PPSN2000 :The Sixth International Conference
on Parallel Problem Solving from Nature

Paris, September 18-20, 2000

Interacting Trajectories in Design Space and Niche Space

(A Philosopher Speculates About Evolution)

AARON SLOMAN

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

**Ideas developed
in collaboration with
Steve Allen, Luc Beaudoin, Darryl Davis,
Catriona Kennedy, Brian Logan,
Matthias Scheutz, Ian Wright,
and others in the**

Cognition and Affect Project

SCHOOL OF COMPUTER SCIENCE
THE UNIVERSITY OF BIRMINGHAM
http://www.cs.bham.ac.uk/research/cogaff/

I HAVE ALSO LEARNT FROM MANY OTHERS.

COULD EVOLUTION BE A PURPOSEFUL WATCHMAKER?

TRADITIONAL DARWINIAN VIEW:

Evolution is a “blind” process using only random processes which may or may not produce good solutions to biological problems.

Consider a search space with 1000 binary steps, explored 1 every thousand millionth of a second.

Exhaustive search requires this number of years:

(21000)/(1000000000*60*60*24*365.0), which is:**

339773150426898200000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000 years!

Estimated age of the earth: 4500,000,000 years, approximately

How many design decisions are there in a typical organism, e.g. a flea or a human?

What is the search space like?

What proportion of the space has to be explored ?

Does it have a structure that facilitates the search for good solutions?

How? Is co-evolution part of the answer?

Mutually interacting trajectories vs mere parallelism.

Is the intelligence of the organisms produced by evolution part of the answer?

CONJECTURE:

There is some higher level architecture at work which supports something more "purposive".

It doesn't involve conventional intelligence, goals, plans, etc.

Perhaps a kind of feedback mechanism:
e.g. more like an amplifier.

It may be either

- **a contingent consequence of the details of evolution on earth,**

or

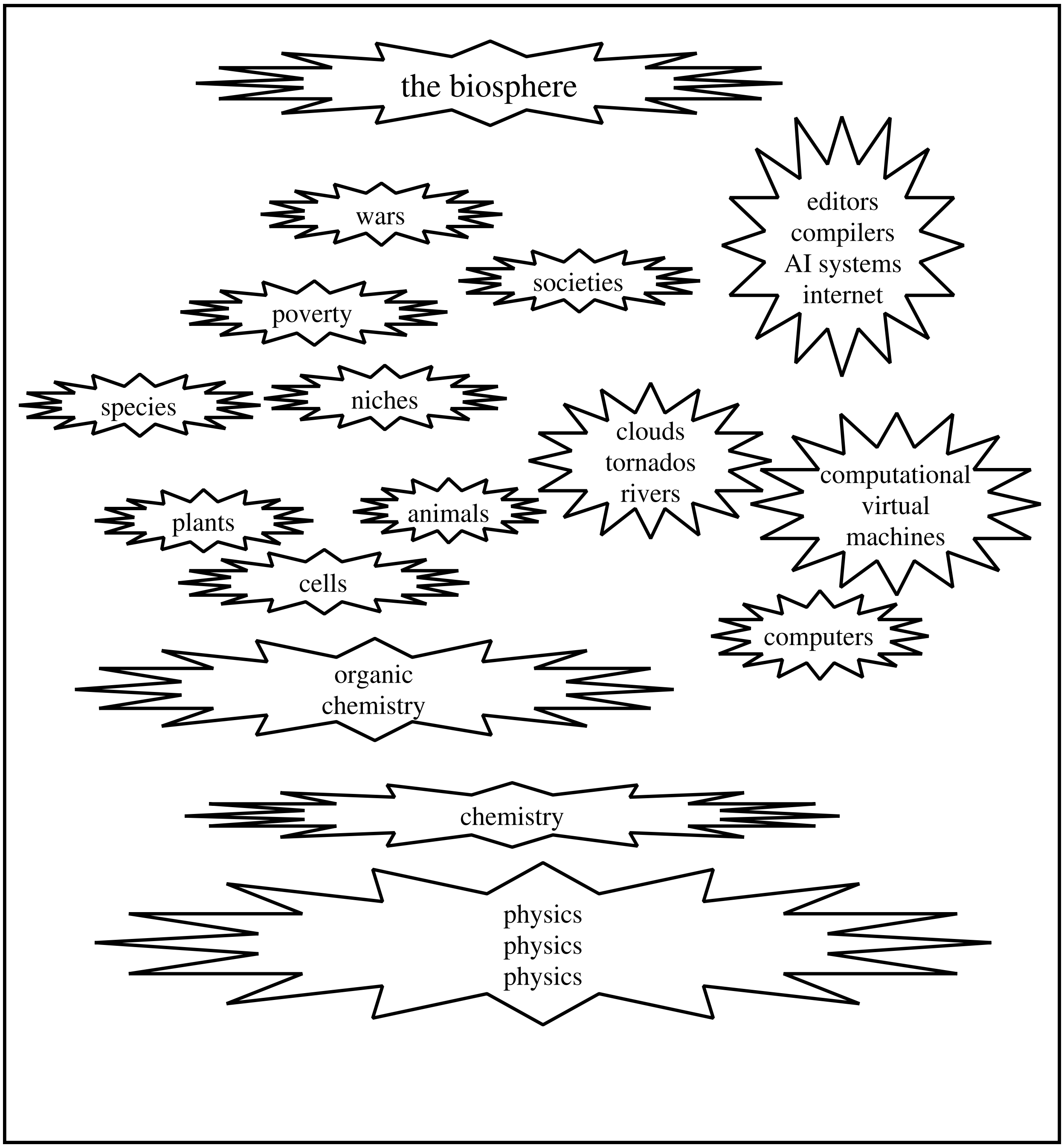
- **a necessary (mathematically demonstrable) consequence of some of the properties of all evolutionary processes in some interesting class of processes.**

This is not a new idea:

Compare 'Laws of form'

e.g. Kauffman.

A SPECIAL CASE OF MULTI-LEVEL ONTOLOGY AND CAUSAL PROCESSES



Many kinds of causally efficacious "emergent" virtual machines, operating at different scales and different levels of abstraction.

- **Is there a well-defined "bottom level" ? We don't know whether physics has a well defined bottom level. It seems to have several levels.**
- **Causality operates at all levels. (Poverty can cause crime)**
- **Causality can be "circular".**

  (E.g. physical processes in a computer cause events in a computational virtual machine, and non-physical events in the virtual machine, e.g. a queen capturing a bishop, a theorem prover finding a contradiction, cause physical events, in the computer, and on the screen.)
- **Levels need not form a rigid hierarchy.**
- **The biosphere is a richly interacting collection of myriad physical and virtual machines running concurrently.**

DIGRESSION: PHILOSOPHY AND ENGINEERING

PHILOSOPHERS: talk about "supervenience"

ENGINEERS: talk about "implementation"

MAYBE THEY SHOULD GET TOGETHER AND PRODUCE A NEW CONCEPTUAL FRAMEWORK FOR MULTI-LEVEL ONTOLOGIES

But most philosophers are ignorant of engineering, especially software engineering.

And most engineers despise philosophy!

HOW COULD MULTI-LEVEL DARWINIAN EVOLUTION OPERATE?
PERHAPS DARWINIAN EVOLUTION USES HIGH LEVEL STRUCTURES AND PROCESSES

POSSIBLE MECHANISMS?

- **Feedback loops between niches and designs in co-evolution**
- **Mate selection**
- **Social mechanisms (Sparta)**
- **Advances in knowledge leading to genetic engineering.**
- **A hidden global architecture in the total evolutionary system, permitting a mixture of reactive and deliberative mechanisms (e.g. experimenting to solve a problem?)**

IS THIS TOTALLY CRAZY ????

Two forms of answer:

- **Experimental simulations showing the possibility of the mechanisms**
- **Mathematical analysis showing the inevitability of the mechanisms (Based on a new conceptual framework)**

TOWARDS A CONCEPTUAL FRAMEWORK: A FIRST PASS

• A *design* specifies a class of architectures

(each of which can have many instances)

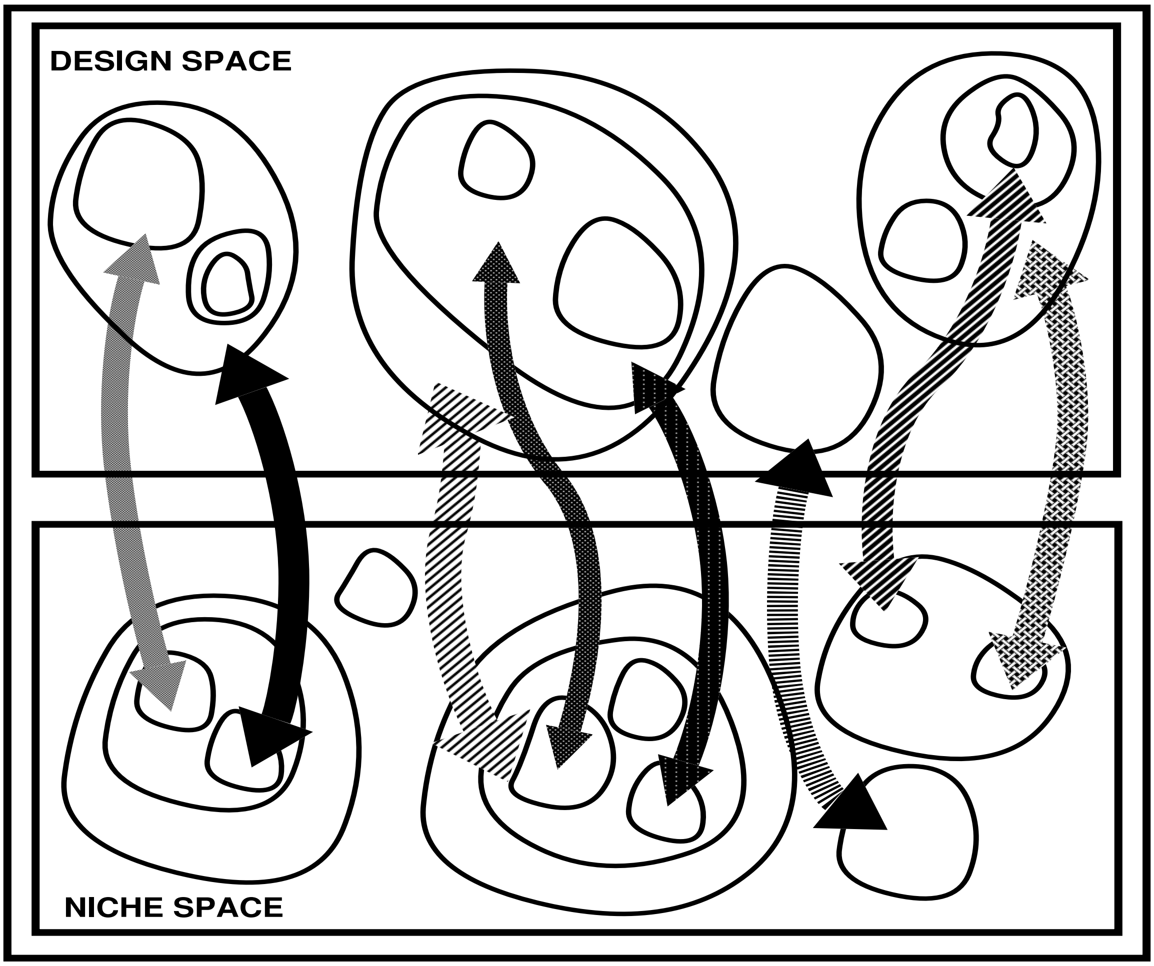
• A *niche* specifies a class of environments and modes of evaluation for designs.

• Each combination of architecture and environment (or design and niche?) determines a class of processes with multiple evaluation critera.

We need good ways of thinking of possible trajectories in these spaces.

Perhaps we can then begin to understand the dynamics at higher levels of abstraction.

Design space, niche space and evaluation relations



Relations between designs and niches are complex and varied.

Instead of 'fitness' VALUES ***we need fitness*** DESCRIPTIONS**.**

Different sorts of arrows represent different types of fitness relations.

One sort of description might be a vector of values (or labels), as in Consumer Association reports.

(TRAJECTORIES ARE NOT SHOWN HERE.)

DIFFERENT SORTS OF TRAJECTORIES

I-TRAJECTORY: possible for an individual machine or organism, whether a result of learning or pre-programmed development. **(e.g. growing from an acorn to an oak, etc.)**

E-TRAJECTORY: possible by evolution across generations, but not for an individual **(e.g. from one species to another)**

S-TRAJECTORY: possible for societies of organisms
May depend on division of labour, etc.

R-TRAJECTORY: changes to individuals possible only through external intervention by another agent, e.g. performing repairs or extensions, or debugging software. All the others require viable organisms at every stage, whereas r-trajectories don't. **(Are r-trajectories for humans imminent?)**

NOTE 1: WHAT IS AN E-TRAJECTORY OR S-TRAJECTORY FROM ONE VIEWPOINT MAY BE AN I-TRAJECTORY FROM ANOTHER: VIEWING A SPECIES OR A SOCIETY AS AN INDIVIDUAL.

NOTE 2: TRAJECTORIES OCCUR IN NICHE SPACE AND IN DESIGN SPACE.

NOTE 3: UNDERSTANDING THE DYNAMICS OF ECOSYSTEMS REQUIRES UNDERSTANDING THE FEEDBACK LOOPS BETWEEN DIFFERENT TRAJECTORIES AT DIFFERENT LEVELS OF ABSTRACTION IN BOTH SPACES.

NOTE 4: WE CAN ALSO CONSTRUE A COMPLEX ORGANISMS AS A SORT OF ECOSYSTEM, IN WHICH PARTS, OR RATHER THEIR DESIGNS, CO-EVOLVE.

TOWARDS A TOPOLOGY FOR DESIGN SPACE AND NICHE SPACE

Both spaces are mathematically very complex

- **Both may have some continuous regions and also many discontinuities.**

  (Remember Darwinian evolution is inherently discontinuous)

- **Regions in both spaces are defined by abstract specifications or collections of specifications.**
- **The sizes of discotinuities vary according to level of abstraction.**

  (Two more legs, vs slightly longer legs)

- **Trajectories in both spaces can be defined at different levels of abstraction.**
- **In many cases the specification of a niche depends in part on the specification of a design. E.g. food-getting requirements depend on existing food-digesting mechanisms**
- **This leads to the view of an organism as itself some kind of ecosystem, composed of co-evolved sub-organisms.**

ASPECTS OF THE LOGIC OF 'BETTER'
THE IMPORTANCE OF TRADE-OFFS

In general there will not be an answer to the question: does this change (e.g. a mutation) lead to higher fitness?

Rather there will be trade-offs: the new variant is better in some respects and worse in others, and for a given respect the order may depend on circumstances.

This should not be confused with a "neutral" change which makes no relevant difference to the individual's abilities to fit the niche.

Sometimes there is a partial ordering of the fitness descriptions, and sometimes not even that, because there is no way to combine the different dimensions of comparison.

Design A might be better than B in one respect, B better than C in another and C better than A in a third.

Tradeoffs between dimensions in fitness vectors may be exploited by the formation of cooperative behaviours and division of labour.

Two individuals that excel in different ways (e.g. hunting and farming) may together be more competent than two with equal but intermediate levels of expertise.

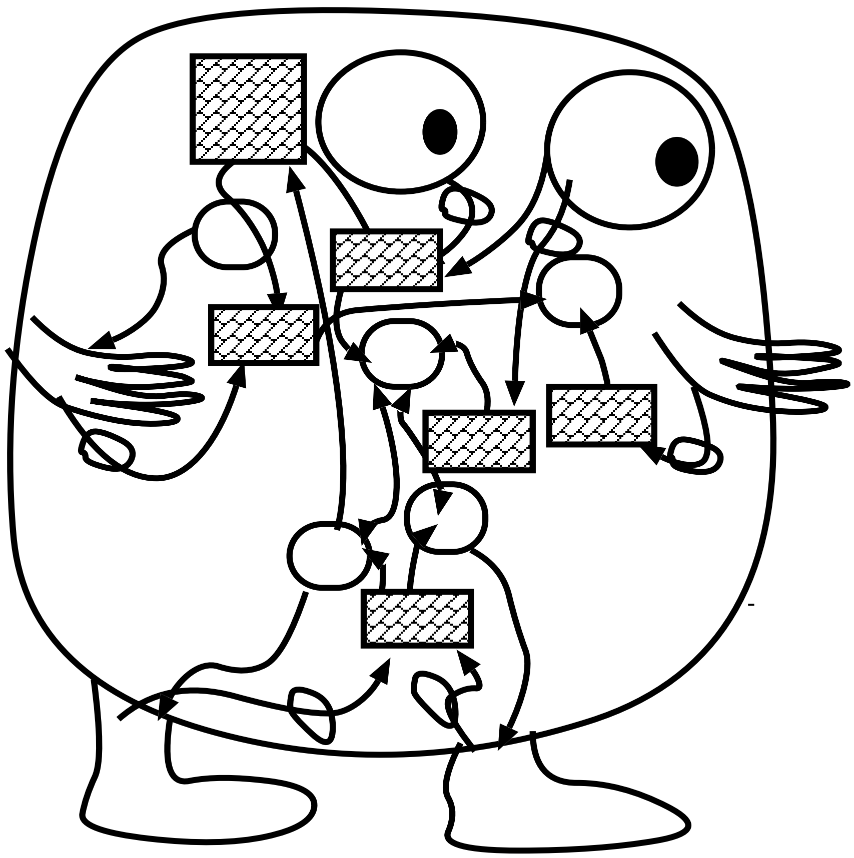
This is particularly true of individuals in a social group requiring many kinds of expertise.

Useful division of labour can also occur across species.

SYMBIOTIC GENES MAY INHABIT THE SAME SPECIES OR DIFFERENT SPECIES.

HOW COULD A HUMAN-LIKE ARCHITECTURE EVOLVE?

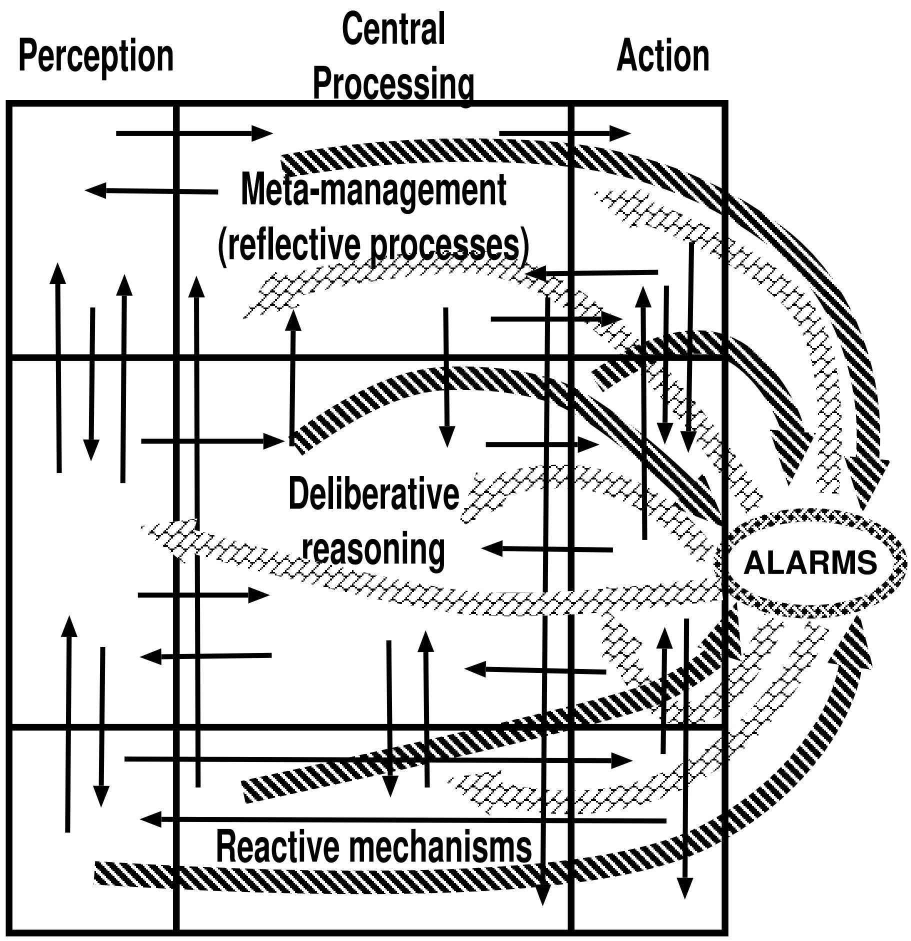
MUST IT BE AN UNINTELLIGIBLE MESS?



Yes, in principle.

However, it can be argued that evolution could not have produced a totally non-modular yet highly functional brain.

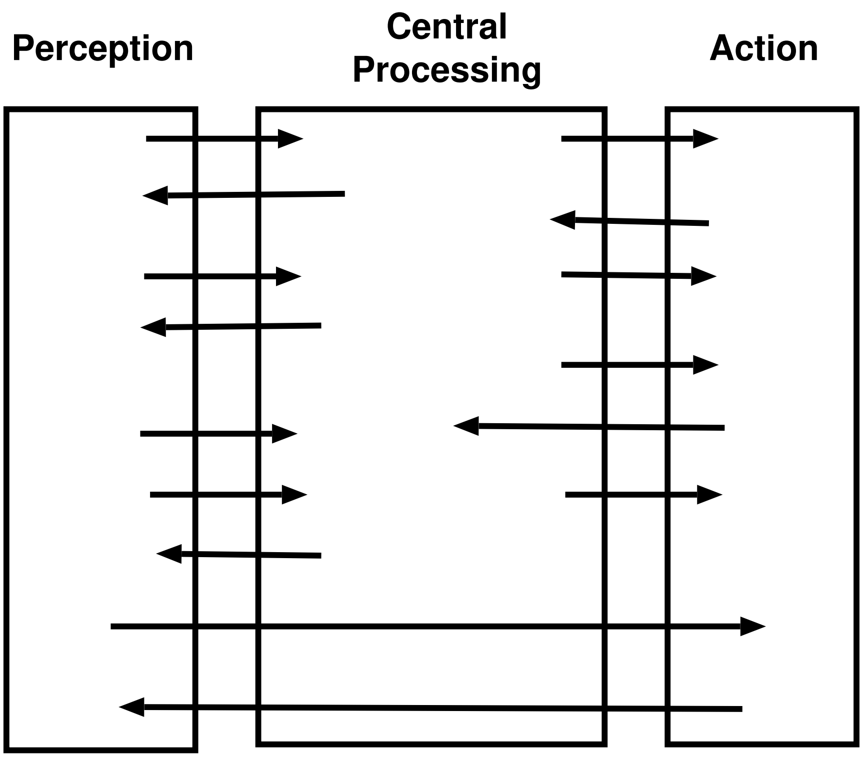
The (Birmingham) ‘CogAff’ Architecture
(A partial view)



This view of the architecture is motivated by superimposing the ‘triple tower’ (input-central-output) and ‘triple layer’ (three stages of evolution) views depicted below.

Missing additional components are described later.

The “triple tower” View



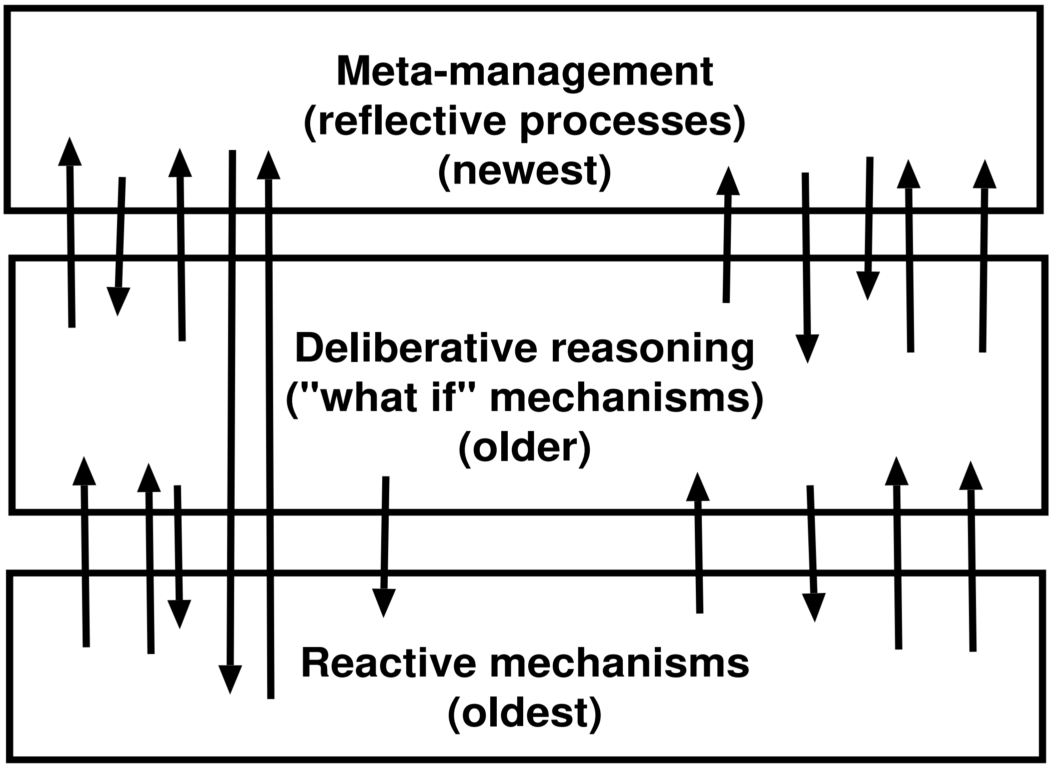
There are many variants: (NILSSON, ALBUS)

We need to understand the design options and the requirements.

Systems can be “nearly decomposable”.

Boundaries can change with learning and development.

ONE OF MANY LAYERED VIEWS
Compare: triune brain: reptilian, old mammalian, new mammalian.



Many variants, but with different subdivisions and interpretations of subdivisions

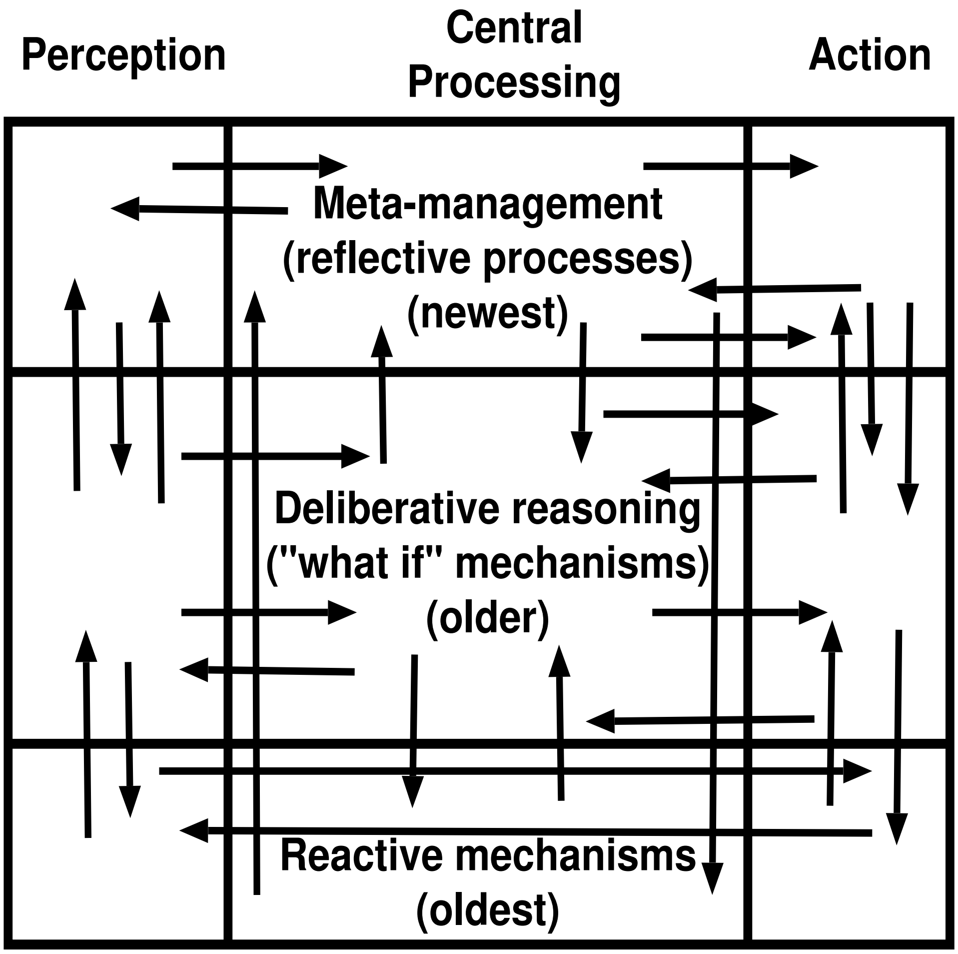
Different principles of subdivision

- **evolutionary stages**
- **levels of abstraction,**
- **control-hierarchy,**
- **information flow**

  (e.g. the popular 'Omega' Ω model of information flow)

COMBINING THE VIEWS: LAYERS + PILLARS = GRID

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

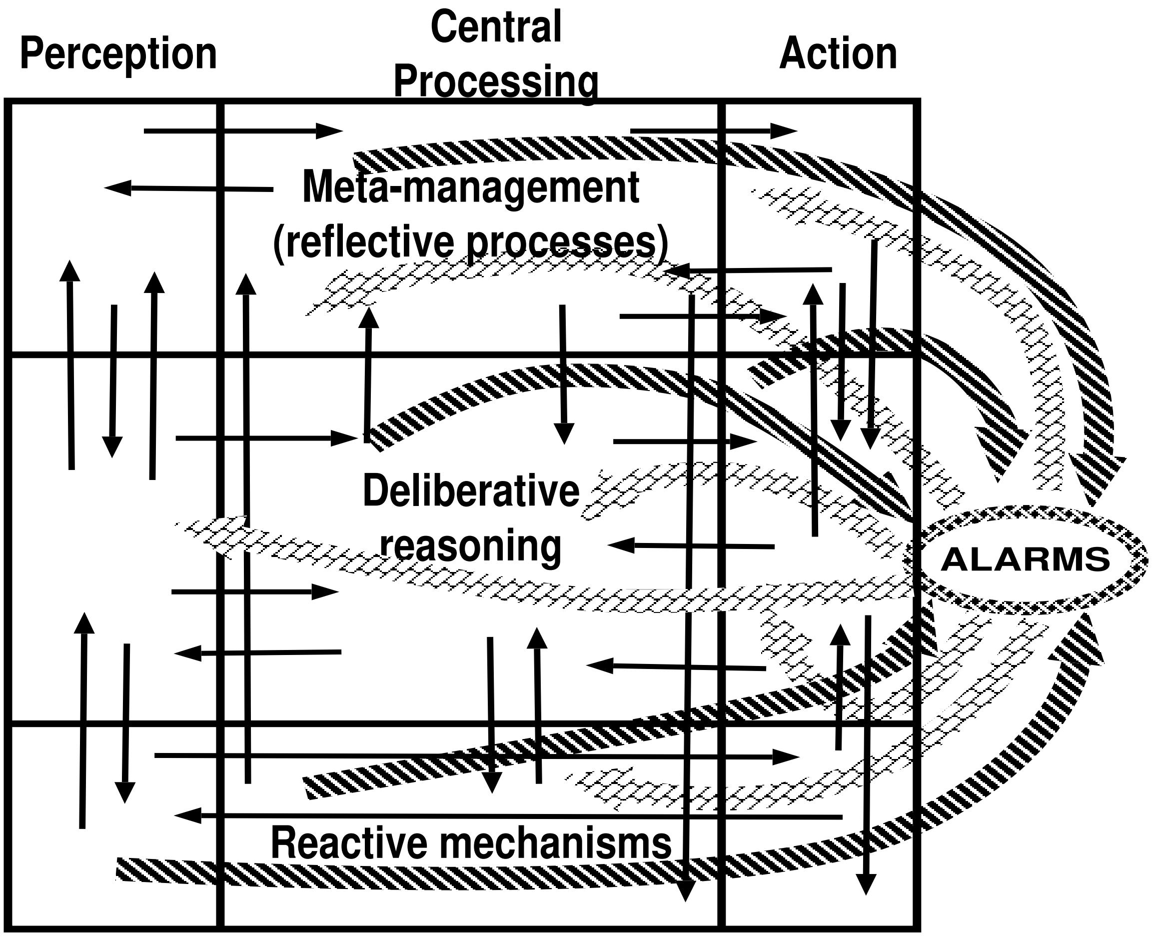


Evolution of different layers might use the familiar *copy then differentiate* process.

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

FAST "GLOBAL ALARM SYSTEM(S)" NEEDED

IT WILL INEVITABLY BE STUPID!

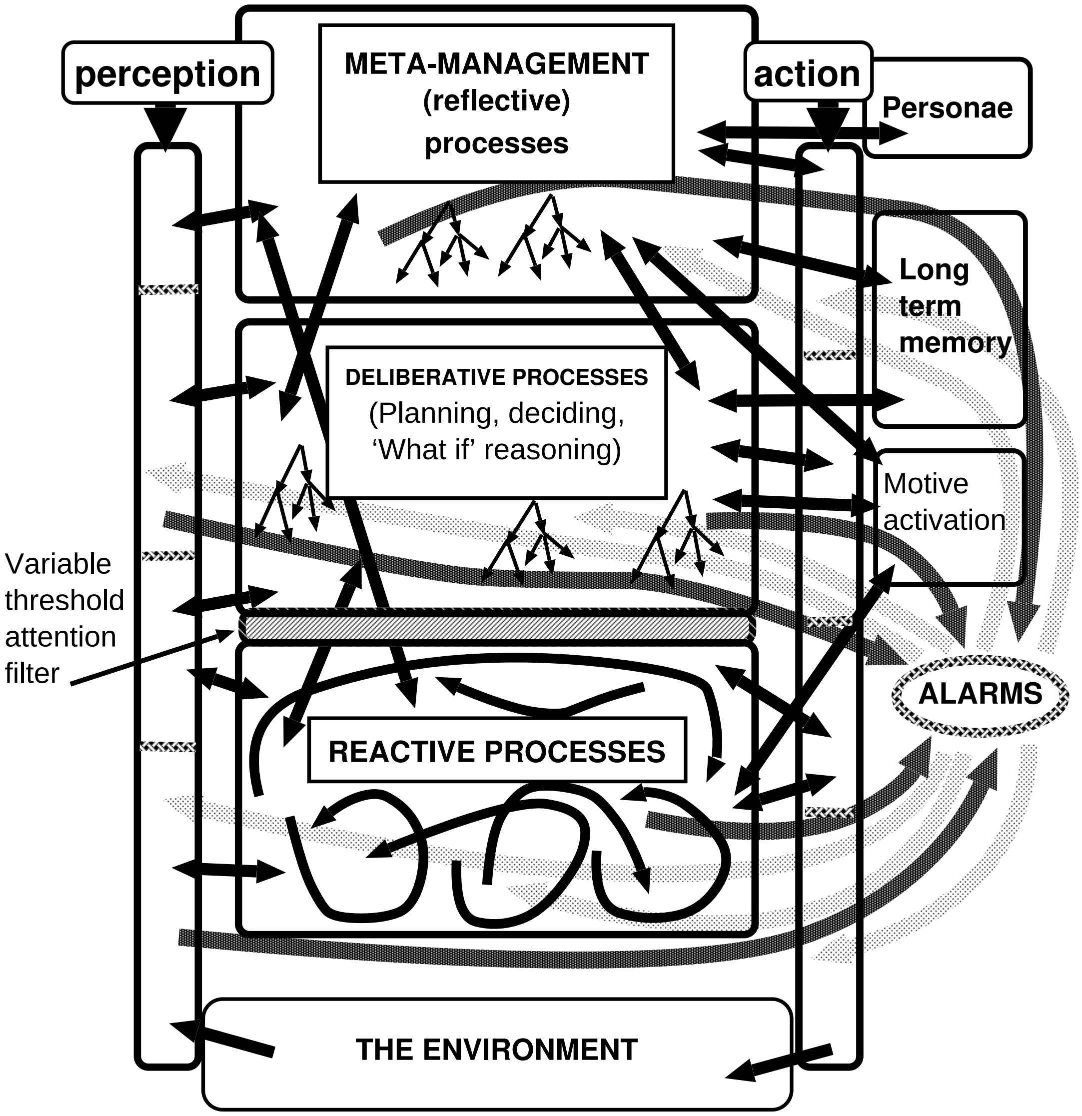


Many variants possible.

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

NOTE: WE CAN EXPLAIN THE EVOLUTION OF DIFFERENT CLASSES OF EMOTIONS IN THIS FRAMEWORK: PRIMARY EMOTIONS, SECONDARY EMOTIONS, TERTIARY EMOTIONS, AND MORE.

ANOTHER VIEW



This provides a framework for explaining many familiar and not so familiar aspects of human mentality.

Reminder: it's as much an *ecosystem* of mind as a *society* of mind (Minsky).

ADDITIONAL COMPONENTS

EXTRA MECHANISMS NEEDED

personae (variable personalities)

attitudes standards & values

formalisms categories descriptions

moods (global processing states)

motives motive comparators

motive generators (Frijda's "concerns")

Long term associative memories

attention filter skill-compiler

MANY PROFOUND IMPLICATIONS

e.g. for kinds of development
kinds of perceptual processes
kinds of brain damage
kinds of emotions

HOW TO MAKE PROGRESS?

Let’s do lots of experiments,

• including designing systems ourselves with different sorts of capabilities,

• in order better to understand what might need to evolve and what the obstacles might be.

Let’s avoid silly disputes between different approaches,

e.g.

• designed vs evolved or learnt,

• simulated vs ‘real’,

• symbolic vs neural, reactive vs deliberative,

• new vs old AI.

The debates are stupid and get in the way of collaborative research.

Making simplifications for the sake of advancing scientific understanding is not ‘cheating’.

WE NEED TO DEVISE GOOD TEACHING SCENARIOS

WE NEED

TO STRIVE TOWARDS

A FAR MORE POWERFUL

ECOSYSTEM OF IDEAS

READ (AND CRITICISE) OUR PAPERS

http://www.cs.bham.ac.uk/research/cogaff/

TRY (AND CRITICISE) OUR TOOLS

http://www.cs.bham.ac.uk/research/poplog/freepoplog.html

http://www.cs.bham.ac.uk/~axs/cog_affect/sim_agent.html