Response to Reviewers (Revision 2)

May 28, 2020

Reply to Reviewer 1

Q1: The authors should formally define what an "NTIC process" is (l. 155, l. 182). I take it to be "NTIC > 0" but I am not clear on whether the condition $J_t = 0$ also applies at this point. I do understand now that $J_t = 0$ is implied by a C-process.

A: The reviewer is correct. We refereed NTIC processes as processes with positive NTIC. To avoid potential ambiguity, we added a sentence to make this explicit at

> **Line 118:**
> *"For simplicity, in what follows, we refer to NTIC processes as those processes with positive NTIC."*

Furthermore, we also replaced "NTIC process" with "C-process" in many places in this revision especially when the contexts involve conscious experience and informational closure.

Q2: There are still some language issues and typos, especially in the new sections (2.1 etc).

A: We have read through this revision. We hope that we have corrected the language issues especially in Sec. 2.1.

Q3: l. 266-277: This paragraph is a bit confusing. There are two factors mixed together that influence the level of consciousness: the relative size of S and E, and then the coarse-graining scale of Y. Fig. 4 is only about the latter. In understand that the potential for consciousness is highest if X is split into equal parts S and E, but the relative size of S and E is not about coarse-graining per se. This could be made more explicit. Also, it seems like this paragraph contradicts a statement in l. 434-437 where it says that an NTIC micro process is sufficient for consciousness.

A: The reviewer is correct. In our current setup, the relative size of the state space $\mathcal{S}$ of S and the state space $\mathcal{E}$ of E is intertwined with the scale of coarse-graining. This is because the scale of coarse-graining a coarse-grained variable is quantified by the size of its state space. This means that we consider a coarse-grained variable is finer when we maps a microscopic variable to a coarse-grained variable with a larger state space.

In our current setup, the relationship among the state spaces $\mathcal{S}$, $\mathcal{E}$, and $\mathcal{X}$ are defined by the Cartesian product $\mathcal{X} = \mathcal{S} \times \mathcal{E}$. As described in Footnote 3, this setup implies that increasing the size of $\mathcal{S}$ inevitably decreases the size of $\mathcal{E}$. Consequently, increasing the relative size of $\mathcal{S}$ moves S (so as Y) and E toward a finer and a coarser scale, respectively.

We understand that this is not an ideal construction of ICT and is one of the current limitations of ICT (which is also related to an added paragraph in Limitation at Line 593 in our last revision). To make this point clear, we made the two following changes:

Line 264:
Original:
"At the finest scale we consider the whole universe X as the process Y"

Revised:

"Since we can quantify the scale of a coarse-grained variable by the size of its state space, therefore, at the finest scale we consider the whole universe X as the process Y."

A footnote is added at

> **Line 280:**
> *"In our current setup, the size of the state space $\mathcal{S}$ and $\mathcal{E}$ correspondingly determines the scale of coarse-graining of S and E. Further research is needed to reveal the relationship among NTIC, scales of coarse-graining, and different constructions of S and E. "*

We also agree that Line 434-437 is not well expressed. We did not mean to refer to the microscopic process as a process at the finest scale. We revised it as follows:

Line 438:
Original:
" However, ICT suggests that coarse-graining is necessary only when the microscopic processes are not NTIC (e.g. the neural system). An NTIC process can be formed in a noise-free deterministic system without coarse-graining. According to ICT, this NTIC process is sufficient for consciousness."

Revised:
" However, ICT suggests that coarse-graining is necessary only when a process is not informationally closed. Therefore, if a C-process is formed at a microscopic scale (e.g. the scale of individual neurons), according to ICT, this C-process is sufficient for consciousness."

Q4: l. 388-391: "Consequently, mental simulation, imagination, computing alternative realities, and generating counterfactuals often come with conscious experience" I'm not sure this statement fits with ICT. It seems to me that the internal process $Y_t \to Y_{t+1}$ has to be deterministic, or at least any uncertainty cannot be uncertainty that could be resolved at a lower level (because then it wouldn't be a C-process, as it would not be informationally closed with respect to S or X). This means that the C-process needs to be a good model of the environment, meaning that the correlation between Y_{t+1} and E_{t+1} needs to be very strong. So I don't see there being room for modeling "alternative realities". Because imagining counterfactuals would mean that Y_{t+1} ultimately does not fit with E_t and thus $I(Y_{t+1}; Y_t|E_t)$ would be much larger than if Y_{t+1} actually fit with E_{t+1}. So correct predictions should be good, but imagining counterfactuals seems bad from an ITC point of view. No?

A: If we understand the reviewer's question correctly, this question includes two contradictions to ICT from the paragraph about mental simulations:

1. The correlation Y_{t+1} and E_{t+1} needs to be strong for a C-process to be a good model of the environment.

2. During imagining counterfactuals and modelling alternative realities, Y_{t+1} ultimately does not fit with E_t. However, ICT hypothesises that the level of consciousness is contributed by the degrees of mutual information $I(Y_{t+1} : E_t)$.

To the first point, we can consider both NTIC processes (including C-processes) and the "synchronised" processes, which have high mutual information $I(Y_{t+1} : E_{t+1})$ are both good models of the environment. However, to be informationally closed, i.e. $I(Y_{t+1} : E_t|Y_t) = 0$, NTIC more emphasises how a process models the influence from the environment to its future states. We consider this as a critical feature for a system which models system-environment interactions.

Further more, $Y_t \to Y_{t+1}$ does not have to be deterministic ($H(Y_{t+1}|Y_t) = 0$) to be a C-process. A C-process needs to be informationally closed meaning that the transfer entropy term

$$\begin{aligned} I(Y_{t+1} : E_t|Y_t) &= H(Y_{t+1}|Y_t) - H(Y_{t+1}|E_t, Y_t) \\ &= 0 \end{aligned}$$

Therefore,

$$H(Y_{t+1}|Y_t) = H(Y_{t+1}|E_t, Y_t)$$

This indicates that when information about Y_{t+1} given E_t is involved in information about Y_{t+1} given Y_t, Y is informationally closed with respect to E even though Y is not deterministic.

Regarding the second point, the reviewer is correct that these mental simulations seem to be disentangled from the environment. However, other processes inside the bodies should also be considered as part of the environment of C-processes. Therefore, similar to our argument about dreaming, we consider that conscious experience happening in these internal simulations is contributed by the mutual information between the C-process and other processes in the neural system.

We agree that our original statement is too strong. We have revised this paragraph as follows:

Line 392:
Original:
"Cognitive functions involving simulations are expected to involve NTIC processes. Consequently, mental simulation, imagination, computing alternative realities, and generating counterfactuals often come with conscious experience. "

Revised:
"Cognitive functions involving internal simulations about agent-environment interactions (e.g. imagination, computing alternative realities, and generating counterfactuals) are expected to involve NTIC processes. We speculate that, these internal simulations may involve interactions between C-processes and other processes in the neural system. Therefore, they often come with conscious experience."

Q5: l. 602: if you mean integration in the IIT sense, IIT should be cited here.

A: Thank the reviewer for clarifying this. We do not intend to refer to the specific notion of integration defined by IIT. More generally, here we refer to any high-order dependencies. This is one of our work in progress and we are still working on potential informational objects that can be well adapted in our current framework.

We have added a footnote to disambiguate this concept at

Line 606:
"Integration here refers to any high-order dependencies."