Computational Neurodynamics

Topic 13
Consciousness

Murray Shanahan

Overview

- Recap
- Global workspace theory
- The connective core hypothesis

Recap

- We have looked at
  - The brain's fundamental building blocks, neurons, and how to simulate them
  - How large numbers of neurons are connected together in complex networks,
- We examined various phenomena exhibited by such networks, such as
  - Competition
  - Dynamical complexity
  - Oscillation and synchrony

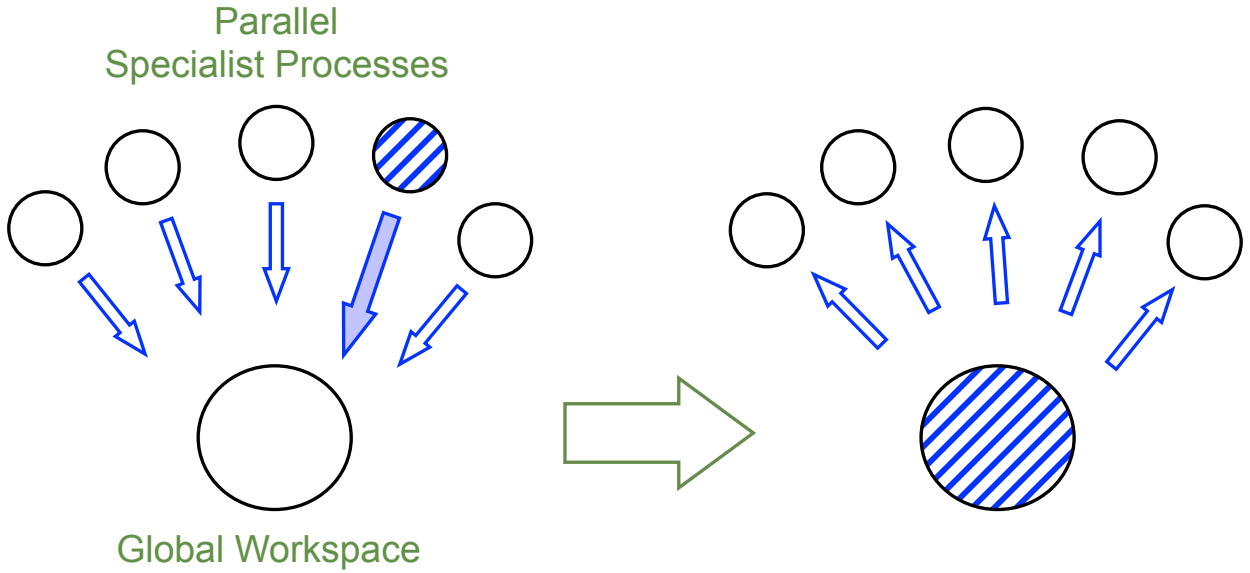
Long-term Aims

- How do all the elements of the course hang together?
- The ultimate aim of work in this area is to *reverse engineer* the biological brain
- And to do so by distilling the *deep principles* of its organisation and operation
- These principles should apply to *any* biological brain capable of sophisticated cognition
- And they should be applicable to the construction of *artefacts* capable of sophisticated cognition

Towards a Theory

- But neuroscience still lacks a fundamental theory of the brain, something comparable to the theory of evolution in biology
- Among the things this theory will have to account for are
  - The flexibility of cognition. That is, the ability to vary and combine the elements of an existing behavioural repertoire in novel ways
  - Consciousness. Or more precisely, the distinction between consciously-mediated behaviour and unconscious (automatic or habitual) behaviour

Global Workspace Theory 1



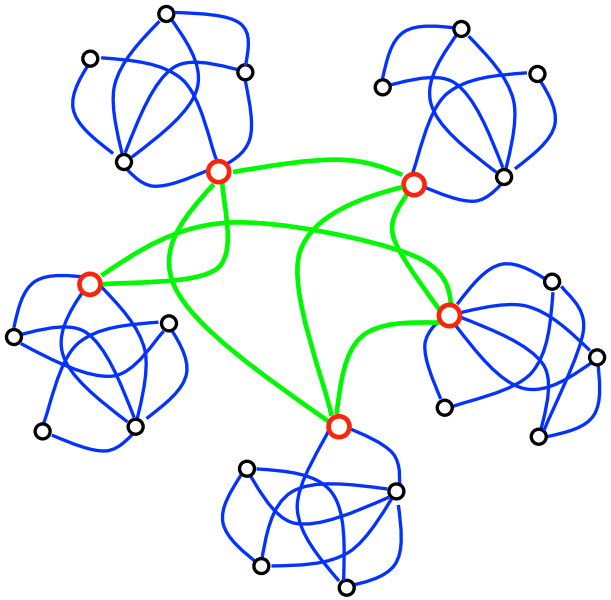
- Multiple parallel *specialist* processes compete and co-operate for access to a *global workspace*
- If granted access to the global workspace, the information a process has to offer is *broadcast* back to the entire set of specialists

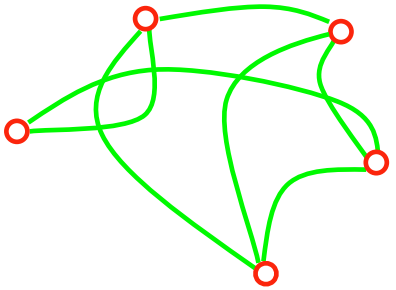
Global Workspace Theory 2

- The global workspace architecture harnesses massively parallel computation
- The global workspace itself exhibits a *serial* procession of states
- Yet each state-to-state transition is the result of filtering and integrating the contributions of huge numbers of *parallel* computations
- According to global workspace theory (Baars)
  - The human brain instantiates such an architecture
  - Information processing in the parallel specialists is non-conscious
  - Only information that is broadcast is consciously processed

Connective Cores (Again)

- Recall that extracting the connector hubs and their inter-connections from a modular network yields its *connective core*

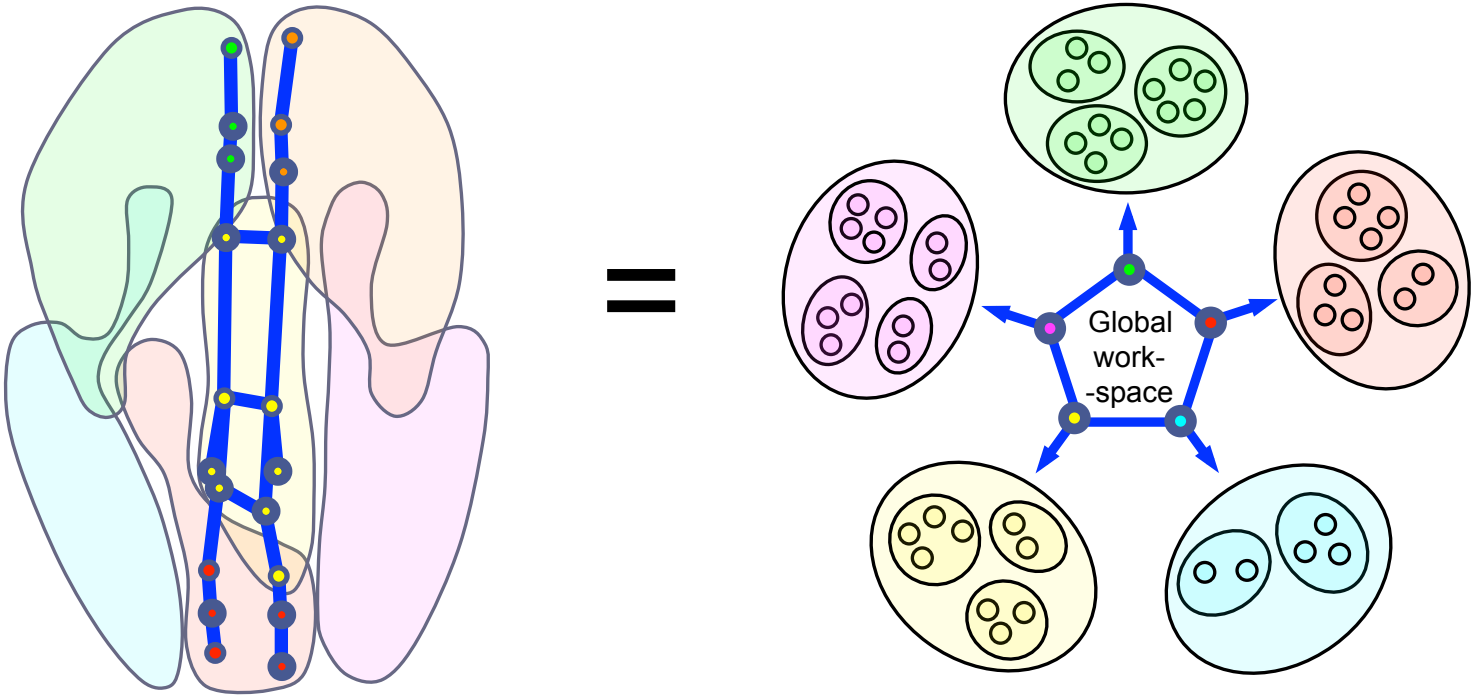




The connective core

The Locus of a GNW

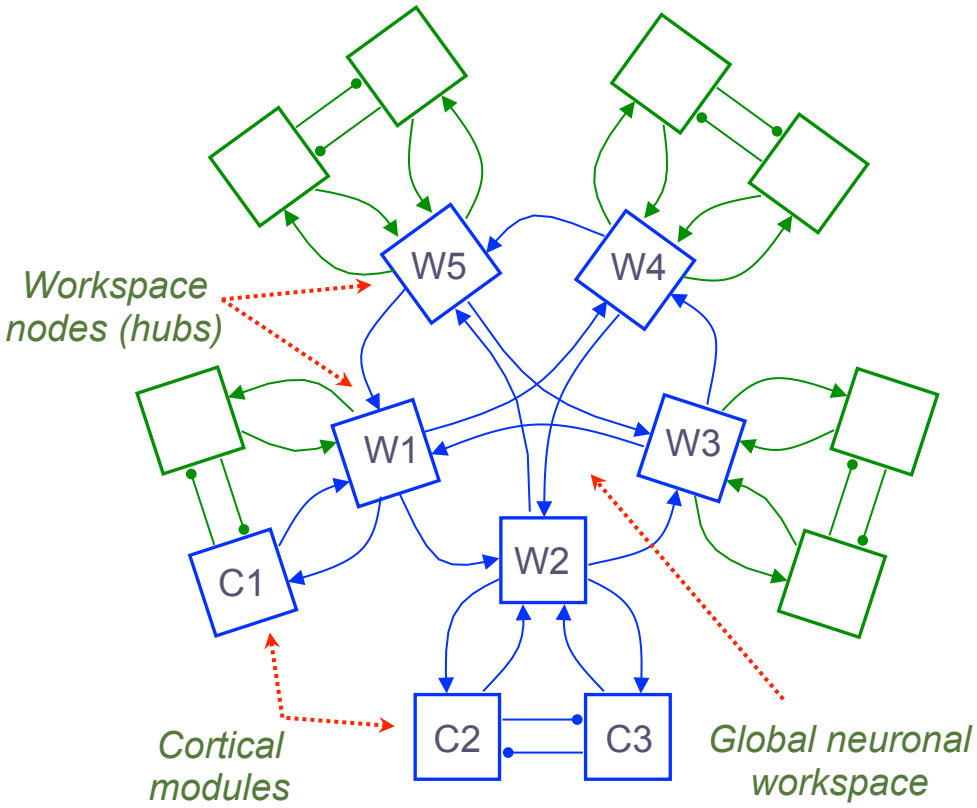
- The connective core of the human brain is a good candidate for the locus of a *global neuronal workspace* (GNW)



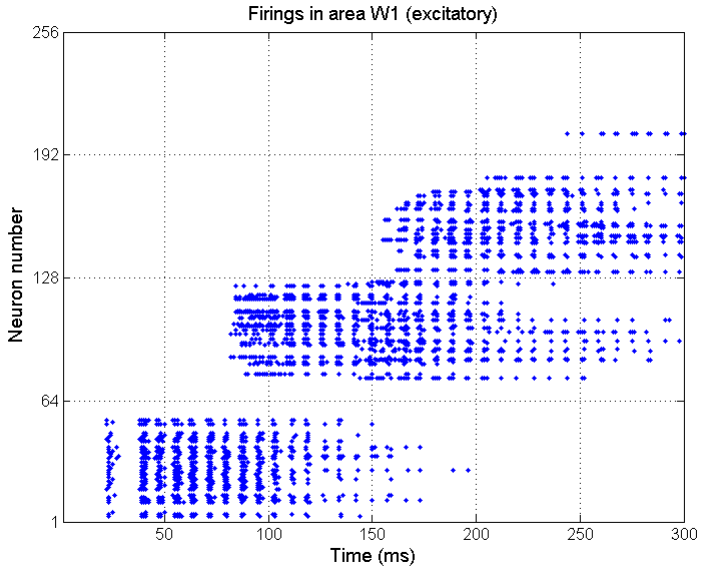
Workspace Connectivity

- A network that realises a GNW is likely to be (hierarchically) modular and small-world
  - The parallel specialist processes will be densely intra-connected modules
  - But to enable the brain-wide dissemination of the influence from any specialist process, there must be a web of inter-module connections
- It's also likely to have a pronounced connective core
  - Information and influence *funnels into* and *fans out from* the connective core
  - This imposes a *limited capacity* on the communications infrastructure
  - It encourages a degree of *serial processing*
  - And promotes winner-takes-all *competition*

A Computer Model

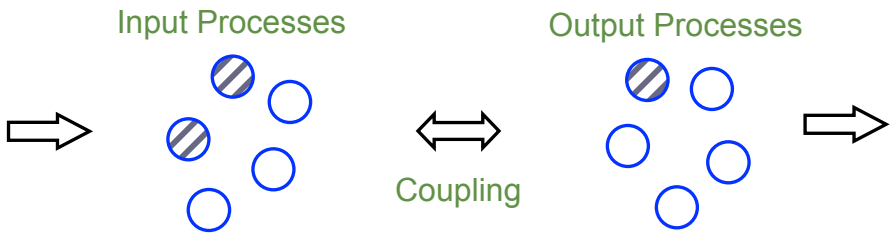


- This model (Shanahan, 2008) includes a connective core
- It generates a sequence of reverberating broadcast states

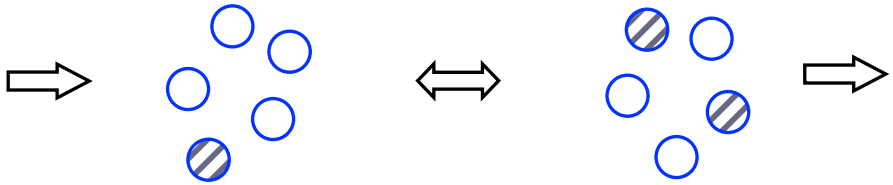


Coalition Formation

- Behaviour is generated by *coalitions* of coupled brain processes (sensory, motor, memory, affective)



- Coalition formation is part of a *cycle*. As time passes, coalitions form, then break up, then new coalitions form, and so on



- It is also *competitive*. Only one coalition can eventually take charge of a shared resource (eg: hand position, gaze direction)

The Connective Core Hypothesis

- Hypothesis: The blueprint for a cognitively capable brain must include a connective core
- The connective core is a limited capacity, highly connected communications infrastructure that provides
  - a locus of broadcast,
  - an arena for competition among coalitions of brain processes, and
  - a medium for coupling and communication between coalition members
- As well as supporting a conscious / unconscious distinction, it promotes *cognitive integration*

Cognitive Integration

- Cognitive integration is achieved when the *full resources of the brain* are brought to bear on the ongoing situation
  - Specialist motor skills in diverse domains
  - Episodic memory of diverse past experiences
  - Working memory
- This requires the availability of the whole *combinatorial repertoire* of process coalitions
- Total cognitive integration is an ideal, and failures are commonplace (eg: the U-bend of my sink)
- But even non-human animals manage impressive feats of cognitive integration (eg: innovative tool manufacture in corvids)

The Need for Broadcast

- How does broadcast subserve *open-ended*, *combinatorial* coalition formation?
  - In order to maximise the availability of a system's resources, to allow the largest repertoire of combinations, it must be possible for each of its components to influence and be influenced by the whole system
  - That is to say, the system must exhibit dynamical complexity, a balance of integrated and segregated activity
  - This demands a global communications infrastructure to disseminate influence and information among otherwise independent processes

The Need for a Bottleneck

- Why does the broadcast mechanism have to be a limited capacity “bottleneck”?
  - Only one coalition at a time can take charge of what the animal or robot does next
  - So coalition formation is competitive, and there can only be one winner at a time
  - The connective core enforces this winner-takes-all rule because only one coalition at a time can dominate it
  - This also enforces serial processing, which is vital for the *sequential chaining* of mental operations
  - The connective core is a bottleneck, in a good way

Conclusion: Seriality and Unity

- Information and influence funnel in to and fan out from the connective core (GNW), which acts as a limited bandwidth processing bottleneck, allowing for serial mental operations

Serial from Parallel

- It also promotes integration across the brain, so that a coherent response to the ongoing situation can be orchestrated from its full resources

Unity from Multiplicity

Related Reading

Baars, B.J. (1997). *In the Theater of Consciousness: The Workspace of the Mind*. Oxford University Press.

Shanahan, M. (2008). A Spiking Neuron Model of Cortical Broadcast and Competition. *Consciousness and Cognition* 17, 288–303.

Shanahan, M.P. (2010). *Embodiment and the Inner Life: Cognition and Consciousness in the Space of Possible Minds*. Oxford University Press.

Shanahan, M. (2012). The Brain's Connective Core and its Role in Animal Cognition. *Philosophical Transactions of the Royal Society B* 67(1603), pp.2704–2714.