Coalescing Minds and Personal Identity

Harri Valpola & Kaj Sotala

Accepted as contributed paper to the Toward a Science of Consciousness 2015 conference.

Recent work has discussed the possibility of merging human consciousnesses together by artificial means. Hirstein (2012; Mindmelding: Consciousness, Neuroscience, and the Mind's Privacy) argued that it is possible in principle for the subjective experience of one person to be experienced by another, and Grau et al. (2014; Conscious Brain-to-Brain Communication in Humans Using Non-Invasive Technologies. PLoS ONE) demonstrated a simple case of direct brain-to-brain communication over a great distance.

We focus on elaborating on Sotala & Valpola (2012; Coalescing minds: brain uploading-related group mind scenarios. International Journal of Machine Consciousness), who argued for the possibility for melding two consciousnesses together in a "reverse split-brain operation". They started by reviewing evidence for three properties of the brain that are related to consciousness:

1. to first approximation, the whole human neocortex is based on the same underlying algorithm which can learn to process many different kinds of input
2. global attention emerges from and the contents of consciousness is determined by biased competition, a distributed process in which representations compete locally for the right to broadcast information to other parts of the brain and where the competition in each area is biased according to information received from other areas
3. the cortical algorithm has an inbuilt ability to transfer information such as memories between cortical areas

From these premises, they proposed an exocortex, a prosthetic extension of the brain which employed the general cortical algorithm and which would integrate with a user's existing brain by using the same kinds of rules of biased competition and cortical memory transfer. Critically, the exocortex could then interface with other exocortices, connecting several brains together.

These could either meld together into a single mind with single conscious process, or have the exocortex filter the connections, so that both biological brains maintained their own global workspace (and thus consciousness) while still sharing information. Although the consciousnesses might not be capable of directly communicating with each other, memory formation would still work normally. In several mammal species, one of the cortical hemispheres is capable of sleeping while the other is awake, and the memories of what happened while the other hemisphere was asleep become available when both hemispheres are awake at the same time. As both conscious processes could have access to the synaptic weights storing the different memories, both could make use of the memories of what either processes had done earlier.

An exocortex would give rise to a number of scenarios touching upon personal identity, which had so far remained mere thought experiments. Two previously separate people could merge into a single consciousness for a while, then disconnect and split back into two. Possibly the more interesting question, however, would be the identity of a connected mind that shared two separate consciousnesses that could access each other's knowledge and memories, but would still remain somewhat separate. In a world with exocortices, individuality would become (even more of a) matter of degree, with two or more people having the possibility to meld themselves to varying extents.