A neural coarse graining theory of consciousness

Acer Y.C. Chang*[1], Martin Biehl†[1], and Ryota Kanai‡[1]

[1]ARAYA, Inc., Tokyo, Japan

October 26, 2018

Contents

*acercyc@araya.org
†martin@araya.org
‡kanair@araya.org

Abstract

Neural systems process information through different levels of organisation in a hierarchical manner. Information at lower levels is finer-grained and can be coarse-grained for higher level computation. However, one is aware of information processed only at specific levels. Theorists have addressed this issue. For example, the intermediate level theory of consciousness suggests that the intermediate level seems to be privileged with respect to consciousness. It is true that we do not experience information processed by individual neurons which is always highly noisy. Besides, we have no conscious experience from interpersonal activities albeit massive interactions among individuals. Instead, neurophysiological evidence has been showing that conscious experience tends to co-vary with information encoded in coarse-grained neural states such as neural population codes. We argue that the neural states within the scope of the informational closure determine the contents of consciousness and brain processes outside apart from the representations of that level remain unconscious. This argument suggests a distinction between conscious and unconscious processing and provides a generic computational framework. Finally, using the deep learning network, we can measure informational closure in deep hidden layers. Our preliminary results show that informational closure representation emerged after learning. We further decoded the information from the representation and compared it with human conscious perception.

Keywords:

Keywords: theory of Consciousness, informational closure, neural coarse-graining, level of analysis

1 Introduction

Imagine you are a neuron in Alice's brain. Your daily work is to collect neurotransmitters through dendrites from other neurons, accumulate membrane potential, and finally send signals to other neurons through action potentials along axons. However, you have no idea that you are one of neurons in Alice's supplementary motor area and involved in many motor control processes for Alice's action, for example, grabbing a cup. You are ignorant of intentions, goals, and motor plans that Alice has at every moment even though you are part of the physiological substrate responsible for all those actions. A similar story also applies to Alice's conscious mind. To grab a cup, for example, Alice is conscious of her intention and visuosensory experience of this action. However, her conscious experience does not show the dynamic of your membrane potential and about action potentials you send to other neurons every second. That is, information in Alice's conscious experience does not contain information involved in your information processing.

[[maybe need a transition sentence here]]

It seems to be true that we are unable to consciously access information processed at every scale in the neural system. Only information at certain levels of coarse-graining can become conscious contents. At microscopic levels of neural systems, information processing in individual neurons is highly noisy and state transitions of neurons are extremely stochastic (Goldwyn & Shea-Brown, 2011; White *et al.*, 2000). However, instead of stochastic action potentials and fluctuation of random noise at microscopic levels, what we are aware of in our conscious mind shows astonishing stability and robustness against the ubiquitous noise in the neural system (Mathis & Mozer, 1995). Likewise, conscious experience does not associate with states of the whole brain level because we have known that some parts of the neural system contribute very little to conscious experience (the cerebellum for example (Lemon & Edgley, 2010)). Additionally, we do not experience "super-consciousness" when we have massive inter-individual interactions with other human-beings. Even though evidence has shown several types of collective information processing at human social groups levels exist (Malone & Bernstein, 2015), information processed at this level seems to depart from our individual conscious contents.

In this article, we propose a new theory of consciousness, calledThe Neural Coarse-graining Theory of Consciousness (NCGT). This proposal states that consciousness does not rest on particular neural architectures or locations in the neural systems. Instead, we argue that consciousness correlates with information processing at a certain coarse-grained level where the neural systems create a high degree of non-trivial informational closure (NTIC). We further postulate that the *level* of consciousness corresponds to the degree of NTIC and the *contents* of consciousness corresponds to the state of the NTIC process (for the level versus content of consciousness see Laureys (2005); Overgaard & Overgaard (2010)).

Our theory rests on two core concepts: non-trivial informational closure and neural coarse-graining. In this article, we first introduce non-trivial informational closure and its importance to information processing for human scale agents. We next argue that through coarse-graining the human neural system can achieve non-trivial informational closure at a specific coarse-grained level. We then propose that the state of the informational closure correlates with the state of consciousness for both levels and contents which can be explained, quantified, and predicted computationally by this theory. We also address how this theory can concisely explain empirical results from previous studies and reconcile several current major theories of consciousness. Finally, we provide predictions and testable approaches for this theory.

[[Do I need a figure?, Something like this one?]]

To Be Modified

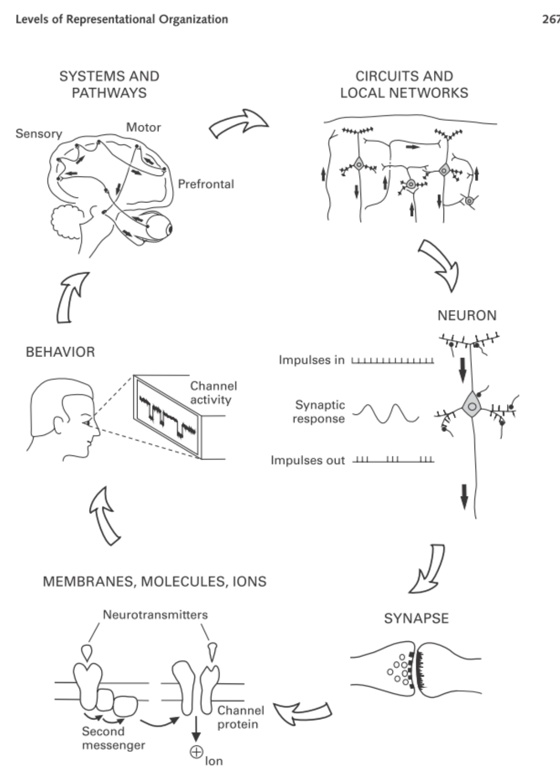


Figure 10.1
Brain function is characterized by several aggregate levels of organization, rather loosely defined at different scales of measurement, beginning at the molecular level of ion channels and other proteins up to the highest level of cognition and behavior mediated by connected, macroscopic brain systems. Adapted from Shepherd (1990). Note the circularity in a person looking at his own ion channel activity displayed on an oscilloscope.

Pennartz (2017)

Figure 1: xxxx

2 Non-trivial informational closure

I am still not sure we should go for NTIC without describe my intuition about consciousness and closure? Or we should start with my intuition. Or we should postpone to explain the intuition in Chapter 4?

The notion of non-trivial informational closure (NTIC) is introduced by Bertschinger *et al.* (2006). The concept of closure is closely related to system identification in systems theory. One can define a system from its environment by computing the closureness of the system (Maturana & Varela, 1991; Rosen, 1991; Pattee, 2012; Luhmann, 1995). The closureness can be further quantified by information theory.

Writing Materials

- Consciousness as an intrinsic information. Therefore, information should be self-define. This means the boundary of conscious content should be related to informational closure.
- [[should we describe the intuition of boundary problem here?]]
- ==
- *" The notion of closure plays a prominent role in systems theory where it is used to identify or define the system in distinction from its environment"*
  Bertschinger *et al.* (2006)
- *" Our theoretical interest concerns the type of system that is a unity for and by itself and*

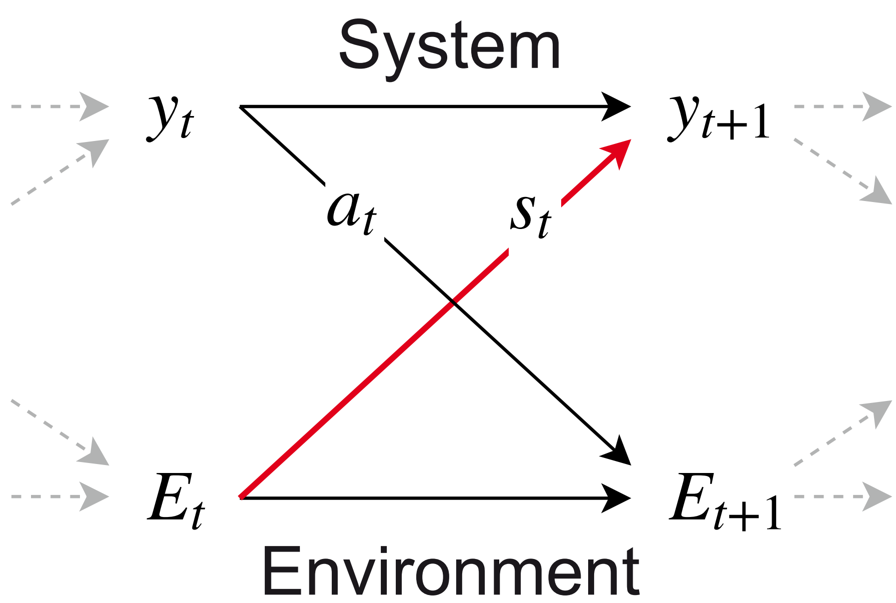


Figure 2: Figure adapted from Bertschinger *et al.* (2006)

not only for an external observer distinguishing some entity from the rest of the world. This requires a system that can be described as a whole without reference to its environment. In systems theory, this property is usually referred to as closure."
(Bertschinger *et al.*, 2006)

- *" These concepts of closure play an important role in the architecture of systems theory, because they are used to 1. define the system (in distinction to its environment) and to 2. explain the autonomy of the system."*
(Bertschinger *et al.*, 2006)
- *" Informational closure: The higher process is informationally closed, i.e., there is no information flow from the lower to the higher level. Knowledge of the microstate will not improve predictions of the macrostate."*
(PFANTE *et al.*, 2014, p. 4)

Writing Materials

Information flow J_t from the environment E to a system Y at time t can be defined as the conditional mutual information I between the system future state Y_{t+1} and the current environment current state E_t given the current system state Y_t

$$\begin{aligned} J_t(E \to Y) &:= I(Y_{t+1}; E_t | Y_t) \\ &= I(Y_{t+1}; E_t) - (I(Y_{t+1}; Y_t) - I(Y_{t+1}; Y_t | E_t)) \end{aligned} \tag{1}$$

Bertschinger *et al.* (2006) defines that a system is informational closured when information flow from the environment to the system is close to zero.

$$J_t(E \to Y) = 0 \tag{2}$$

Achieving information closure (minimising J_t) could be trivial, which means the environment and the system are entirely independent of each other.

$$I(Y_{t+1}; E_t) = 0 \quad \Rightarrow \quad J_t(E \to Y) = 0 \tag{3}$$

However, informational closure can be achieved non-trivially. In the non-trivial case, the system trying to encode information about the environmental dynamics can also be informational closured. This means

$$I(Y_{t+1}; E_t) > 0 \tag{4}$$

This also implies

$$I(Y_{t+1}; Y_t) - I(Y_{t+1}; Y_t|E_t) > 0 \tag{5}$$

And, non-trivial informational closure can be defined as

$$\begin{aligned} NTIC \quad &:= I(Y_{t+1}; Y_t) - I(Y_{t+1}; Y_t|E_t) \\ &= I(Y_{t+1}; E_t) - I(Y_{t+1}; E_t|Y_t) \end{aligned} \tag{6}$$

And, Maximising NTIC is equivalent to

$$\begin{aligned} &\text{maximise} \quad I(Y_{t+1}; Y_t) \qquad \text{and} \\ &\text{minimise} \quad I(Y_{t+1}; Y_t|E_t) \end{aligned} \tag{7}$$

This implies the system contains in itself the information of its own future and the self-predictive information is gained from the information about the environment. Therefore, to achieve NTIC, the system can internalizes and synchronises the dynamics of the environment, i.e. modelling the environment. Furthermore, having high degrees of NTIC entails having high predictive power about the environment. This gives biological agents a great functional and evolutional advantage to achieve NTIC in their information processing systems. Practically, Guttenberg *et al.* (2016) proposed that NTIC can be achieved by maximising the predictive power about the environment and the self-predictive information of the system concurrently.

Though the significant advantage for biological systems to achieve NTIC, to our knowledge, studies directly examining NTIC in biological systems are limited. However, recent studies have showed relevant properties in the visual system of salamanders. By examining the salamander retina, in two studies (Palmer *et al.*, 2015; Sederberg *et al.*, 2018) the authors found that the a large groups of neural populations of retinal ganglion cells encoded predictive information about external stimuli also had high self-predictive information about their own future states (see Figure 3), therefore, in line with the characteristic of NTIC. Therefore, the self-predictive information of the neural populations can guides prediction about external stimuli without reference to stimulus parameters explicitly. Moreover, the internal predictability can be generalised to different stimulus classes suggesting the tendency to internalise the information of environmental dynamics in the neural populations. **[[Add Some ending sentences]]**

To Be Modified

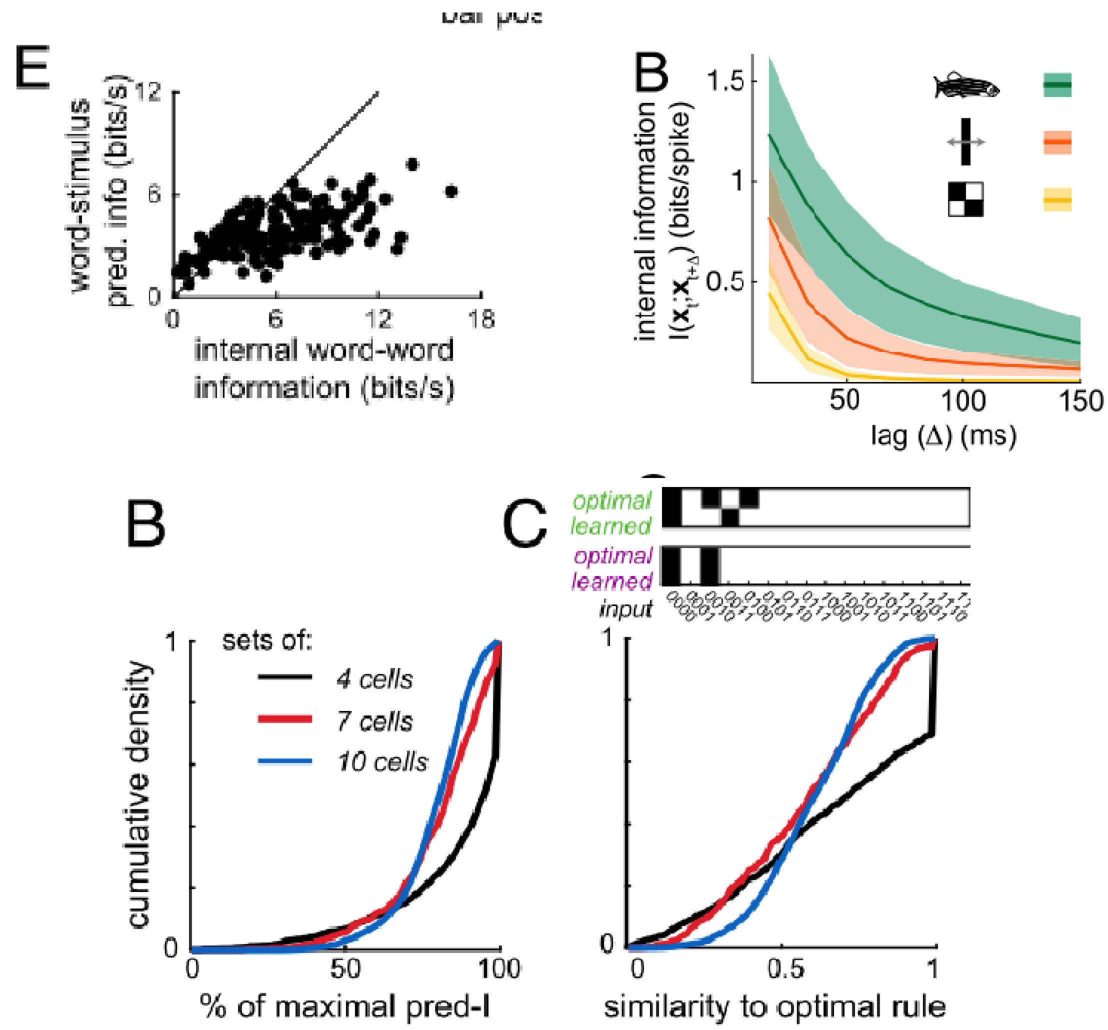


Figure 3: Figure adapted from Sederberg *et al.* (2018)

2.1 Informational closure and level identification

maybe not necessary

2.2 Biological Creature may try to achieve closure even the environment is not

Maybe not necessary

3 Neural coarse-graining

However, to achieve information closure (so as NTIC) in highly stochastic systems (e.g., the neural system) is challenging. The current states of the system have low predictability of its future states due to the stochasticity. Information processing in the neural system is suffered from multiple noise sources throughout, including sensory, cellular, electrical, and synaptic noises. Previous studies have shown that states of neurons have large trial-to-trial variability at the cellular level. That is, the same pattern of activity never occurs twice even when the same stimulus is presented. Neurons are inevitably subject to thermal fluctuations and other physical noises (Faisal *et al.*, 2008). Furthermore, environmental information is always only partial observable to individual sensors and neurons which only receive little amount of information from the environment. Therefore, self and environmental predictive information is difficult to be represented at the cellular level.

@ We argues that the human neural system to achieve NTIC through coarse-graining.

Writing Materials

- One way to encode a stable information in neural system is to encode the information in the coarse-grained variables.
- *" Coarse-graining: a method of reducing the complexity of a system by treating groups of atoms/molecules as single quasi-particles CG: coarse-grained"*
- *" a number of different microstates,First, typically a number of different microstates, will realize the same value of the macro-variables of upper level causal theories."*
- This is a common strategy that we can see in large number of previous physiological studies.
- neural system use population coding and temporal coding to fight against noise.
- Therefore, even though the low level information processing is noisy, it is possible that the neural system achieve NTIC at a coarse-grained level.
- This implies that the information about the environmental dynamic should be represented at a certain coarse-grained level rather then other levels.
- It is important to note that NTIC should not be a monotonic function of scale of coarse-graining. For the extreme case xxxxx
- We argues that the human neural system to achieve NTIC at a certain level of coarse-graining.
- The environment that human scale agents interact with is approximately govern by Newton's law and therefore the state transitions are nearly deterministic.

[Call For Help]: Is there any good ref I can cite?

- [[Need something here]]
- However, information processing at the microscopic levels in the neural system is noisy.
- *" neurons are noisy: the same pattern of activity never occurs twice, even when the same stimulus is presented."*
- And every individual sensor only receives a little amount of information about the environment.
- Therefore, the dynamic of microscopic level is highly stochastic
- Hence, to coarse-grain states of microscopic levels is necessary to represent and model the human-scale environment state and dynamics.
- Coarse-graining is necessary to resist noise
- *" The population will therefore be relatively insensitive to the loss of cells or noise in a small number of cells"*
  Eurich & Wilke (2000)
- More importantly, the information processing related to consciousness should not be situated at this level. Otherwise our conscious percept should be very noisy and unstable.

Writing Materials

3.1 Coarse-graining in neural system

- Coarse-graining can be defined as a function that maps multiple microscopic states to a macroscopic states.
- One way to encode a stable information in neural system is to encode the information in the coarse-grained variables.
- Studies have shown that coarse-graining is a common way that the neural system cope with inevitable noise at microscopic levels.
- We have known that the neural system uses various coding strategies to encode information at coarse-grained levels.
- *"Because the sequence of action potentials generated by a given stimulus varies from trial to trial, neuronal responses are typically treated statistically or probabilistically"*

- In the temporal domain, rate coding can be seem as coarse-graining on temporal events at the neuronal level Stein *et al.* (2005).
- contained in the firing rate of the neuron
- Firing rate of sensory neurons correlate stimulus intensities. (Kandel *et al.*, 2000)
- This coding strategy is highly robust against noise causing variation of interspike intervals.
- Another example is population coding.
- Population coding is
  the idea that nervous systems signal and compute parameters using large populations of imprecise neurons, rather than one or a few precise neurons.
- *" Population code (also ensemble code) denotes a code by which neural information is encoded in the spatiotemporal activity patterns of many neurons. "*
  Binder *et al.* (2009)
- In such case, single neurons are not informative. Instead, coarse-grained variables such as population averaging encode the information.
- Encoding information at coarse-grained levels allow the neural system ignores certain causal interactions relevant to fine-grained level of description. Price & Corry (2007)
- Note that, information processing related to consciousness should not be situated at this level. Otherwise our conscious percept should be very noisy and unstable.

Writing Materials

- Population coding
  - review ref
    - Stanley (2013)
    - Quian Quiroga & Panzeri (2009)
  - *" Averaging is used in many neural systems in which information is encoded as patterns of activity across a population of neurons that all subserve a similar function (for example, see REFs 140,141)"*
    140. Georgopoulos, A., Schwartz, A. & Kettner, R. Neuronal population coding of movement direction. Science 233, 14161419 (1986). 141. Lee, C., Rohrer, W. H. & Sparks, D. L. Population coding of saccadic eye movements by neurons in the superior colliculus. Nature 332, 357360 (1988).
  - *" Electrophysiological studies in the monkey have shown that behaviourally relevant signals are averaged not only across neuronal populations but also over time in the formation of a behavioural decision 156 ."*
    156. Gold, J. I. & Shadlen, M. N. Representation of a perceptual decision in developing oculomotor commands. Nature 404, 390394 (2000).
  - *"A distributed representation of information of this type is more robust to the effects of noise."*
  - *" Neural population code: the set of response features of a population of neurons that carry all information about the considered stimuli. These features consist of spatio-temporal sequences of action potentials distributed across neurons and/or time."*
  - The diverse response selectivity of sensory neurons
  - *" How a neural population represents information is partly determined by the diverse selectivity of individual neurons Shamir (2014)"*
- sensory hierarchy
  - sensory hierarchy also shows characteristic of coarse-graining
  - Due to partial information that sensors can receive
  - To form complex representation like shape, objects, coarse-graining is necessary.
  - Through coarse-graining, higher level cortical areas can integrate those partial informa-

might be confusing, need to think about the terminology here

To Be Modified

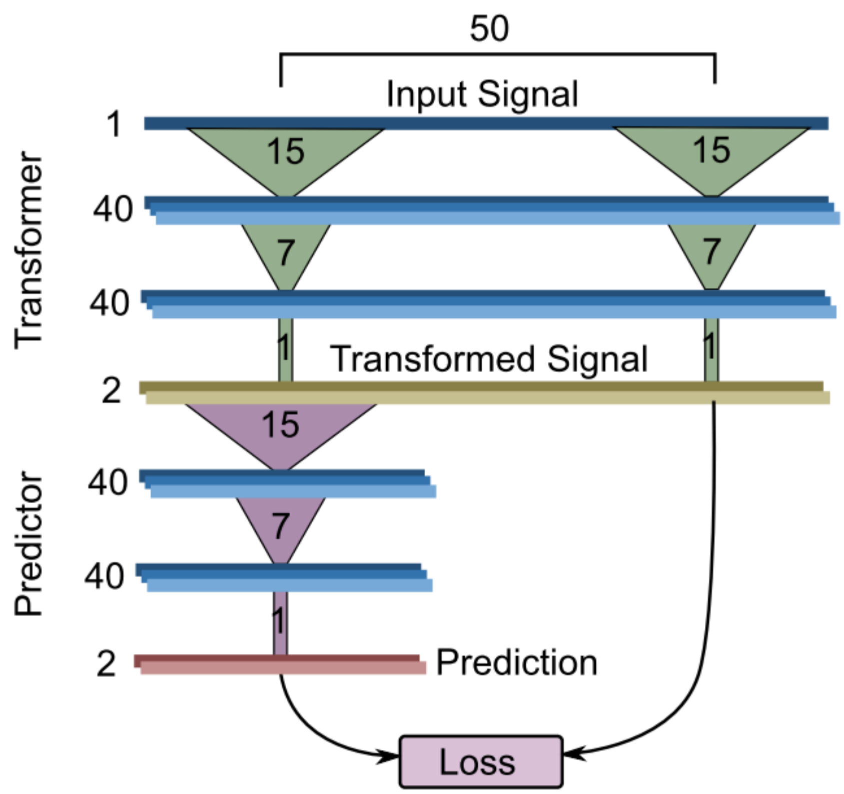


Figure 4: NCG [[I need to simplify this figure]](Guttenberg *et al.*, 2016)

tion to infer hidden causes.

Writing Materials

3.2 Neural coarse-graining

- NTIC can be more easily achieved by coarse-graining.
- (Guttenberg *et al.*, 2016) proposed an architecture, called neural coarse-graining (NCG), which aims to captures the emergent large-scale dynamics in the environment.
- Achieving NTIC is the key objective to to extract out latent control parameters in the environment.
- Through two objectives during training, the model can achieve NTIC.
- this can be achieved by maximising the self-predictive power of deep layers and its entropy through all training data.
- With the objectives, NCG network learns to coarse-grain input signals and extract latent parameters which are both predictive and predictable in the input signal and therefore achieves NTIC by building a higher-level representation of the underlying dynamics in the deep layers.
- Agents can process information and act at abstract levels

3.3 The advantage of NCG

- At the physical scale which human being live, the environmental dynamics is nearly deterministic following the laws of classical physics. This means that it is possible for agents living at this physical scale to build NTIC process internally.

- When the information in individual sensory channel is noisy and partial, coarse-graining becomes necessary for achieve NTIC.

3.4 NTIC at specific levels of neural coarse-graining

[Call For Help]: any reference?

- It is important to note that the relationship between degrees of NTIC and coarse-graining is not monotonic.
- Overly coarse-graining a system can lead to low degrees of NTIC.
- Imagining an extreme scenario that all microscopic states map to a single macroscopic state, the macroscopic process losses all the predictive information about the environment.
- This is evidenced by empirical data as well.
- Study by Sederberg *et al.* (2018) found that self-predictive information of 4-cell sets readout by an optimal perceptron is higher than 7 and 10-cell sets.
- The authors indicated that

  > reading out from groups of four cells can be accomplished with high absolute efficiency. Perhaps the best way to combine more than four cells is to break down the readout into indivisible units of four cells, which are later recombined in subsequent processing. ((Sederberg *et al.*, 2018))

  > reading out from groups of four cells can be accomplished with high absolute efficiency. Perhaps the best way to combine more than four cells is to break down the readout into indivisible units of four cells, which are later recombined in subsequent processing.

(Sederberg *et al.*, 2018)

4 A neural coarse graining theory of consciousness

Need to describe my intuition more concretely

[[Do I need to start with my philosophical intuition?]]

- Through coarse-graining, we can define a new set of variables.
- If the new set of variables are informational closured, then this creates a new "reality" for all the variables inside.
- Reality means the all the information (include variables and the relationships) within the boundary of causal or informational influence between variables.
- This is the same as "if a tree falls in a forest" thought experiment.
- Therefore, to understand this system, you don't need more information from external environment.
- It also means everything in this system is self-defined.
- ============= reasoning ============
- Information processing in the neural system exists at different level of abstraction.
- The physical environment dynamic at the human scale are nearly deterministic.

  Ref: Do we need ref here?
- To maximise fitness for individual human-being, neural system should try hard to infer and model the deterministic physical rules.
- This drives the neural system to create a non-trivial informational closure.
- [[okay, I need to be very careful about the causal relation here.]]
- Therefore, coarse-graining becomes necessary because every individual sensor and neuron at the microscopic scale can only access and process a small part of information.
- However, at a macroscopic scale, information processing can achieve non-trivial informational closure (Fig. 5)
- ============ our claim ============
- We claim that the state of a coarse-grained scale which realises non-trivial informational closured correlates the state of consciousness.

- We further postulate that level of consciousness corresponds to informational closureness of processes.
- The contents of consciousness correspond to the states of the non-trivial informational closure.
- I need to somehow acknowledge Pennartz (2017)
  - *" In conclusion, for a well-behaved representational system, it does not matter whether a neural activation pattern represents an external world feature truthfully or not, as long as the system represents the feature consistently over time and in sensory space."*
  - pennartz levels are defined by functional elements call "functional ensembles". Pennartz (2017) *"This delineation is different from a scale-based distinction of levels, because different levels are distinguished here based on function , i.e., what each level can accomplish in terms of computation and representation, culminating in perception at the highest level."*
- In the following part, we.... todo

To Be Modified

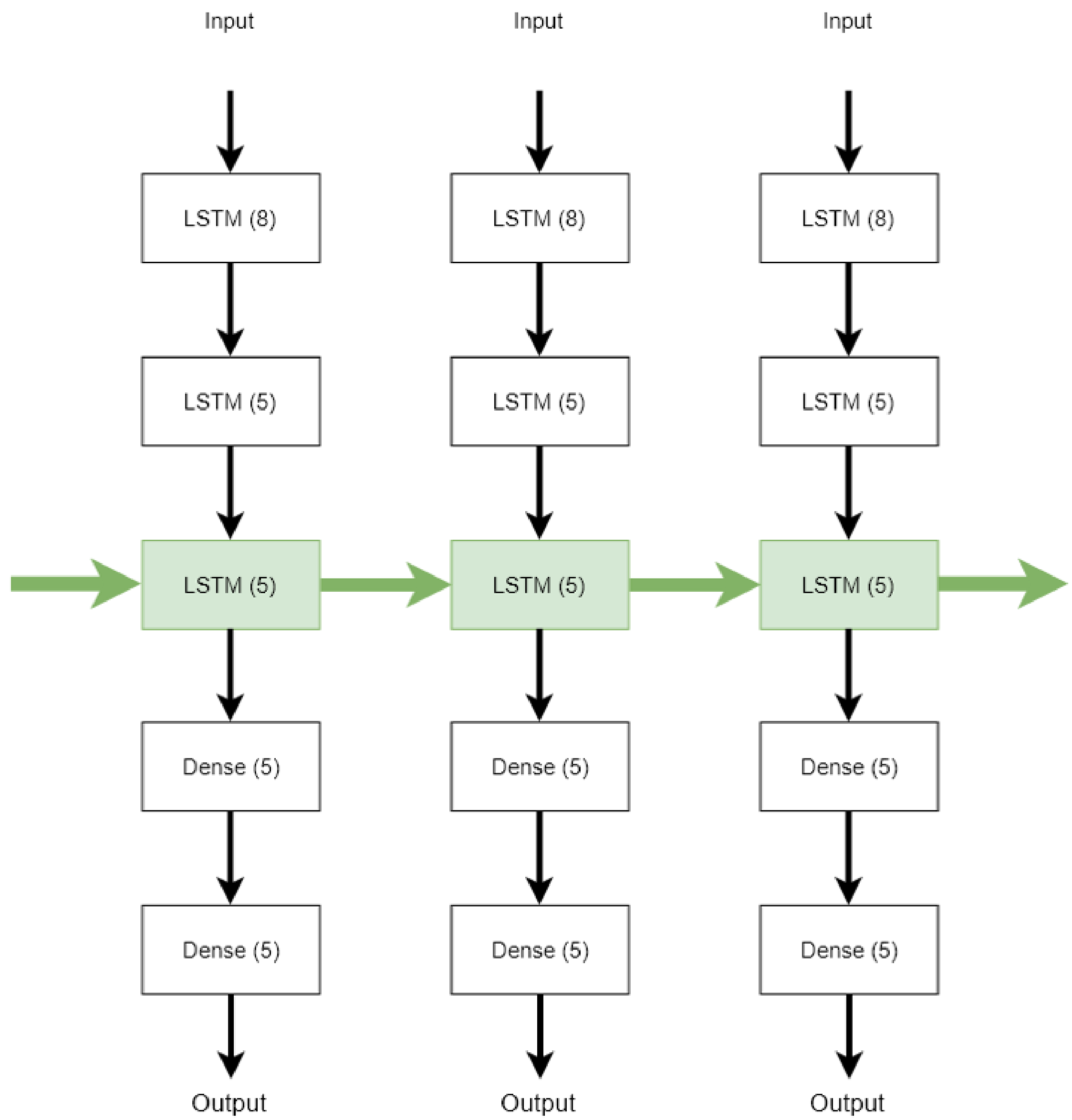


Figure 5: informational closure through neural coarse-graining

- *" Multilevel views on consciousness and cognition have been expressed before by many other theorists, such as Oppenheim and Putnam (1958), Attneave (1961), Hofstadter (1979, 1985), Churchland and Sejnowski (1992), Wimsatt (1994), and Lycan (1996)."*
Pennartz (2017)
  - Oppenheim , P. , & Putnam , H. (1958). Unity of science as a working hypothesis . In H. Feigl , G. Maxwell , & M. Scriven (Eds.), Concepts, theories, and the mind body problem (pp. 3 36). Minneapolis : University of Minnesota Press .
  - Attneave , F. (1961). In defense of homunculi . In W. A. Rosenblith (Ed.), Sensory communication (pp. 777 782). Cambridge, MA : MIT Press .
  - Hofstadter , D. R. (1979). Godel, Escher, Bach, an eternal golden braid . New York : Basic Books .
  - Churchland , P. S. , & Sejnowski , T. J. (1992). The computational brain . Cambridge, MA : MIT Press .

  - Wimsatt , W. C. (1994). The ontology of complex systems: levels of organization, perspectives, and causal thickets. Canadian Journal of Philosophy , 20 , 207 274 .
  - Lycan , W. G. (1996). Consciousness and experience . Cambridge, MA : MIT Press .
- It's important that not the neural states but the coarse-grained state determine contents of consciousness.
- Consciousness is information
  - Information is level-dependent.
  - Information is different at microscopic and macroscopic levels
  - Consciousness must be level-dependent
- *" This illustrates how coarse-graining allows the formulation of incomplete generalizations that are relatively invariant, although at the cost of predictive precision regarding fine-grained details."*
  Price & Corry (2007)
  - conscious perception are very stable and time invariant.
  - Conscious perception is informative but we loss all the precise information about the states at microscopic level.

4.1 Measure of conscious level

- The level of consciousness can be computed from how much degree of the non-trivial informational closure in a coarse-grained scale.
- Based on Equation xxx, the measure of non-trivial informational closure can be decomposed to XXX and XXX. equation
- use

  Since, in the following, we are mainly interested in the case of non-trivial closur

$$\begin{aligned} NTIC_m &:= MI(S_{n+1}; E_n, \ldots, E_{n-m}) - MI(S_{n+1}; E_n, \ldots, E_{n-m}|S \\ &= MI(S_{n+1}; E_n, \ldots, E_{n-m}) - MI(S_{n+1}; E_n|S_n) \end{aligned}$$

- or use

$$\begin{aligned} MI(S_{n+1}; E_n) &= 0 \quad \text{and} \\ MI(S_{n+1}; S_n) - MI(S_{n+1}; S_n|E_n) &= 0 \end{aligned}$$

- To have high level non-trivial informational closure, the states of the coarse-grained scale need to
- First, maximise the mutual information between environment and the representation. This term implies that the agent need to maximise predictive power of environmental state. This also suggest that the environmental complexity is a crucial factor to the level of consciousness. The more rich and complex environment the agent try to predict the higher level of consciousness the agent has.
- Second, to minimise the mutual information between environment and the representation conditional on the past state of the representation. This term suggests that the representation is self-predicable and it mimics the environmental transition.
- Therefore, this theory predicts that level of consciousness determined by the environment that the agent adapts to. To form a high non-trivial informational closure, the agent will need to live in a complex environment and model the environmental transition precisely.
- [[REWORDING | Note that, our claim is that informational closure is a more fundamental property that correlates consciousness. A species acquiring advantage to model a complex environment through evolution and , therefore, achieving a high-level non-trivial informational closure has a high level of consciousness, is the consequence of the evolutional process.]]

4.2 Conscious contents

- We claim that conscious contents correspond to the state of the informational closure.
- [[= not sure I want to say this here =]]
- Consciousness is informative. A conscious percept including volition and intention is a unique state that narrowing down all possible states into one. Therefore, a conscious percept reduces a huge amount of uncertainty.
- This is the same claim as the third axiom called "information" in IIT.

Ref: information axiom and need to say more about this.

- ====================================
- Note that because the closure is at a coarse-grained scale one state of the informational closure can be mapped to several states at microscopic scale.

4.3 Sensory hierarchy and neural coarse-graining

This is super important. I need to deal with this very carefully. Need to completely rewording

To Be Modified

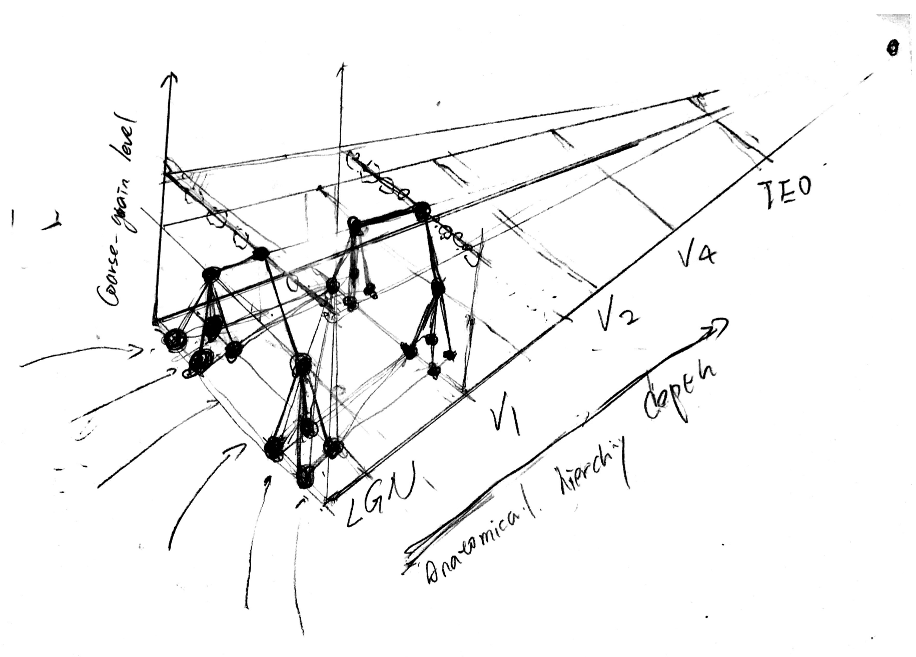
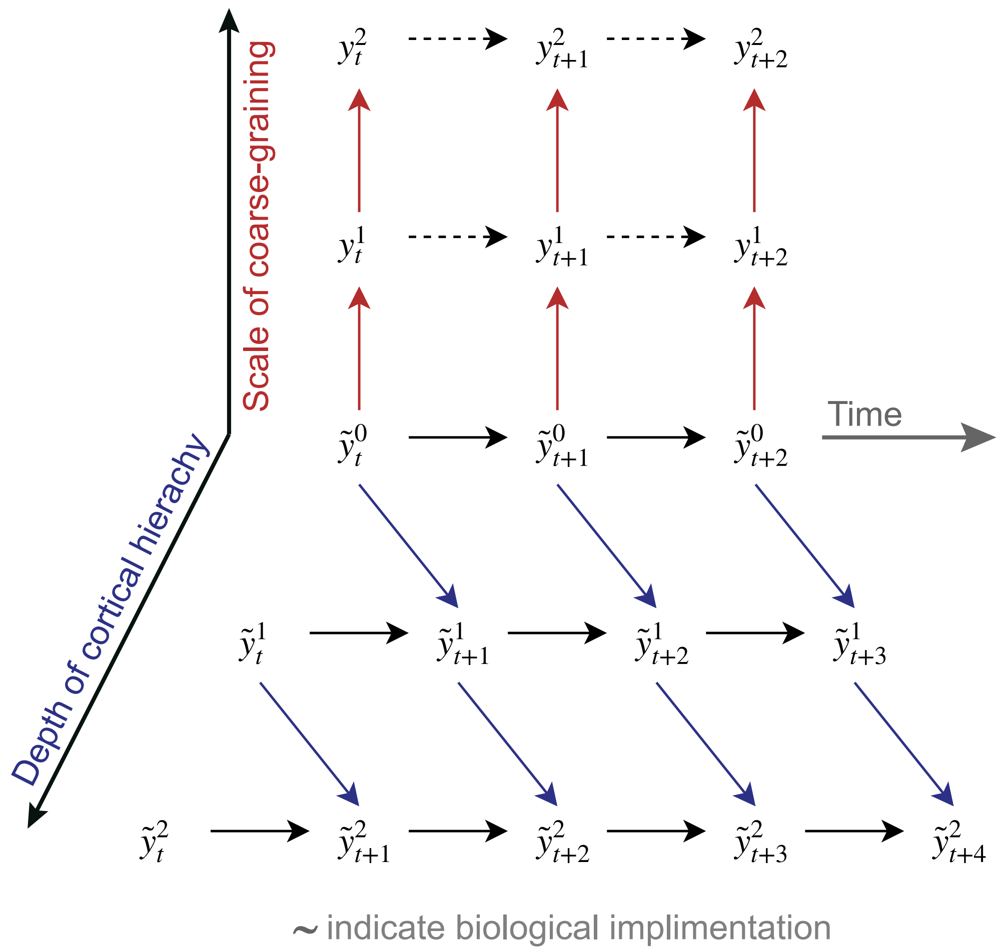


Figure 6: Coarse-grain and sensory anatomical hierarchies



- It's important to note that the coarse-graining direction is not necessary to be align with anatomical sensory hierarchy in the neural system.
- the coarse-grained information is needed and computed only when the information is required for microscopic scale processing.
- **[[I need a strong example here]]**
- Of course, from evolutional perspective, neural system should develop a better hardware to support neural coarse-graining
- For example, pooling layers in CNN is an implementation of coarse-graining function. It' is very plausible visual processing hierarchy in the neural system realises pooling layers.
- It is possible that the neural system represents higher-level information processing by low level physical subtract. For example, studies have found representation for summery statistics of signal and neural. population.
  From this point of view, sensory hierarchy may extend along with levels of coarse-graining.

Ref: neural representation for summery statistics

Writing Materials

more information theory and philosophy relevance

- One consequence of coarse-graining is that it makes it permissible to ignore certain causal factors that would be relevant at more fine-grained level of description. Price & Corry (2007)
- This means the fine-grain variable does not exist at the coarse-grained causal chain.

Writing Materials

4.4 Information and system theory

[[A closure can define a system, consciousness is information]]

[[Could be linked to system theory, but may need help]]

" This self-referential distinction from its environment therefore gives rise to the specific autonomy of such a system. Consequently, in systems theory, closure properties and autonomy are considered to be closely related concepts which are both at the heart of defining the system itself."

4.5 Empirical approaches of measuring neural informational closure

5 Conscious vs unconscious processing

Now we can combine "Relation to empirical findings in neuroscience and studies of consciousness" and "About Phenomenology (conscious processing)" in to this

6 Comparison with other theories

Main points should be same/diff, our advantages, and predictions

6.1 Theories about level of abstraction and coarse-graining

- Before the current theory, several theories and perspectives of consciousness have proposed the notion that links level of abstraction or coarse-graining in the neural system to consciousness.
- We shortly describe two theories which is related to our current proposal.

XXX and XXX

6.1.1 Intermediate Level Theory

- Intermediate Level Theory (ILT) proposes that conscious experience only correlates specific levels of representations in the sensory processing the hierarchy. (Prinz, 2007; Jackendoff, 1987). This theory claims that consciousness only associates with the representation at intermediate levels rather than the ones at lower-level or higher-level.
- A critical difference between ILT and our theory is that our theory doesn't assume the direction of coarse-graining should align with the direction of anatomical processing levels along cortical hierarchies.

6.1.2 The difference between ILT and our theory

Prinz suggested that *"as. Cells in V1 are not promising candidates, because they do not reliably respond in ways that are consistent with features that we experience consciously. "*
Prinz (2007)

* Our prediction is, if information that is carried by v1 neurons are necessary to complete informational closure, the states of V1 neurons should also co-vary with conscious contents.

6.2 First-order theories

Writing Materials

" The rst-order view maintains that conscious awareness is determined by early sensory activity alone, independently of higher-order representations. Thus the crucial difference between the rst-order and higher-order views is that the latter but not the former predict that conscious awareness is determined at least in part by prefrontal and parietal activity. Therefore, in the group of rst-order theorists we include here not only those commonly labeled as such in the philosophical literature [11] but also theorists, such as Ned Block [12,13], who hold that the phenomenal qualities of awareness depend on the biological substrate rather than merely the content of the rst-order representations, as well as scientists who hold that visual awareness depends on activity in content-specic regions of the extrastriate cortex [24,25] or on feedback from these regions to the primary visual cortex [26]. Most rst-order theories explain the difference between awareness and unawareness by positing

that the latter is associated with weak information [29] or with representations in alternative (e.g. subcortical) sensory pathways. Thus this view suggests that conscious and unconscious processing will have different functional consequences. Whereas some rst-order theorists hold that some higher cognitive functions can be performed without conscious awareness [3032], most rst-order theorists take a difference in perceptual task performance (e.g. hits vs misses in a detection task) as evidence for a difference in conscious awareness [12,13,25,26,33]. In other words, they typically associate task performance with awareness.."
Lau & Rosenthal (2011)

  - Among those are first-order theories, on which a mental state is conscious if being in that state results in ones being conscious of something;

- According to first-order views, phenomenal consciousness consists in analog or fine-grained contents that are available to the first-order processes that guide thought and action. So a phenomenally-conscious percept of red, for example, consists in a state with the analog content red which is tokened in such a way as to feed into thoughts about red, or into actions that are in one way or another guided by redness.

Writing Materials

6.3 Global workspace theory

- Global workspace theory (GWT)Baars (1988, 1997, 2002) or Global Neuronal Workspace theory (GNWT)Dehaene *et al.* (1998); Dehaene & Naccache (2001); Dehaene & Changeux (2011) states that the neural system consists of many separate specialised modules and a central global workspace (GW) which integrates and passes information amount the modules. Only the information that is accessed by the global workspace can be conscious aware of. The modules competes with each other to gain the access to the global workspace and the information from the winner can trigger an all-or-none "ignition" in the global workspace and be globally broadcasted to other modules. Conscious contents then associates information which gains access to the internal global workspace Dehaene *et al.* (2017a).
- GW in previous data is the place that form NTIC.
- non-linearity and emergence
- what's the different prediction?
- We predict that when the new

6.3.1 Global availability and Broadcasting

6.3.2 Code translation

- neural coarse-graining can natually solve this problem.
- coarse-graining between levels can be thought as code translation

is this true?

6.3.3 Stability

- the presence of contexts as stable coalitions shaping access to the workspace.
- e all-or-none behavior, whereby the incoming evidence either quickly dies out (corresponding to subliminal processing) or is accumulated and amplied non-linearly into a full-blown state of high-level activity. This global ignition has been proposed as a marker of conscious perception [70]. Indeed, empirically, when stimulus strength is varied, the early stages of non-conscious processing typically show a linear variation in activation, whereas conscious access is often characterized by a late non-linear amplication of activation which invades a distributed set of parietal, prefrontal and cingulate areas

6.3.4 Broadcasting

6.3.5 Ignition

Writing Materials

- Conscious access as an accumulation of evidence leading to an all-or-none ignition What, if anything, remains unique to conscious processing? Although many cognitive operations can be partially launched non-consciously, these operations rarely if ever run to completion in the absence of consciousness. A subliminal stimulus may induce above-chance performance, behavioural priming, and a small amount of brain activity in narrowly dened brain networks, but these measures often increase dramatically as soon as the subject reports seeing the stimulus, especially in high-level areas [46 60,61]. Accumulation of evidence has been demonstrated with non-conscious stimuli [59], but only conscious stimuli cross the threshold beyond which an overt strategy can be exibly deployed

Writing Materials

- GWT is one of the current popular consciousness theories.
- *"AI system and corresponds to our first functional definition of consciousness: global availability"*
  Dehaene *et al.* (2017a)
- C1: Global availability of relevant information The need for integration and coordination Dehaene *et al.* (2017a)
- Evidence for integration and broadcasting
- Stability as a feature of consciousness *" meta-stability seems to be necessary for the nervous system to integrate information from a variety of modules and then broadcast it back to them, achieving flexible cross-module routing."*
- *"integration of computational resources in a large-scale coordination and for the exchange of information among processors"*
- *"Grounded on the distinction between conscious and unconscious processes, Bernard Baars global workspace theory is one of the most influential cognitive theories of consciousness. This theory relies on the metaphor of a theater. In this theater, unconscious specialized processors (equivalent to modules) are assumed to be the actors and the audience. While the audience represents the set of passive processors, actors represents active processors playing on the stage of the theater (i.e., the workspace). These actors are engaged in a competition for being seen by the audience: by broadcasting their information they actually compete for more broadcasting. Active processors with the highest coherent activity can form local coalitions that strengthen them in this competition process. The strongest coalition in this competition dominates the workspace, in a winner-take-all fashion, and corresponds to the content of consciousness. The workspace is equated by Baars to working-memory, in which only the most active content becomes conscious. Additionally, the dominant coalition benefits from global broadcasting, which allows it to recruit new processors from the audience in order to form a global coalition. Here, consciousness allows for the integration of computational resources in a large-scale coordination and for the exchange of information among processors that would otherwise remain separated. In this theory, each processor can operate in the conscious mode if it benefits from global broadcasting through the workspace, or it can operate in the unconscious mode when disconnected from the workspace. An important feature of the global workspace theory is the presence of contexts as stable coalitions shaping access to the workspace. Contexts are constituted of unconscious processors reflecting, in a hierarchical manner, our expectations, our beliefs, our goals, and ultimately our self. In particular, attention is implemented as a goal context in this theory. It is described as a mechanism that controls access to the workspace, acting as a filter and biasing the competition process toward a particular set of actors. At any given moment, the dominant coalition is under the spotlight of attention, and its informational content becomes the content of conscious experience. A crucial aspect of Baars theory is that it avoids the problem of the homunculus by reducing it to an audience of multiple unconscious processors. Here, there is no need for a hypothetical single conscious observer in the system, and thus there is no issue of infinite regression with a homunculus inside another homunculus. Instead, consciousness is considered to reflect the global broadcasting of information to an audience of unconscious processors. As the audience is unconscious, unsupervised, and receptive rather than attending to the information, it does not constitute an internal homunculus."*

6.4 Higher Order Thought theory of consciousness

[[People think a high order representation receives information from the first order representation. This may be a misunderstanding. This may be the result of neural coarse-graining. Because contents in informational closure look like watching fine-grained information, people may have an wrong impression about it]]

==

- Information closure directly is critical for other systems to construct forward models.
- Therefore, most of the high level cognition can utilise the state of informational closure.
- For example, mental planning needs a precise forward model to infer the environmental dynamics.
- [[This is way HOT always involve consciousness]]
- Rosenthal (2005)

" Higher-order theories of consciousness argue that conscious awareness crucially depends on higher-order mental representations that represent oneself as being in particular mental states."
Lau & Rosenthal (2011)

" According to this view, humans not only have first-order non-conceptual and/or analog perceptions of states of their environments and bodies, they also have second-order non-conceptual and/or analog perceptions of their first-order states of perception. And the most popular version of higher-order perception theory holds, in addition, that humans (and perhaps other animals) not only have sense-organs that scan the environment/body to produce fine-grained representations that can then serve to ground thoughts and action-planning, but they also have inner senses, charged with scanning the outputs of the first-order senses (i.e. perceptual experiences) to produce equally fine-grained, but higher-order, representations of those outputs (i.e. to produce higher-order experiences). A version of this view was first proposed by the British Empiricist philosopher John Locke (1690). In our own time it has been defended especially by Armstrong (1968, 1984) and by Lycan (1996)."
https://plato.stanford.edu/entries/consciousness-higher/#SelRepHigOrdThe

" neutral about whether conscious awareness adds signicant utility or immediate impact on behavior and task performance [1,27]. This is because the view assumes that task performance in most perceptual and cognitive tasks depends mainly on rst-order rather than higher-order representations. Because conscious awareness can differ even if all rst-order representations remain completely unchanged, such awareness itself might serve little function [1,27]."
Lau & Rosenthal (2011)

My theory would have similar prediction. Because conscious contents are coarse-grained results from low-level information. It is possible that more than two low-level neural states all coarse-grained to a high-level state.

6.4.1 Metacognition

" Metacognition: cognition that is about another cognitive process as opposed to about objects in the world. In this article we use it mainly to refer to sensory metacognition: a cognitive process that concerns the quality or efcacy of a perceptual process. The capacity of sensory metacognition is sometimes empirically assessed by the correspondence between subjective report and task performance: how closely are they associated with each other on a trial-by-trial basis"
Lau & Rosenthal (2011)

6.5 Sensorimotor contingency

- Sensorimotor contingency (SMC) theory of consciousness is proposed to account of conscious perception. It suggests that different types of SMCs give rise to different characteristics of conscious perception O'Regan & Noë (2001).
- The theory radically rejects the view that conscious contents is associated with information encoded in internal representations in the neural system.
- Rather, the quality of conscious perception depends on agents' mastery of SMCs.
- It emphasizes the importance of the interactions between a system (the neural system) and its.
- Counter-intuitively, because consciousness inhabits in interactions between agents and environments, this theory suggests that consciousness is not driven by brain activities. Instead, consciousness should be shaped by the way of interaction, i.e. SMCs, between agents and environments.

- Similar to the SMC theory, our theory predicts that the interaction between the neural system and the environment shapes the quality of conscious perception.
- This is because to achieve NTIC the neural system should maximise the conditional mutual information between the future state of the NTIC process and the environment given the agent's current action (see equation XXX). I NEED A EQUATION!
- This suggests the to maximise NTIC the NTIC process needs to involve the information about the environmental transition and dynamic after agent preforms an action.
- Therefore, as a consequence of maximising NTIC, the representation forming NTIC implicitly encodes information about SMCs.
- In the long-term, once NTIC neural dynamic interoperates the environmental state transition given actions via maximising NTIC, the state space of XXX then is shaped by the new learned SMC
- Accordingly, the way how the agent interacts with the environment influences the structure of the NTIC representation.
- Need to involve some sentence from
- [[Do I need a figure here?]]
- However, in contrast to the SMC theory, our theory suggests the state of NTIC process correlate conscious contents. Therefore, if the NTIC process is formed inside the neural system, we should have correspondent conscious contents in our conscious experience without interacting with the environment.
- [[Not sure here]] This can easily explain that the conscious experience of dream is just induced by self-evolved state changes of the NTIC process.
- Further more the SMC and our theories make complete different prediction about the relationship between SMC and consciousness.
- Contrary to the SMC theory which suggests that consciousness resides on SMCs, our theory predicts that pure S-R contingency between the neural system and the environment can not achieve informational closure within the neural system and therefore is not conscious (please see section 8.2)

6.6 Predictive Processing

- The predictive processing is a powerful framework (PP) which integrates several influential ideas in the history of neuroscience including "perception as unconscious inference" from Helmholtz Helmholtz (1866), "analysis by synthesis", predictive coding, "Bayesian brain hypothesis"
- is a powerful framework
- a recent emerged theoretical framework of perception and brain function, which suggests that the brain continuously generates and updates predictions about incoming sensory signals, represents a general principle of neural functioning.
- According to predictive processing (PP) account of perception, the neural system constantly performs unconscious statistical inference about hidden causes in the external environment. The perceptual contents are the "best guess" about the environment Clark (2013) Hohwy (2013)
- Bperceptual content is the result of probabilistic inference on the hidden causes of sensory signals.
- Because it is well integrated with Bayesian brain hypothesis, the PP framework has been used to interpret conscious perception in many perceptual domain. Hohwy (2013) Seth (2014)
- However, to link PP to consciousness, there are still two major gaps.
  - First, previous studies have demonstrated many brain processes following the principle of Bayesian inference and Bayesian integration. But not all these processes are conscious. This indicates that PP is not a sufficient mechanism for conscious perception. Some other factors may be involved to determine which PPs can be conscious or not.
  - Second, it has been emphasized that only the results of the inference or prediction can be consciously accessed. The process of inference in completely unconscious. So WHY?

- Even though PP has been acquiring attention in different discipline in neuroscience, So far, to our knowledge, no detailed explanation has been given to account for the two questions.
- We first hypothesize that predictive processing stays unconscious is because it is not able to achieve NTIC.
  - [[Example from PID controller?]]
- Second, according to our theory, it is obvious that the computation of predictive processing is unconscious and only the result is conscious. Because the computational details is not at the coarse-grained level which achieves NTIC.
  - So far, many neuronal models and coding schema have been proposed to explain how to implement Bayesian inference and prediction at the neuronal level.
  - However?
- We hypothesize that the distinction between conscious and unconscious predictive mechanisms is whether they build up on the information from the NTIC process.

Writing Materials

- Lamme (2015)
  - To my knowledge, this is the first time this has been so explicitly laid outwriters on predictive coding thus far have always stayed a little vague on where exactly consciousness sits in the Bayesian framework.
  - Yet the two points seem contradictory. In the Bayesian predictive coding framework, consciousness is the result of the unconscious inferential processes.
  - However, these authors have so far been very reticent about how exactly consciousness ts into the PC framework. For now, there is an emerging idea that understands consciousness as the nal result of unconsciousness processing. According to Hohwy (2013), perceptual consciousness ts in the PC theory as the upshot or conclusion of unconscious perceptual inferences. Melloni (2014) also expects that inferential processes are conducted behind the curtain of consciousness and what we experience are the outcomes or results of an unconsciou inferential process. Lamme (2015) agrees with the idea that consciousness is based on the relation or more precisely conjunction of current inputs and priors, which together produce posterior beliefs. Karl Friston agrees to a certain extent with this interpretation of consciousness. According to Hobson and Friston (2016), consciousness is not a phenomenon but a process, a process of optimizing beliefs through inference (consciousness, dreams, and inference; the Cartesian theater revisited; a response to our theater critics).
- Link PP and c
  - Hobson & Friston (2016)
  - Melloni (2015)
- Seth (2014)
- When uncertainty is high, we should expect an increase in the ability of ever higher-level prior beliefs to determine conscious experience. I
- determine conscious experience.
- PP combines and builds upon previous ideas about the role of unconscious inference in perception (Helmholtz, 1925; Barlow, 1961; Gregory, 1970), the process of analysis by synthesis in psychology (Neisser, 1967), the predictive coding approach in neuroscience (Srinivasan et al., 1982; Rao and Ballard, 1999; Huang and Rao, 2011) and generative models and related probabilistic computational principles (MacKay, 1956; Mumford, 1992; Dayan et al., 1995; Hinton, 2007a,b; Tenenbaum et al., 2011).
- *"The brain is constantly confronted with a wealth of sensory information that must be processed efficiently to facilitate appropriate reactions. One way of optimizing this processing effort is to predict incoming sensory information based on previous experience so that expected information is processed efficiently and resources can be allocated to novel or surprising information. Theoretical and computational studies led to the formulation of the predictive coding framework (Friston 2005, Hawkins and Blakeslee 2004, Mumford 1992, Rao and Ballard 1999). Predictive coding states that the brain continually generates models of the world based on context and*

information from memory to predict sensory input. In terms of brain processing, a predictive model is created in higher cortical areas and communicated through feedback connections to lower sensory areas. In contrast, feedforward connections process and project an error signal, i.e. the mismatch between the predicted information and the actual sensory input (Rao & Ballard, 1999). The predictive model is constantly updated according to this error signal. "

- *"Predictive inference and perception The concept of PC overturns classical notions of perception as a largely bottom-up process of evidence accumulation or feature detection, proposing instead that perceptual content is specied by top-down predictive signals that emerge from hierarchically organized generative models of the causes of sensory signals. According to PC, the brain is continuously attempting to minimize the discrepancy or prediction error between its inputs and its emerging models of the causes of these inputs via neural computations approximating Bayesian inference (Figure 1 and Box 1). Importantly, prediction errors can be minimized either by updating generative models (perceptual inference and learning; changing the model to t the world) or by performing actions to bring about sensory states in line with predictions (active inference; changing the world to t the model)."*

- *"In terms of brain function, Hermann von Helmholtzs 19th Century notion of perception as unconscious inference has found new momentum in modern formulations of predictive processing, the Bayesian brain, and the free energy principle (Clark, 2013; Friston, 2010; Hohwy, 2013). On these formulations, which can all be subsumed under the general term prediction error minimization (PEM, (Hohwy, 2013)) perceptual content is the result of probabilistic inference on the hidden causes of sensory signals. "*

- *"Generative models encoded by brain structure and dynamics make predictions about sensory inputs, based on prior expectations. Action (e.g., body movements) and perception conspire to reduce sensory prediction errors, giving rise to perceptual content and"*

- *"More informally, what we perceive is the brains best guess of whats out there (or in here, for perceptions of body and self)."*

- *" predictive processing (PP) 1 or the Bayesian brain this holds that perceptual content is determined by hierarchically organized generative (i.e., predictive) models (HGMs) of the external causes of sensory signals, induced by a process approximating Bayesian inference (Friston, 2009)."*

- Not all predictive mechanism are conscious.
- We hypothesize that the distinction between conscious and unconscious predictive mechanisms is whether they build up on the information from the NTIC process.
- we can also explain why only the result (posterior) of predictive processing is conscious.
- Because of this XXXX, the representation has the maximal non-trivial informational closure (ntic) also has the maximal predictive power of the external environment.
  I need math here!!
- Therefore, it's reasonably assume that if evolution wants to build predictive mechanisms, the mechanisms should be settled on the information at the coarse-grained (cged) level.
- More specifically, a forward model, which received state and action as input and compute transition and then output next states, should be build on the representation that forms the informational closure.
- Therefore, the state of predictive model can be conscious while it is linked to informational closure representation.
- Note that, a system showing some predictive power is not sufficient to be conscious.
- For example, the proportionalintegralderivative controller (PID) controller shows predictive behaviours due to its derivative which computes and predict error value in the future. Derivative action predicts system behaviour and thus improves settling time and stability of the system.

- However, the state of the whole PID system still be determined by the inputs, i.e. the states of the controlled process, thus cannot complete informational closure.
  . It can compute position, velocity, and acceleration of objects, and output control signals (action). However, the state of the PID controller is not information-closed. The state still fully determined by the input signals. Therefore, it fails to reach informational closure (ic). Hence, according to our theory, the PID system is not conscious.
- Similarly, the neural system also has many circuits that show predictive power but do not have ic.

Writing Materials

6.6.1 statistical learning/prediction is not conscious

- considers the brain as a predictive machine
- According to predictive processing account, conscious contents are the inference results (posterior from Bayesian inference framework) of predictive mechanism.
- However, we can see that not all predictive mechanisms in the neural system can be consciously aware of.
- So what's the critical difference between them in terms of consciousness?
- We argue that when predictive mechanisms utilise the information in the informational closure, the predictive result is conscious.
- Because the forward model build on the informational closure representation can be seen as part of the closure.

IS THIS TRUE? NEED DISCUSSION

- we can consider two scenarios
  - Predictive information is encoded in neuron populations but there is high entropy of conditional probability which means it's not closure. Therefore, the conscious level could be very low even though the it still has some degree of predictability.
  - Another scenario is a forward model but its state fully depends on external sensory input. This system is not informational closured and therefore is not conscious.
  - Based on this point of view, predictive power for a system does not guarantee consciousness.
  - For example, in active inference , action is guided by prediction error so the state of the system is non-closured.

Ref: active inference

6.7 Integrated information theory

[Call For Help]: Help from Jun?

- **[[Information integration without awareness]]**Mudrik *et al.* (n.d.)
- *"consciousness is necessary for high-level but not low-level semantic integration"*
  Mudrik *et al.* (n.d.) I would argue an inverse relationship:"High-level information processing is necessary for consciousness." This is because higher-level information processing is more likely to create an informational closure representation.
- Similarly, it is not *"consciousness is necessary for multisensory integration"*
  . It should be multisensory integration is more likely to create an informational closure representation. Because in most situations, multisensory signals are generated from hidden causes which may have deterministic dynamic and relation with each other. Therefore, to infer the hidden causes and represent the deterministic relations can form informational closure representation which is conscious.

6.7.1 Hoel's theory

[Call For Help]: Help from Martin?

This is very important. Need to be careful here.

- Hoel *et al.* (2016)
- causal emergence

- Hoel *et al.* (2013) examined causal power at different coarse-grained scale and show that causal power can be stronger at macro rather than micro levels.
- To answer the question about what is the best scale to compute Phi, Hoel *et al.* (2016) examined causal power at different coarse-grained scale.
- Further instigation is needed to elucidate the relationship between our theory and causal emergence

6.8 Internal simulation and self-modelling theory

7 Future works

- Lack of direct evidence because no one focus on informational closure

7.1 About location of NTIC

7.2 Attention

Attention can be when the neural want to output a specific information in NTIC.

7.3 Finding coarse-grain function

- It's not clear what the coarse-graining function is.

this should be explained by environmental deterministic

- *" not all causal relationships (or relationships of nomological dependency) among micro-events aggregate up to causal relationships among coarse-grained macro-events that are constituted by those micro-events. Instead, whether one gets causation at the macroscopic level will depend (among other things) on the particular coarse-graining that is chosen."*
  Price & Corry (2007)
- *" It may seem surprising, even counterintuitive, that causal and statistical dependence relationships involving fine grained microscopic variables do not automatically show up in causal and dependence relationships among the macroscopic variables that are realized by the fine grained variables."*
  Price & Corry (2007)
- Jonas & Kording (2017) well demonstrated that it is difficult to understand higher-level information processing by the current neuroscience methods.

8 Conclusion

- Why is this theory better than other theories of consciousness
- NCC: To find NCC, it's not about where and when, it's about scale and level of description

Funding

Acknowledgements

Supplemental Data

References

Asplund, Christopher L, Fougnie, Daryl, Zughni, Samir, Martin, Justin W, & Marois, René. 2014. The attentional blink reveals the probabilistic nature of discrete conscious perception. *Psychological science*, **25**(3), 824–831.

Baars, Bernard J. 1988. *A cognitive theory of consciousness*. Cambridge University Press.

Baars, Bernard J. 1997. In the theatre of consciousness. Global workspace theory, a rigorous scientific theory of consciousness. *Journal of Consciousness Studies*, **4**(4), 292–309.

Baars, Bernard J. 2002. The conscious access hypothesis: origins and recent evidence. *Trends in cognitive sciences*, **6**(1), 47–52.

Bertschinger, Nils, Olbrich, Eckehard, Ay, Nihat, & Jost, Jürgen. 2006. Information and closure in systems theory. *Pages 9–21 of: Explorations in the Complexity of Possible Life. Proceedings of the 7th German Workshop of Artificial Life.*

Binder, Marc D, Hirokawa, Nobutaka, & Windhorst, Uwe. 2009. Encyclopedia of neuroscience.

Clark, Andy. 2013. Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, **36**(3), 181204.

Dehaene, Stanislas, & Changeux, Jean-Pierre. 2011. Experimental and theoretical approaches to conscious processing. *Neuron*, **70**(2), 200–227.

Dehaene, Stanislas, & Naccache, Lionel. 2001. Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition*, **79**(1-2), 1–37.

Dehaene, Stanislas, Kerszberg, Michel, & Changeux, Jean-Pierre. 1998. A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the national Academy of Sciences*, **95**(24), 14529–14534.

Dehaene, Stanislas, Lau, Hakwan, & Kouider, Sid. 2017a. What is consciousness, and could machines have it?

Dehaene, Stanislas, Lau, Hakwan, & Kouider, Sid. 2017b. What is consciousness, and could machines have it? *Science*, **358**(6362), 486–492.

Eurich, Christian W, & Wilke, Stefan D. 2000. Multidimensional encoding strategy of spiking neurons. *Neural Computation*, **12**(7), 1519–1529.

Faisal, A Aldo, Selen, Luc PJ, & Wolpert, Daniel M. 2008. Noise in the nervous system. *Nature reviews neuroscience*, **9**(4), 292.

Goldwyn, Joshua H., & Shea-Brown, Eric. 2011. The what and where of adding channel noise to the Hodgkin-Huxley equations. *PLoS Computational Biology*, **7**(11).

Guttenberg, Nicholas, Biehl, Martin, & Kanai, Ryota. 2016. Neural Coarse-Graining: Extracting slowly-varying latent degrees of freedom with neural networks. *arXiv preprint arXiv:1609.00116.*

Helmholtz, H von. 1866. Concerning the perceptions in general. *Treatise on physiological optics,.*

Hobson, J Allan, & Friston, Karl J. 2016. A response to our theatre critics. *Journal of Consciousness Studies*, **23**(3-4), 245–254.

Hoel, Erik P, Albantakis, Larissa, & Tononi, Giulio. 2013. Quantifying causal emergence shows that macro can beat micro. *Proceedings of the National Academy of Sciences*, **110**(49), 19790–19795.

Hoel, Erik P, Albantakis, Larissa, Marshall, William, & Tononi, Giulio. 2016. Can the macro beat the micro? Integrated information across spatiotemporal scales. *Neuroscience of Consciousness*, **2016**(1), niw012.

Hohwy, Jakob. 2013. *The Predictive Mind.* Oxford University Press. 187 cites:.

Jackendoff, Ray. 1987. *Consciousness and the computational mind.* The MIT Press.

Jonas, Eric, & Kording, Konrad Paul. 2017. Could a neuroscientist understand a microprocessor? *PLoS computational biology*, **13**(1), e1005268.

Kandel, Eric R, Schwartz, James H, Jessell, Thomas M, of Biochemistry, Department, Jessell, Molecular Biophysics Thomas, Siegelbaum, Steven, & Hudspeth, AJ. 2000. *Principles of neural science.* Vol. 4. McGraw-hill New York.

Lamme, Victor. 2015. *Predictive coding is unconscious, so that consciousness happens now.* Open MIND. Frankfurt am Main: MIND Group.

Lau, Hakwan, & Rosenthal, David. 2011. Empirical support for higher-order theories of conscious awareness. *Trends in cognitive sciences*, **15**(8), 365–373.

Laureys, Steven. 2005. The neural correlate of (un) awareness: lessons from the vegetative state. *Trends in cognitive sciences*, **9**(12), 556–559.

Lemon, RN, & Edgley, SA. 2010. Life without a cerebellum. *Brain*, **133**(3), 652–654.

Luhmann, Niklas. 1995. Probleme mit operativer Schließung. *Soziologische Aufklärung*, **6**, 12–24.

Malone, Thomas W, & Bernstein, Michael S. 2015. *Handbook of collective intelligence*. MIT Press.

Mathis, Donald W, & Mozer, Michael C. 1995. On the computational utility of consciousness. *Pages 11–18 of: Advances in neural information processing systems*.

Maturana, Humberto R, & Varela, Francisco J. 1991. *Autopoiesis and cognition: The realization of the living*. Vol. 42. Springer Science & Business Media.

Melloni, Lucia. 2015. Consciousness as Inference in Time. *Open MIND*, **22**.

Moreno-Bote, Rubén, Knill, David C, & Pouget, Alexandre. 2011. Bayesian sampling in visual perception. *Proceedings of the National Academy of Sciences*, **108**(30), 12491–12496.

Mudrik, Liad, Faivre, Nathan, & Koch, Christof. Information integration without awareness. *Trends in Cognitive Sciences*.

O'Regan, J Kevin, & Noë, Alva. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, **24**(5), 939–973.

Overgaard, Morten, & Overgaard, Rikke. 2010. Neural correlates of contents and levels of consciousness. *Frontiers in psychology*, **1**, 164.

Palmer, Stephanie E., Marre, Olivier, Berry, Michael J., & Bialek, William. 2015. Predictive information in a sensory population. *Proceedings of the National Academy of Sciences*, **112**(22), 6908–6913.

Pattee, Howard Hunt. 2012. Evolving self-reference: matter, symbols, and semantic closure. *Pages 211–226 of: Laws, language and life*. Springer.

Pennartz, Cyriel MA. 2017. Consciousness, Representation, Action: The Importance of Being Goal-Directed. *Trends in cognitive sciences*.

PFANTE, OLIVER, BERTSCHINGER, NILS, OLBRICH, ECKEHARD, AY, NIHAT, & JOST, JÜRGEN. 2014. COMPARISON BETWEEN DIFFERENT METHODS OF LEVEL IDENTIFICATION. *Advances in Complex Systems*, **17**(02), 1450007.

Price, Huw, & Corry, Richard. 2007. *Causation, physics, and the constitution of reality: Russell's republic revisited*. Oxford University Press.

Prinz, Jesse. 2007. The intermediate level theory of consciousness. *The Blackwell companion to consciousness*, 257–271.

Quian Quiroga, Rodrigo, & Panzeri, Stefano. 2009. Extracting information from neuronal populations: Information theory and decoding approaches. *Nature Reviews Neuroscience*, **10**(3), 173–185.

Rosen, Robert. 1991. *Life itself: a comprehensive inquiry into the nature, origin, and fabrication of life*. Columbia University Press.

Rosenthal, David M. 2005. *Consciousness and mind*. Oxford University Press.

Sederberg, Audrey J, MacLean, Jason N, & Palmer, Stephanie E. 2018. Learning to make external sensory stimulus predictions using internal correlations in populations of neurons. *Proceedings of the National Academy of Sciences*, 201710779.

Seth, Anil K. 2014. A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive Neuroscience*, **5**(2), 97–118. 51 cites:.

Shamir, Maoz. 2014. Emerging principles of population coding: In search for the neural code. *Current Opinion in Neurobiology*, **25**, 140–148.

Stanley, Garrett B. 2013. Reading and writing the neural code. *Nature Neuroscience*, **16**(3), 259–263.

Stein, Richard B, Gossen, E Roderich, & Jones, Kelvin E. 2005. Neuronal variability: noise or part of the signal? *Nature Reviews Neuroscience*, **6**(5), 389.

Vul, Edward, Nieuwenstein, Mark, & Kanwisher, Nancy. 2008. Temporal selection is suppressed, delayed, and diffused during the attentional blink. *Psychological Science*, **19**(1), 55–61.

Vul, Edward, Hanus, Deborah, & Kanwisher, Nancy. 2009. Attention as inference: selection is probabilistic; responses are all-or-none samples. *Journal of Experimental Psychology: General*, **138**(4), 546.

White, John A., Rubinstein, Jay T., & Kay, Alan R. 2000. Channel noise in neurons. *Trends in Neurosciences*, **23**(3), 131–137.

=== backup ===

Biological evidence of informational closure in neural system

- So far, not so much research on non-trivial informational closure in biological neural systems.
- However, [[REWORDING | XXX]]
- Recent biological evidence suggests that low-level neurons encodes predictive information about external and
- **[[need to say this is exactly the non-trivial informational closure, and refer to equation XXX]]**
- Sederberg *et al.* (2018)
- *" in recordings of populations of RGCs of salamander retina driven by a simple stimulus with partially predictable dynamics, joint activity patterns transmit information that is predictive of future stimuli "*
- *" We examined whether the temporal correlations within populations of RGCs can be used to guide the search for readouts of predictive information."*
- *" We showed that groups of cells with high external, stimulus-predictive information also had high internal predictive information on two classes of stimuli, "*
- *" Very simple learning rules could find near-optimal readouts of predictive information without any external instructive signal. This suggests that bottom-up prediction may play an important role in sensory processing"*
- *" Internal predictive information can guide stimulus prediction without explicit reference to stimulus parameters."*
- *" Wordword internal predictive information is correlated with wordstimulus predictive information across sets of four cells."*
- They also showed generalization of internal predictability across stimulus. This suggests .... what?
- *" The generalization of internal predictability across stimulus classes partially reflects the ability of the retina to adapt to the statistics of stimuli"*
- Palmer *et al.* (2015)
- *" Groups of cells in the retina carry information about the future state of their own activity, and we show that this information can be compressed further and encoded by downstream predictor neurons that exhibit feature selectivity that would support predictive computations."*

- *" It seems natural to phrase the problem of prediction in relation to the visual stimulus, as in Fig. 3, but the brain has no direct access to visual stimuli except that provided by the retina. Could the brain learn, in an unsupervised fashion, to predict the future of retinal outputs? More precisely, if we observe that a population of retinal ganglion cells generates the word w t at time t, what can we say about the word that will be generated at time t + t in the future? The answer to this question is contained in the conditional distribution of one word on the other, P ð w t + t j w t Þ "*
  http://tinyurl.com/y9zmhlln
- *" larger groups of cells carry predictive information for hundreds of milliseconds, "*

Relation to empirical findings in neuroscience and studies of consciousness

8.1 Unconscious Processing

- According the theory, there are two factor that renders neural processes unconscious.
- First, processes are reflexive. This causes non-closured
- Second, the abstraction level of the information is below the non-trivial informational closure scale.

8.2 reflexive behaviours is unconscious (or environment-state dependent processes)

8.2.1 Reflection

- Reflexive systems is not informational closured.
- Base d on our theory,
- x -> s -> a
  - the system state is determined by environmental state, then it is not conscious.
- Therefore, in general, a simple policy network is not conscious.

8.2.2 Blindsight

- This also blindsight.
- Blindsight patients and still make above chance judgement without conscious awareness of visual targets.
- The stimulus-reaction policy network still intact, but it cannot form informational closure and therefore, patients has no conscious awareness of visual targets
- Patient B.F. reactions are stimulus driven.

8.2.3 procedure memory

8.2.4 cerebellum

- cerebellum feed-forward connections, maybe just a policy network.

Ref: cerebellum

8.3 Under the level of informational closure

8.3.1 statistical inference

- The computation of Bayesian brain is unconscious.

8.3.2 subliminal perception

- Above chance performance can be performed through non-closure pathway (reflexive behaviours)
- Due to coarse-graining, mapping between changes of microstate

8.3.3 Libet Experiments

8.4 Deterministic vs probabilistic

Dehaene *et al.* (2017b) Vul *et al.* (2008); Moreno-Bote *et al.* (2011); Asplund *et al.* (2014); Vul *et al.* (2009)

About Phenomenology (conscious processing)

8.5 Perceptual overflow

8.6 Volition and freewill

- Volition is a result of coarse-graining
- Lebel's experiment can be well explained by the experiment.

8.7 Planning

8.8 Imagination

8.9 Dream vs non-dream sleep

8.10 Binocular rivalry

8.11 Hallucinations

8.12 Synaesthesia

Theory predictions (Counter-intuitive prediction)

8.13 informational closure

- We predict that through coarse-graining, conscious agents create

8.14 Prediction on split-brain (vs IIT)

* IIT suggests that cutting, two subparts may both still have MIP, therefore, have two conscious mind.

* However, our theory suggests that if the cut largely destroy the ic, then both part may not be conscious.