# Coalescing Minds and Personal Identity
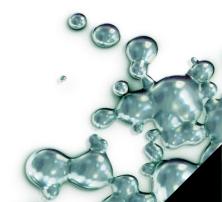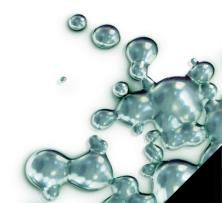
Harri Valpola

Kaj Sotala

# Overview

- How to coalesce minds using an exocortex (Harri Valpola)
  - "Reverse split-brain operation"
  - Neuroscience: properties of cortex making exocortex possible
  - Technology: different levels of integration (memories, consciousness)
- Implications to personal identity (Kaj Sotala)

**ZEN**robotics®

# Neuroscience Behind Exocortex

- Exocortex: a prosthetic extension of the brain (esp. cerebral cortex)

- Properties of mammalian cerebral cortex making exocortex possible:

1. Uniform cortical algorithm

2. Emergent attentional process through biased competition

3. Inherent information transfer between cortical areas
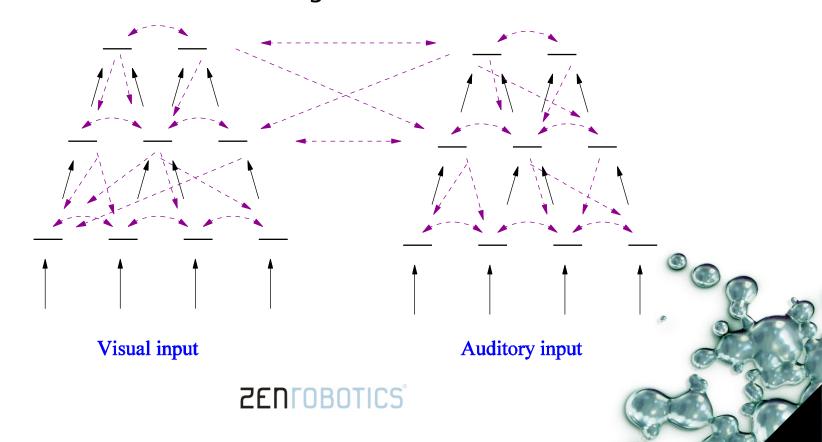
ZENROBOTICS®

# Uniform Cortical Algorithm

- To first approximation, the whole mammalian neocortex in based on the same underlying algorithm
  - Evolution and development: starts very uniform, specialises later
  - Hemispherectomy and language
  - Experiments with kittens: visual representations develop in "auditory cortices" if visual input is rewired there
  - Mice get trichromatic colour vision simply by adding a third type of cone cell on retina (no changes to cortex or other parts of the brain)
  - The brain can adapt to slowly occurring damage (such as hydroencephaly)
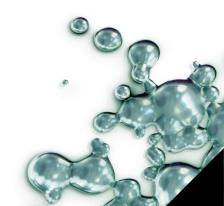
ZENROBOTICS

# Attention Emerges Through Biased Competition

- Each cortical area selects information to be represented
- Biased-competition model of attention: attention emerges from local selection and global communication



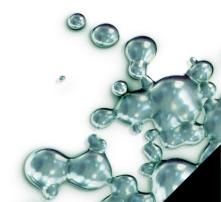Visual input                    Auditory input

# Information Transfer in the Cortex

- Representations on the cortex are plastic. They can change, e.g., in response to intensive training. Yet, memories are preserved.

- This is possible because one part of the cortex can learn the memories of another. (Cf. consolidation of memories.)

- In some mammalian species, one hemisphere is asleep while the other is awake. Memory transfer occurs while both are awake at the same time.

ZENROBOTICS

# How to Build an Exocortex

- Step 1: Make enough two-way connections with the cortex
  - Though straight-forward conceptually, the most difficult part in practice
  - 1K connections: clearly not enough
  - 100M connections: clearly enough (number of connections in corpus callosum)
- Step 2: Run the cortical algorithm long enough for the cortex and exocortex to learn each others representations
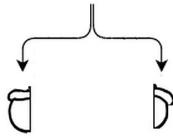  - Comparable to learning a new language, takes time

ZEΠ⌈OBOTICS

# Levels of Integration: Memories vs. Consciousness

- Exocortex makes adding new inputs and outputs to cortex trivial → opens up many possibilities for, e.g., augmenting senses

- Connect two exocortices together = reverse split-brain operation, coalesce two minds

- Coalesce individual exocortices with a standard interface exocortex: it will become possible to combine minds without any further learning

- Two main connection modes possible through a standard interface exocortex:
  - Shared synaptic changes = shared memories
  - Shared neural activations = unified conscious process
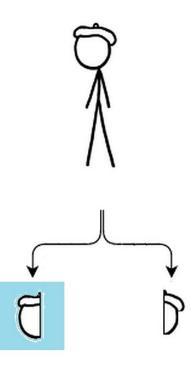
ZENROBOTICS®

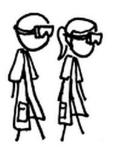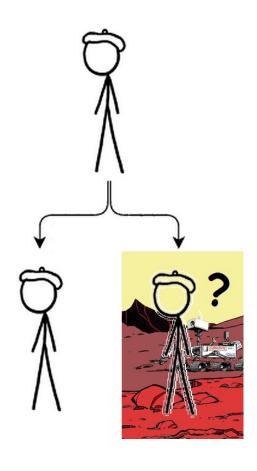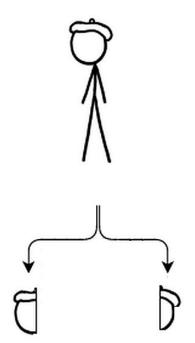# Parfit's split-brain: Lefty and Righty

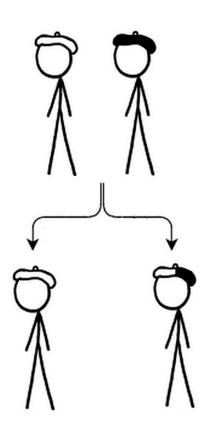# The freezing objection

# Parfit: Teletransporter

# Michael Cerullo (2015): Branching identity

"There will be continuity of consciousness between any conscious entity P1 at time T1 and P2 at time T1 if P2 contains half or more of the psychological structure of P1 at time T1 and is activated at time T1."
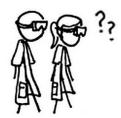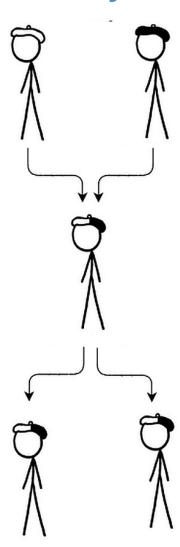
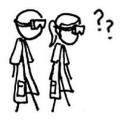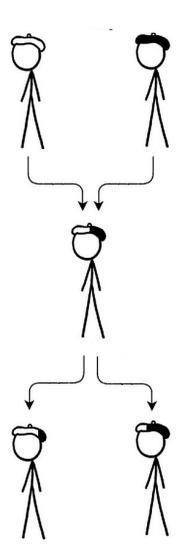# Branching identity (Cerullo): Memory transfer

# Branching & merging identity

# Branching & merging identity with partial memory transfer

# Identity as a non-binary condition

# Identity as a non-binary condition

- Recall Cerullo: "There will be continuity of consciousness between any conscious entity P1 at time T1 and P2 at time T1 if P2 contains half or more of the psychological structure of P1 at time T1 and is activated at time T1."

- Has an implicit binary condition: there either is continuity of consciousness or there is not

- 'Half or more' sounds a little arbitrary…

# Identity as a non-binary condition

- We propose: "There will be a partial continuity of consciousness between any conscious entity P1 at time T1 and P2 at time T1 to the extent that P2 contains psychological structure of P1 at time T1 and is activated at time T1."
- Fuzzy condition: e.g. 20% continuity of consciousness
- Could be specified further…