Biologically Inspired Cognitive Architectures 2011
Washington DC, November 4-6

A Model of Emotion as Patterned Metacontrol

R. Sanz, G. Sanchez-Escribano, C. Herrera

UPM Autonomous Systems Laboratory

Abstract

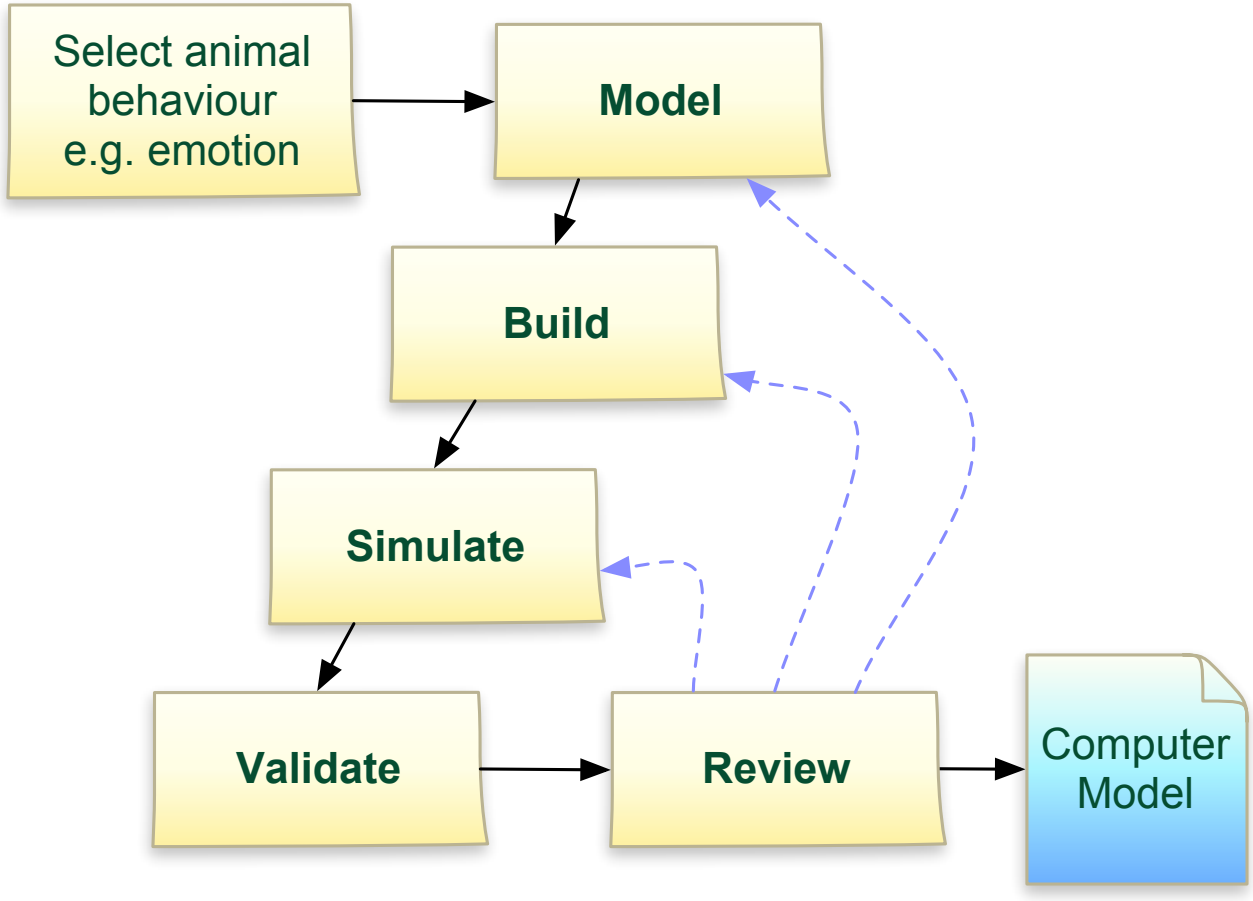
Adaptive systems use feedback as a key strategy to cope with uncertainty and change in their environments. The information fed back from the sensorimotor loop into the control architecture can be used to change different elements of the controller at four different levels: parameters of the control model, the control model itself, the functional organization of the agent and the functional components of the agent. The complexity of such a space of potential configurations is daunting. The only viable alternative for the agent –in practical, economical, evolutionary terms– is the reduction of the dimensionality of the configuration space.

This reduction is achieved both by functionalisation –or, to be more precise, by interface minimization– and by patterning, i.e. the selection among a predefined set of organisational configurations. This last analysis let us state the central problem of how autonomy emerges from the integration of the cognitive, emotional and autonomic systems in strict functional terms: autonomy is achieved by the closure of functional dependency.

In this paper we will show a general model of how the emotional biological systems operate following this theoretical analysis and how this model is also of applicability to a wide spectrum of artificial systems.

Keywords: Cognitive architecture, Functional organisation, Emotion, Adaptive Systems, Autonomous Systems.

Modeling Flow



Motivation

- Objective: Produce technology for **improving mission-level resilience** in technical systems in broad domains (ASys - robust autonomy).
- This is an on-going work:
  - Looking for architectural inspiration in emotional mechanisms
  - Not trying to build models to replicate biological behavior but to create universal architectural assets

Contents

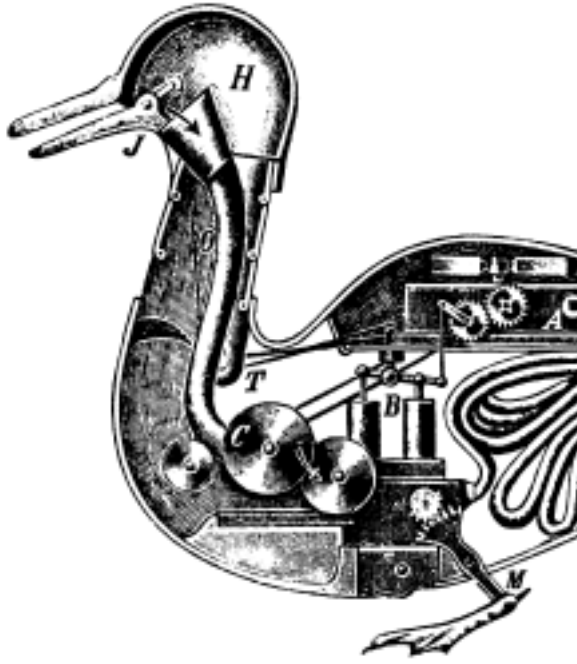
- Emotion in vivo and in silico
- Adaptive systems
- Emotion as pattern-based metacontrol
- Science and technology

Emotion in vivo and in silico

The many facets of emotion

- Affective states, emotional states, cognitive impact, physiological impact, evolutionary development, emotional expression, emotional appraisal, etc.

- Are all those many aspects to be considered in machines?



Four core aspects of emotion

- Behavioral/expressive
  - Acting and externalisation of state
- Experiential/subjective
  - State valuation - qualia
- Somatic/neurophysiological
  - Bodily state
- Cognitive/interpretive
  - State representation – mental model

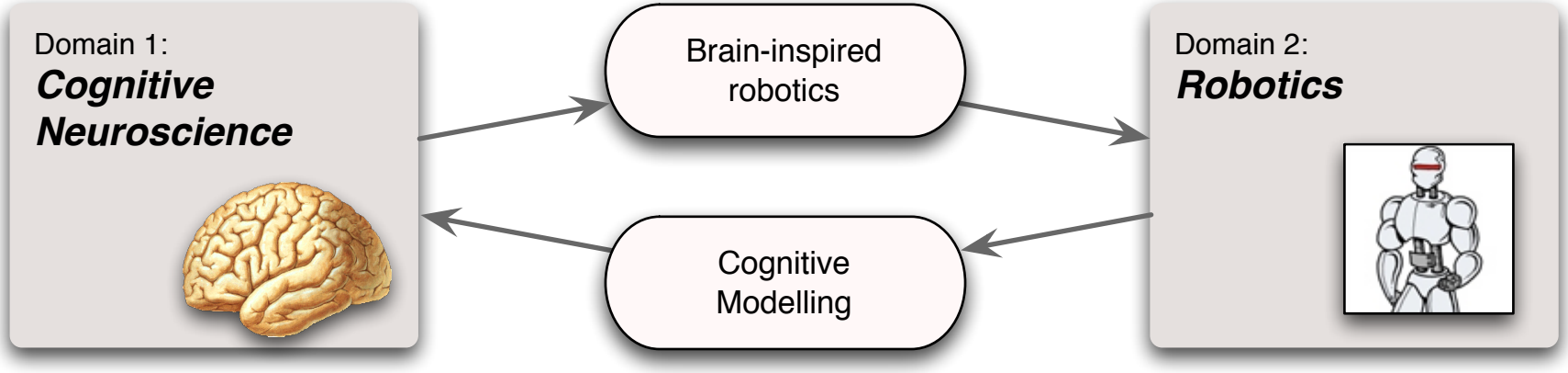
[Hudlicka]

Emotion Theories

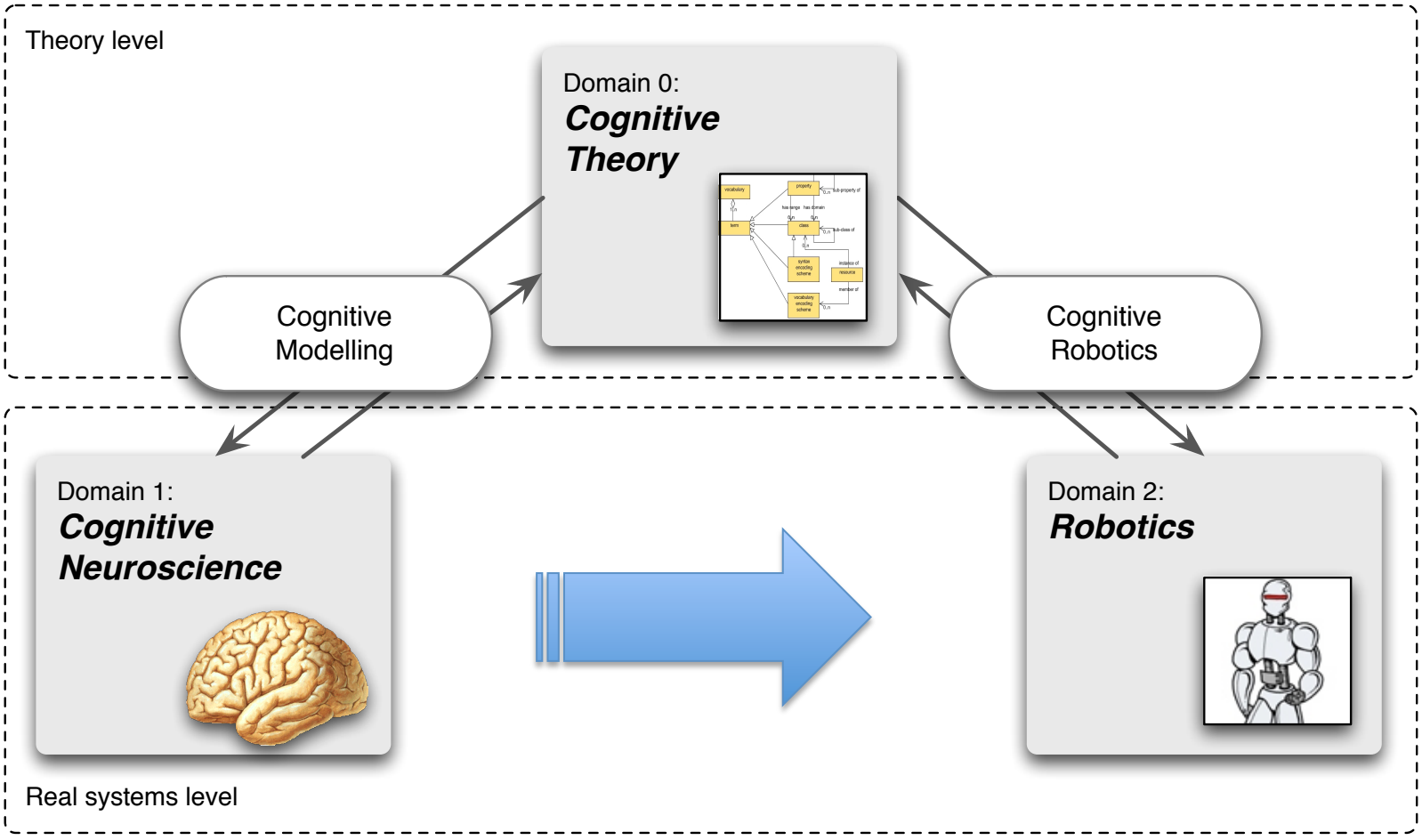
- Spreading activation models
  - Networks of nodes of cognitive-emotional bundles
- Componential models
  - Organisational variations in cognitive processes
- Parameter-based models
  - Appraisal dimensions that map to expressions

- From appraisal to external effects

Emotion beyond biology



Emotion beyond biology



A Stepped Program

- Find **fundamental architectural organisations**
- Eliminate unnecessary details in the modeling of biological emotion
- Consolidate in **domain independent models**
- Transfer to domains of technological use

Adaptive systems

Emotion-Cognition-Autonomy

- In the pursue of autonomy we investigate emotion-cognition integration architecture.
- Emotions are **control architecture mechanisms** directly **serving the goals** of the autonomous agent.
- They do so by providing real-time adaptation mechanisms to the agent.

The issue of adaptation

- Autonomous systems both natural and artificial do adapt:
  - Animals – as species and individuals
  - Machines – as objects and product lines

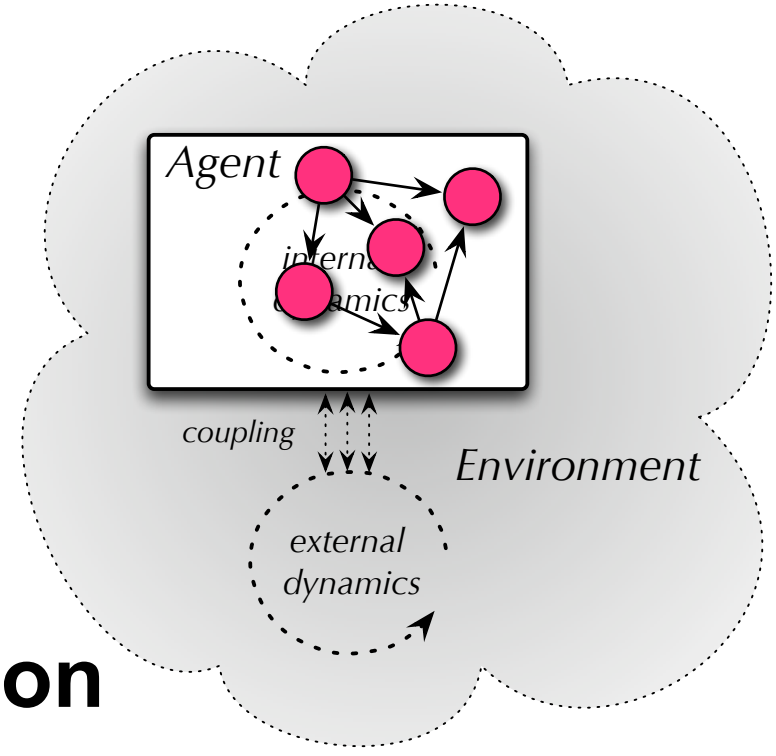
- Adaptation shall be framed in the context of autonomy:

  Autonomy (agent, environment, task)

[Sanz]

Change and Timescale

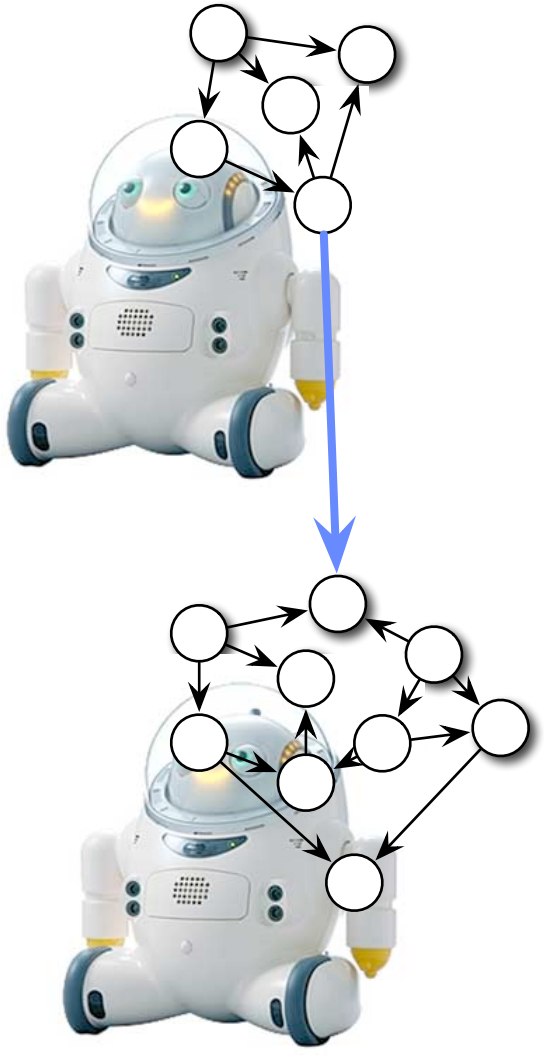
- Different timescales in dynamic change in agents:
  - React/Control
  - Learn/Adapt
- Two levels of changes:
  - State
  - Organisation
- Key: **Structural adaptation**



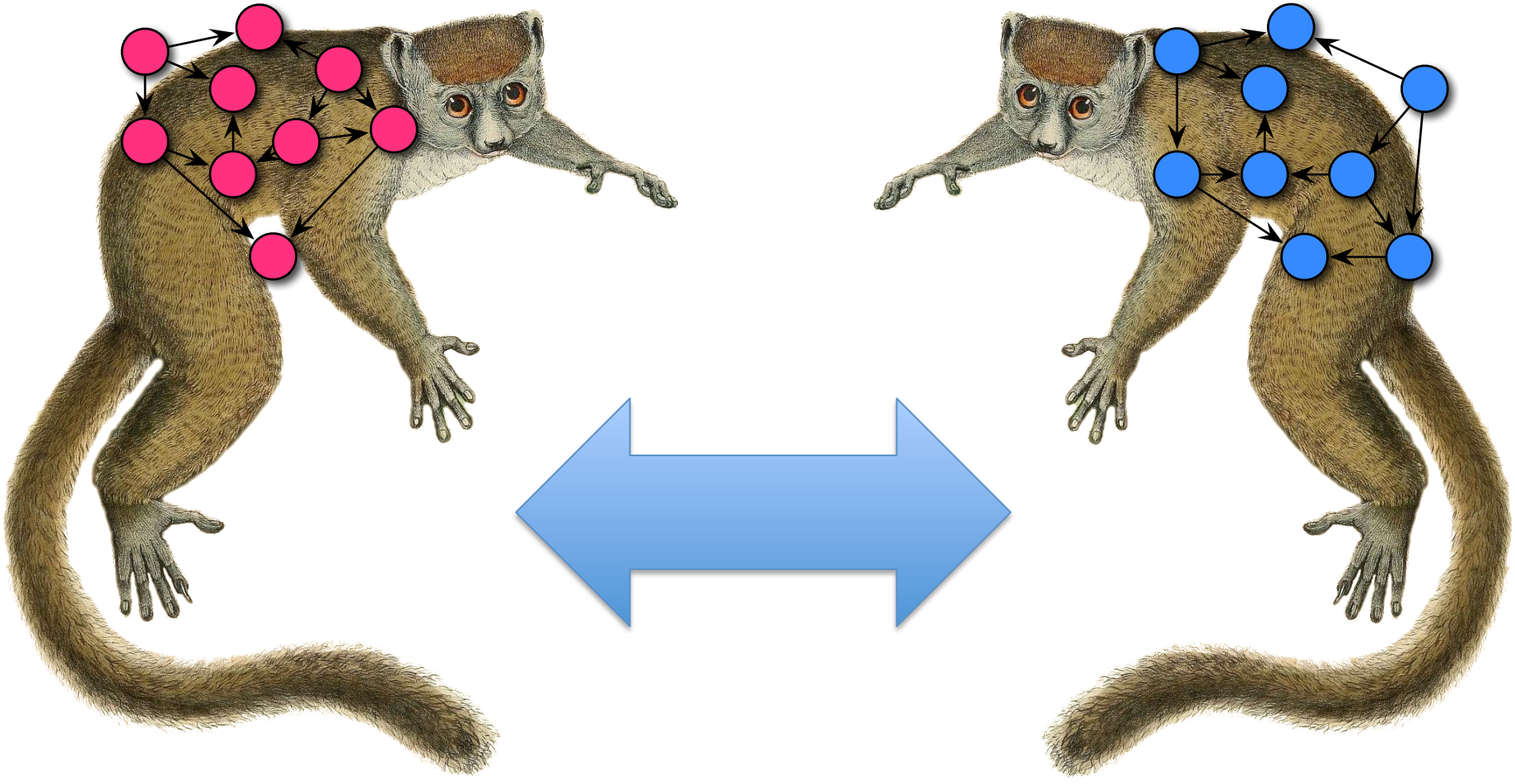
[Sanz]

Structural Adaptation

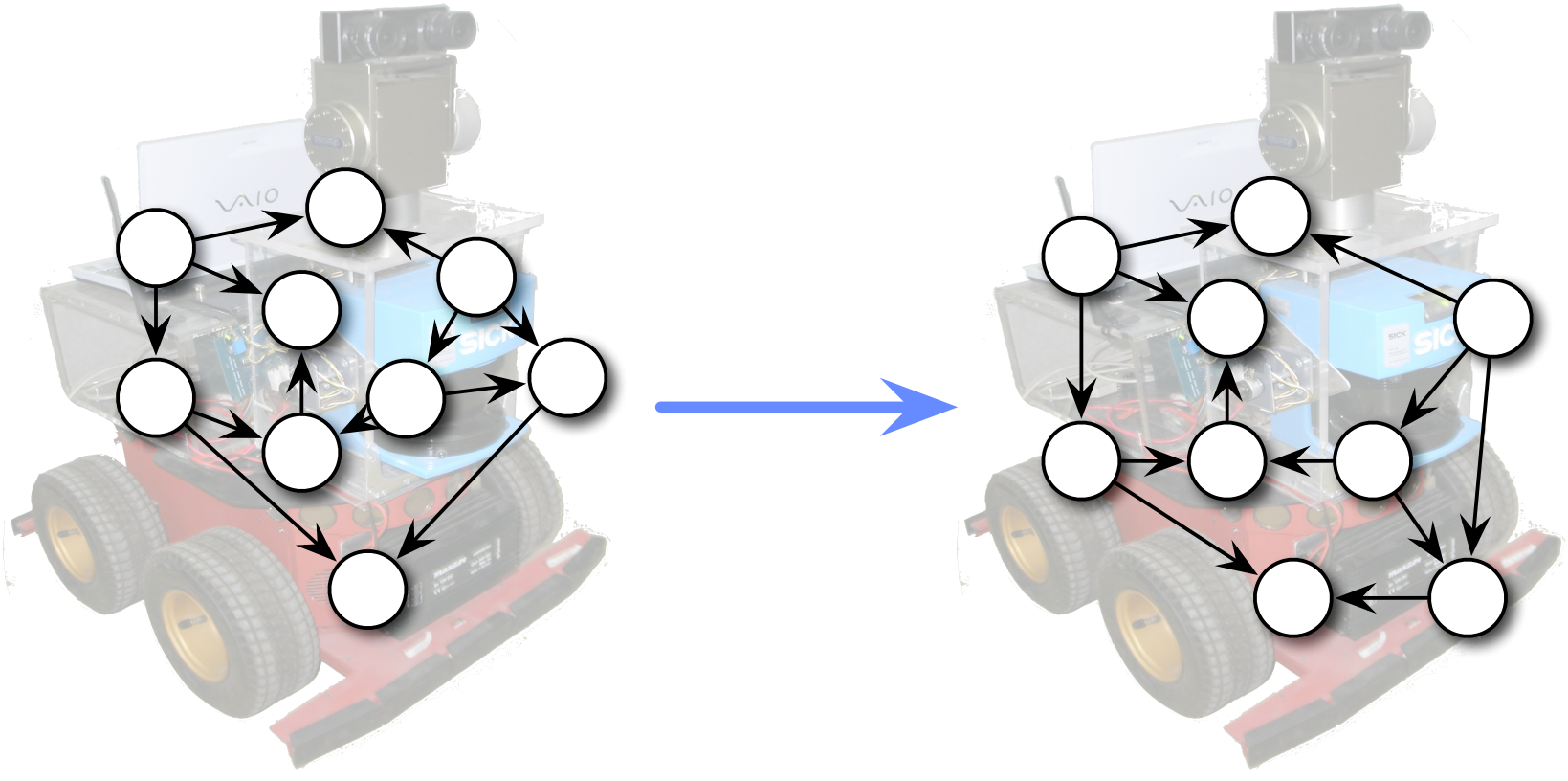
- By structural adaptation the agent changes its systemic structure:
  Agent = {Parts,Relations}
- The changes may affect both the **body** and the **mind**
- Mind structural changes are easier and more flexible due to the plastic, informational nature of minds



Fight or Flight

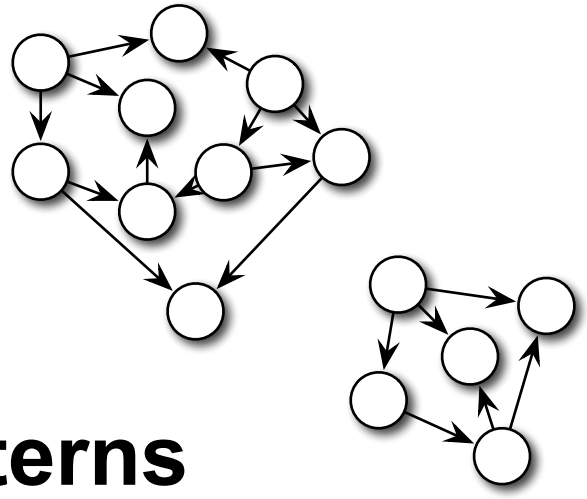


Architectural Change



What are the modules ? What granularity?

Architectural Patterns



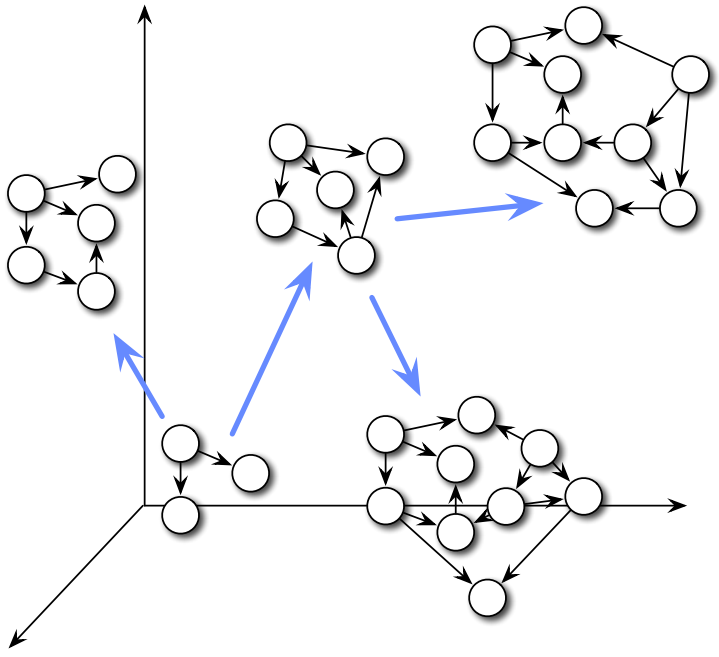
- Design patterns
  - Functional organisations
  - Reusable strategies
- **Architectural design patterns**
- We look at emotion as an architectural pattern operating over over other architectural patterns – a pattern-based pattern

Taming the adaptation problem

- There's a big difficulty of a general adaptation strategy based on structural change:

  the dimension of the space of organisational possibilities is **too big** (an inverse problem)

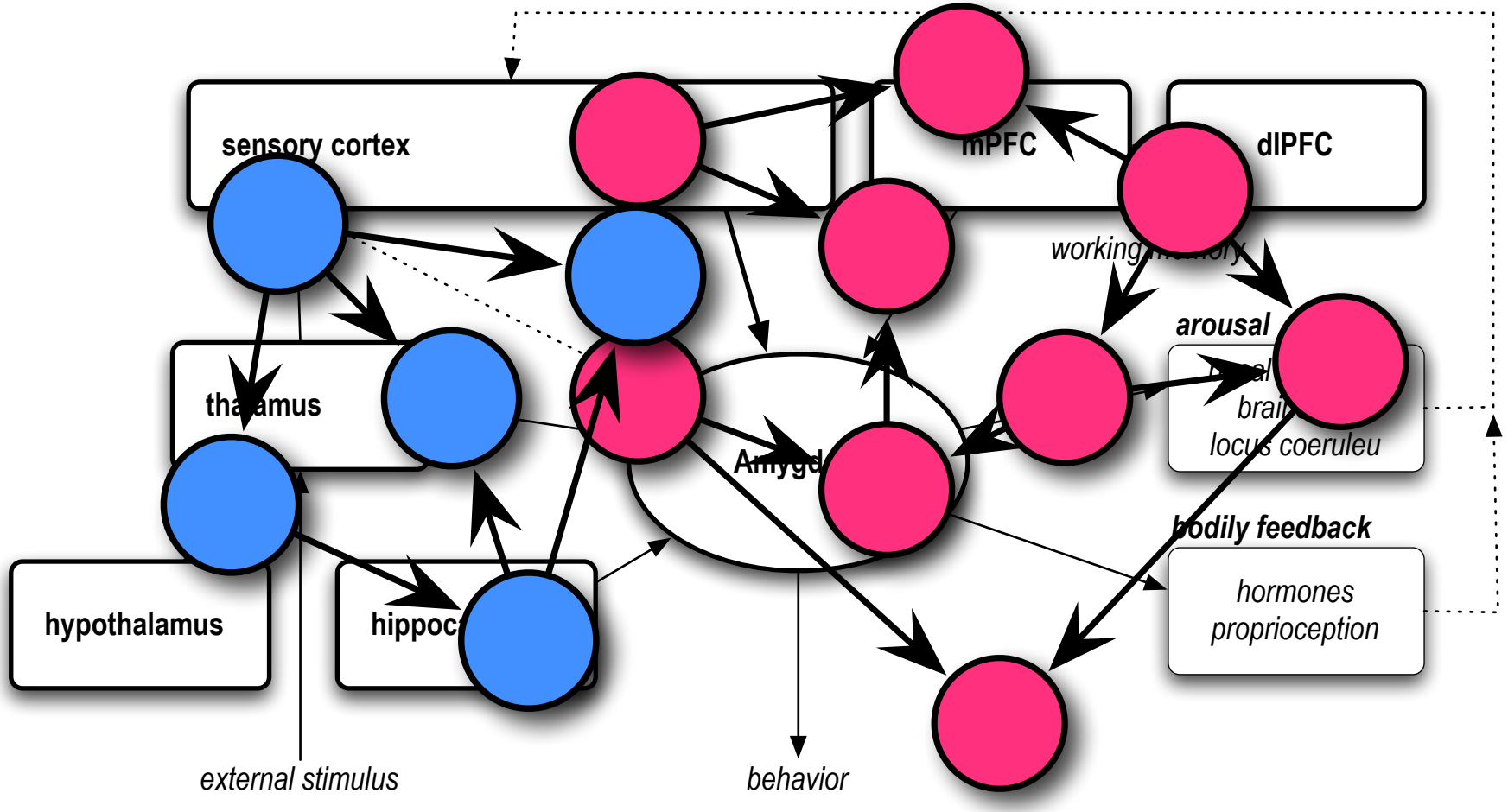
- It is possible to simplify the problem by means of the generation of **operational patterns**: forms of agent components' organisations that are tuned to specific niches/ tasks



Two basic processes

- Functionalisation
  - Aggregation and modularisation of subsystems (e.g. organs): Interface minimisation
  - Homogenisation: use common integration infrastructure
- Patternning
  - At the realisation –anatomical– level
  - At the organisation –physiological– level

Anatomy/Physiology



Emotion as pattern-based metacontrol

What patterns to activate?

- The decission concening what control patterns shall be activated depends on an appraisal of the agent-environment situation in terms of agent goals.
- *“interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism.”*

[Sander]

A View on Emotion

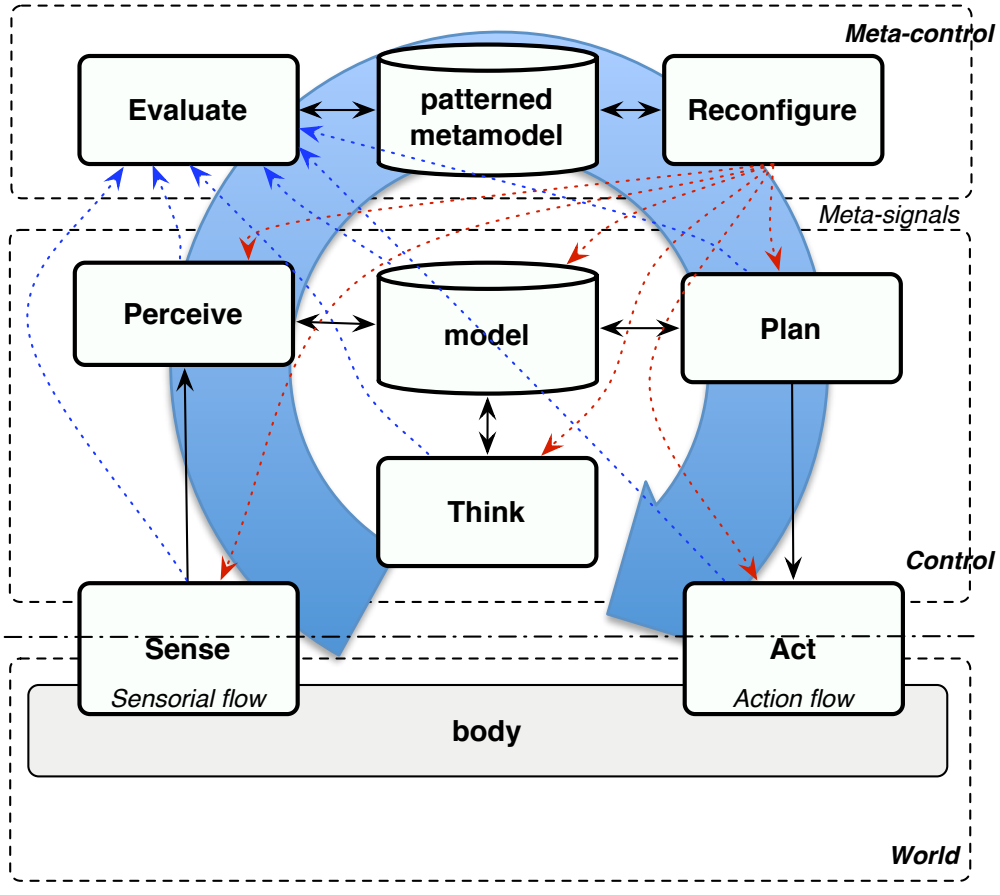
- We understand **emotion** as a real-time adaptation mechanism based on the activation of predefined patterns of organization of the agent.
- These patterns have been produced (evolutionarily) for specific activities/niches
- In a sense, emotion selects the best available agent architecture for the present moment (esp. the control, informational part).

Emotion and Feedback

- Emotion implements feedback from the evaluation of the degree of fitness between the functional state of the agent and the current state of affairs.
- This is broadcast as a **synthetic global state**
- These signals –in several appraisal dimensions– perform system-wide adaptation
- *"to achieve a multi-level communication of simplified but high impact information"* [Fellous]

The Emotion Bus

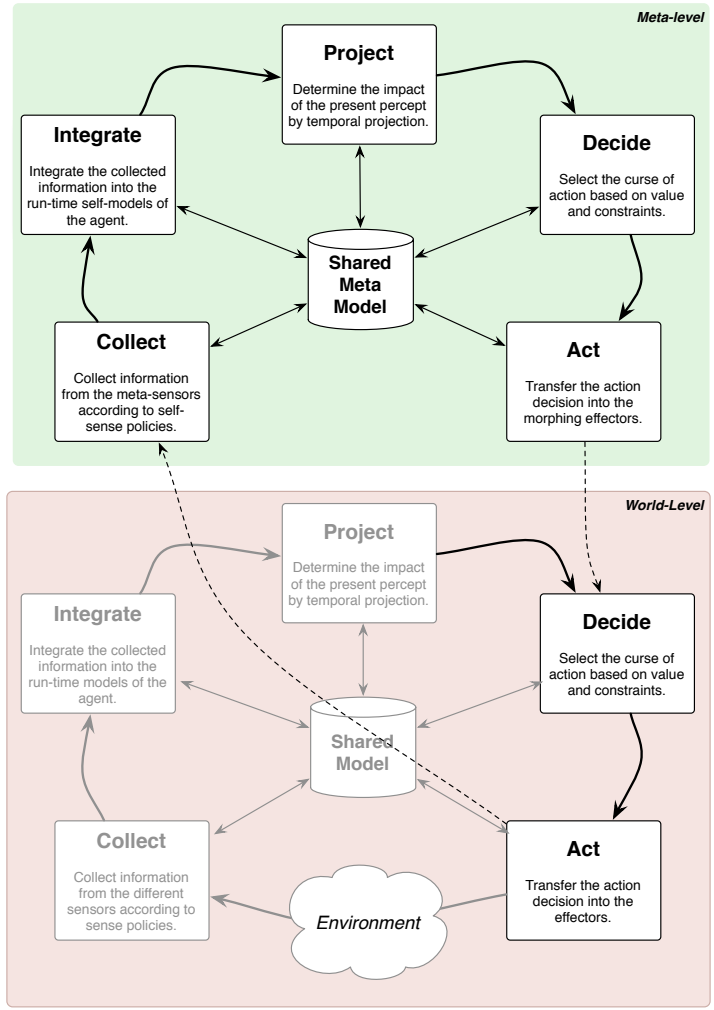
- Emotions are system-wide signals implementing this reconfiguration process – an **emotional bus.**
- When represented at the level of awareness they constitute the emotional **feelings.**



Two Cognitive Loops

- The control level
- The metacontrol level

- The control implements a target closure: controls domain magnitudes
- The metasystem implements a **functional closure**: perceives and controls *function*



Heterophenomenology

- Emotion capture the pursue of goals by leveraging phenomenality.
- In building experiencing machines, engineering the right phenomenological mechanism is crucial because it will be the origin of the intrinsic motivations of the agents.
- We shall adopt an heterophenomenological engineering approach to build phenomenologies into machines to match our very own needs.

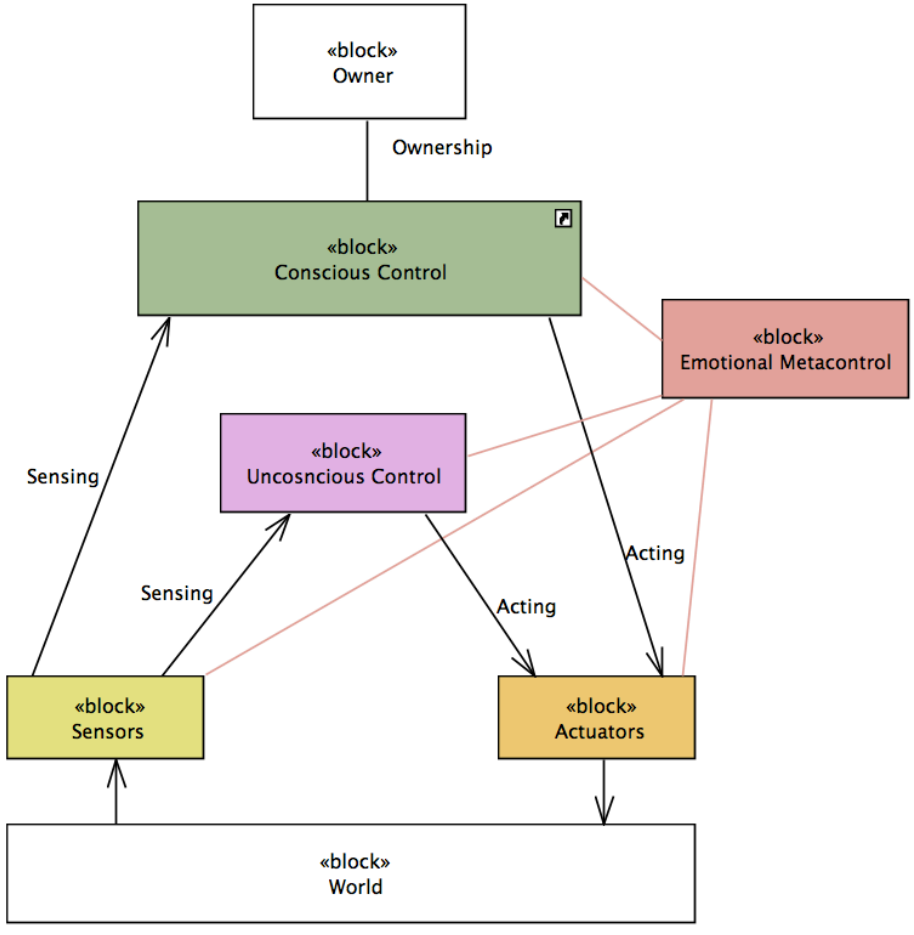
Science and technology

Evaluation

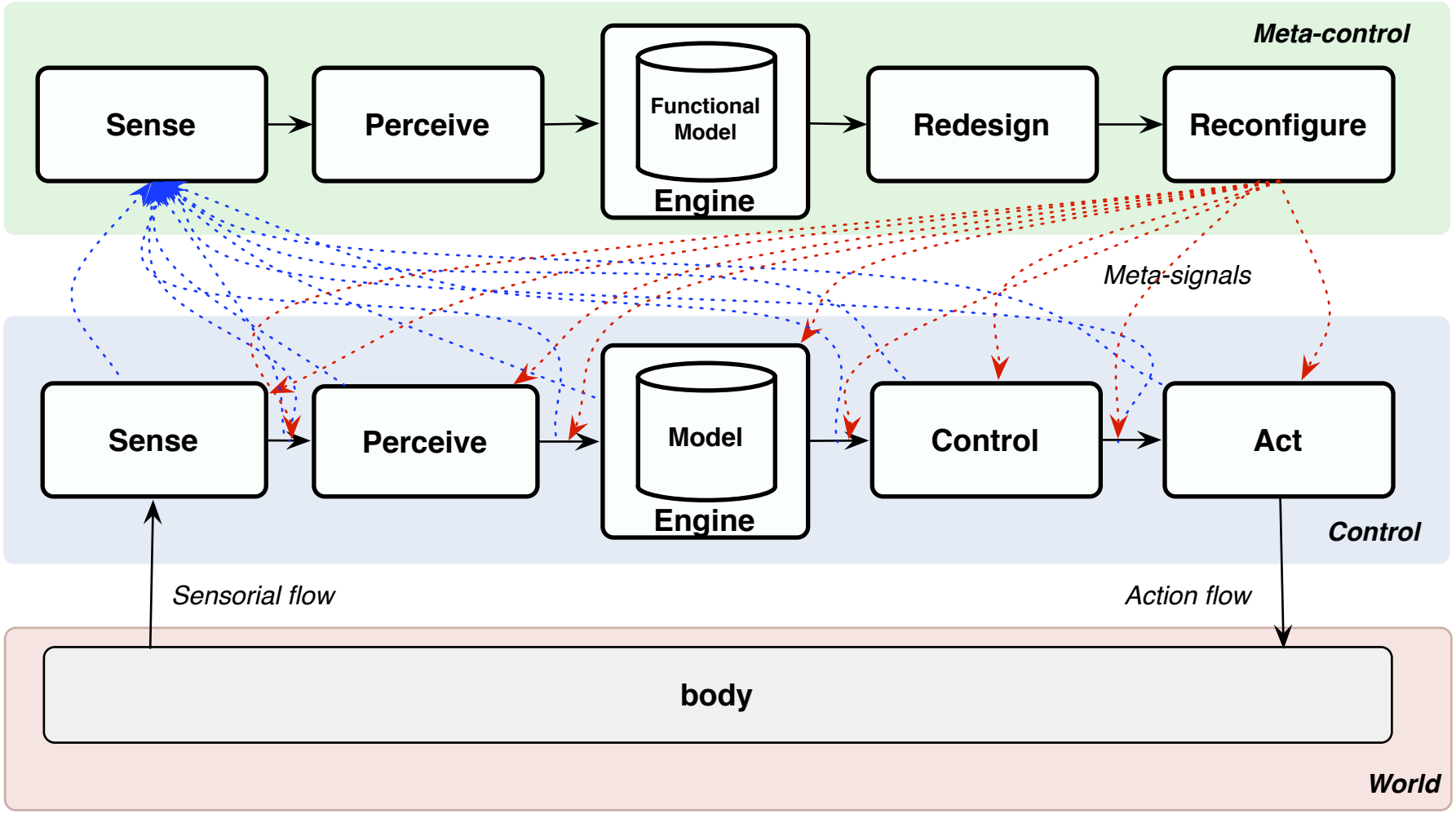
- This model can be evaluated in terms of both:
  - Its potentiality as an architectural asset for building better machines
  - Its explanatory power of the natural emotion
- For the first it must offer a path into technological consolidation as designs and reusable assets.
- For the second it must address specific traits of human emotion: appraisal, cognitive integration, emotion regulation, conscious and unconscious emotions, etc. Lots of painstaking work needed.

Mapping into Machines

- Mapping the model into machine realisations imply the construction of modular block-based cognitive systems – issues of granularity
- Shall be deployed over a plastic integration infrastructure
- Modern software architectural platforms helps with this (esp. distributed, real-time system platforms)

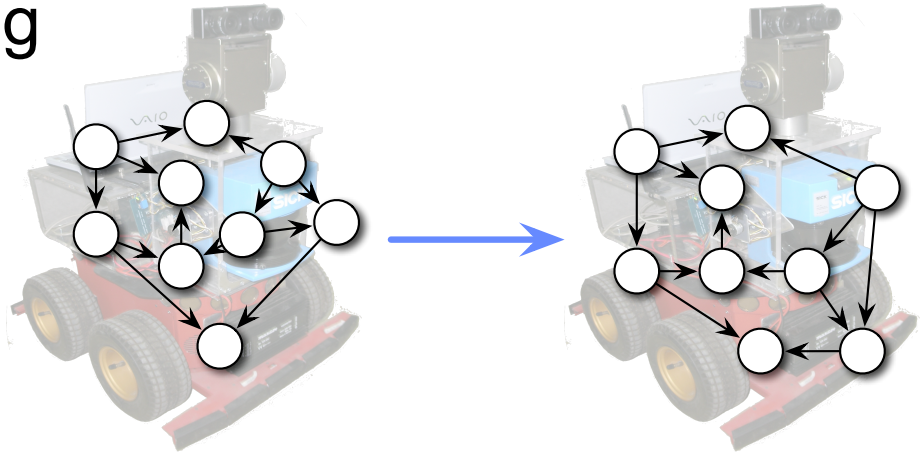


Cognitive Functional Metacontrol



Technological path

- Leveraging a technology of artificial emotion
- Two major issues:
  - The organisational patterns:
    - Persistent patterns and dynamic patterns
    - Some may be reverse-engineered from biological/brain systems
  - The plastic engineering infrastructure:
    - Componential
    - Self-referential
    - Run-time changeable
    - Real-time
    - Multigranular



Current state

- Done: Phase 0 ad-hoc implementation of rule-based metacontrol over rigid control patterns. Modular controllers based on instrumented widely available robot software.
- Ongoing: Phase 1
  - Explicit function representation and model-based reasoning
  - Bus-based emotion architecture
  - Rigorous ontology for cognitive-emotional systems
- Future: Pattern learning + Consciousness

Problems and issues

- Dual **representation of function** in both implementational and goal terms
- Emotion **dynamics** – pattern change
- Concurrent emotions and emotion **composition** – pattern coexistence, integration and interference
- **Heterogeneous** agent emotions – ToM, societies
- Perception of the emotional state – persistent questions of **self** and **qualia**

Conclusions

- Emotions are seen as **processes of goal-centric metacontrol of patterned control architectures** of autonomous agents.
- These processes happen at many levels –e.g. at the conscious and unconscious levels. The agent can perceive its very own emotional processes – having emotional qualia.
- This model can be implemented and leveraged in technical systems – esp. in complex control systems.
- Current robotic architectures and infrastructures are in general too simple to fully leverage this vision.
- Several theoretical issues still open.

Thanks !!

References

- Fellous,J.M. (2004) From Human Emotions to Robot Emotions, AAAI Spring Symposium on Architectures for Modeling Emotion, Stanford University, CA.
- Hudlicka, E. (2009). Challenges in developing computational models of emotion and con- sciousness. International Journal of Machine Consciousness, 1(1).
- Sander,D., Grandjean,D., Scherer, K.R. (2005) A systems approach to appraisal mechanisms in emotion. Neural Networks 18.
- Sanz, R., Matia, F., and Galan, S. (2000a). Fridges, elephants and the meaning of autonomy and intelligence. In IEEE International Symposium on Intelligent Control, ISIC'2000, Patras, Greece.
- Sanz, R. and Lopez, I. (2000). Minds, MIPS and structural feedback. In Performance Metrics for Intelligent Systems, PerMIS '2000, Gaithersburg, USA.