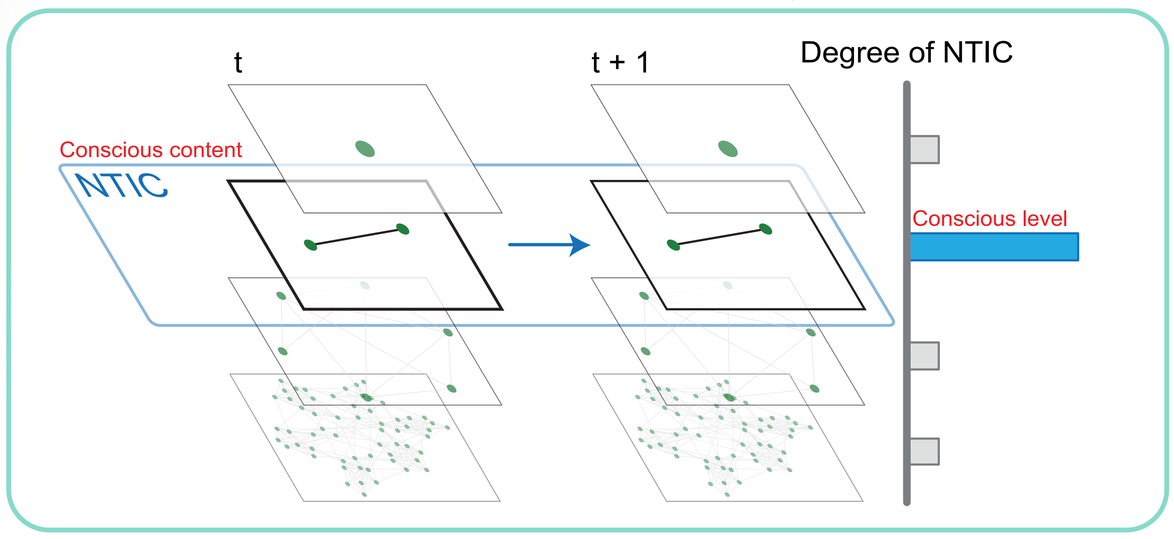
A neural coarse-graining theory of consciousness

Acer Y.C. Chang, Martin Biehl, Ryota Kanai
acercyc@araya.org
ARAYA, Inc., Tokyo, Japan

Graphical summary



Background

We are unable to consciously access information processed at every scale in the neural system. Only information at certain levels of coarse-graining can be parts of conscious contents.

Information processing in individual neurons is highly noisy. However, what we are aware of in our conscious contents shows astonishing stability and robustness against noise in the neural system.

Meanwhile, we don't experience "super-consciousness" when we have massive inter-individual interactions with other human-beings.

We propose that consciousness is associated with information processing at a certain coarse-grained level where the state dynamics forms a high degree of non-trivial informational closure (NTIC). The state of consciousness corresponds to the state of the NTIC process.

A neural coarse-graining theory of consciousness

The environmental dynamics for human scale agents is nearly deterministic.

However, at the microscopic levels, neurons or neural populations can only receive partial information about the environment and the information processing is highly noisy.

The human neural system may model the environment through coarse-graining —thereby achieving non-trivial informational closure.

Practically, this can be achieved by maximising the predictive power about the environment and the self-predictive information of the system concurrently (Guttenberg et al. 2016).

We claim that consciousness is associated with information processing at a specific coarse-grained level which achieves high NTIC in the neural system.

We postulate that the level of consciousness corresponds to the degree of NTIC.

$$C_t^{Level} = NTIC_t$$

The content of consciousness corresponds to the state of the NTIC process.

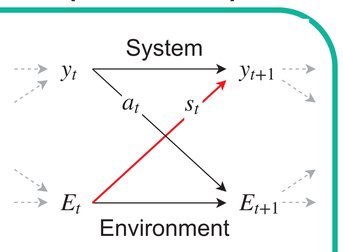
$$C_t^{Content} = Y_t$$

Based on Eq.(5), the self-predictive information about the future states that the current state holds and the information about the environment encoded in the process determines the level of consciousness.

This also suggests that the richness of the environment being modelled by the NTIC process has a direct contribution to the level of consciousness.

Non-trivial informational closure (NTIC)

Information closure can be defined by the near zero information flow J_t between the environment E and the system Y (Bertschinger et al. 2006).



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

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

$$I(Y_{t+1}; E_t) = 0 \quad \Rightarrow \quad J_t(E \to Y) = 0 \quad (2)$$

In the non-trivial case, the system trying to encode environmental dynamic can also achieve informational closure. This means

$$I(Y_{t+1}; E_t) \neq 0, \quad (3)$$

And, non-trivial informational closure can be defined as

$$\begin{aligned} NTIC &:= I(Y_{t+1}; Y_t) - I(Y_{t+1}; Y_t | E_t) \\ &= I(Y_{t+1}; E_t) - I(Y_{t+1}; E_t | Y_t) \end{aligned} \quad (4)$$

Maximising NTIC is equivalent to

$$\begin{aligned} &\text{maximise} \quad I(Y_{t+1}; Y_t) \quad \textit{and} \\ &\text{minimise} \quad I(Y_{t+1}; Y_t | E_t) \end{aligned} \quad (5)$$

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.

Conscious vs. Unconscious Processing

Our theory indicates that conscious information needs to be:

(1) processed at a certain coarse-grained level
(2) encoded in the NTIC process

Therefore, we predict that:

For conscious processing:

When a process achieves NTIC, the current state of the process contains the information about its future state (Eq. (5)). Therefore, the neural system can build forward models to predict the future state of the environment using the information encoded in the NTIC process. Cognitive functions recruiting forward models built upon it are conscious. (E.g., planning and imagination)

For unconscious processing:

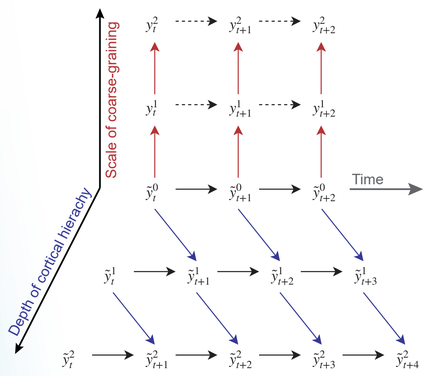
When processes fail to satisfy one of the two conditions, they remain unconscious.

~(1) Neural activities at microscopic levels (e.g. single cell activities, random neural fluctuation in Libet's free will experiment.

~(2) Processes do not achieve high NTIC (e.g. reflexive neural circuit, blindsight, procedure memory)

Theories about level of processing

Some theories also address the relationship between level of processing and consciousness. (e.g. Intermediate Level Theory by Jackendoff, 1987). However, it's crucial to note that our theory claims only the level of coarse-graining achieving NTIC matters to consciousness rather than the anatomical level of sensory hierarchy.



Predictive Processing (PP)

PP claims that our conscious percepts are the prediction about the external environment. However, PP does not explain why only the results of PP is conscious but not the underlying computation. Furthermore, PP does not explain why not all the predictive mechanisms in the neural system are conscious (e.g. mismatch negativity).

We explain and predict that

For those PPs of which the results are conscious, the conscious results are the coarse-grained information at the level achieving NTIC. The underlying computational processing is implemented at lower levels which are low informational closure (see Conscious vs. Unconscious Processing ~(1))

For those PPs that operate unconsciously, we predict that the states of those PPs are highly stimulus-dependent and act more like reflexive behaviors, and, therefore, have low NTIC. (see Conscious vs. Unconscious Processing ~(2)

Global workspace theory (GWT)

Our theory can well accommodate several claims by GWT.

Conscious perception necessitates representational stability: Our theory rests upon self-predictability which formally encompasses stability.

Broadcasting: When new sensory information from a modality is involved in NTIC, the information can be preserved in the following dynamic of NTIC. The neural system can extract the information for specific task goals. This is equivalent to the function of broadcasting. However, the current version of our theory does not specify the precise way to decode information from NTIC. More work is needed to describe broadcasting thoroughly.

Non-linear ignition: Because mapping microscopic to macroscopic states is many-to-one, it commonly shows a nonlinear relationship between microstates and macrostates. In our theory, new information presents in conscious perception corresponds to a transition in the state space of the NTIC process. This can cause a significant change of the brain states and resembles non-linear ignition of brain activities.

Integrated information theory (IIT)

Our theory is different from IIT which proposes that consciousness is information integration in a system.

For level of consciousness, IIT predicts a system with higher level of integration produces a higher level of consciousness. In our theory, the level of consciousness corresponds to the level tof NTIC, which suggests that level of consciousness is determined by self-predictive information and the information about the environment encoded in the process.

In terms of conscious content, IIT suggests that the causal structure of a system determines the quality of conscious contents. In the current version of our theory, how the state of NTIC process maps to conscious content has not been sophisticatedly specified. More work is needed to understand the relationship between the state of NTIC process and conscious contents.

Our theory is closely related to ***causal emergence*** described by Hoel et al. (2013) which was examined by a following work using integrated information theory (Hoel et al. 2016). More research on the precise mathematical relationship between causal emergence and our theory is needed.

Bertschinger, Nils, Olbrich, Eckehard, Ay, Nihat, & Jost, Jurgen. 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.

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.

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.

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