdependent, it involves both perceived and perceiving aspects, making subjective properties dependent of others interactions and thereof not independent. The assumption of objectivity as primitive or fundamental is deeply grounded in classical neuroscience, as well as other scientific fields [24–26]. Contemporary theories of consciousness tend to focus on the physical parts from which, for example, the unity of experience would emerge as a whole. The parts are considered cells, neurons, brain regions, and the whole being the unified conscious experience. This is called building blocks models [5] or reductionist approaches [25].

Nevertheless, there is an epistemic issue: "our knowledge is limited to the realm of our own subjective impressions, allowing us no knowledge of objective reality as it is in itself" [23]. One alternative to deal with that issue is to remove the assumption of objectivity and take consciousness as a primitive property of the world. One theoretical example is the conscious agent model [3,27], where the world consists of conscious agents and their experiences. Once the emergence of subjectivity is solved, now the inverse problem comes into play: how does objective phenomenon such as quantum physics or relativity arise from? Thus, the aim of such models is recovering fundamental physics from the agent's interactions, for example, quantum mechanics [27]. Ontologically, conscious agent model is different than current scientific approaches to consciousness and cognition. Moreover, there is still much work to satisfactorily reach that goal, and it is not so evident that the current versions of conscious agent models are capable to recover the entire objective realm (see objections and replies section in [27]). Through these pages, we propose some new concepts toward answering these questions, as well as starting form the idea that consciousness and subjectivity are fundamental notion of reality.

2.2. Yogacara Philosophy and Phenomenology

Starting from subjective aspects of reality may sound new to modern science, but the discussion of epistemic restrictions have been part of millenary traditions such as Buddhism and its Yogacara school, long before phenomenology appears as the science of phenomena and experience. Yogacara (Sanskrit for *Yoga Practice*), also called Vijnanavada (*Doctrine of Consciousness*) or Vijnaptimatra (*Consciousness*

Only), is one of the two main branches of Mahayana (*Great Vehicle*) Buddhism (the other being Madhyamaka, *Middle way*). All the alternative names of Yogacara philosophy involve the key concept of consciousness, and specifically, consciousness-only. This concept is sometimes wrongly interpreted. Nevertheless, the meaning behind is closer to epistemic limitations mentioned in modern phenomenology than variants of philosophical idealism [1,28].

To understand consciousness-only, another concept from the Yogacara philosophy is needed: Trisvabhāva or the three natures. Trisvabhāva is the premise that all the possible forms of existence are divided into three types: i) Parikalpita-svabhāva, the *fully conceptualized* nature, ii) Paratantra-svabhāva, the *other dependent* nature, and iii) Parinișpanna-svabhāva, the *perfect-accomplished-real* nature. As explained by [2]: "The first nature is the nature of existence produced from attachment to imaginatively constructed discrimination. The second nature is the nature of existence arising from causes and conditions. The third nature is the nature of existence being perfectly accomplished (real)". The third nature of existence is "the ultimate reality, something that never changes". An important remark is that this nature does not correspond to mind or the "ultimate mind" from which everything would originate. The ultimate reality is invariant and can not be directly depicted, it is neither objective nor subjective.

Interestingly, these three natures are inseparable from the mind and its attributes (*Citta-Caittas*), as mentioned in Cheng Weishi Lun [29] and translated to English by [30]: "The mind and its attributes (Citta-Caittas), together with the manifestations produced by it (darsana and nimittabhaga), are engendered through numerous conditioning factors, and are thus like the phenomena produced by a magician's tricks, which, not really existing though they seem to exist, deceive the ignorant. All this is called the *nature of dependence on others* (Paratantra). The ignorant thereupon perversely believe in them as Atman and as dharmas, which exist or do not exist, are identical or different, are inclusive or exclusive, etc. But, like *flowers in the sky*, etc., they are non-existent both in inner nature and external aspect. All this is called the *nature of mere-imagination* (Parikalpita). These things, which are thus dependent on others and are wrongly regarded as Atman and as dharmas, are in reality, all void (sunya). The genuine nature of consciousness thus revealed by this *voidness* is called the *nature of ultimate reality* (Parinispanna). Thus, these three natures are all inseparable from mind...". One can observe from the above citation, that consciousness as process is actually of the second nature of existence: the other dependent nature. Therefore, one main feature of consciousness processes is this "other dependent", unlike fundamental physical particles, whose existences are considered independent of others.

This remark might become clear when the mind is defined as possessed by sentient beings. The second nature or the other dependent nature is what Yogacara refers to the mind and its attributes. On that framework, the mind, as part of sentient beings, is divided into eight types of consciousnesses, what in modern science one would call senses or ways of perceiving: the five **sense-consciousnesses** (eye or visual, ear or auditory, nose or olfactory, tongue or gustatory, body or tactile consciousnesses), **mental consciousness**, **manas consciousness** (the seventh or thought-centre consciousness), and **alaya consciousness** (the eighth or storehouse consciousness). Each type of consciousness manifests itself in two forms: the perceived division (nimittabhaga in Sanskrit) and the perceiving division (darsanabhaga in Sanskrit). Here, mental consciousness becomes relevant because it is closer to modern notions of awareness. Finally, the mind is not related to an invariant nature, but indeed, it is the major mechanism why illusions appear to us, sentient beings [1].

Contrary to dualism, the notions above deny any conceptual duality (e.g. physical-non-physical, external-internal) regarding the *perfect-accomplished-real* nature. Different than idealism [28], the mind is not seen as cause effective of the rest of the world, by only of the illusion of distinctions on that world. Consciousness is essential because everything considered, affirmed or denied, even the idea of objectivity, occur to us only in consciousness. However, consciousness is not the ultimate reality. Therefore, the ontological query is suspended while an epistemic caution is reinforced: "all our efforts to get beyond ourselves are nothing but projections of our consciousness"[1]. In modern words, consciousness-only would be better understood as a claim of awareness-only, or perception-only, much closer to phenomenology.

3. Compositional Model of Consciousness

As presented in the previous section, inspired by the philosophy of the Yogacara school, the key feature of consciousness is other-dependent. Following this idea, it is natural to model consciousness in a categorical framework where morphisms are composed from a given set of generators and each generator is specified by relations with other generators. In this section, we introduce such a framework based on the theory of ZX-calculus invented by Bob Coecke and Ross Duncan [31] as a graphical language for qubit quantum theory. This diagrammatic language is mathematically rigorous [8] and has proven useful to reconstruct different aspects of physical theories. However, our framework is much more general: all finite dimensional ZX-calculus are unified in a single one, thus called qufinite ZX_Δ-calculus, and the parameters take values in an arbitrary commutative semiring, rather than complex number only.

In the sequel, we first give an introduction to the basic concepts in category theory and the concept of commutative semiring, then we present all generators and rewriting rules between them for the qufinite ZX_Δ-calculus.

3.1. Preliminaries

Category

A category $\mathfrak{C}$ consists of:

- a class of objects $ob(\mathfrak{C})$;
- for each pair of objects A, B, a set $\mathfrak{C}(A, B)$ of morphisms from A to B;
- for each triple of objects A, B, C, a composition map

$$\begin{array}{ccc} \mathfrak{C}(B,C) \times \mathfrak{C}(A,B) & \longrightarrow & \mathfrak{C}(A,C) \\ (g,f) & \mapsto & g \circ f; \end{array}$$

- for each object A, an identity morphism $1_A \in \mathfrak{C}(A, A)$,

satisfying the following axioms:

- associativity: for any $f \in \mathfrak{C}(A,B), g \in \mathfrak{C}(B,C), h \in \mathfrak{C}(C,D)$, there holds $(h \circ g) \circ f = h \circ (g \circ f)$;
- identity law: for any $f \in \mathfrak{C}(A,B), 1_B \circ f = f = f \circ 1_A$.

A morphism $f \in \mathfrak{C}(A,B)$ is an *isomorphism* if there exists a morphism $g \in \mathfrak{C}(B,A)$ such that $g \circ f = 1_A$ and $f \circ g = 1_B$. A *product category* $\mathfrak{A} \times \mathfrak{B}$ can be defined componentwise by two categories $\mathfrak{A}$ and $\mathfrak{B}$.

Functor

Given categories $\mathfrak{C}$ and $\mathfrak{D}$, a functor $F : \mathfrak{C} \longrightarrow \mathfrak{D}$ consists of:

- a mapping

$$\begin{array}{ccc} \mathfrak{C} & \longrightarrow & \mathfrak{D} \\ A & \mapsto & F(A); \end{array}$$

- for each pair of objects A, B of $\mathfrak{C}$, a map

$$\begin{array}{ccc} \mathfrak{C}(A,B) & \longrightarrow & \mathfrak{D}(F(A),F(B)) \\ f & \mapsto & F(f), \end{array}$$

satisfying the following axioms:

- preserving composition: for any morphisms $f \in \mathfrak{C}(A,B), g \in \mathfrak{C}(B,C)$, there holds $F(g \circ f) = F(g) \circ F(f)$;
- preserving identity: for any object A of $\mathfrak{C}$, $F(1_A) = 1_{F(A)}$.

A functor $F : \mathfrak{C} \longrightarrow \mathfrak{D}$ is *faithful (full)* if for each pair of objects A, B of $\mathfrak{C}$, the map

$$\begin{array}{ccc} \mathfrak{C}(A,B) & \longrightarrow & \mathfrak{D}(F(A),F(B)) \\ f & \mapsto & F(f) \end{array}$$

is injective (surjective).

Natural transformation

Let $F, G : \mathfrak{C} \longrightarrow \mathfrak{D}$ be two functors. A natural transformation $\tau : F \to G$ is a family $(\tau_A : F(A) \longrightarrow G(A))_{A \in \mathfrak{C}}$ of morphisms in $\mathfrak{D}$ such that the following square commutes:

$$\begin{array}{ccc} F(A) & \xrightarrow{\tau_A} & G(A) \\ {\scriptstyle F(f)}\downarrow & & \downarrow{\scriptstyle G(f)} \\ F(B) & \xrightarrow{\tau_B} & G(B) \end{array}$$

for all morphisms $f \in \mathfrak{C}(A,B)$. A natural isomorphism is a natural transformation where each of the τ_A is an isomorphism.

Strict monoidal category

A strict monoidal category consists of:

- a category $\mathfrak{C}$;
- a unit object $I \in ob(\mathfrak{C})$;
- a bifunctor $- \otimes - : \mathfrak{C} \times \mathfrak{C} \longrightarrow \mathfrak{C}$,

satisfying

- associativity: for each triple of objects A, B, C of $\mathfrak{C}$, $A \otimes (B \otimes C) = (A \otimes B) \otimes C$; for each triple of morphisms f, g, h of $\mathfrak{C}$, $f \otimes (g \otimes h) = (f \otimes g) \otimes h$;
- unit law: for each object A of $\mathfrak{C}$, $A \otimes I = A = I \otimes A$; for each morphism f of $\mathfrak{C}$, $f \otimes 1_I = f = 1_I \otimes f$.

Strict symmetric monoidal category

A strict monoidal category $\mathfrak{C}$ is symmetric if it is equipped with a natural isomorphism

$$\sigma_{A,B} : A \otimes B \to B \otimes A$$

for all objects A, B, C of $\mathfrak{C}$ satisfying:

$$\sigma_{B,A} \circ \sigma_{A,B} = 1_{A \otimes B}, \ \ \sigma_{A,I} = 1_A, \ \ (1_B \otimes \sigma_{A,C}) \circ (\sigma_{A,B} \otimes 1_C) = \sigma_{A,B \otimes C}.$$

Self-dual strict compact closed category

A self-dual strict compact closed category is a strict symmetric monoidal category $\mathfrak{C}$ such that for each object A of $\mathfrak{C}$, there exists two morphisms

$$\epsilon_A : A \otimes A \to I, \;\; \eta_A : I \to A \otimes A$$

satisfying:

$$(\epsilon_A \otimes 1_A) \circ (1_A \otimes \eta_A) = 1_A, \;\; (1_A \otimes \epsilon_A) \circ (\eta_A \otimes 1_A) = 1_A.$$

Commutative Semiring

A commutative semiring is a set $\mathcal{S}$ equipped with addition $+$ and multiplication $\cdot$, such that:

- $(\mathcal{S}, +)$ is a commutative monoid with identity element 0:

$$(a+b)+c = a+(b+c), \quad 0+a = a+0 = a, \quad a+b = b+a$$

- $(\mathcal{S}, \cdot)$ is a commutative monoid with identity element 1:

$$(a \cdot b) \cdot c = a \cdot (b \cdot c), \quad a \cdot b = b \cdot a, \quad 1 \cdot a = a \cdot 1 = a$$

- Multiplication left and right distributes over addition:

$$a \cdot (b+c) = (a \cdot b) + (a \cdot c), \quad (a+b) \cdot c = (a \cdot c) + (b \cdot c)$$

- Multiplication by 0 annihilates elements in $\mathcal{S}$:

$$0 \cdot a = a \cdot 0 = 0$$

3.2. Qufinite ZX_Δ-calculus as a Compositional Model of Consciousness

In this section, we give a graphical calculus for processes which we call qufinite ZX_Δ-calculus which has a presentation in terms of diagrammatic generators and rewriting rules. Throughout this section, $\mathbb{N} = \{0, 1, 2, \cdots\}$ is the set of natural numbers, $2 \leq d \in \mathbb{N}$, $\oplus$ is the modulo d addition, $\mathcal{S}$ is an arbitrary commutative semiring. All the diagrams are read from top to bottom.

3.2.1. Generators of Qufinite ZX_Δ-calculus

First we give all the generators for the qufinite ZX_Δ calculus.

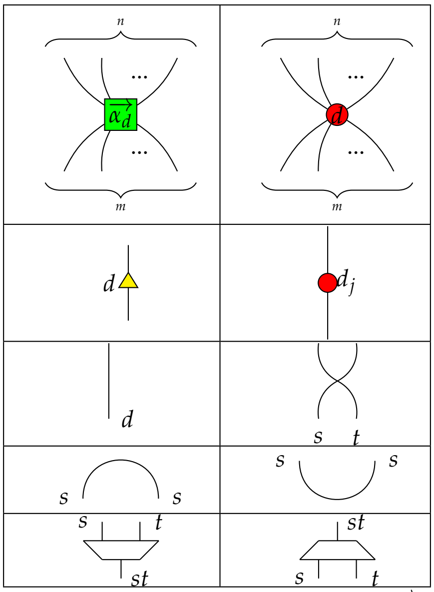
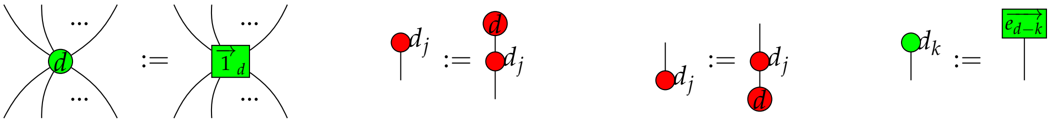
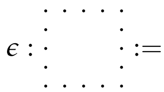


Table 1. Generators of qufinite ZX_Δ-calculus, where $m, n \in \mathbb{N}, \overrightarrow{\alpha_d} = (a_1, \cdots, a_{d-1}), a_i \in \mathcal{S}, i \in \{1, \cdots, d-1\}, j \in \{0, 1, \cdots, d-1\}, s, t \in \mathbb{N}\backslash\{0\}$.

For simplicity, we make the following conventions:



and



where $\overrightarrow{1}_d = (\overbrace{1, \cdots, 1}^{d-1}), j \in \{0, 1, \cdots, d-1\}, k \in \{1, \cdots, d-1\}, \overrightarrow{e_{d-k}} = (\overbrace{\underbrace{0, \cdots, 1}_{d-k}, \cdots, 0}^{d-1})$, ϵ represents an empty diagram.

3.2.2. Rules of Qufinite ZX_Δ-calculus

Now we give some rewriting rules for qufinite ZX_Δ-calculus which specify the generators.

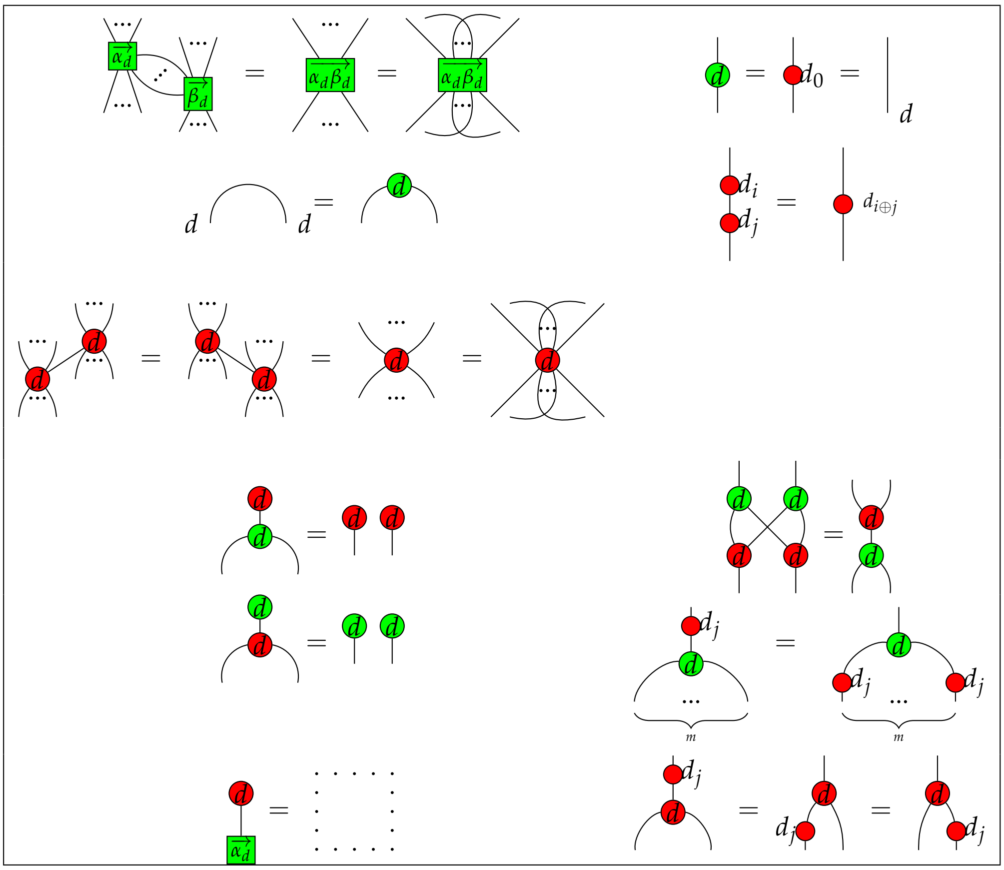


Figure 1. Qufinite ZX_Δ-calculus rules I, where $\overrightarrow{\alpha_d} = (a_1, \cdots, a_{d-1}), \overrightarrow{\beta_d} = (b_1, \cdots, b_{d-1}), \overrightarrow{\alpha_d \beta_d} = (a_1 b_1, \cdots, a_{d-1} b_{d-1}), a_k, b_k \in \mathcal{S}, k \in \{1, \cdots, d-1\}, j \in \{0, 1, \cdots, d-1\}, m \in \mathbb{N}$.

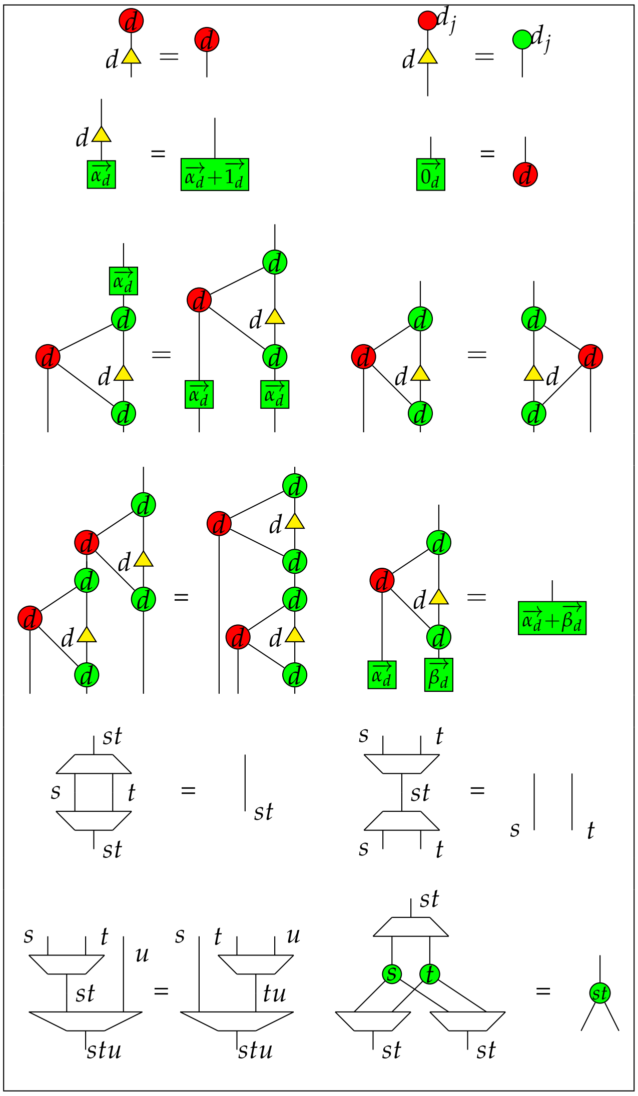
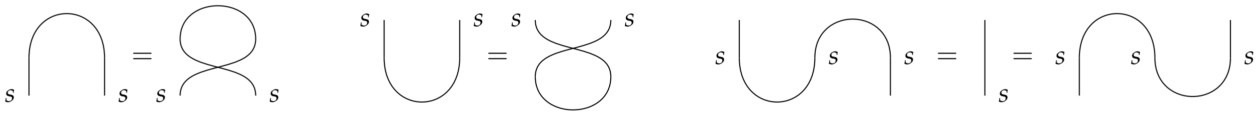
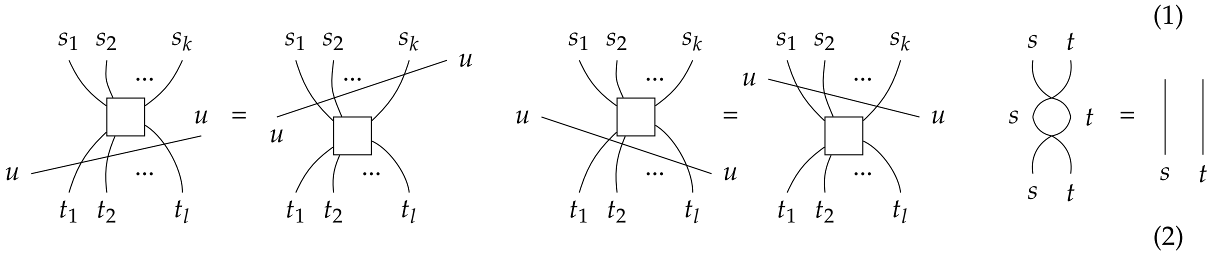


Figure 2. Qufinite ZX_Δ-calculus rules II, where $\overrightarrow{1}_d = \overbrace{(1,\cdots,1)}^{d-1}$, $\overrightarrow{0}_d = \overbrace{(0,\cdots,0)}^{d-1}$, $\overrightarrow{\alpha_d} = (a_1,\cdots,a_{d-1})$, $\overrightarrow{\beta_d} = (b_1,\cdots,b_{d-1})$, $a_k, b_k \in \mathcal{S}$, $k \in \{1,\cdots,d-1\}$, $j \in \{1,\cdots,d-1\}$, $s,t,u \in \mathbb{N}\backslash\{0\}$.

Also we have the structure rules for a self-dual compact closed category:

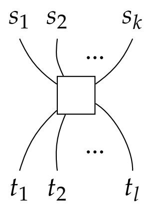


(1)



(2)

where

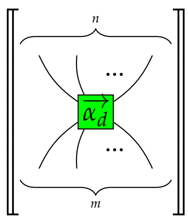


is an arbitrary diagram in the qufinite ZX_Δ-calculus.

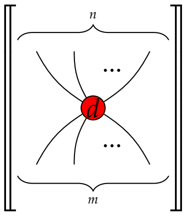
From that, the strict compact closed category $\mathfrak{C}$ is defined. The objects of $\mathfrak{C}$ are all the positive integers, and the monoidal product on objects are multiplication of integer numbers. Denote the set of generators listed in Table 1 as G. Let $\mathcal{C}[G]$ be a free monoidal category generated by G in the following way: any two diagrams D_1 and D_2 are placed side-by-side with D_1 on the left of D_2 to form the monoidal product on morphisms $D_1 \otimes D_2$, or the outputs of D_1 connect with the inputs of D_2 when their types all match to each other to form the sequential composition of morphisms $D_2 \circ D_1$. The empty diagram is a unint of parallel composition and the diagram of a straight line is a unit of the sequential composition. Denote the set of rules listed in Figure 1, Figure 2, (1) and (2) by R. One can check that rewriting one diagram to another diagram according to the rules of R is an equivalence relation on diagrams in $\mathcal{C}[G]$. We also call this equivalence as R, then the quotient category $\mathfrak{C} = \mathcal{C}[G]/R$ is a strict compact closed category. The qufinite ZX-calculus is seen as a graphical calculus based on the category $\mathfrak{C}$.

3.2.3. Standard interpretation of qufinite ZX_Δ-calculus

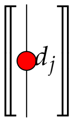
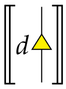
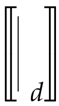
Let $\mathbf{Mat}_{\mathcal{S}}$ be the category whose objects are non-zero natural numbers and whose morphisms $M : m \to n$ are $n \times m$ matrices taking values in a given commutative semiring $\mathcal{S}$. The composition is matrix multiplication, the monoidal product on objects and morphisms are multiplication of natural numbers and the Kronecker product of matrices respectively. We give a standard interpretation $[\![\cdot]\!]$ for the qufinite ZX_Δ-calculus diagrams in $\mathbf{Mat}_{\mathcal{S}}$:

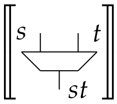
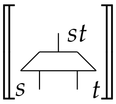
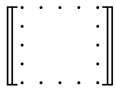


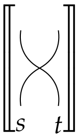
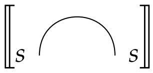
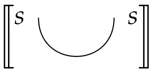
$$= \sum_{i=0}^{d-1} a_j \left|i\right\rangle^{\otimes m} \left\langle i\right|^{\otimes n}, a_0 = 1, a_i \in \mathcal{S},$$



$$= \sum_{\substack{0\le i_1,\cdots,i_m,j_1,\cdots,j_n\le d-1\\ i_1+\cdots+i_m\equiv j_1+\cdots+j_n (mod\ d)}} |i_1,\cdots,i_m\rangle\langle j_1,\cdots,j_n|,$$

 $= \sum_{i=0}^{d-1} |i\rangle\langle i\oplus j|,$  $= |0\rangle\langle 0| + \sum_{i=1}^{d-1}(|0\rangle+|i\rangle)\langle i|,$  $= \sum_{i=0}^{d-1}|i\rangle\langle i|,$

 $= \sum_{k=0}^{s-1}\sum_{l=0}^{t-1}|kt+l\rangle\langle kl|,$  $= \sum_{k=0}^{st-1}\left|[\frac{k}{t}]\right\rangle\left|k-t[\frac{k}{t}]\right\rangle\langle k|,$  $=1,$

 $= \sum_{k=0}^{s-1}\sum_{l=0}^{t-1}|kl\rangle\langle lk|,$  $= \sum_{i=0}^{s-1}|i\rangle|i\rangle,$  $= \sum_{i=0}^{s-1}\langle i|\langle i|,$

$$[\![D_1\otimes D_2]\!] = [\![D_1]\!]\otimes[\![D_2]\!], \quad [\![D_1\circ D_2]\!] = [\![D_1]\!]\circ[\![D_2]\!],$$

where $s,t\in\mathbb{N}\backslash\{0\}$, $\langle i| = \overbrace{(\underbrace{0,\cdots,1}_{i+1},\cdots,0)}^{d}$, $|i\rangle = (\overbrace{(\underbrace{0,\cdots,1}_{i+1},\cdots,0)}^{d})^T$, $i\in\{0,1,\cdots,d-1\}$, and $[r]$ is the integer part of a real number r.

One can verify that the qufinite ZX_Δ-calculus is sound in the sense that for any two diagrams $D_1, D_2\in\mathfrak{C}$, $D_1 = D_2$ must imply that $[\![D_1]\!] = [\![D_2]\!]$.

3.2.4. Interpretation in terms of Consciousness model

The qufinite ZX_Δ-calculus, as a compositional theory for processes, actually has the "other dependent" feature which is one of the key feature of consciousness described above: the generators of qufinite ZX_Δ-calculus are not specified by themselves. In contrast, each of them are specified by the others, as it is possible to observe in the rewriting rules. The non-generator diagrams (as processes) are produced by thus being dependent on the generators. Furthermore, each diagram with output but without input will represent a consciousness state, and a general diagram represents some sort of consciousness process. Compositionally, sequential composition of two diagrams represents two successive consciousness processes happened one after another, while parallel composition of two diagrams represents two successive consciousness processes happened simultaneously. This justifies our use of the framework of qufinite ZX_Δ-calculus as a compositional model for consciousness under the approach/assumption of "consciousness as fundamental".

4. The Feature Binding

4.1. Unbinding the Binding Problem

Unity of consciousness is commonly associated with, but not reduced to, the binding properties of perception, neurons and brain regions. The so-called binding problem [32,33]. In a materialistic and reductionist formulation, the problem is stated as the need of a neural mechanism from which unified experiences emerge by a combination of separated elements, e.g. how to bind different features of a perceived object, such as colour and shape (Figure 3A). This corresponds to a low level of the **combination aspect** of the binding problem [33]. A high level would be another instance where these combined objects are thought to be bound with other background features, as well as emotional feelings to create one single unified phenomenal subjective experience [34]. This is the phenomenal unity of the combination problem [32,34], the intuition that regardless of the distinct neural paths, our experiences are integrate-wholes. This is the subjective or **phenomenal binding problem**. According to this construction, the subjective unified experience seems inconsistent with the many separate brain activities from which the whole experience is thought to emerge, there is not a single module or region where that integration may take place [33,35]. Furthermore, there is a **segregation aspect** of binding, i.e. having a blue square and a red triangle how one can recognize that the blue belongs to the square and the red to the triangle and not vice versa [33,35,36]. In other words, how sensory inputs are allocated to recognize "discrete objects" and not just a collection of separated colours and shapes (Figure 3B). Mechanistically, the question is how cells and neurons recognize that they are being activated either by different objects or by only one complex object. This is a discrimination issue, the **feature binding problem**, associated with distinctive properties of experience. Both, the combination and segregation, are considered part of one and the same binding problem. Hence, the question becomes to understand how properties of objects are first combined, then segregated to later being recombined or unified in one whole experience together with all the extra features of the experienced context.

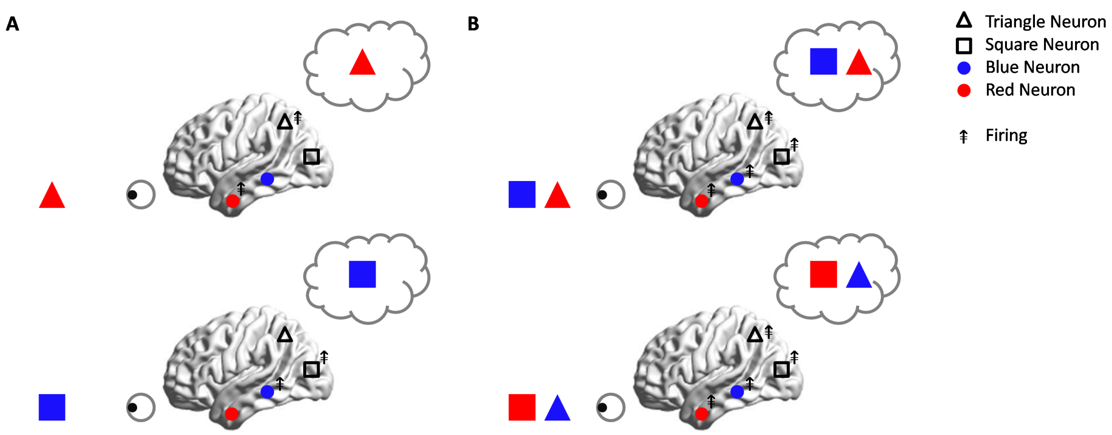


Figure 3. The Feature Binding Problem. A) The combining aspect of binding is about how cells and neurons integrate different features, for instance, shape and colour. Based on the assumption of independent neurons or modular brain regions processing different features, the integration may take place if neurons corresponding to each feature are simultaneously activated. In the upper figure, a red triangle activates the triangle shape neuron and the colour red neuron, in lower figure an example for blue square. B) The problem arises when the triangle and the square are presented simultaneously, first with one combination of colours (top) and then inverting them (bottom). In both cases, all the neurons or regions are activated at the same time. Therefore, the question becomes how the brain can segregate each colour to the corresponding shape. One alternative is a combination coding, such as new layers of neurons would bind the previous ones. Another is binding by synchrony, i.e. neurons with correlated firing would bind features together. Unfortunately, for these and other possible solutions, there are theoretical and empirical concerns. The main objection is indeed the original assumption of independent processing features or modular paths.

Unfortunately, these two aspects and versions of the problem are not always differentiated, making the discussion sometimes ambiguous [37]. In this line, Revonsuo and others clearly stated different related binding problems, some associated with consciousness and others not. At least three levels are distinguished: phenomenal, neural and cognitive [32]. In turns that Feldman describes four binding problems [33]: Coordination, Subjective, Feature and Variable binding, all of them regarded to different tasks, time scales and brain circuits. According to these definitions, different models are trying to solve the questions about combination and segregation, mainly regarding feature binding. Some of them are combination coding, population coding, binding by synchrony [38], and feature integration model [39]. However, none of them is theoretically and empirically satisfactory, they seem acting at different hierarchical levels and stages of perception [35,36], leaving important open questions that need to be reconsidered in light of new compositional and diagrammatic concepts presented in previous and next section (Section 4.2). To avoid any confusion, our focus will be the feature binding regarding the combination and segregation aspects, while the subjective/phenomenal problem is leaved for a future discussion.

4.2. A Toy model for Feature Binding

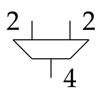
The original version of feature binding problem comes from the apparent modular codification observed in neurons of the primary visual cortex, which seemed to respond selectively to single features, such as colour or shape. It creates the paradox that any original combination or relationships between stimuli features are lost when decomposed into independent modules, and the need of recombination somewhere later [37]. Nevertheless, this modular independence is misleading and disconfirmed by modern experiments [40–42]. The same neuron is activated by multiple stimuli and features, and indeed it is also concurrently selective to combinations of features [37]. The brain works in parallel where different circuits and tasks are performed simultaneously. Therefore, the unbinding, the separations of the causes of an input, seems more relevant than binding itself [33].

Unity of consciousness may be naturally described as a result of composition. In this sense, an application of our framework is given as a tentative solution for a toy model of the feature binding problem, which additionally is seen as a part of that unity. Assume there are two choices for colour: green and red; and two choices for shape: square and triangle. The scenario of the feature binding is as follows: given a shape and a colour at the same time, say, square and red, one can perceive a combined object– red square; given two combined objects, say green square and red triangle, one can perceive the two objects simultaneously. Then the binding problem is simply restated as: what is the mechanism/transformation for realising the above scenario?

One alternative to solve this question borrows an idea from quantum theory. Firstly, the two shapes are encoded into a two-state system A_2: $square \mapsto |0\rangle, triangle \mapsto |1\rangle$. Secondly, the two colours into another two-state system B_2: $green \mapsto |0\rangle, red \mapsto |1\rangle$. Thirdly, the combined objects are described as a four-state system C_4: $green\ square \mapsto |0\rangle, red\ square \mapsto |1\rangle, green\ triangle \mapsto |2\rangle, red\ triangle \mapsto |3\rangle$. Then the binding mechanism is realised by the following linear map:

$$\begin{array}{ccc} L : A_2 \otimes B_2 & \longrightarrow & C_4 \\ |00\rangle & \mapsto & |0\rangle \\ |01\rangle & \mapsto & |1\rangle \\ |10\rangle & \mapsto & |2\rangle \\ |11\rangle & \mapsto & |3\rangle \end{array}$$

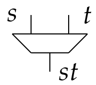
Here two combined objects presented at the same time are modelled by the superposition of the two states representing the two objects. For example, a green square and red triangle shown simultaneously are represented as $|00\rangle + |11\rangle$. Then one can check that the linear map L is the mechanism that realises the binding: given green square and red triangle simultaneously, a green square and a red triangle is obtained simultaneously via L; the other cases are similar. As one can see clearly from section 3.2.3, the linear map L is just the standard interpretation of the following generator in the ZX_Δ-calculus:



This toy model is generalised to a generic situation:

$$\begin{array}{ccc} L : A_s \otimes B_t & \longrightarrow & C_{st} \\ |ij\rangle & \mapsto & |it + j\rangle \end{array}$$

diagrammatically represented by the generator:



where $0 \leq i \leq s-1, 0 \leq j \leq t-1$.

5. Conclusions

Across these pages, the framework of qufinite ZX_Δ-calculus was introduced and a preliminary application is presented by providing a solution to a toy model of the feature biding problem. The framework is based on arbitrary commutative semirings as a compositional model of consciousness, making the emphasis on its potential use for mathematical and structural studies of consciousness. The philosophy behind our framework is taken from the Yogacara school of Buddhism, assuming consciousness as fundamental and characterizing the main feature of consciousness as other-dependent. Therefore, generators and processes become abstract mathematical structures, independent of their realizations. Moreover, our approach is related, almost tautologically, to quantum theory, since the qufinite ZX_Δ-calculus is a unification of all dimensional qudit ZX-calculus, which are graphical languages for quantum theory when interpreted in Hilbert spaces. Thus, part of the reconstruction goal pursued by conscious agent model is reached here for free, only invoking phenomenal aspects. In other words, our approach to consciousness processes and quantum theory share a similar mathematical structure. Because of its other-dependent feature and sufficient generality, our framework may pave a good way for further research on scientific study of consciousness. One obvious further step would be to tackle the phenomenal binding problem, as well as developing a comparison with the conscious agent model [3,27]. Furthermore, it is worth trying to generalise the qufinite ZX_Δ-calculus to infinite dimensional case, from which standard quantum mechanics might be recovered. These conclusions and interpretations may also inspire great debate and we are willing to motivate that discussions.

Author Contributions: Conceptualization, CMS and QW; investigation CMS and QW; writing-original draft preparation, CMS; writing-review and editing, CMS and QW; visualization, CMS and QW.

Funding: CMS is funded by Comisión Nacional de Investigación Ciencia y Tecnología (CONICYT) through Programa Formacion de Capital Avanzado (PFCHA), Doctoral scholarship Becas Chile: CONICYT PFCHA/DOCTORADO BECAS CHILE/2016 - 72170507. QW is supported by AFOSR grant FA2386-18-1-4028.

Acknowledgments: The authors appreciate valuable feedback and discussions from XX, XX2 and XX.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lusthaus, D. Buddhist Phenomenology, first ed.; Routledge Curzon, 2002; p. 632. doi:10.4324/9781315870687.
2. Makeham, J. Introduction. In Transforming Consciousness: Yogacara Thought in Modern China; Makeham, J., Ed.; Oxford University Press, 2014. doi:10.1093/acprof:oso/9780199358120.001.0001.
3. Fields, C.; Hoffman, D.D.; Prakash, C.; Singh, M. Conscious agent networks: Formal analysis and application to cognition. Cognitive Systems Research **2018**, 47, 186–213. doi:10.1016/j.cogsys.2017.10.003.
4. Evan Thompson. Mind in Life; Harvard University Press, 2007.
5. Searle, J.R. Consciousness. Annual Review of Neuroscience **2000**, 23, 557–578. doi:10.1146/annurev.neuro.23.1.557.
6. Coecke, B., Ed. New Structures for Physics, lectures n ed.; Springer Berlin Heidelberg, 2011; p. 1034. doi:10.1007/978-3-642-12821-9.
7. Abramsky, S.; Coecke, B. A categorical semantics of quantum protocols. 19th Annual IEEE Symposium on Logic in Computer Science (LICS'04), 2004, pp. 415–425.
8. Coecke, B.; Kissinger, A. Picturing Quantum Processes. A first Course in Diagrammatic reasoning; Cambridge University Press: Cambridge, UK, 2017. doi:10.1017/9781316219317.

9. Kissinger, A.; Uijlen, S. A categorical semantics for causal structure. 32nd Annual ACM/IEEE Symposium on Logic in Computer Science (LICS), 2017, pp. 1–12. doi:10.1109/LICS.2017.8005095.
10. Pinzani, N.; Gogioso, S.; Coecke, B. Categorical Semantics for Time Travel **2019**. [1902.00032].
11. Kissinger, A.; Hoban, M.; Coecke, B. Equivalence of relativistic causal structure and process terminality **2017**. [1708.04118].
12. Coecke, B.; Sadrzadeh, M.; Clark, S. Mathematical Foundations for a Compositional Distributional Model of Meaning. Linguistic Analysis **2010**, 36, 345–384, [arXiv:1003.4394v1].
13. Bolt, J.; Coecke, B.; Genovese, F.; Lewis, M.; Marsden, D.; Piedeleu, R. Interacting Conceptual Spaces I : Grammatical Composition of Concepts. ArXiv **2017**, [arXiv:1703.08314].
14. Signorelli, C.M.; Dundar-Coecke, S.; Wang, V.; Coecke, B. Cognitive Structures of Space-Time.
15. Coecke, B. An Alternative Gospel of Structure: Order, Composition, Processes. In Quantum Physics and Linguistics:A Compositional, Diagrammatic Discourse; Heunen, C.; Sadrzadeh, M.; E. Grefenstette., Eds.; Oxford University Press, 2013; [1307.4038]. doi:10.1093/acprof:oso/9780199646296.003.0001.
16. Coecke, B.; Duncan, R.; Kissinger, A.; Wang, Q. Generalised Compositional Theories and Diagrammatic Reasoning. In Quantum Theory: Informational Foundations and Foils. Fundamental Theories of Physics.; Chiribella, G.; Spekkens, R., Eds.; Springer, 2016; Vol. 181, pp. 309–366, [1506.03632]. doi:10.1007/978-94-017-7303-4_10.
17. Prentner, R. Consciousness and topologically structured phenomenal spaces. Consciousness and Cognition **2019**, 70, 25–38. doi:10.1016/j.concog.2019.02.002.
18. Seth, A.K. Consciousness: The last 50 years (and the next). Brain and Neuroscience Advances **2018**, 2, 239821281881601. doi:10.1177/2398212818816019.
19. Thomas Nagel. What is it like to be a bat? The Philosophical Review **1974**, 83, 435–450.
20. Chalmers, D. The puzzle of conscious experience. Scientific American **1995**, 273, 80–86.
21. Crick, F.; Koch, C. Consciousness and neuroscience. Cerebr **1998**, 8, 97–1007. doi:10.1093/cercor/8.2.97.
22. Vacariu, G. The Mind-Body Problem Today. Open Journal of Philosophy **2011**, 01, 26–34. doi:10.4236/ojpp.2011.11005.
23. Mulder, D.H. Objectivity.
24. Anderson, P.W. More Is Different. Science **1972**, 177, 393–396. doi:10.1126/science.177.4047.393.
25. Mazzocchi, F. Complexity in biology. EMBO reports **2008**, 9, 10–14.
26. beim Graben, P. Contextual Emergence in Neuroscience. Closed Loop Neuroscience **2016**, pp. 171–184. doi:10.1016/B978-0-12-802452-2.00012-3.
27. Hoffman, D.D.; Prakash, C. Objects of consciousness. Frontiers in Psychology **2014**, 5, 1–22. doi:10.3389/fpsyg.2014.00577.
28. Kochumuttom, T.A. A buddhist Doctrine of Experience. A New Translation and Interpretation of the Works of Vasubandhu t Motilal Banarsidass: Delhi, 1999.
29. Xuanzang. Cheng Weishi Lun. In no. 1585 of Taisho shinshu daizokyo; Vol. 31.
30. Xuanzang, Tat Wei.; Vasubandhu. Cheng Wei Shi Lun; The Doctrine of Mere-Consciousness.; Ch'eng Wei-shih Lun Publication Committee: Hong Kong, 1973.
31. Coecke, B.; Duncan, R. Interacting quantum observables: Categorical algebra and diagrammatics. New Journal of Physics **2011**, 13, [arXiv:0906.4725v3]. doi:10.1088/1367-2630/13/4/043016.
32. Revonsuo, A.; Newman, J. Binding and consciousness. Consciousness and cognition **1999**, 8, 123–127. doi:10.1006/ccog.1999.0393.
33. Feldman, J. The neural binding problem(s). Cognitive Neurodynamics **2013**, 7, 1–11. doi:10.1007/s11571-012-9219-8.
34. Bayne, T.; Chalmers, D.J. What is the unity of consciousness? The Unity of Consciousness: Binding, Integration, and Dissociation **2012**, pp. 1–41. doi:10.1093/acprof:oso/9780198508571.003.0002.
35. Velik, R. From simple receptors to complex multimodal percepts: A first global picture on the mechanisms involved in perceptual binding. Frontiers in Psychology **2012**, 3. doi:10.3389/fpsyg.2012.00259.
36. Treisman, A. Solutions to the binding problem: Progress through controversy and convergence. Neuron **1999**, 24, 105–125. doi:10.1016/S0896-6273(00)80826-0.
37. Di Lollo, V. The feature-binding problem is an ill-posed problem. Trends in Cognitive Sciences **2012**, 16, 317–321. doi:10.1016/j.tics.2012.04.007.

38. Engel, A.K.; Singer, W. Temporal binding and the neural correlates of sensory awareness. Trends in Cognitive Sciences **2001**, 5, 16–25. doi:10.1016/S1364-6613(00)01568-0.
39. Treisman, A.M.; Gelade, G. A Feature-Integration of Attention. Cognitive Psychology **1980**, 12, 97–136.
40. Gegenfurtner, K.R.; Kiper, D.C.; Fenstemaker, S.B. Processing of color, form, and motion in macaque area V2. Visual Neuroscience **1996**, 13, 161–172. doi:10.1017/S0952523800007203.
41. Mandelli, M.J.F.; Kiper, D.C. The local and global processing of chromatic glass patterns. Journal of Vision **2005**, 5, 405–416. doi:10.1167/5.5.2.
42. Nakayama, K.; Martini, P. Situating visual search. Vision Research **2011**, 51, 1526–1537. doi:10.1016/j.visres.2010.09.003.