Global Workspace Architecture

Linking Cognition and Consciousness

Murray Shanahan

Imperial College London

Department of Computing

Overview

- Three related issues
  - Neural parallelism
  - Modular theories of mind
  - The frame problem
- Global workspace theory (GWT)
- Applying GWT to the three issues
- A spiking neuron model

First Issue

Neural Parallelism

- An animal's nervous system is massively parallel
- Masssive parallelism surely underpins human cognitive prowess
- So how are the massively parallel computational resources of an animal's central nervous system harnessed for the benefit of that animal?
- How can they orchestrate a coherent and flexible response to each novel situation?
- What is their underlying architecture?
- Nature has solved this problem. How?

Second Issue

Modular Theories of Mind (1)

- Many cognitive scientists advocate modular theories of mind (Gardner, Tooby & Cosmides, Fodor, Mithen, Carruthers)
- The mind comprises (or incorporates) an assemblage of distinct specialist *modules*
- Fine-grained horizontally modular theories (eg: Tooby & Cosmides) posit specialists for particular behaviours (eg: foraging)
- More coarse-grained vertically modular theories (eg: Fodor) posit specialists for certain input and output processes (eg: parsing, low-level vision)

Modular Theories of Mind (2)

- In addition to the specialist modules, all modular theories demand (for humans) some overarching faculty, central system, super-module, meta-representational facility, or whatever
- This addition is capable, when required, of transcending modular boundaries to produce flexibly intelligent behaviour rather than an automatic, preprogrammed response to a novel situation
- But nobody has a very convincing account of this

Third Issue

The Frame Problem (1)

- The frame problem orginated in classical AI

  How can we formalise the effects of actions in mathematical logic without having to explicitly enumerate all the trivial non-effects?

- This is tricky, but was more-or-less solved in the mid 1990s
- Our concern is the wider interpretation given to the frame problem by philosophers, notably Dennett and Fodor

The Frame Problem (2)

- Fodor's version:

  How do *informationally unencapsulated* cognitive processes manage to select only the information that is relevant to them without having to explicitly consider everything an agent believes ?

- A cognitive process is *informationally unencapsulated* if it has the potential to draw on information from any domain
- Analogical reasoning is the epitome of informational unencapsulation

Computational "Infeasibility"

- Fodor claims that informationally unencapsulated cognitive processes are computationally infeasible

> " The totality of one's epistemic commitments is *vastly* too large a space to have to search … *whatever* it is that one is trying to figure out. "
>
> (Fodor, 2000)

- Fodor believes that this is a fatal blow for cognitive science as we know it because it entails we cannot find a computational explanation of the human mind's "central systems"

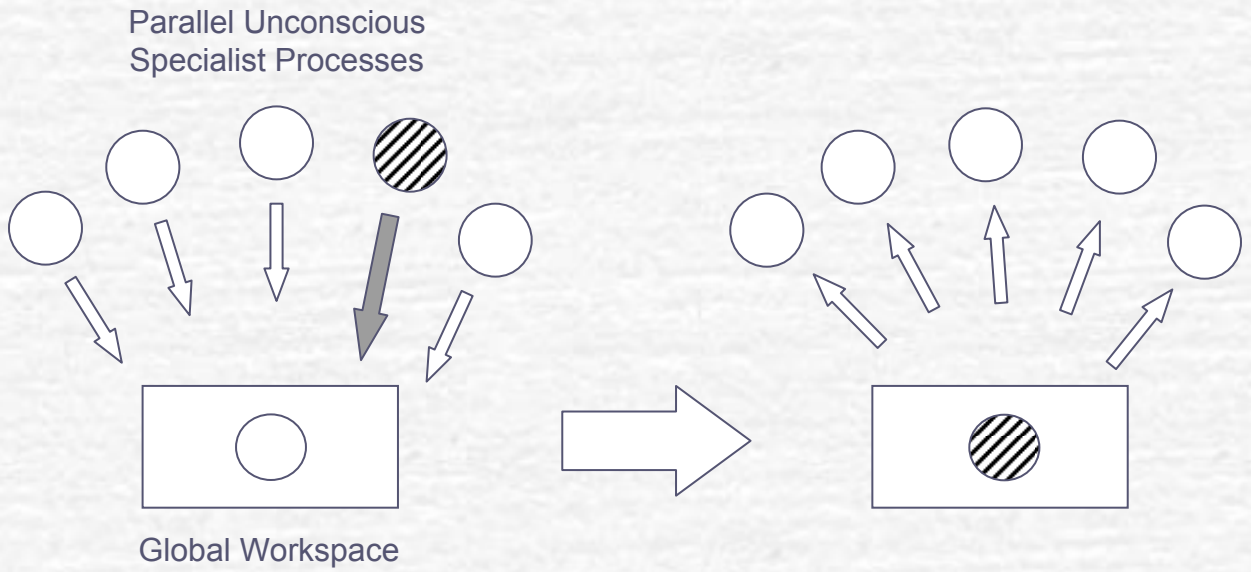
Fodor's Modularity of Mind

- The mind's *peripheral* processes are special purpose, do things like parsing and low-level vision, and are computational
- The mind's *central* processes are general purpose, do things like analogical reasoning, are informationally unencapsulated, and (probably) *aren't computational*

> "… it probably isn't true that [all] cognitive processes are computations. … [so] it's a mystery, not just a problem, what model of the mind cognitive science ought to try next. " (Fodor, 2000)

The Solution

Global Workspace Architecture



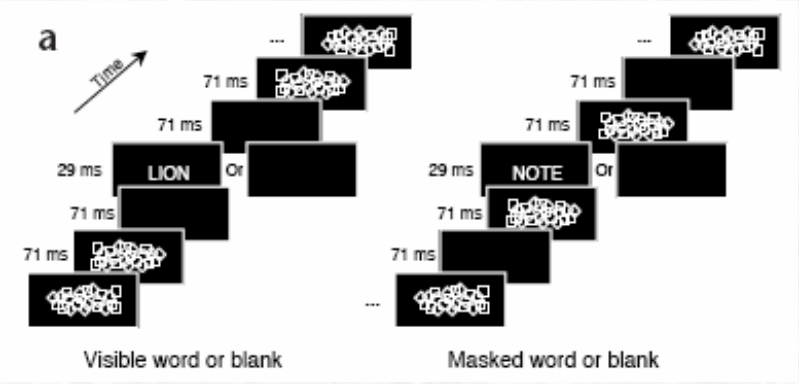
- Multiple parallel *specialist* processes compete and co-operate for access to a *global workspace*
- If granted access to the global workspace, the information a process has to offer is *broadcast* back to the entire set of specialists

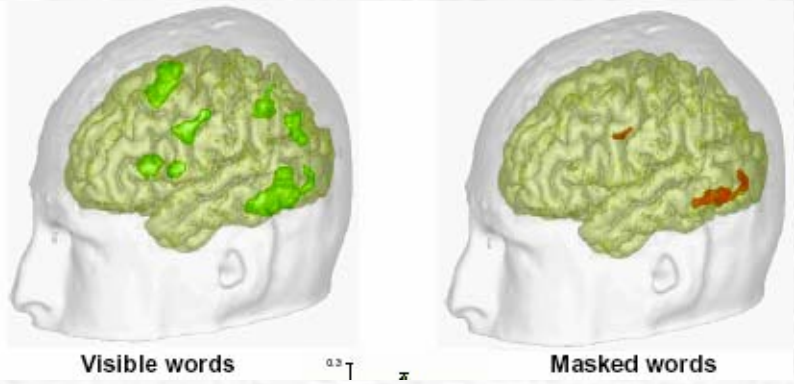
Conscious vs Non-Conscious

- Global workspace theory (Baars) hypothesises that the mammalian brain instantiates such an architecture
- It also posits an empirical distinction between conscious and non-conscious information processing
- Information processing in the parallel specialists is non-conscious
- Only information that is broadcast is consciously processed

Empirical Evidence

- Contrastive analysis compares and contrasts closely matched conscious and unconscious brain processes
- Dehaene, *et al*. (2001)
  - Imaged subjects being presented with “masked” words
  - Masked and visible conditions compared



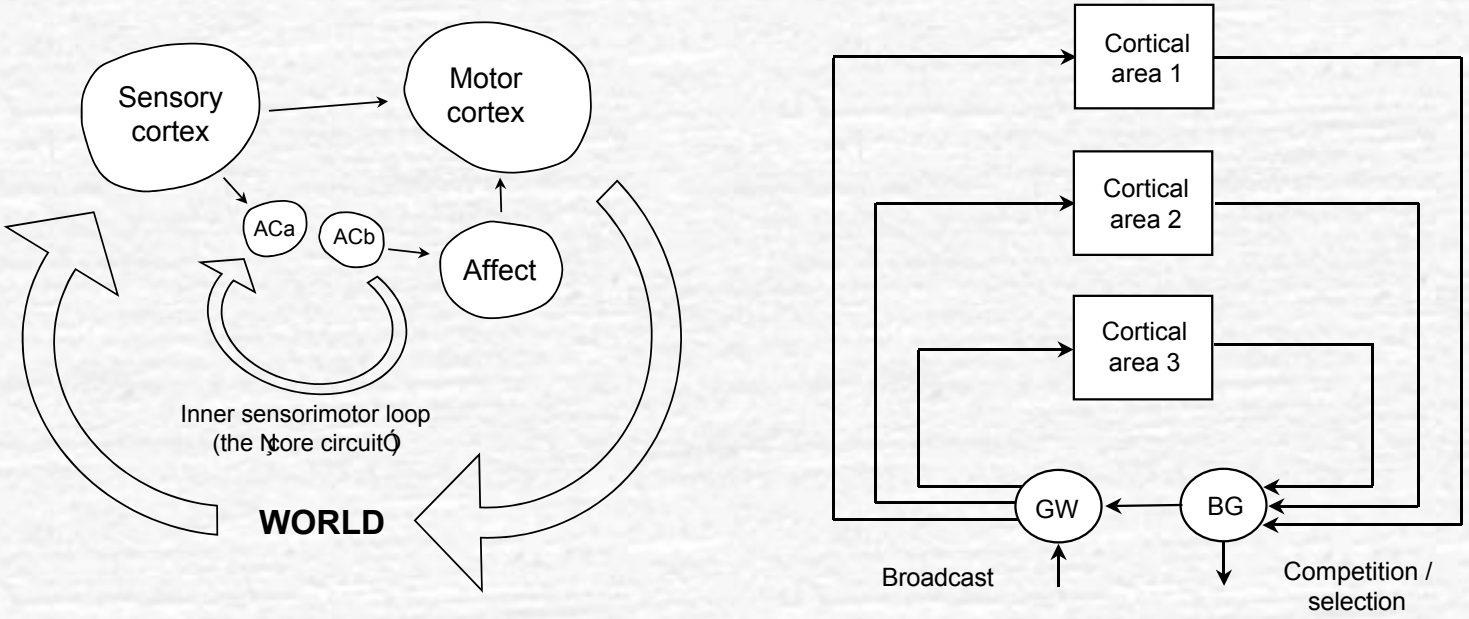


- Such experiments suggest that conscious information processing recruits widespread brain resources while unconscious processing is more localised

Embodiment

- According to GWT, only something that instantiates a global workspace architecture is capable of conscious information processing
- But this is a necessary not a sufficient condition
- I have argued (Shanahan, 2005) that the architecture must direct the actions of a spatially localised body using a sensory apparatus fastened to that body
- This allows the set of parallel specialists a shared viewpoint, from which they can be indexically directed to the world and fulfil a common remit

Combining GWT with Internal Simulation



An internal sensorimotor loop can be combined with mechanisms for broadcast and competition (Shanahan, 2006), thereby marrying the *simulation hypothesis* (Cotterill, Hesslow) with *global workspace theory* (Baars)

Applying the Solution

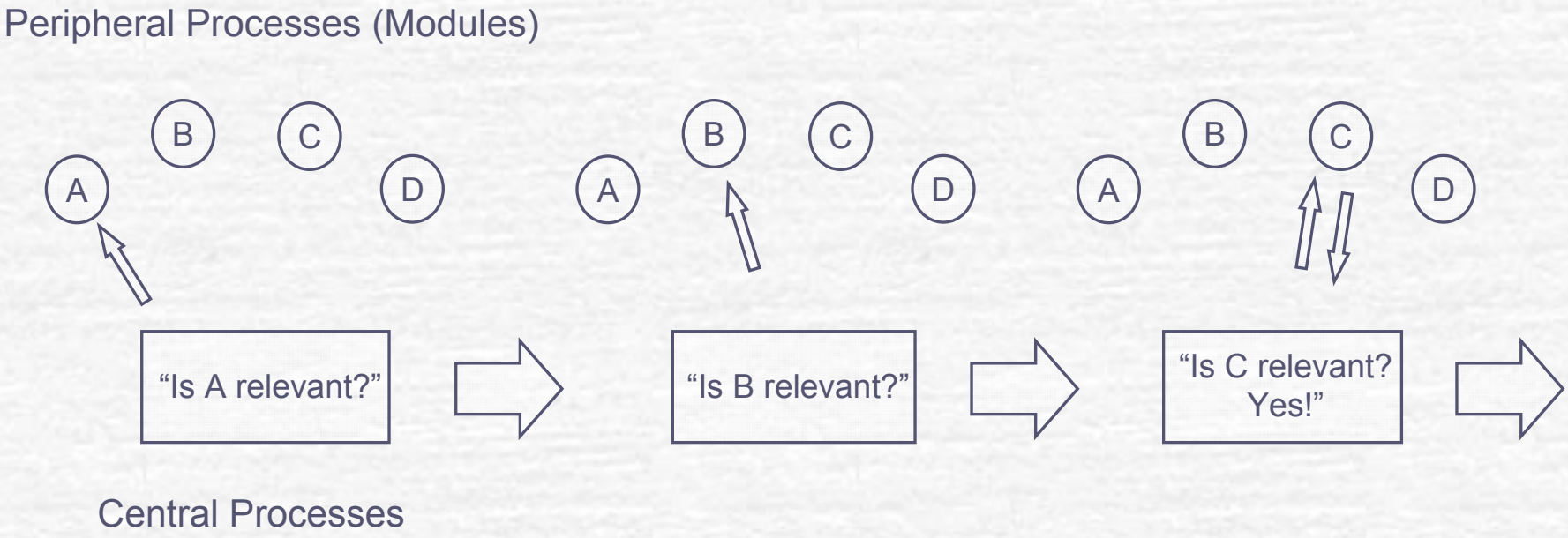
Remember This

Serial from Parallel / Unity from Multiplicity

- The global workspace architecture harnesses the power of 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
- The global workspace architecture thereby distils unity out of multiplicity
- This is perhaps the essence of consciousness, of what it means to be a singular, unified subject

GWT and the Frame Problem (1)

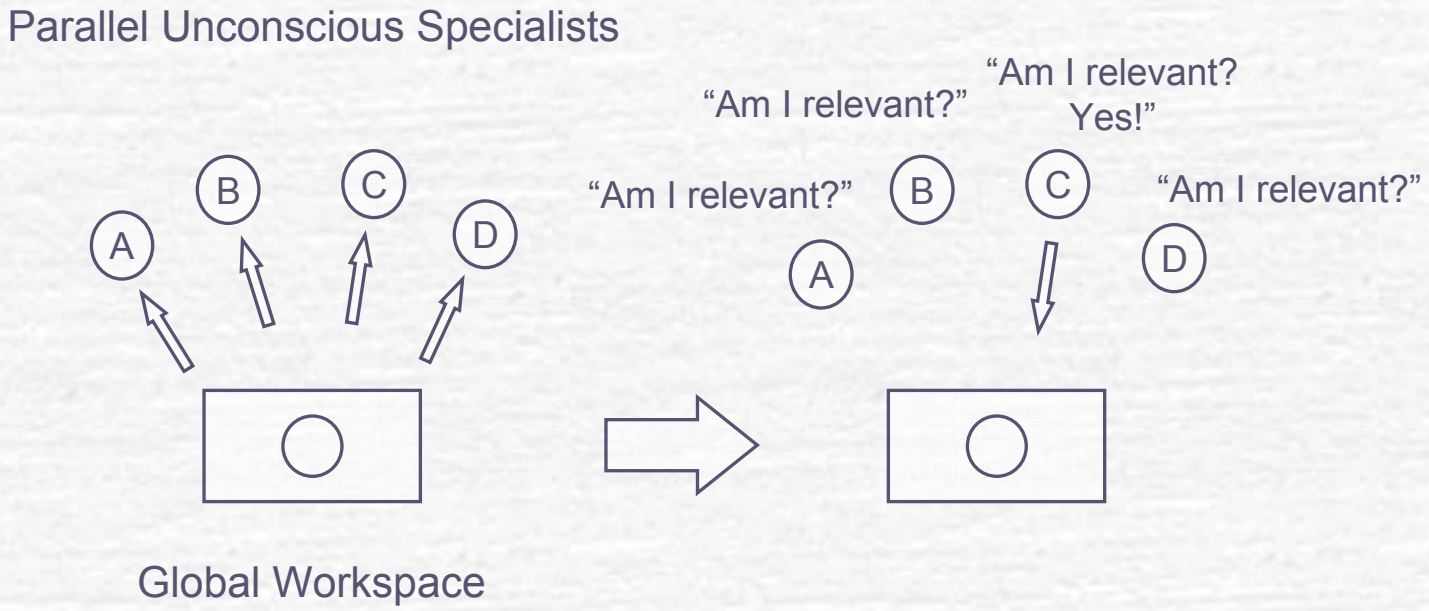
- Both Fodor and Dennett seem to have a strictly serial architecture in mind when they characterise the frame problem



- This certainly looks computationally infeasible

GWT and the Frame Problem (2)

- But global workspace architecture offers a parallel alternative



- In the context of an appropriate parallel architecture, the frame problem looks more manageable (Shanahan & Baars, 2005)

Analogical Reasoning (1)

- Analogical reasoning is informational unencapsulation in its purest form

  "Analogical reasoning depends precisely upon the transfer of information among cognitive domains previously assumed to be irrelevant " (Fodor)

- Computational models of analogical reasoning distinguish between
  - *retrieval* – the process of finding a potential analogue in long-term memory for a representation already in working memory – and
  - *mapping* – the subsequent process of finding correspondences between the two

Analogical Reasoning (2)

- Retrieval is the locus of the frame problem in analogical reasoning
- The most psychologically plausible computational model is currently LISA (Hummel & Holyoak), which mixes serial and parallel computation, and also fits a global workspace architecture very closely

LISA		GWT
Serial presentation in working memory	⟷	Serial presentation in global workspace
Active proposition in WM	⟷	Contents of GW
Parallel activation in long-term memory	⟷	Broadcast of GW contents
Distributed propositions in LTM	⟷	Unconscious parallel processes

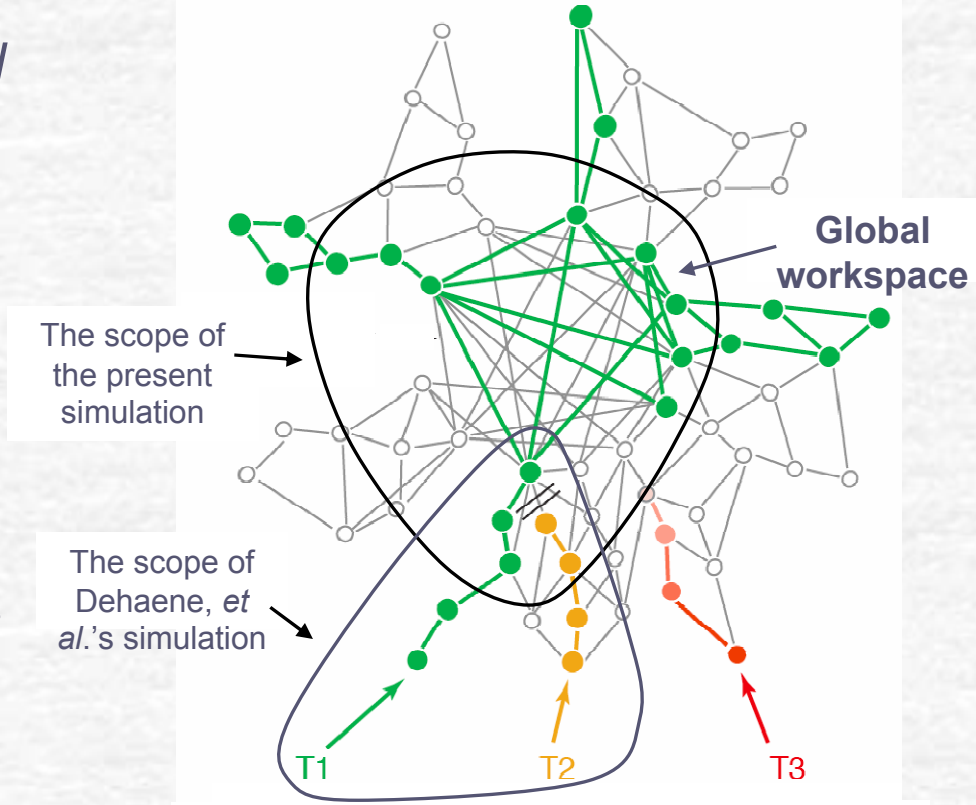
Rescuing Modularity

- The global workspace architecture can be appropriated by any of the modular theories of mind
- It potentially supplies the means of transcending modular boundaries required to realise human-level, flexible, creatively intelligent cognition
- Its application to the frame problem in general, and to analogical reasoning in particular, is an example of this

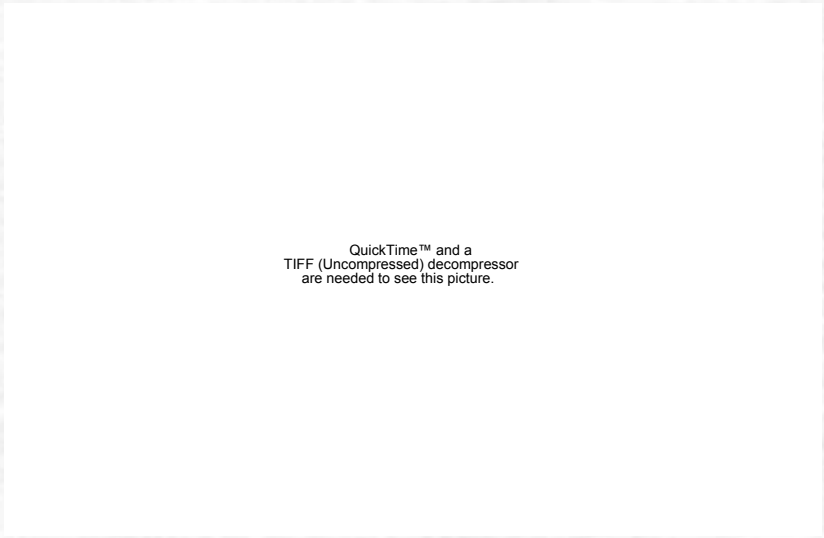
A Spiking Neuron Model

A Global Neuronal Workspace

- Dehaene, Changeux, *et al* postulate a global *neuronal* workspace as the neural substrate for GWT
- Workspace realised by long-range cortico-cortical fibres
- They have built computer models of various aspects of the GNW hypothesis
- But these do not model broadcast



White Matter Substrate

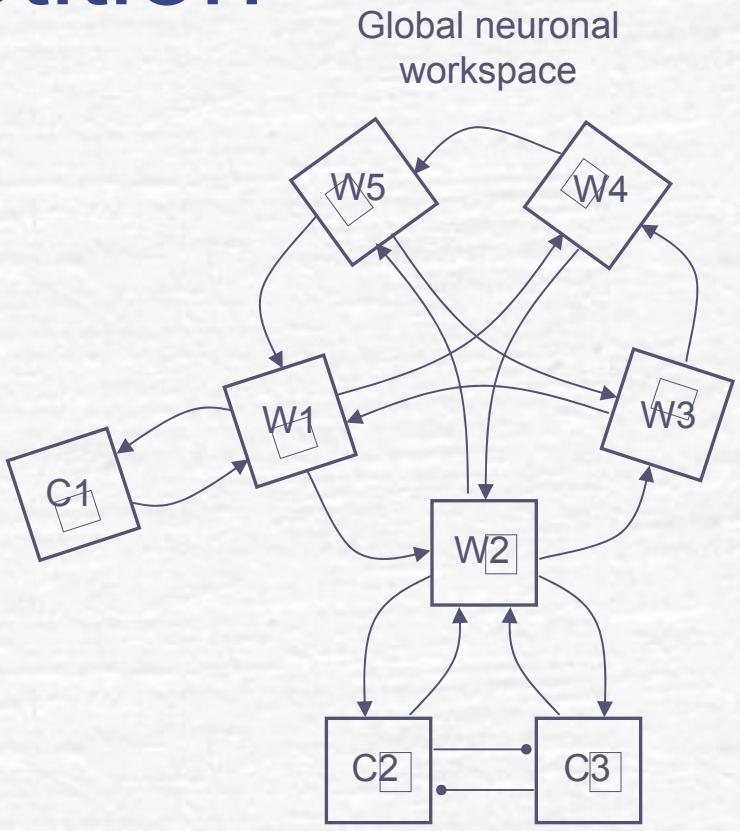


Corticocortical fibres
(Wakana, *et al*., 2004)

Corticocortical and thalamocortical fibres
(O'Donnell & Westin, 2006)

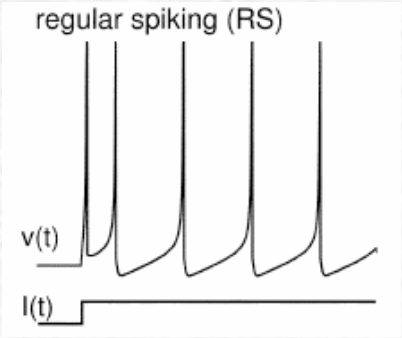
Modelling Broadcast and Competition

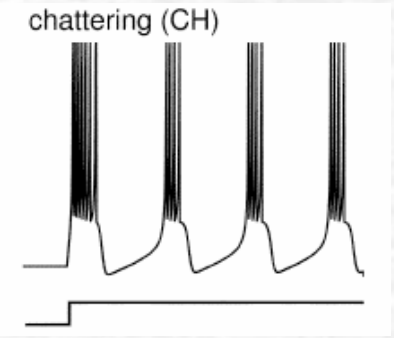
- The model comprises five workspace nodes (W1 to W5) and three further cortical columns (C1 to C3)
- Columns C2 and C3 are assumed to be close to each other, while C1 is remote
- C2 and C3 compete for access to node W2
- A pattern that appears on one workspace node spreads (is broadcast) to the rest and sustained through reverberation



The Spiking Neuron Model

- Izhikevich has devised a simple mathematical model of a spiking neuron with favourable computational properties





- By varying four parameters a, b, c, and d, it can be made to exhibit a range of behaviours comparable to the Huxley-Hodgkin model

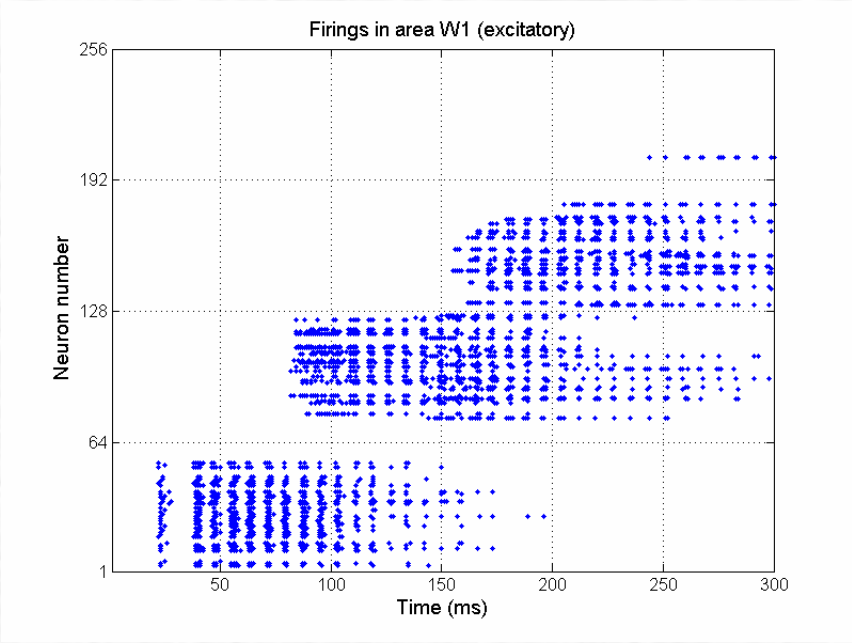
$$\dot{v} = 0.04v^2 + 5v + 140 - u + I$$

$$\dot{u} = a(bv - u)$$

$$\text{if } v \geq 30 \text{ then} \begin{cases} v \leftarrow c \\ u \leftarrow u + d \end{cases}$$

A Sequence of Workspace States

- When the whole thing runs, the workspace exhibits a sequence of distinct broadcast states
- In this run, we have three states, supplied successively by C1 and C3
- In this run C2 loses the competition. In others it wins



Conclusion

There seems to be a fundamental link between cognition and consciousness

So perhaps an understanding of cognition has to go hand-in-hand with an understanding of consciousness

And perhaps cognitive architectures need to incorporate global workspaces

References

Shanahan, M.P. & Baars, B.J. (2005). Applying global workspace theory to the frame problem, *Cognition* 98 (2), 157–176.

Shanahan, M.P. (2005). Global access, embodiment, and the conscious subject, *Journal of Consciousness Studies* 12 (12), 46–66.

Shanahan, M.P. (2006). A cognitive architecture that combines inner rehearsal with a global workspace, *Consciousness and Cognition* 15, 433–449.

Shanahan, M.P. (2007). A spiking neuron model of cortical broadcast and compeition, *Consciousness and Cognition*, in press.