BRITISH HCI GROUP ONE-DAY MEETING
UNIVERSITY COLLEGE LONDON
AFFECTIVE COMPUTING: THE ROLE OF EMOTION IN HCI
10th April 1999

WHY CAN'T A GOLDFISH LONG FOR ITS MOTHER?

Architectural prerequisites for various types of emotions.

AARON SLOMAN

http://www.cs.bham.ac.uk/~axs/
A.Sloman@cs.bham.ac.uk

Ideas developed

in collaboration with

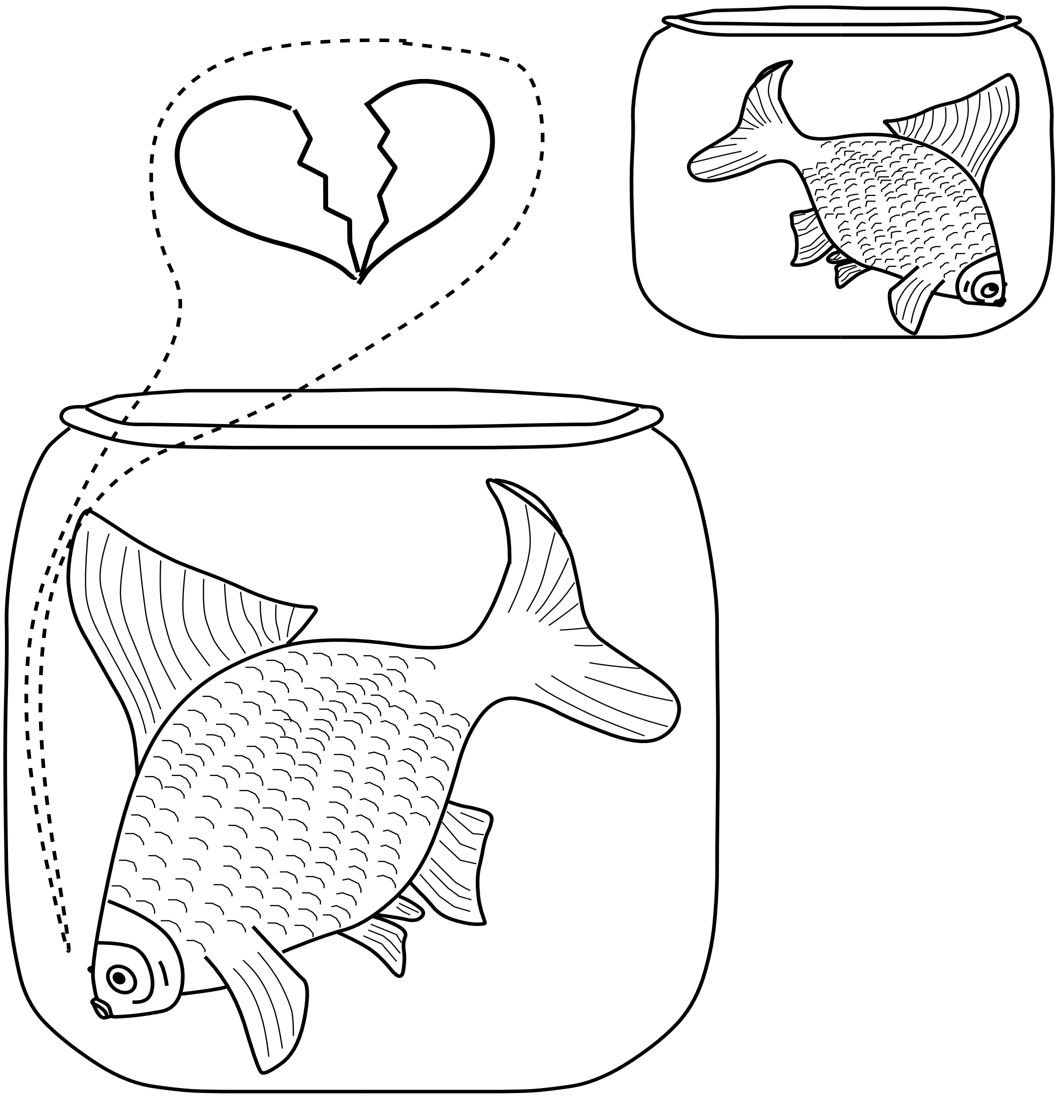
Steve Allen, Luc Beaudoin,
Brian Logan, Riccardo Poli,
Ian Wright,

and others in the

COGNITION AND AFFECT PROJECT

SCHOOL OF COMPUTER SCIENCE
THE UNIVERSITY OF BIRMINGHAM

WHY CAN'T A GOLDFISH LONG FOR ITS MOTHER?



- **Because it cannot make its mouth droop?**
- **Because it lacks tear glands to make it weep?**
- **Because it cannot sigh....?**
- **Because it lacks our proprioceptive feedback...??**

No, because:

1. **it lacks the appropriate information processing architecture**
2. **including representational mechanisms, concepts and knowledge.**

WHAT KIND OF MACHINE CAN HAVE EMOTIONS?

PROBLEM:

Umpteen different definitions of "emotion".

in psychology, philosophy, neuroscience . . .

REPHRASE:

What are the architectural requirements

for human-like mental states and processes?

(Never mind the definitions)

I.e. collect examples of many types of real phenomena.
Try to build a theory which explains them all!
Subject to constraints from neuroscience, psychology, biological evolution, feasibility, tractability, etc.

ALLOW FOR VARIATION:

- **Across species,**
- **Within species,**
- **Within an individual during normal development**
- **After brain damage**
- **Across planets (grieving, infatuated, Martians?)**
- **Across the natural/artificial divide**

PAY LEAST ATTENTION TO EXPERIMENTAL PSYCHOLOGY

(Shallow vs Deep science)

Which human-like states and processes?
Which real phenomena?

Consider the following cases:

- YOU ARE:
  startled by a loud noise,
  frozen in terror as boulder crashes towards you,
  nauseated by a horrible smell
- YOU ARE:
  afraid the bridge you are crossing may give way
  relieved that you got to the far side safely
  afraid the bridge your child is crossing may give way
  worried about what to say during your interview
  undecided whether to cancel your holiday in ...
- YOU ARE:
  infatuated with someone you met recently,
  overwhelmed with grief,
  riddled with guilt about betraying a friend
  full of excited anticipation of a loved one's return
  full of longing for your mother,
  basking in a warm glow of pride after winning an election.

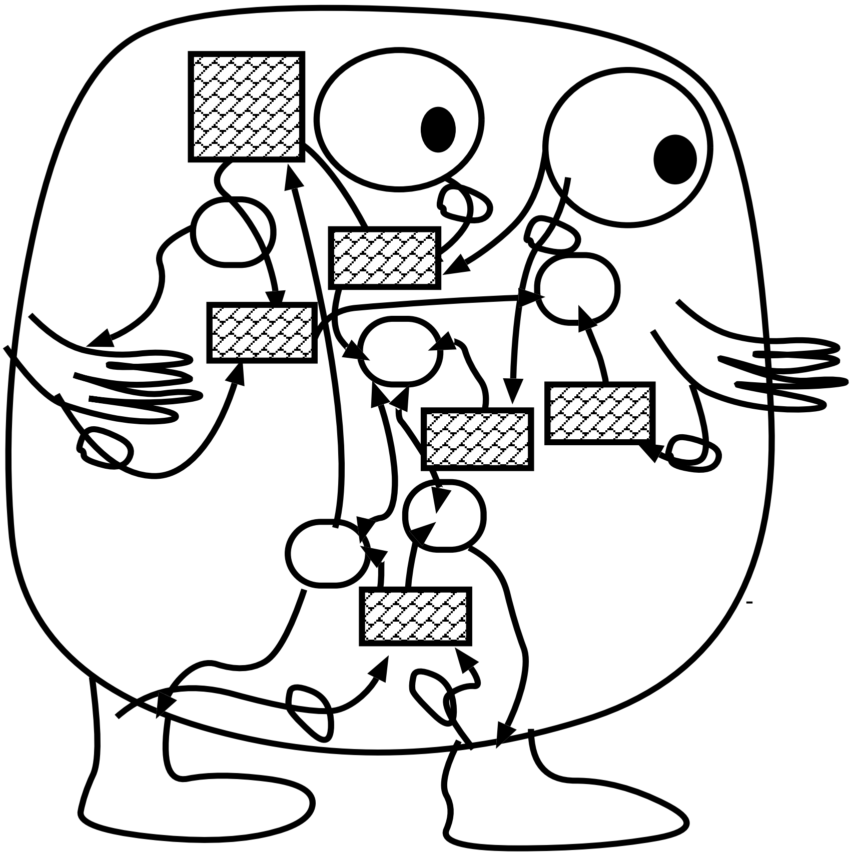
I'll describe different architectural underpinnings for

PRIMARY EMOTIONS
SECONDARY EMOTIONS **(central and peripheral)**
TERTIARY EMOTIONS **(with and without peripheral effects)**

All have many variants, there's no time to discuss.

WHAT SORT OF ARCHITECTURE CAN ACCOUNT FOR SUCH PHENOMENA?

COULD IT BE AN UNINTELLIGIBLE MESS?

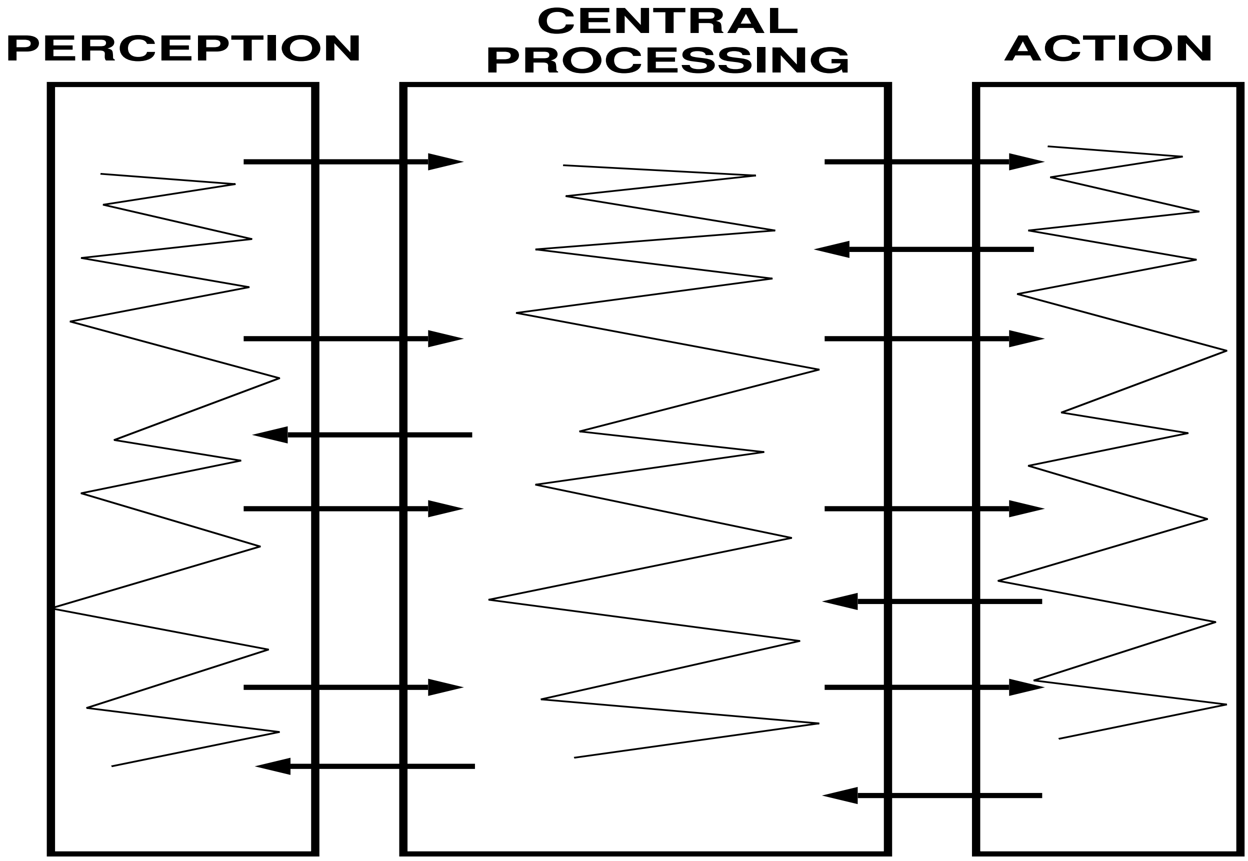


Yes, in principle.
However, it can be argued that evolution could not have produced a totally non-modular yet highly functional brain.
(Compare Nilsson, and Wittgenstein on ‘sawdust’)

TOWARDS A UNIFYING MODULAR THEORY OF BRAIN AND MIND: A BIRD'S EYE VIEW

One perspective:

THE "TRIPLE PILLAR" MODEL



(many variants)
(Nilsson, Albus)

MODULAR **does not mean** RIGID **or** INNATE

Systems can be "nearly decomposable". Boundaries can change with learning and development.

SENSING AND ACTING CAN BE ARBITRARILY SOPHISTICATED

- **On simple models sensors and motors are mere transducers.**

- **More realistically, they can have sophisticated information processing architectures.**

  E.g. perception and action can be hierarchically organised with concurrent interacting sub-systems.

- **Perception goes far beyond segmenting, recognising, describing what is "out there". It includes:**

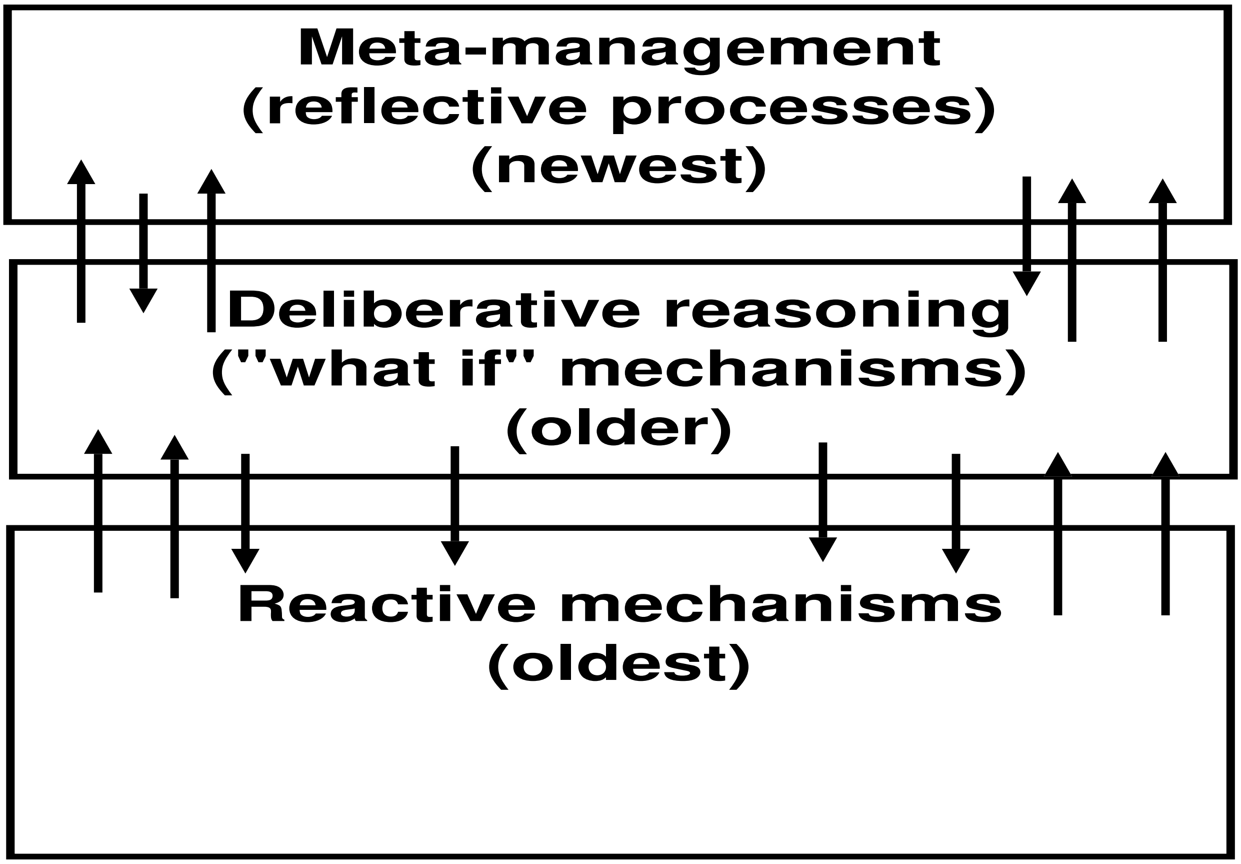
  - **providing information about *affordances* (Gibson, not Marr),**
  - **directly triggering physiological reactions (e.g. posture control, sexual responses)**
  - **evaluating what is detected,**
  - **triggering new motivations**
  - **triggering "alarm" mechanisms**
  - **. . . . .**

AN EXTENSION OF GIBSON'S THEORY:

Different sub-systems use different affordances, and different ontologies. (Evidence from brain damage.)

ANOTHER COMMON ARCHITECTURAL PARTITION
(functional, evolutionary)

THE "TRIPLE LAYER" MODEL

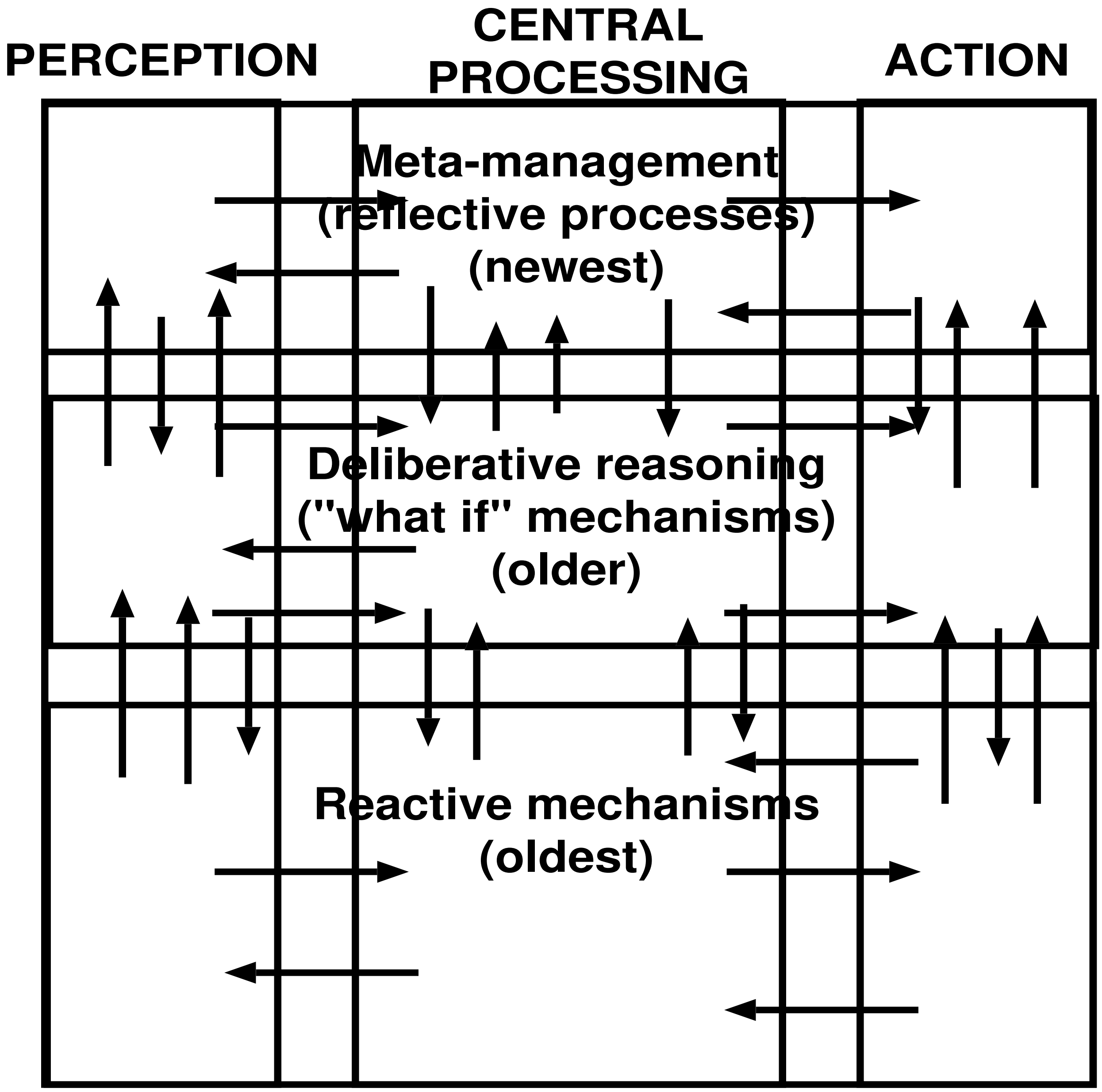


(many variants – especially third layer)

Reactive systems can be highly parallel, very fast, and use analog circuits.
Deliberative mechanisms are inherently slow, serial, knowledge-based, resource limited.

COMBINING THE VIEWS: LAYERS + PILLARS = GRID

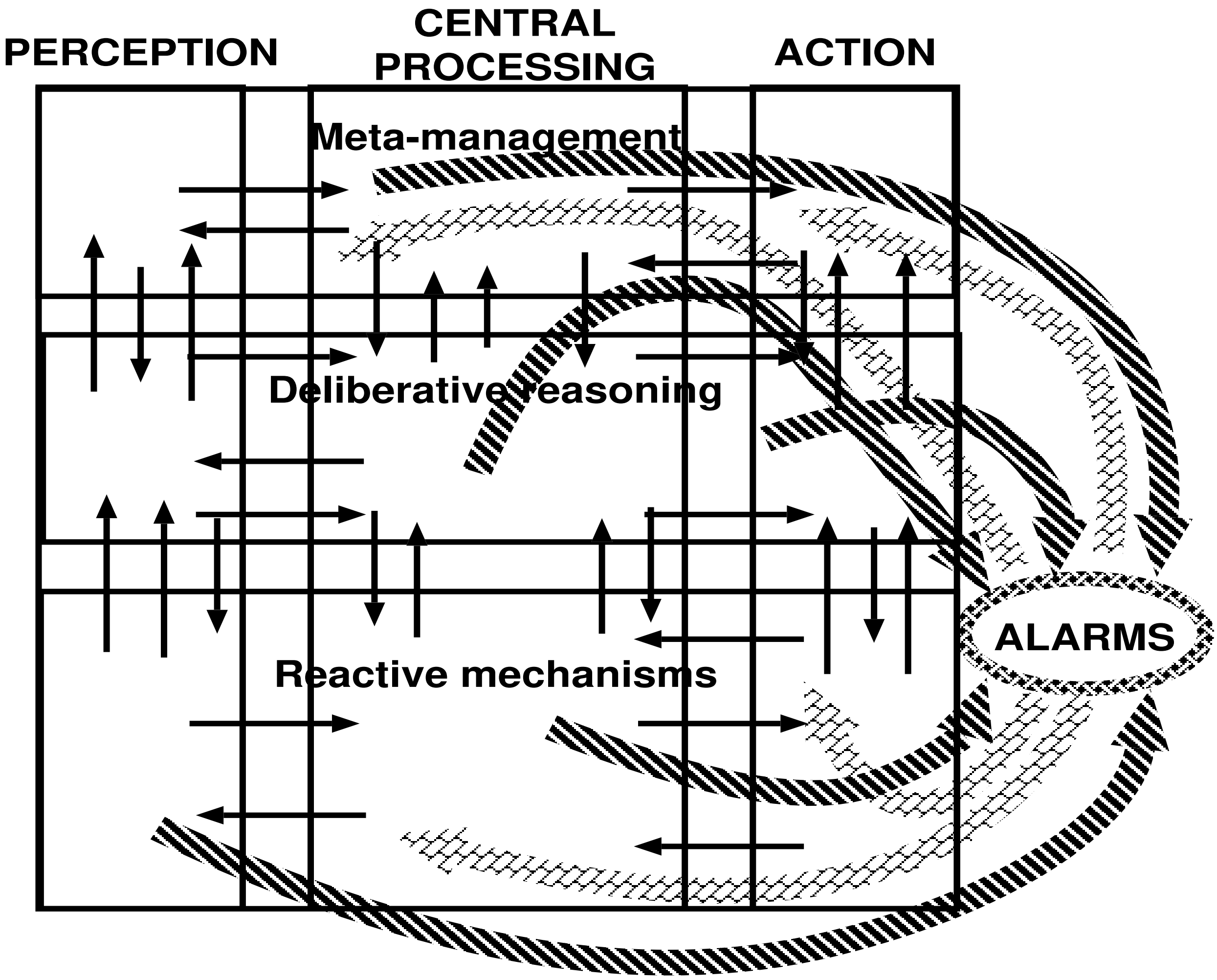
A grid of co-evolving sub-organisms, each contributing to the niches of the others.



As processing grows more sophisticated, so it can be come slower, to the point of danger.

FAST, POWERFUL,
“GLOBAL ALARM SYSTEM”
NEEDED

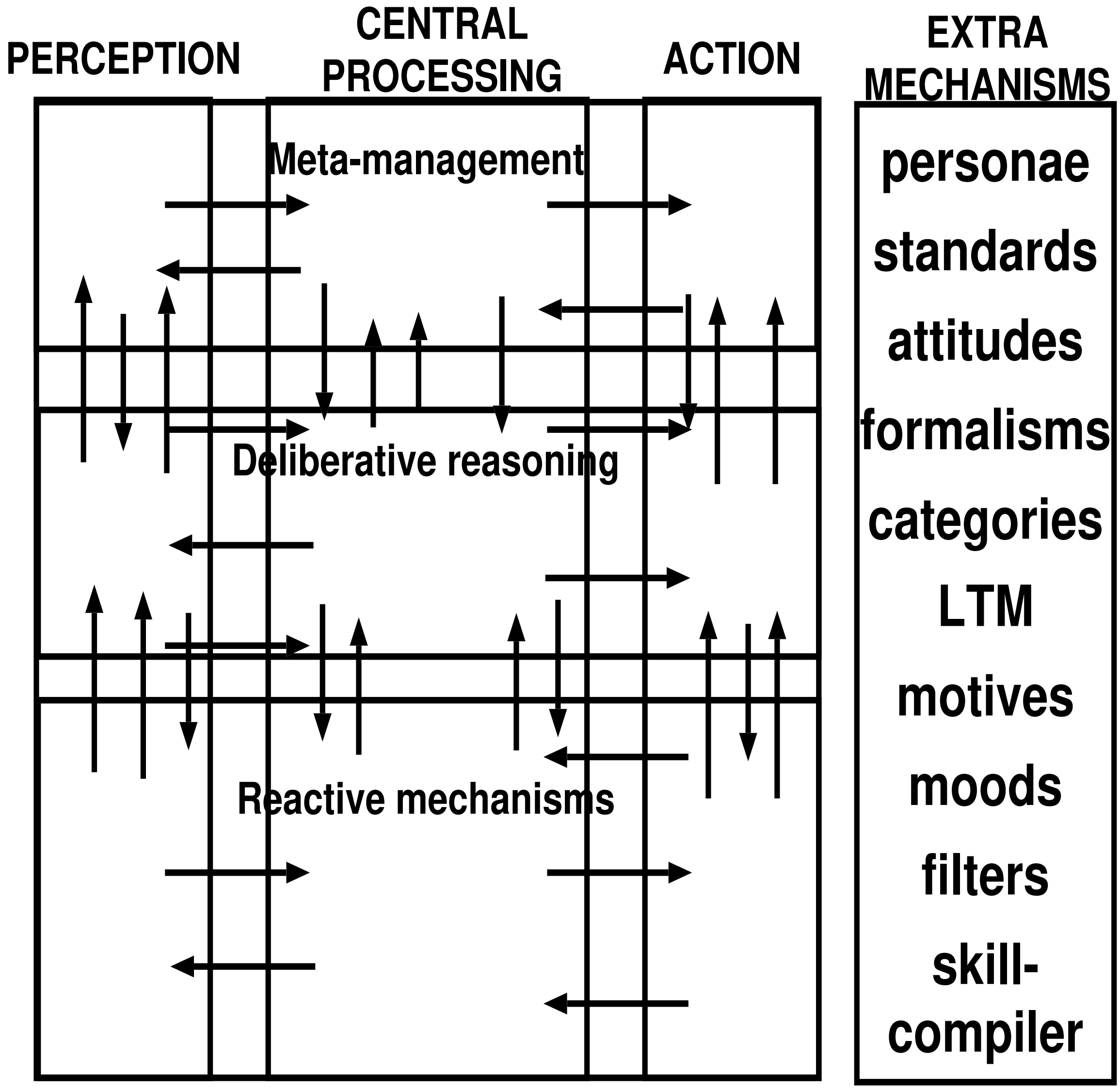
IT WILL INEVITABLY BE STUPID!



MANY VARIANTS POSSIBLE.

E.g. one alarm system or several?
(Brain stem, limbic system, ...???)

ADDITIONAL COMPONENTS
(No time to discuss)



MANY PROFOUND IMPLICATIONS

e.g. for kinds of development

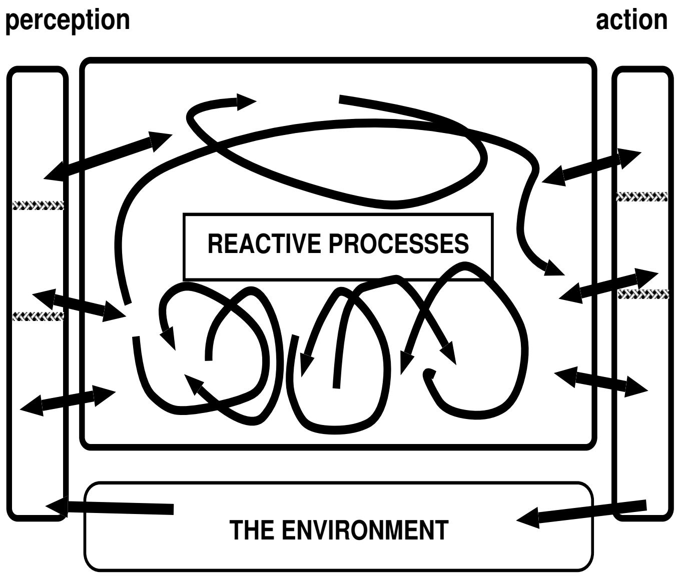
kinds of perceptual processes

kinds of brain damage

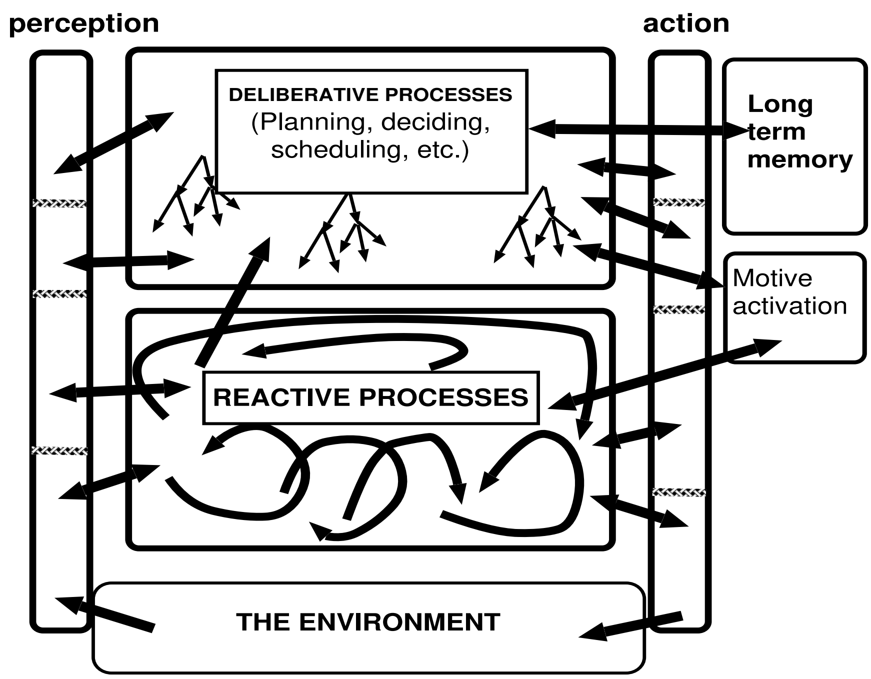
kinds of emotions

NOT ALL PARTS OF THE GRID ARE PRESENT IN ALL ANIMALS

How to design an insect?



Add a deliberative layer, e.g. for a monkey?



THESE LAYERS EXPLAIN

PRIMARY, SECONDARY, TERTIARY EMOTIONS

Different architectural layers support different sorts of emotions, and help us define

AN ARCHITECTURE-BASED ONTOLOGY OF MIND

Different animals will have different mental ontologies

Humans at different stages of development will have different mental ontologies

The REACTIVE **layer with** GLOBAL ALARMS **supports "primary" emotions:**

- **being startled**
- **being disgusted by horrible sights and smells**
- **being terrified by large fast-approaching objects?**
- **sexual arousal? Aesthetic arousal ?**

etc. etc.

The DELIBERATIVE **layer enables "secondary" emotions (cognitively based):**

- **being anxious about possible futures**
- **being frustrated by failure**
- **excitement at anticipated success**
- **being relieved at avoiding danger**
- **being relieved or pleasantly surprised by success**

etc. etc.

THE THIRD LAYER
enables
SELF-MONITORING,
SELF-EVALUATION
and
SELF-CONTROL

AND THEREFORE ALSO LOSS OF CONTROL (PERTURBANCE)

(and qualia!)

This makes possible "tertiary" emotions, through having and losing control of thoughts and attention:

- **Feeling overwhelmed with shame**
- **Feeling humiliated**
- **Aspects of grief, anger, excited anticipation, pride,**
- **Being infatuated, besotted**

and many more *typically HUMAN* **emotions.**

NOTES:

1. Different aspects of love, hate, jealousy, pride, ambition, embarrassment, grief, infatuation can be found in all three categories.

2. Remember that these are not STATIC **states but** DEVELOPING **processes, with very varied aetiology.**

SOCIALLY IMPORTANT HUMAN EMOTIONS INVOLVE RICH CONCEPTS AND KNOWLEDGE and RICH CONTROL MECHANISMS (architectures)

- **Our everyday attributions of emotions, moods, attitudes, desires, and other affective states implicitly presuppose that people are information processors.**
- **To long for something you need to know of its existence, its remoteness, and the possibility of being together again.**
- **Besides these *semantic* information states, longing also involves *control* states.**

  One who has deep longing for X does not merely occasionally think it would be wonderful to be with X. In deep longing thoughts are often *uncontrollably* drawn to X.

- **Physiological processes (outside the brain) may or may not be involved. Their importance is normally over-stressed by experimental psychologists under the malign influence of the James-Lange theory of emotions. (Contrast Oatley, and poets.)**

CONCLUSION: THE SCIENCE

• Much of this is conjectural – many details still have to be filled in and consequences developed (both of which can come partly from building working models, partly from multi-disciplinary empirical investigations).

• An architecture-based ontology can bring some order into the morass of studies of affect (e.g. myriad definitions of "emotion").

COMPARE THE RELATION BETWEEN THE PERIODIC TABLE OF ELEMENTS AND THE ARCHITECTURE OF MATTER.

• This can lead to a better approach to comparative psychology, developmental psychology (the architecture develops after birth), and effects of brain damage and disease.

• It will provide a conceptual framework for discussing which kinds of emotions can arise in software agents that lack the reactive mechanisms required for controlling a physical body.

CONCLUSION: HCI ENGINEERING

HCI Designers need to understand these issues:

(a) if they want to model human affective processes,

(b) if they wish to design systems which engage fruitfully with human affective processes,

(c) if they wish to produce teaching/training packages for would-be counsellors, psychotherapists, psychologists.

For more details,see the Cognition and Affect project papers

ftp://ftp.cs.bham.ac.uk/pub/groups/cog_affect/0-INDEX.html

COLLABORATORS ALWAYS WELCOME