COGS

Machine Consciousness: A Design Procedure.

Abstract: Whether anyone likes it or not, there is something out there that goes under the heading of 'machine consciousness' or 'models of consciousness'. It leads to gatherings of people who come from different backgrounds, some from AI and some from computer modelling of the brain. I fall into the second group and will talk of possible contributions that understanding brain mechanisms through digital models might make to the question: is it possible to build a conscious machine? I shall give a brief description of the digital methodology and then suggest that five major, personally felt attributes are fundamental: perception, imagination, attention, prediction and emotion. These are due to interlocking mechanisms in the brain which might serve as design models for a conscious machine. If time permits, this will be examined against Chalmers' philosophical objections to the idea of a conscious machine.

Neuro-IT Brussels 2/12/03

Machine Consciousness: A Fragile and Embryonic Paradigm

Igor Aleksander FREng

Leverhulme Fellow and Research Professor

Imperial College, London

Visiting Research Fellow, U. Sussex

Machine Consciousness Conferences

2001: Cold Spring Harbour

2003: Symposium at ASSC 7, Memphis

2003: ESF Workshop, Birmingham

2003: NoE ‘Exyxtence’ Workshop, Turin

"… For twenty years I have mistrusted consciousness .. It is the name of a non-entity and has no right place among first principles ..."

William James

'Does Consciousness exist?' 1904

Consciousness:

Where are the molecules?

Being conscious:

What are the mechanisms?

"… Plato, Descartes and Eccles, make no clear distinction between the terms 'consciousness', 'mind' and 'soul'. But in the modern context these terms have different meanings."

Max Velmans, 'Understanding Consciousness' 2000

What do I mean?

Being conscious means that I am experiencing internally an out-there world, self , past, future, intention and fiction.

My Mind is the sum total of my capacity for having such experiences.

Soul is best left to theologians

WHY MACHINES?

There comes a time when you have to stop talking and MAKE SOMETHING

EVEN IF IT IS A MISTAKE!

The late Tom Kilburn of Manchester University

OBJECTIVES FOR MAKING THINGS

To explore ‘what it is to be conscious’ with ***engineering clarity***

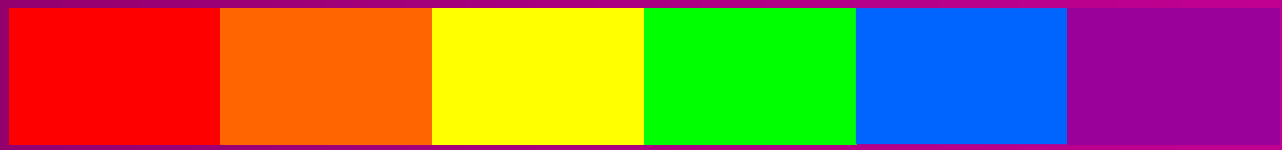
To achieve behaviours in machinery that, in areas of technological need, approach the competence achieved by conscious organisms.

(e.g. an autonomous exploratory robot on Mars)

THERE IS A SPECTRUM OF WAYS IN WHICH PEOPLE APPROACH

BEING CONSCIOUS

IN MACHINES

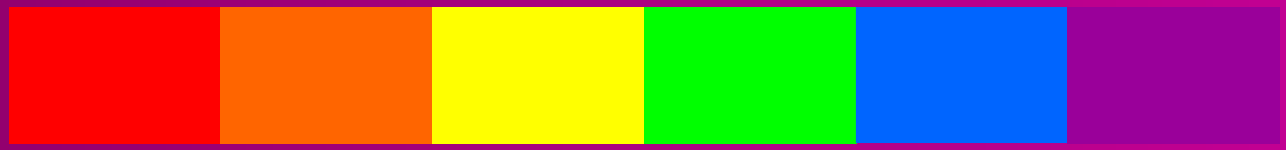


FUNCTIONAL

WHAT IS IT TO BEHAVE IN A CONSCIOUS WAY?

MATERIAL

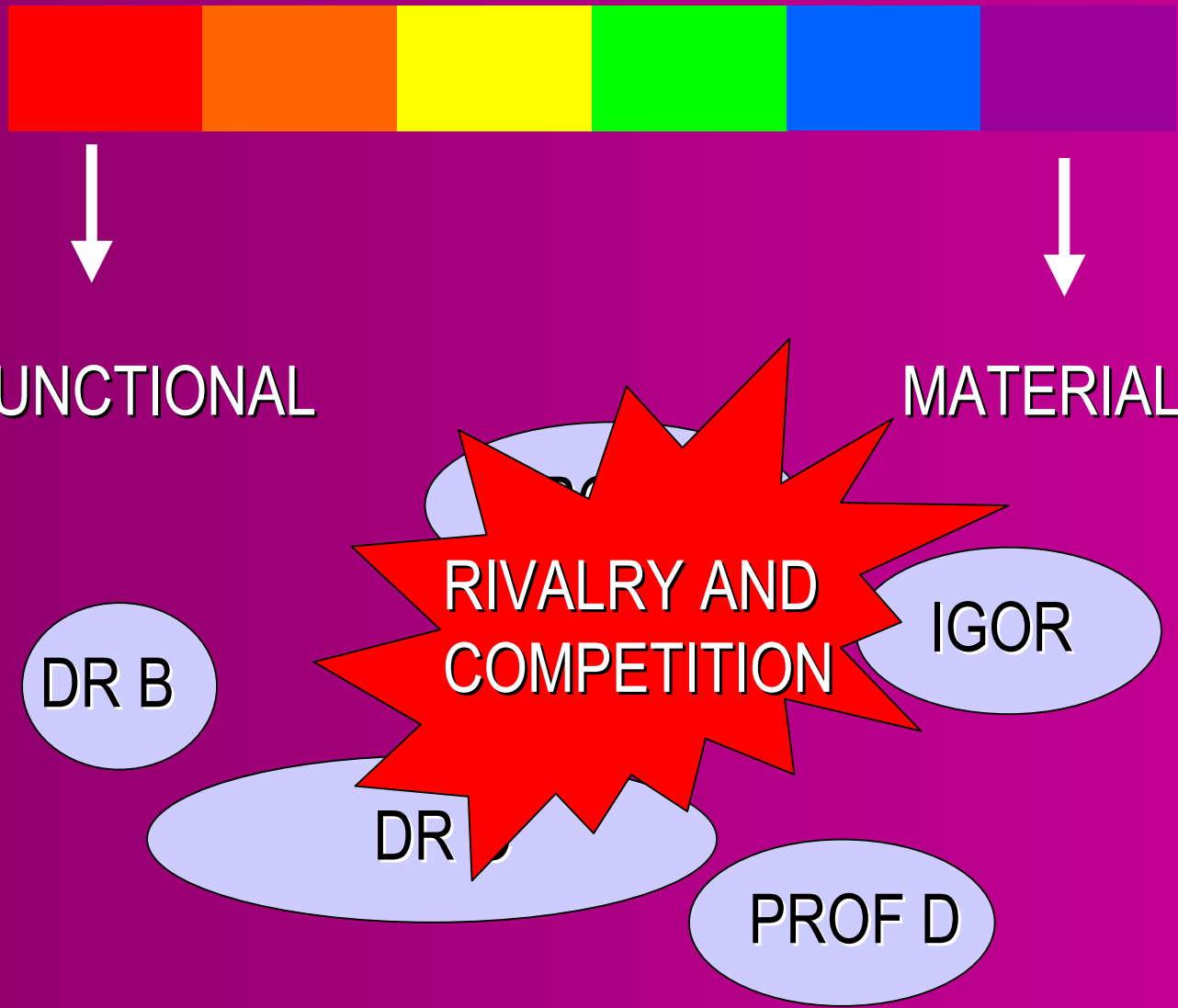
WHAT ARE THE ESSENTIAL MECHANISMS FOR BEING CONSCIOUS?

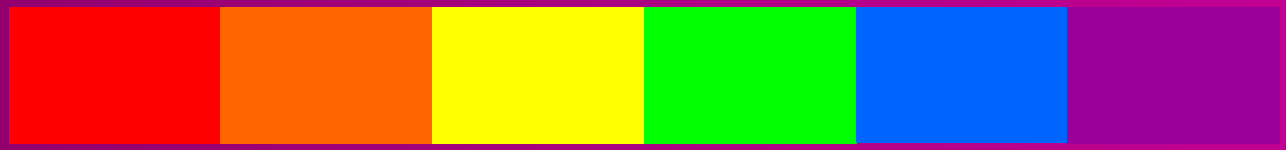


FUNCTIONAL

MATERIAL

MACHINES
PARADIGM


FUNCTIONAL
MATERIAL
RIVALRY AND COMPETITION
DR B
IGOR
PROF D



FUNCTIONAL

MATERIAL

Today's talk ….

IGOR

The Importance of Neural Structures

In retrospect … (where I come from…)

11 years of agonising about the possibility of MC:

From an automata theory/material point of view

1992: The Colin Cherry Memorial lecture

ICANN 92

1996: ‘Impossible Minds’: 1 postulate 11 corollaries

2000: Depictive theory published

2001: Axioms introduced at Skovde

2003: Axiomatic/Depictive theory published in JCS

Key design principle:

Introspection, shunned by psychologists

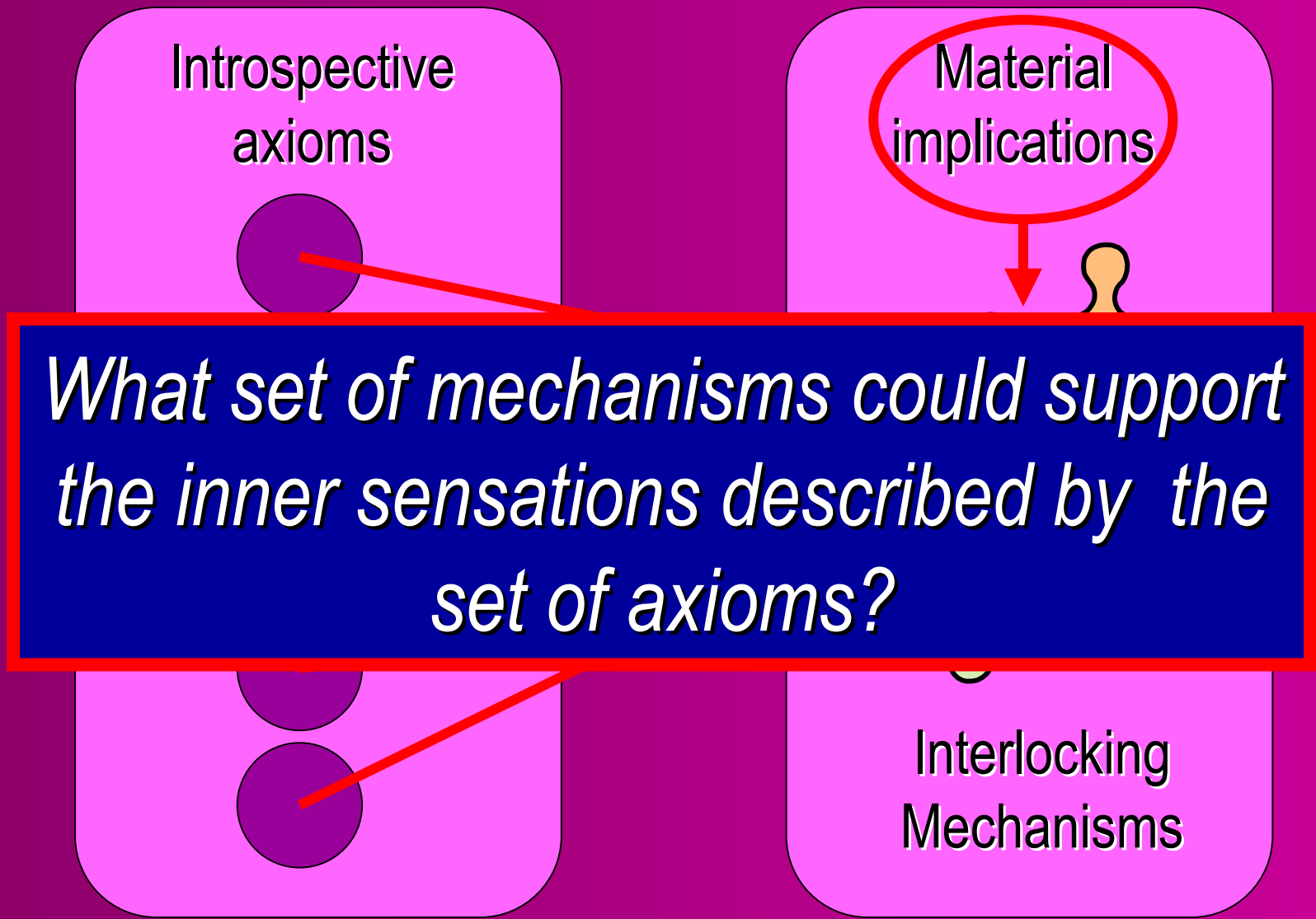
OK for developing designs.

Written as axioms: i.e. features resulting from introspection.

What could be special about mechanisms that can produce for them what for *me* feels that I am conscious ?

"an inner sensation of world, self , past, future, intention and fiction"

WHAT USE ARE INTROSPECTIVE AXIOMS?



THE AXIOMS

Aleksander & Dunmall: *Axioms and Tests for the Presence of Minimal Consciousness in Agents,* Journal of Consciousness Studies June, 2003)

1. **PERCEPTION**
2. **IMAGINATION**
3. **ATTENTION**
4. **PLANNING**
5. **EMOTION**

Axiom

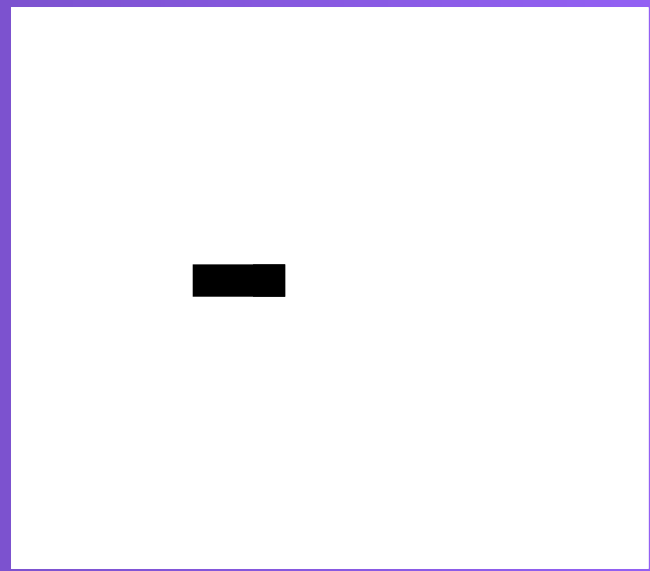
1

First Person Phenomenon:

I feel as if *I* am in the middle of an “out there” world.

Material

Implications of Axiom 1



Material

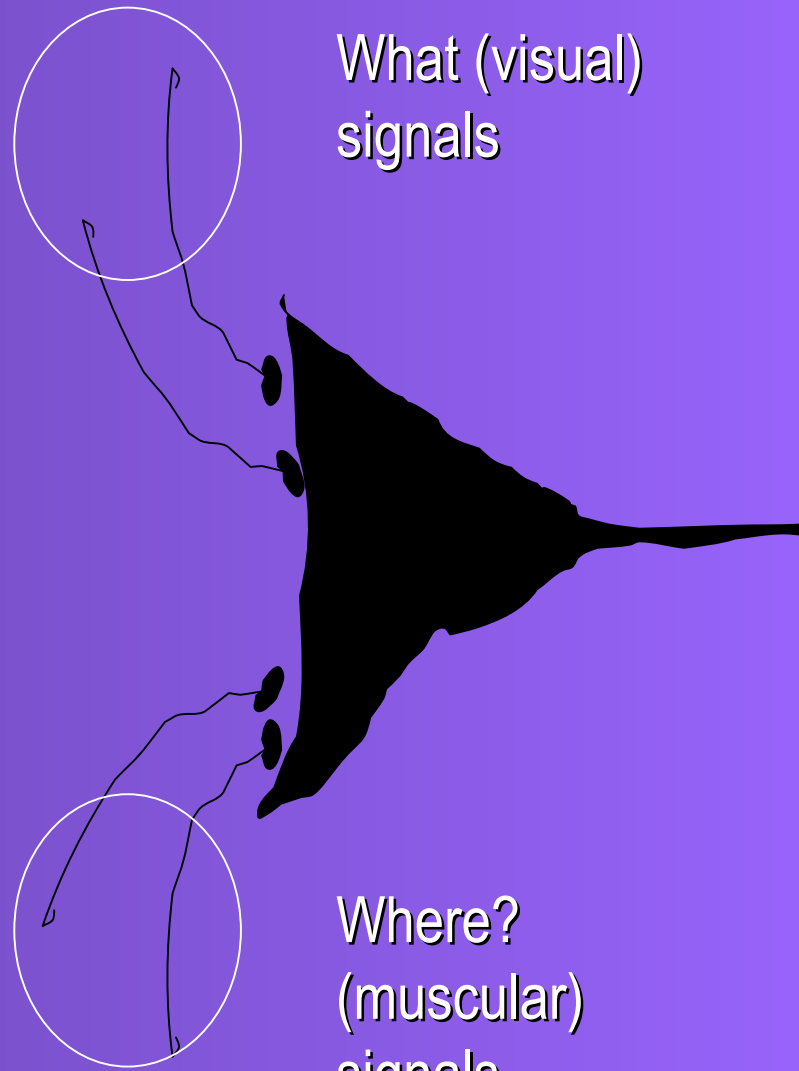
Implications of Axiom 1

Unique minimal world events have a unique coding (easy with a cellular structure)

Minimal events compose complex events both in the world and the mechanism

OUT THERE ??

The material implications of Out-thereness



LOCKING

Evidence:

Galletti and Battaglini, *J Neurosci,*9, 1112-1125, **1989** > **V3 (Gaze-locked)**

Galletti et al, *Europ J Neursci,*9(2),410-413, **1997** >**Arm movment in vis sys V6A.**

Galletti et al, *Exp. Brain Res.* 96 (2): 221-229,**1993** > **Head-indexed in parietal.**

Bender & Youakim, *J Neurophys,*85(1),219-234, **2001**> **Att. Lock V2,V4, 7a.**

Boussaoud et al, *Exp Brain Res,*128 (1-2): 170-180,**1999**> **All visuomotor paths.**

Trotter & Celebrini, *Nature,* 15(5): 398(6724): 239-242,**1999**> **Prestriate V1**

Trotter et. Al, *J Neurophys,*76(5), 2872-2885, **1996** > **View dist. Coded in V1**

Bremmer et al., *Europ J Neursci,*10(1),153-160, **1998** > **V3A, 7a, V6 Post. Pariet**

Duhamel et.al., *Nature*,389 (6653): 845-848, **1997** > **Ventral IntraParietal**

Guo & Li, *Neuroreport,* 8 (6): 1405-1409,**1997** > **All over striate cortex**

Siegel, *Jour of Comp Neursci,* 5(4), 365-381, **1998**> **Gaze locking in parietal 7a**

Gdowski et.al., *Exp. Brain Res.* 135 (4): 511-526, **2000** > **Neck/gaze coding.**

Desimone, *Proc Nat Acad Sci. USA* (93): 13 494-9, **1996**> **Attention in extrastriate**

Aleksander & Dunmall
Proc R. Soc. London B, Jan 22, 2000

Material

Implications of Axiom 1

The design that captures

- Composition from minimal events
- Out-thereness

We shall call

DEPICTIVE

(Alva Noe: enacted/sensorimotor)

Axiom

2

First Person Phenomenon:

I can recall ‘out there’ worlds and imagine worlds *I* have never experienced.

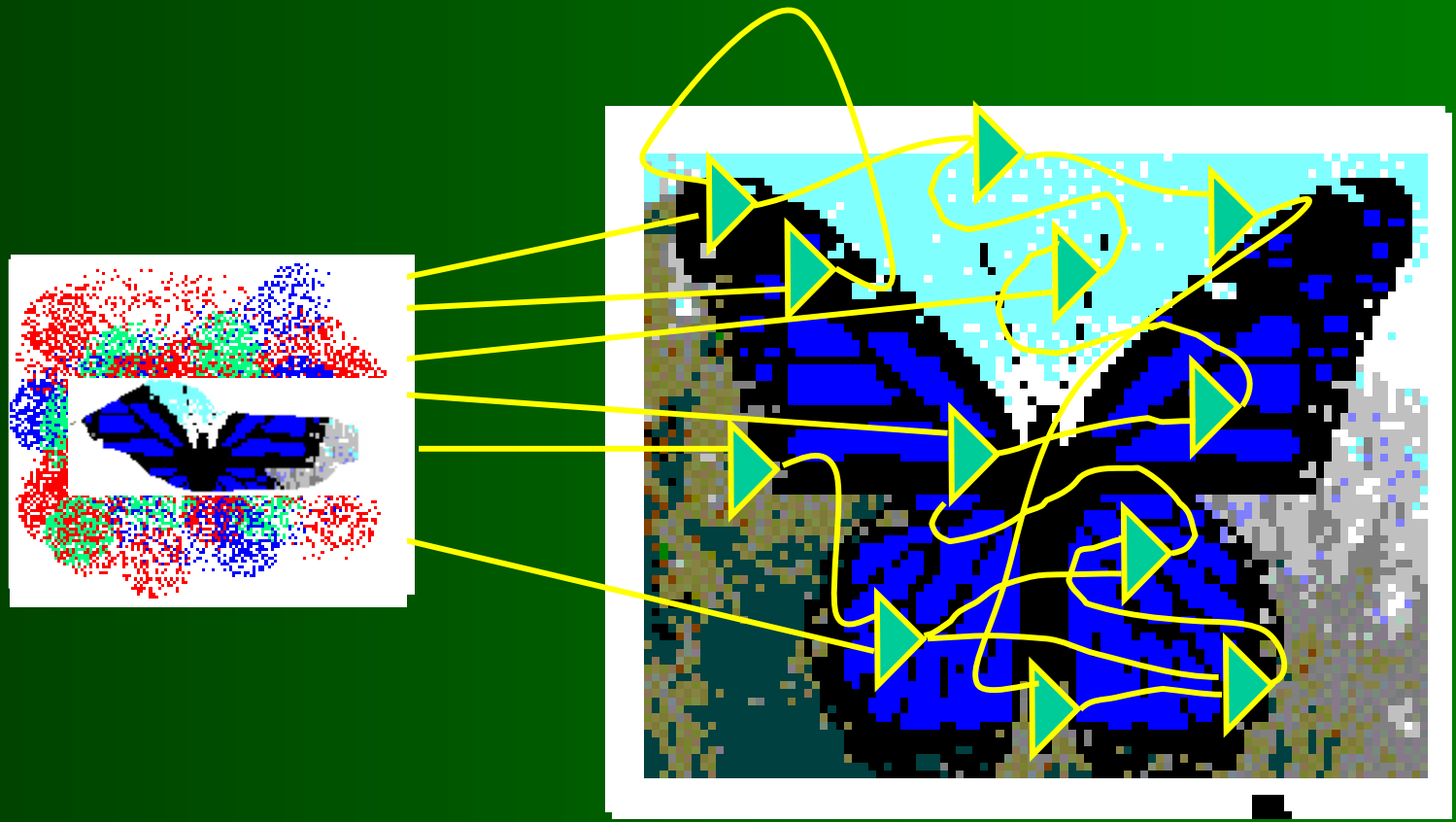
Material

Implications of Axiom 2

Depictive neural firing is sustained due to cells forming closed loops or layers ‘re-entering’:

(Major link with complexity phenomena)

Communicating neurons: memory and imagination



Axiom

3

First Person Phenomenon:

I am conscious of only that to which *I* attend.

Material

Implications of Axiom 3

External Attention:

Is driven by localization signals for important events in the ‘out-there’ world and continuity.

Material

Implications of Axiom 3

Internal Attention:

Is driven by the use of internally driven motor locking signals (inverse to reconstruction)

Or

Informational continuity

Axiom

4

First Person Phenomenon:

I imagine future events in the world and determine my actions to get what *I* want.

Material

Implications of Axiom 4

Depictive states have a chaining property shaped by learning and recalled in imagination. Re - entry is again responsible for this

Axiom

5

First Person Phenomenon:

My emotions affect and are affected by my decisions.

Material

Implications of Axiom 5

Structures continually (and innately?) evaluate perceptual and imagined future events in preparation for action .

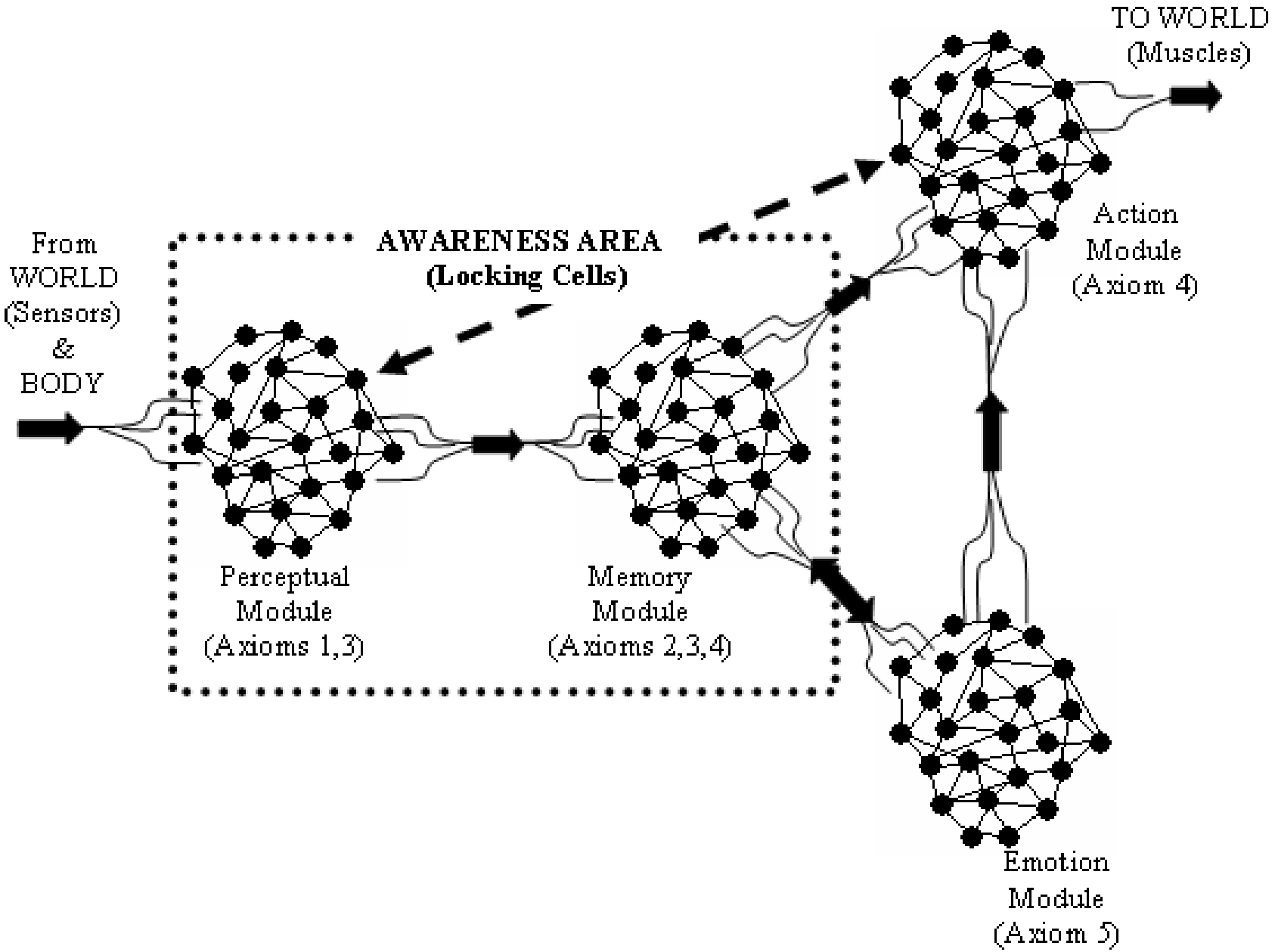
Emotions have a global effect on depictive structures

A vital corollary resulting from the Axioms

SELF

The Emergence of ***SELF***

Combinations of sensory, imaginational attentional and affective depictions lead to descriptions starting with the word *I*


TO WORLD
(Muscles)
Action
Module
(Axiom 4)
AWARENESS AREA
(Locking Cells)
From
WORLD
(Sensors)
&
BODY
Perceptual
Module
(Axioms 1,3)
Memory
Module
(Axioms 2,3,4)
Emotion
Module
(Axiom 5)

Some past work from a MC standpoint

1998: Imagining Colour & Shape (Ax 1,2)

1999: Saccadic Face Perception and Recognition (Ax 1,2,3)

2000: Theory of depiction (Ax 1,2)

2001: Planning and Stacking Action (Ax 1,2,3,4)

2002: Mobile Robot (Ax 1,2,3,4)

2003: Visual Deficits in Parkinson's disease (Ax 1,2,3)

Now

Effect on IT

Creation of an informational concept of neural systems capable of ‘being conscious’

Integration of cognitive faculties in architectures

Achievement of new levels of competence

Contact

i.aleksander@imperial.ac.uk