A NATURALISED DYNAMICAL ACCOUNT OF COGNITION

Xabier E. Barandiaran
xabier@barandiaran.net
http://www.ehu.es/ias-rearch/barandiaran

IAS (Information Autonomy Systems) Research Group
http://www.ehu.es/ias-rearch
Dept. of Logic and Philosophy of Science
UPV-EHU (University of the Basque Country)

this presentation can be found at: http://www.barandiaran.net/textos/alergic_2004

COPYLEFT

OVERVIEW

0. Introduction: the problem, the question
1. Life-as-it-could-be: basic autonomy
2. From life to cognition: the autonomy of the NS
3. Biological Embodiment: more than “just” a physical sensorimotor interface
4. Internal Dynamic Organization: Information is dead, long live to in-formation
5. Naturalizing cognition: recapitulation
6. Postscript: on evolutionary robotics as a theoretical tool (with proposal, ongoing projects and design principles)

INTRODUCTION

the problem:

A NATURALISED ACCOUNT OF COGNITION

Naturalism

- Ontological:
  - Our experience is the result of a unified reality so no specific substances (such as the mental, representations, language, etc.) or *ad hoc* explanations should be admitted to explain it.
- Methodological:
  - Philosophy should go hand by hand with scientific research grounding our understanding of the world on the empirical operations we can inpinge upon it.
- Note:
  - Naturalism should not be judged in itself as a thesis but as a pragmatic proposal evaluated in terms of its achievements...
    - Ultimatelly naturalism should account for itself through naturalist epistemology, i.e. through the scientific understanding of knowledge itself.

The question(s)

- 3 minutes after the Big-Bang there was no cognition and at the scale of 10^{-20} meters there is no cognition...
- How did cognition arise, how is it sustained?
- How can we specify cognition as a natural phenomenon which is distinct from those that surround it, underlay it and preceed it?
- How did the fundamental distinction between subject and object of knowledge appear in the history of nature (where no subject or object as such could be found before)?

Traditional functionalist answer

- The specificity of cognition is given by the representational nature of the functional input-output relationships of certain systems
- Representational means:
  - Causal correlation between internal and external states of affairs (Fodor)
  - Evolutionarily selected according to its correlation (Millikan)

Traditional functionalist answer

- But:
  - Traditional representationalism presuposes distinction between subject and object of representation
  - Requires an external observer or evolutionary history to ground representational content.
  - The fact that an internal state is a representation of states of affairs in the world does not lie on the causal organization of the system: it is an arbitrary choice of the observer

Dynamicism (I)

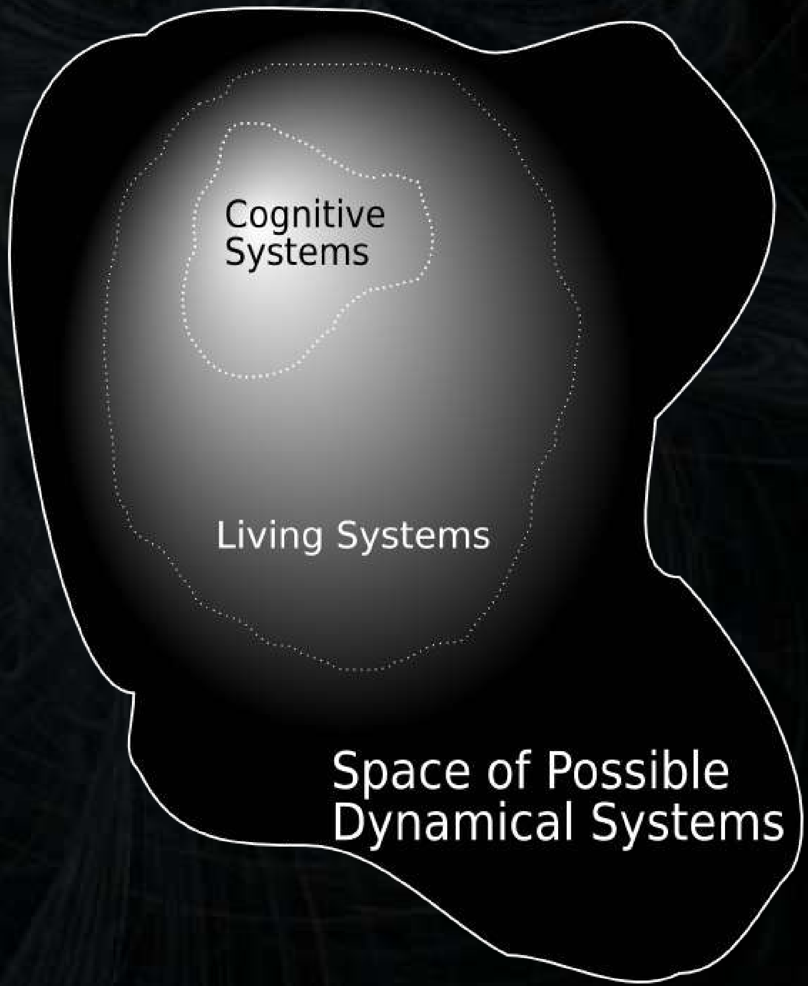
- The dynamical hypothesis:
  - Ontological: cognitive systems are instances of a dynamical causal organization
  - Methodological: cognitive systems are better understood with dynamical system theory
- But:
  - Neither the dynamical hypothesis nor DST offers any criteria to distinguish cognitive from non-cognitive dynamical systems.
    - "This paper simply takes an intuitive grasp of the issue for granted. Crudely put, the question here is not what makes something cognitive, but how cognitive agents work " (van Gelder 1998, p.619).
    - But can we understand how cognitive agents work without knowing what makes them cognitive?

Dynamicism (II)

- Nonetheless dynamicism:
  1. Allows for modelling of underlying mechanisms
  2. Does not presuppose distinction between mind and world: crosses over brain, body and world.
  3. No compromise with representational theoretical primitives.

The question reframed

- From the class of all possible dynamical systems ...
  - Which are the ones we call cognitive?
  - How do we draw the boundaries?
  - If we are not to believe in rigid boundaries still... What specifies the gradient towards the cognitive?
- We are interested in cognition-as-it-could-be independently of particular bio-anatomical structures.



Naturalistic contraints on the answer

- The naturalistic approach we defend should be able to account for cognition in two fundamental aspects:
  - Historic-evolutive: should account for the diachronic emergence of cognition
  - Dynamic-organizative: should account for the synchronic emergence of cognition from the bottom-up,
    - how is cognition sustained and enabled by underlying (more fundamental) processes?
- The answer should be grounded on the available scientific knowledge and provide productive feedback to science both at empirical-analytic and synthetic (constructive) levels.

I. BASIC AUTONOMY

LIFE AS-IT-COULD-BE

Bottom-up constraints for any dynamical system (*that could be*)

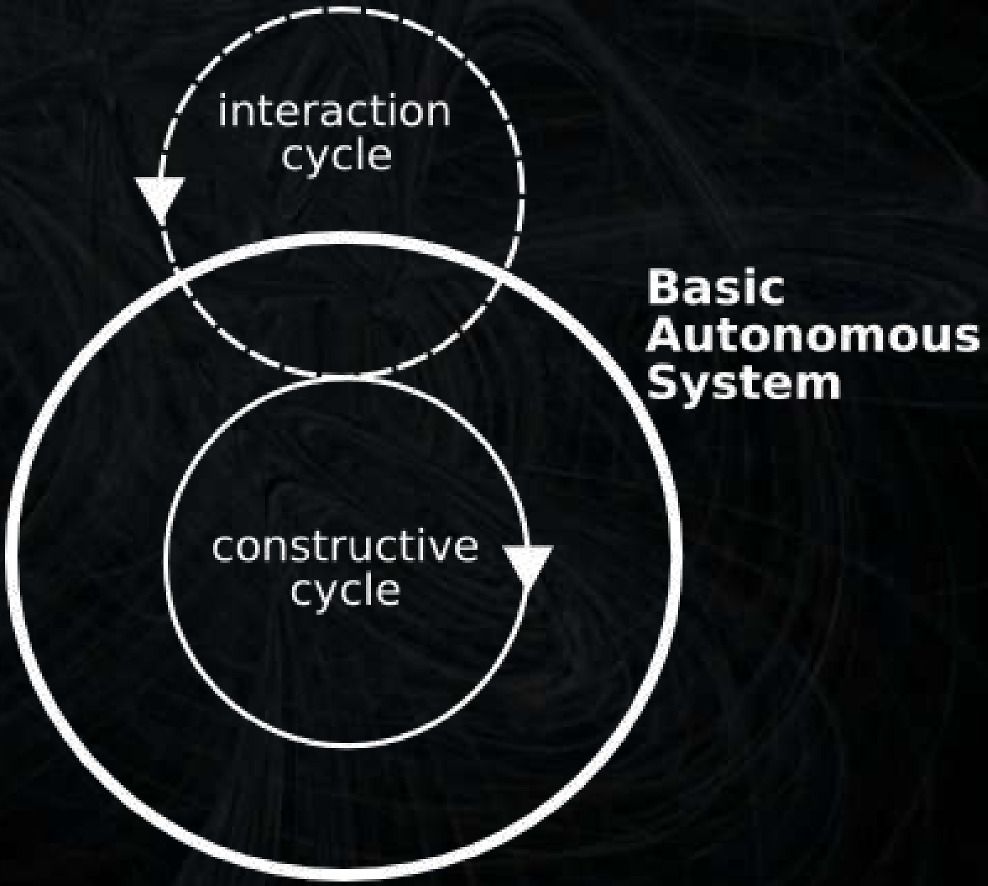
- What-can-be is defined by its stability conditions which act by both constraining and enabling the existence of dynamical domains:
  - Persistence of variables and regular interactions between them that we can operationally isolate and measure.
- Three main kinds of stability in nature:
  - Conservative systems (rocks, atoms, planets): robots and machines in general are conservative systems.
  - Far-from-equilibrium stability (living beings): dissipative structures, thermodynamically open
  - Sequential structures (DNA, replicating templates): require a far-from-equilibrium dynamical system of component production to replicate

Basic autonomy

- Basic autonomy (Ruiz-Mirazo & Moreno 2000) is the organization by which:
  - far-from-equilibrium and thermodynamically open systems
  - adaptively generate internal and interactive constraints
  - to modulate the flow of matter and energy required for their self-maintenance.
- Similar to autopoiesis but thermodynamically open:
  - Interactive dynamics are constitutive of the system (structural coupling with the environment is not something that comes additionally but is essential).

Interaction and construction

- Two cycles:
  - Constructive: generation of internal constraints to control the internal flow of matter and energy for self-maintenance (e.g.: metabolism).
  - Interactive: control of boundary conditions for self-maintenance (e.g.: active transport through membrane, breathing, adaptive behaviour,...)



Functionality and normativity

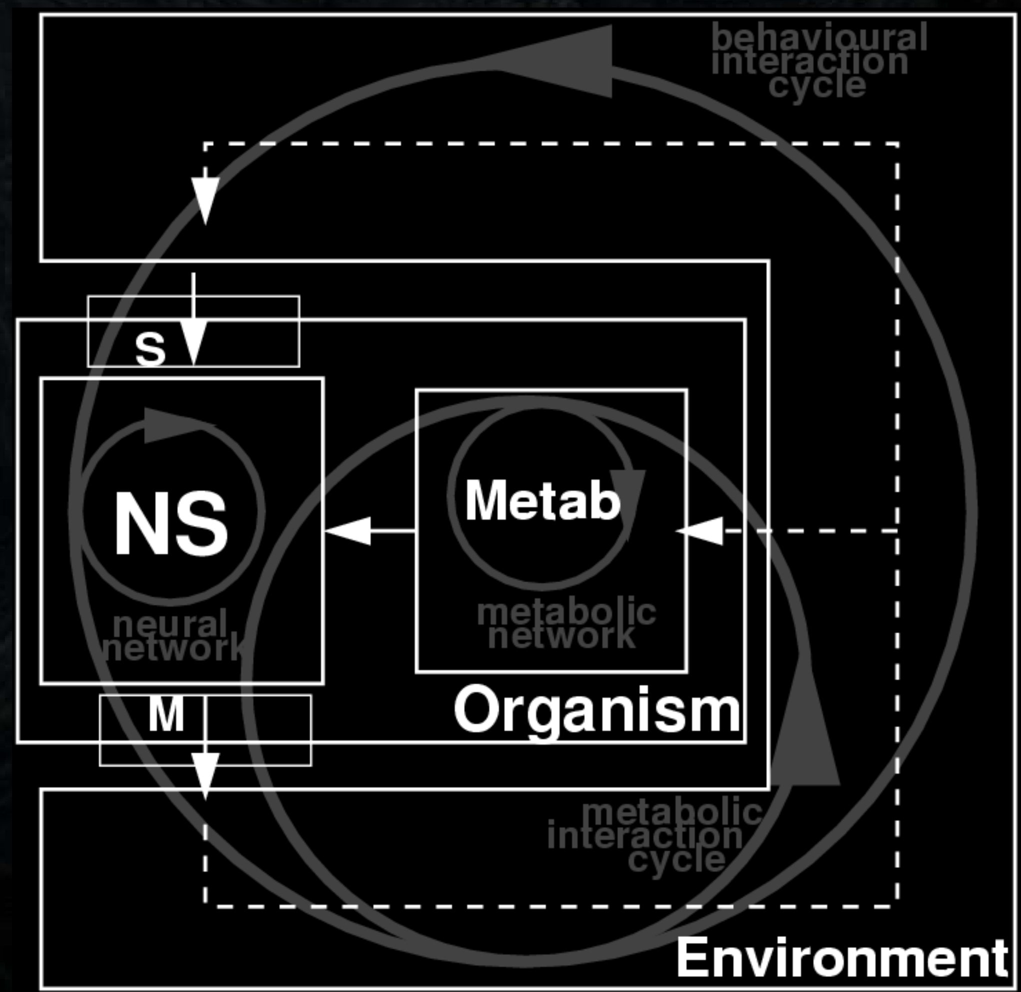
- FUNCTIONALITY: a process is *functional for the system* if it contributes to its self-maintenance
- NORMATIVITY: a process becomes normative if it is dynamically presupossed by other processes in their contribution to the overall self-maintenance.
  - e.g.: the normative (proper, necessary) function of the kidney is to filter blood because the dynamic-metabolic organization of the rest of the organism relies on this blood filtering
- NOTE THAT:
  - No structural decomposition is required.
  - Functional description is not arbitrary (the far-from-equilibrium system) would not exist otherwise.

II. FROM LIFE TO COGNITION

THE AUTONOMY OF THE NERVOUS SYSTEM

Decoupling

- Evolutionarily speaking the appearance of the nervous system (NS) + sensorimotor embodiment implies the decoupling of constructive and interactive cycles
- Solving a bottleneck between body size and interactive opportunities



Hierarchical Decoupling

- Hierarchical Decoupling of the NS from Metabolism:
  - Metabolism generates and sustains a dynamical system (the NS) minimising its local interference with it.
  - Hierarchical: metabolism produces and maintains the architecture of the NS.
  - Decoupling: metabolism underdetermines the activity of the NS (which depends on its internal dynamics and its embodied SM coupling with the environment).
- Operationally:
  - If we are to predict the state of the NS, local states of cell metabolism are not going to be enough: much more important are the electrochemical states of other neurons and the SM-coupling with the environment.

Operational dynamical primitives

- The NS will, in turn, have to be coupled to the global metabolic needs of the organism.
- But the hierarchical decoupling will allows us to specify the operational primitives (dynamical variables) that constitute this domain, mainly:
  1. change of membrane action potential over time (spikes),
  2. synaptic connections (connectivity matrix) and
  3. modulators: synaptic (local and global) and threshold.
- The research for this dynamical primitives and its functional higher level organization constitutes the search for a neural “code”: what kind of local differences can make a global difference (spikes, rates, gas-nets, etc.).

Behavioural Adaptive Autonomy

- The function of the NS in the overall organization of the organism is behavioural adaptivity, dynamically defined as:
  - Homeostatic maintenance of essential variables under viability constraint through the control of the behavioural interactive coupling with the environment
- Now the question becomes:
  - What is the dynamic organization of the NS and how is it related to behavioral adaptivity?

Constraints on the dynamics of the NS

- Two main kinds of external constraint on the NS:
  - Innate constraints (Elman *et al*. 1996):
    - Chronotopic: timing of certain developmental processes
    - Global architectural: global neural pathways, kinds of connectivity, etc.
  - Value constraints:
    - Big perturbations of neural dynamics through specific signals: pain, hunger, pleasure, etc.
- The complexity of the possible neural dynamics is subdetermined by this constraints
- The dynamics of the NS enter a process of local self-organization and historical self-determination through interactions with the environment (internal and external)

Self-organization

- Self-organization:
  - Local non-linear interactions between components generate a global behaviour which is maintained through a certain number of constraints of which at least one is a product of the global pattern.
    - Global pattern is not instructed from outside
    - Global pattern cannot be reduced to any of the local components
  - Example: CPG (Central Pattern Generator), interaction between neurons on a local circuit generate a robust oscillatory pattern(s)

The Autonomy of the NS

- Autonomous systems are dynamical systems defined as a unity by their organization:
  - they produce themselves (their activity is mainly self-determined) and
  - they distinguish themselves from their surroundings

The Autonomy of the NS

- The NS (embodied and situated) is an autonomous systems because:
  - Integrity: The dynamic and far-from-equilibrium structure of the NS is maintained by:
    - the network of processes itself (cohesivelly and recursively)
    - a recursive interaction with the environment
  - Differentiation: The dynamic structure of the nervous system is distinguished from the interactive dynamics with the environment by its functional integration, i.e.:
    - a complexity asymmetry by which the internal processes are more complex that the interactive ones
    - system identity can be maintained across a different range of sensorimotor couplings

Autonomy of the NS

- All the constraints are not self-generated: value and innate constraints are essential but do not completely specify the dynamics of the NS
- Starting with this innate constraint and through its sensorimotor coupling with metabolism and environment the autonomy of the NS is an open historical process of self-determination
    - We could say that the organism (through the hierarchical decoupling of the NS) generates a dynamical domain of a much higher variability (complexity) than its metabolic and genetic structure can control.
- The autonomy of the NS is not an absolute term but a gradual *becoming* (unlike Maturana & Varela's notion of operational closure).

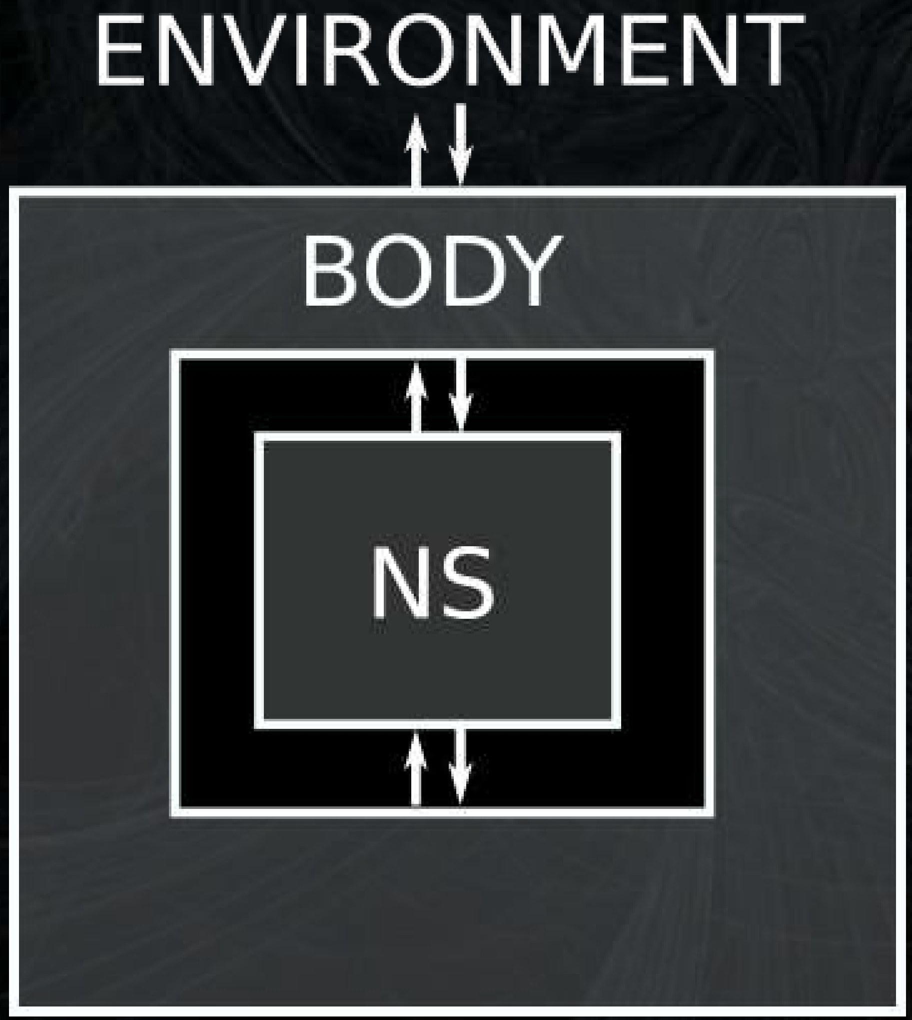
III. BIOLOGICAL EMBODIMENT

MORE THAN "JUST"
A PHYSICAL SENSORIMOTOR INTERFACE

Physical embodiment

- In the dynamical approach to cognition the body is generally conceptualized as the physical interface between the NS and the environment.
- Since cognition is the result of closed sensorimotor loops with the environment (not a set of disembodied abstract computations) then body constraints become crucial to the understanding of behaviour.
- The body becomes like a primary environment for the NS from which the NS cannot decouple (unlike selective engagements with features of the environment).

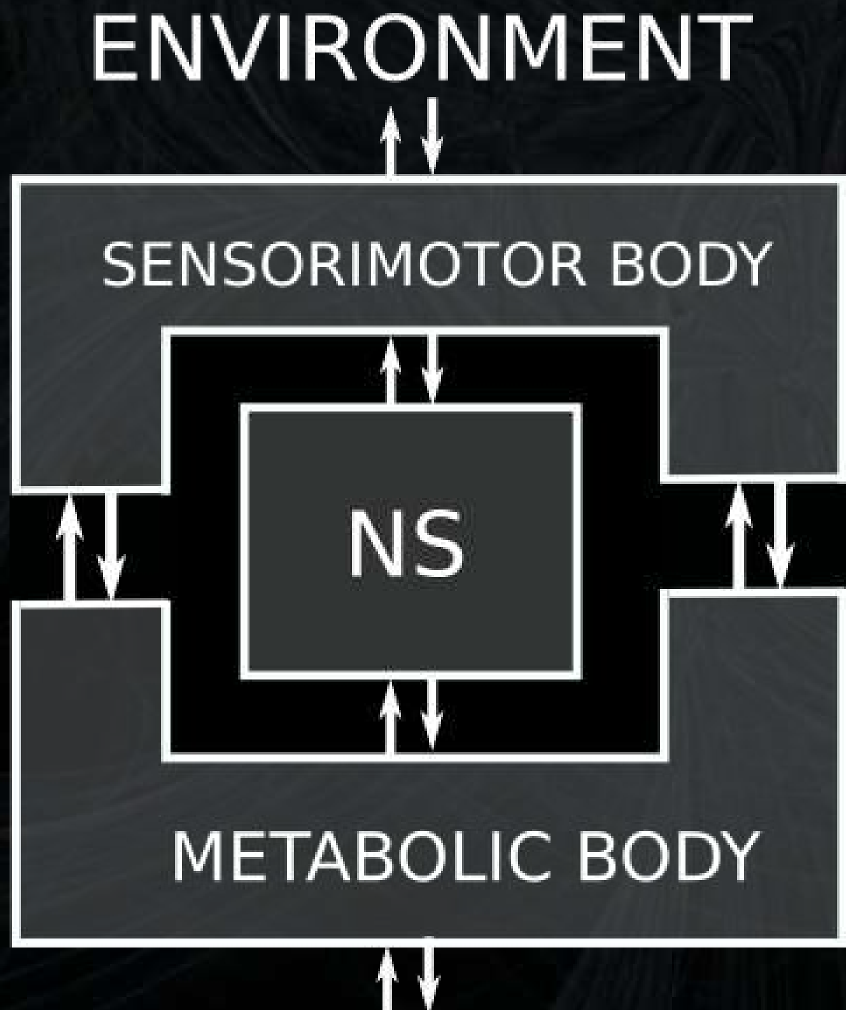
Physical embodiment



Biological embodiment

- The body of the NS is not “just” a physical interface, the (organismic) body, is first of all a biological autonomous (self-sustaining) body.
  - the condition of possibility of the NS as a dynamical system.
- The brain is not just coupled with the environment through the body but also with the body's internal homeostatic dynamics.
- Antonio Damasio: the NS interacts with the environment in terms of the effect of this sensorimotor interactions on the (metabolic) body dynamics.
  - somatic markers
  - internal body landscape

Biological embodiment



Autonomic NS

- Organisms whose adaptive strategies rely on motility (fast displacement) are very constrained in size
- Evolutionary solutions to this problem are vertebrates with endoskeleton and ANS: neural control of metabolism (breathing, blood flow, etc.) ensure metabolic needs of muscles
- Body and ANS as a source of value dynamics
- And finally recruited-for non adaptive sensorimotor evaluation: somatic markers for “higher level cognition” (see the role of emotions in decision making)

IV. INTERNAL DYNAMIC ORGANIZATION

INFORMATION IS DEAD...
LONG LIVE IN-FORMATION!

Hypothesis

- The specificity of cognitive dynamics (what makes it different to other dynamical systems) is given by a particular kind of dynamic organization: in-formation.
- This kind of dynamic organization should account for:
  - intentional and semantic phenomena and
  - the way in which cognitive agents organize their behaviour generating a “world” out of undifferentiated and neutral surroundings

Information is dead...

- Informational accounts of the NS activity rely on statistical measures of stimulus-neural activity correlations (conditional probability of neural activation given stimulus X)
- But:
  - this correlation is not accesible to the system (whose only access to the stimulus is the neural activation itself!)
  - this approach does not provide any criteria for a particular kind of internal dynamic organization but just a kind of system-environment relationship for a particular observer
  - this cannot account for system detectable error

Behaviour -- Structure

- Some preliminary definitions:
  - STRUCTURE: is the subset of internal variables involved in a certain sensorimotor coupling (hyperdescription)
  - STRUCTURAL STABILITY: happens when a subset of internal variables remains stable or invariant during that coupling, allowing the structure to operate without interference
  - STRUCTURAL CHANGE: in given circumstances (different sensorimotor correlations) the structure can change and the old sensorimotor coupling is lost
- So structure sustains behaviour but it can be the case that behaviour sustains structure too because structural stability might depend on a given SM correlation

Example 1: homeostatic adaptation

- Agent performs phototaxis
- Inversion of sensors disrupts phototactic behaviour
- Agent's internal dynamics enter unstable region
- Stabilizes when phototaxis is recovered
- Behavioural stability depends on structural stability

Long life to in-formation !

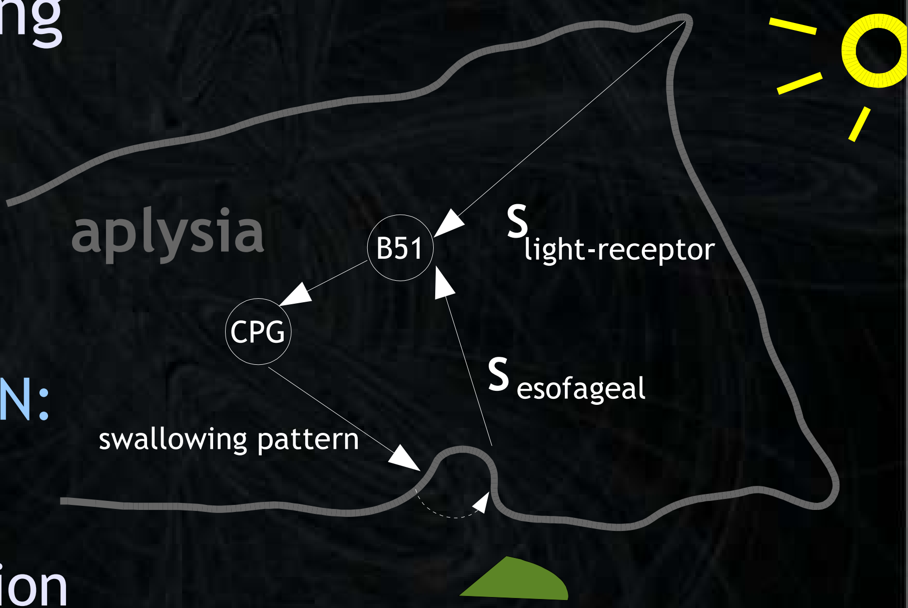
- In-formation: is formation from within of the behavioural coupling organized through the expectancies of the interaction outcomes.
  - Expectancies: can be clearly defined as dynamic counterfactuals (conditionals): if a certain interactive condition is not met during or after a certain behavioural coupling the dynamic structure involved in the coupling dissolves
- The “behaviour sustains structure” bit can be decoupled from immediate SM coupling and become dependent on future SM conditions.

Example 2: Aplysia

- Activity of neuron B51,triggered by light receptors, modulates bucal-motor CPG generating swallowing
  - STRUCTURE: S_{light} -> B51 -> CPG
  - BEHAVIOUR: light-swallowing SM coupling.

Example 2: Aplysia

- Activity of neuron B51,triggered by light receptors, modulates bucal-motor CPG generating swallowing
  - STRUCTURE: S_{light} -> B51 -> CPG
  - BEHAVIOUR: light-swallowing SM coupling.
  - STRUCTURAL STABILITY CONDITION: $S_{esofageal}$ -> B51
  - EXPECTATION: light-food correlation
- Structural stability depends on satisfaction of expectations



In-formational dynamic organization

- Webs of dependencies and transitions can be created between dynamic structures generating an “internal world”
- Affordances: new environmental conditions are “shaped” as possibilities for actions (as a regions of the dynamic structure web)
- Goals: stability condition can be understood as goal states
- Developmental autonomy: the sub-determination of neural dynamics is progressively constrained by the structures stabilized, first through body value signals and then by the already existing dependencies

In-formational dynamic organization

- The gradient towards the cognitive is given by
  1. the time span of the expectancies,
  2. reduction of local-context dependencies and
  3. the complexity of the internal (and external?) web of dynamic dependencies

V. NATURALIZING COGNITION

(RECAPITULATION)

Back to the question

- From the set of all possible dynamical systems what kind of criteria can we offer to distinguish the cognitive from the non-cognitive ones?
- How can we answer the question with what we have seen so far?
- I propose 4 main criteria for naturalizing cognition

4 criteria for naturalizing cognition (I)

1. HIERARCHICAL DECOUPLING (neural dynamics not interefered by local metabolic dynamics) provides domain specificity
2. BIOLOGICAL EMBODIMENT (physical-interactive + metabolic) provides enabling constraints and basic (adaptive) functional feedback

4 criteria for naturalizing cognition (II)

3. AUTONOMY provides identity:
   - integrity through recursivity and functional integration
   - differentiation from environmental dynamics (agency) through complexity asymmetry
4. IN-FORMATIONAL DYNAMIC ORGANIZATION provides dynamic specificity

A naturalized definition of cognition

- Cognition is:
  - a dynamic behaviour
  - generated by an autonomous (holistic, integrated and recurrent) dynamical domain (the NS)
  - in-formationally organized and
  - hierachically decoupled but embodied and situated in its material conditions of possibility

VI. POSTSCRIPT

ON EVOLUTIONARY ROBOTICS
AS A THEORETICAL TOOL
(with suggestions, projects and examples)

Consequences of the 4 criteria for ER (and AI in general)

- Hierarchical decoupling (domain specificity):
  - Hierarchical decoupling justifies the level of abstraction of ER (the modelling of cognition does not require the modelling of all metabolic and anatomical details)
- Biological embodiment:
  - Past-emphasis: brain-body coevolution, embodied dynamics, control of perception, etc.
  - Special lack of metabolic embodiment in current ER models,
  - Metabolism-->Brain interaction providing functional feed-back (far-from-equilibrium and system accessible fitness functions)
  - Brain-->Metabolism interaction control of functional homeostatis (e.g. control of energy rate into motors)

Consequences of the 4 criteria for ER (and AI in general)

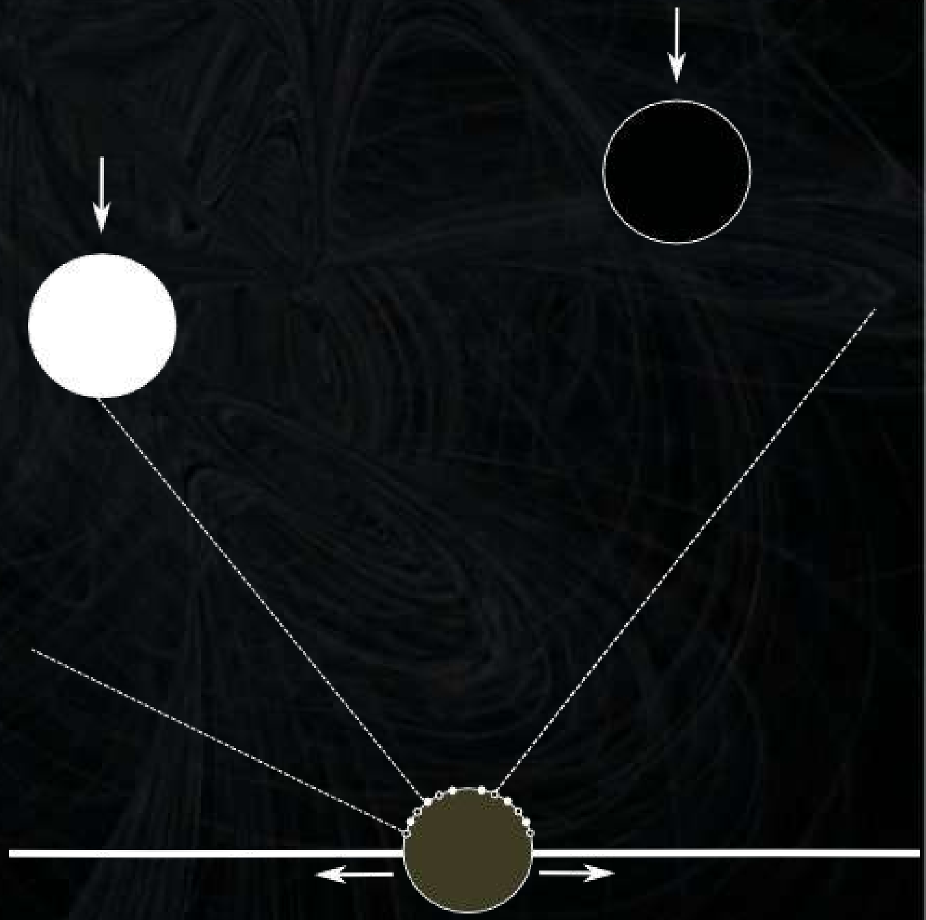
- Autonomy (integrity and differentiation):
  - Synthetically:
    - autonomy is achieved through CTRNNs (recurrency, functional integration, etc.)
    - high environmental variability will force behavioural decoupling from particular agent-environment relationships increasing autonomy
  - Analytically: we could start quantitatively analyzing autonomy with complexity measures
    - recent work by Seth & Edelman (2004) provide interesting analytical tools.

Consequences of the 4 criteria for ER (and AI in general)

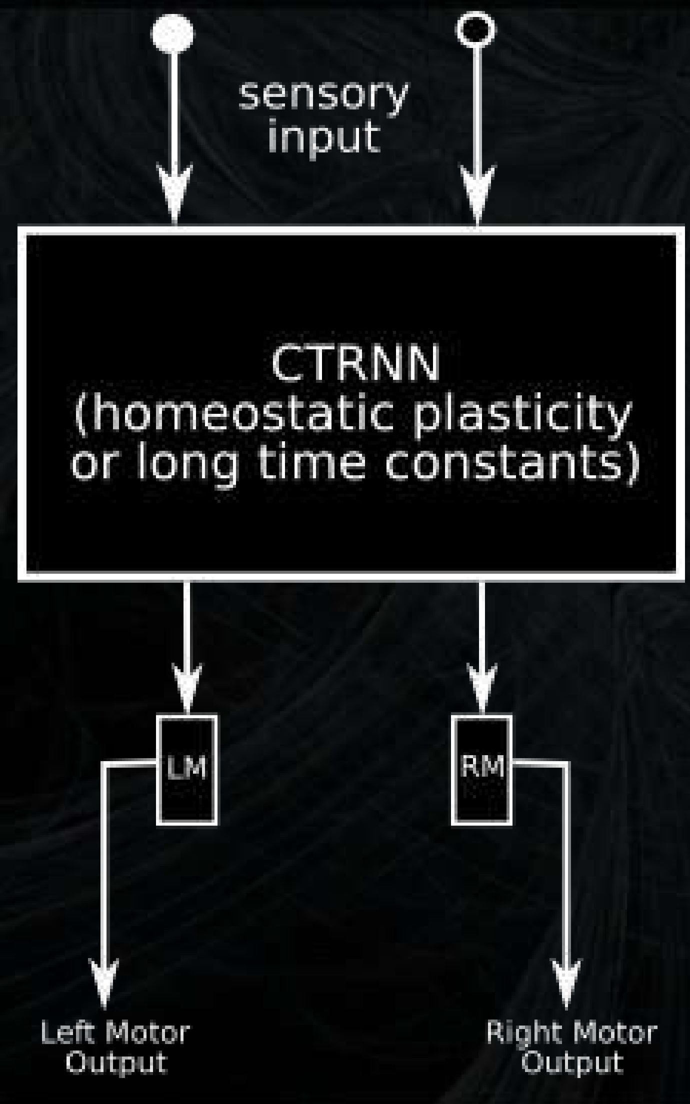
- In-formational dynamic organization:
  - Synthetically: behaviour coupled with internal stability conditions, this can be achieved in several ways:
    - metabolic embodiment is one of them,
    - homeostatic plasticity is another one
  - Analytically: intermediate explanatory patterns in the system relating dynamic structures with behaviour
    - McGregor & Fernando's definition of hyperdescriptions might be useful here

Experimental Design (I): TASC

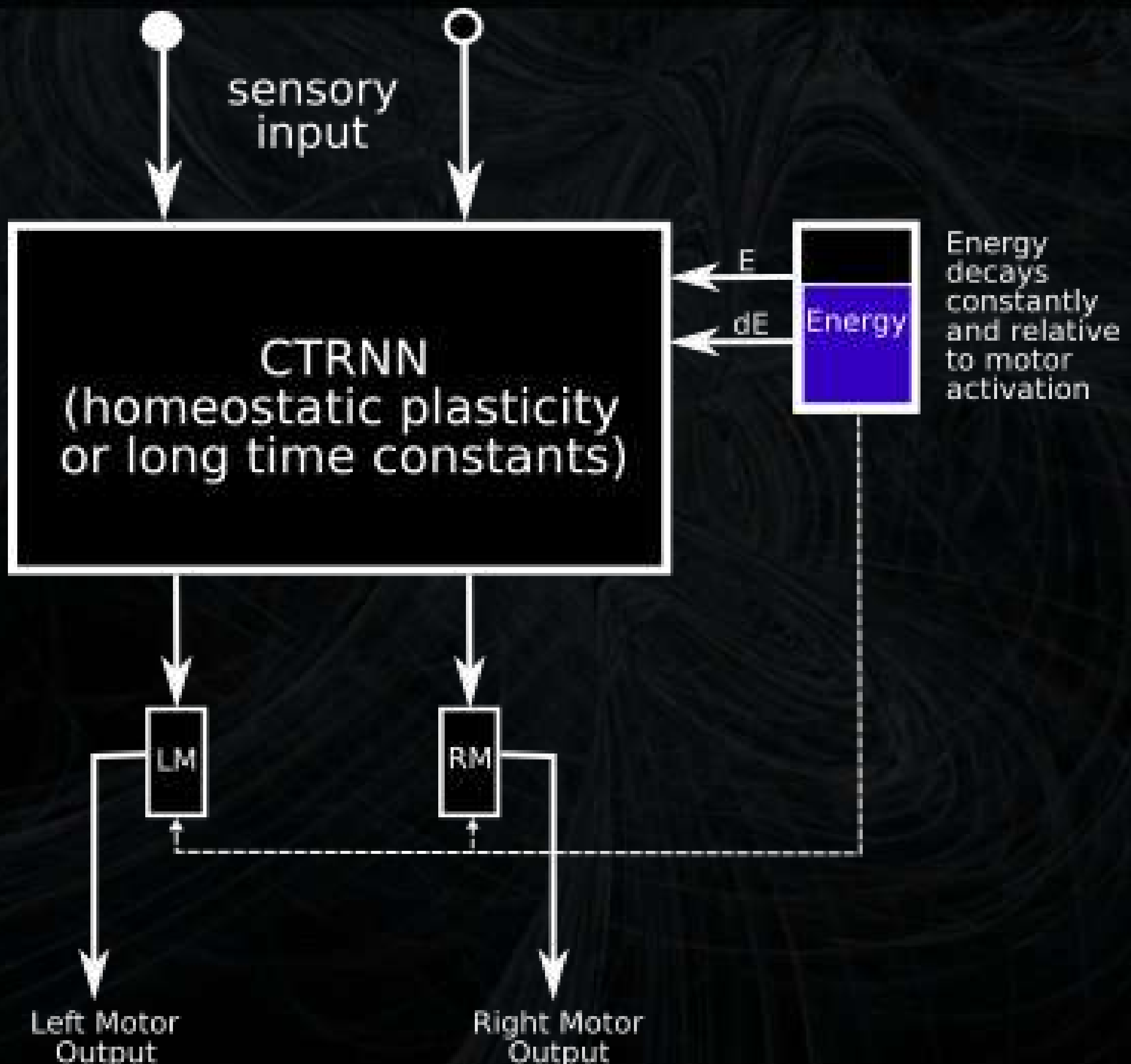
- Two food sources
- Different food profitability
- Agent "eats" food
- Energy based fitness function



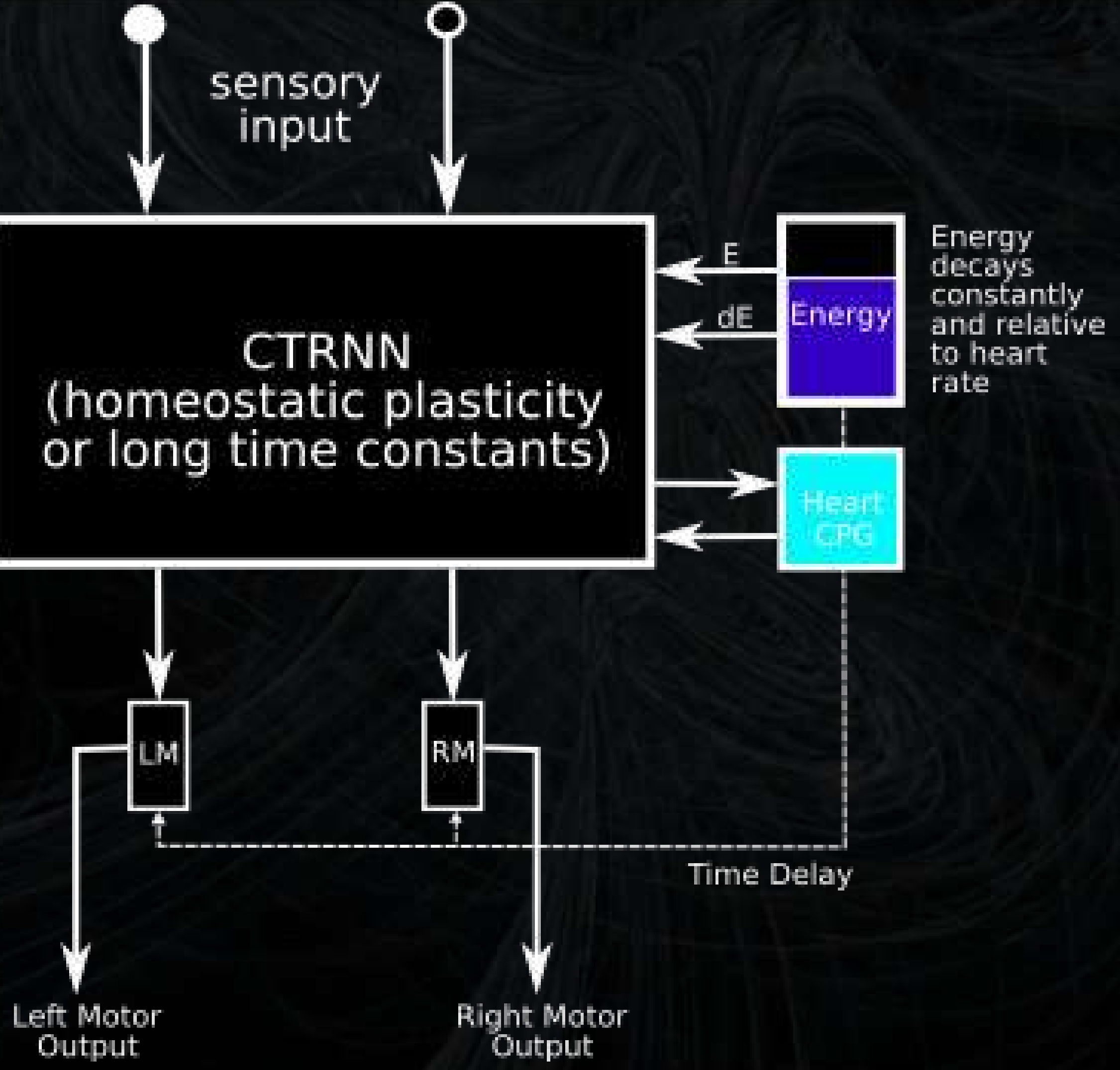
Experimental Design (II): CONTROL ARCHITECTURE



Experimental Design (II): CONTROL ARCHITECTURE



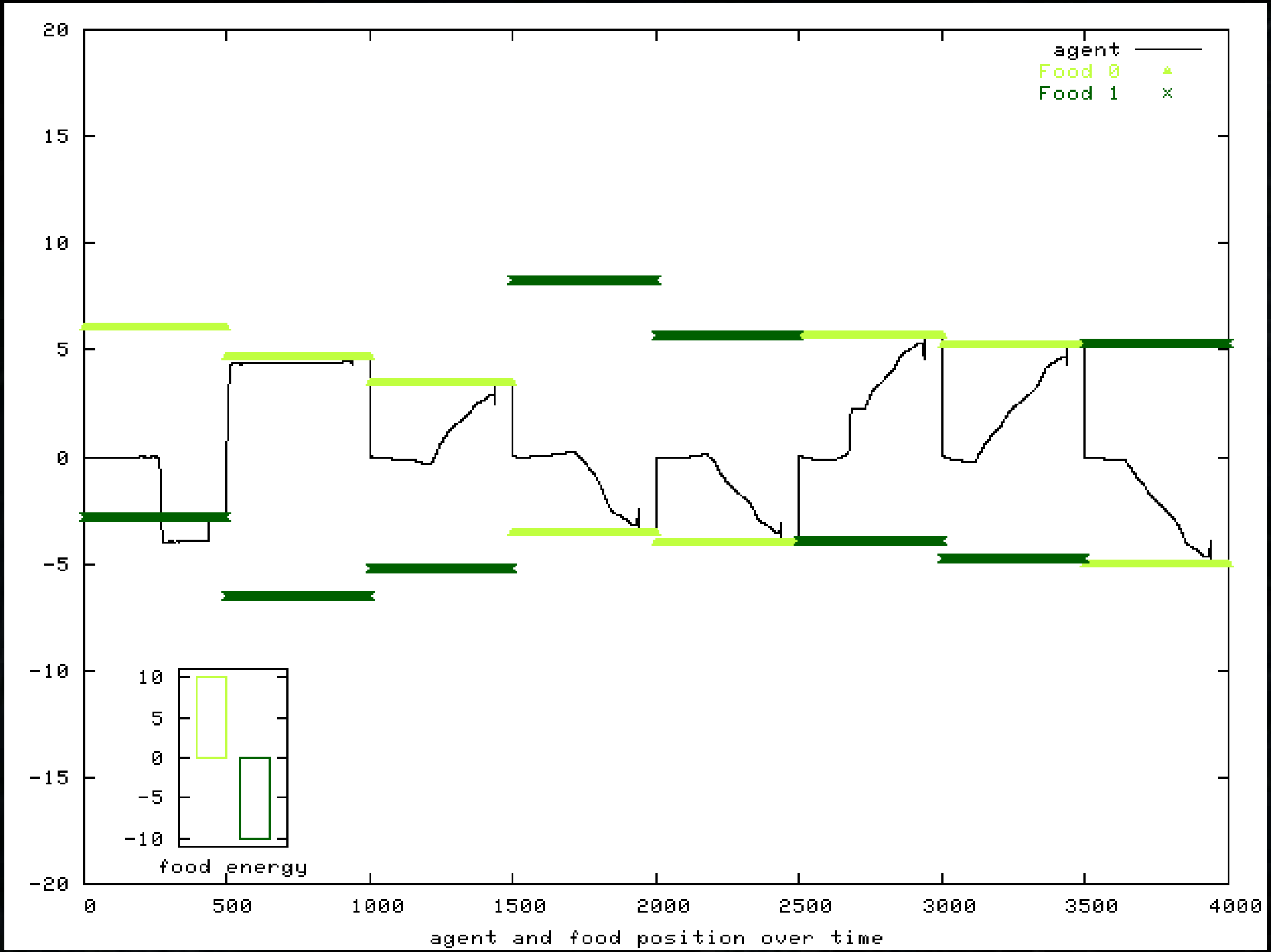
Experimental Design (II): CONTROL ARCHITECTURE



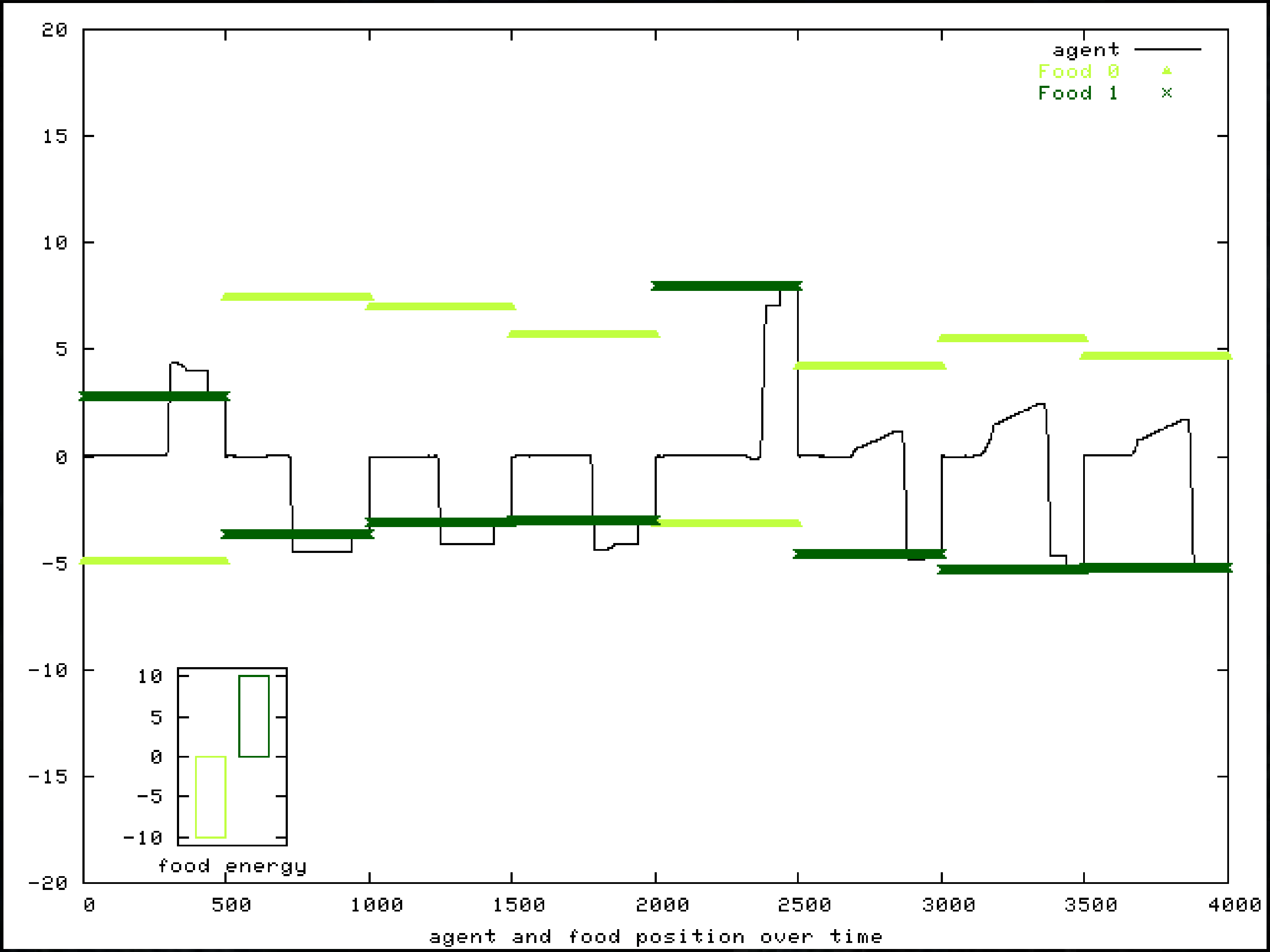
Results so far

- Risk aversion
- Behaviour energy-stability matching
- Learning with TC
- Learning with synaptic plasticity

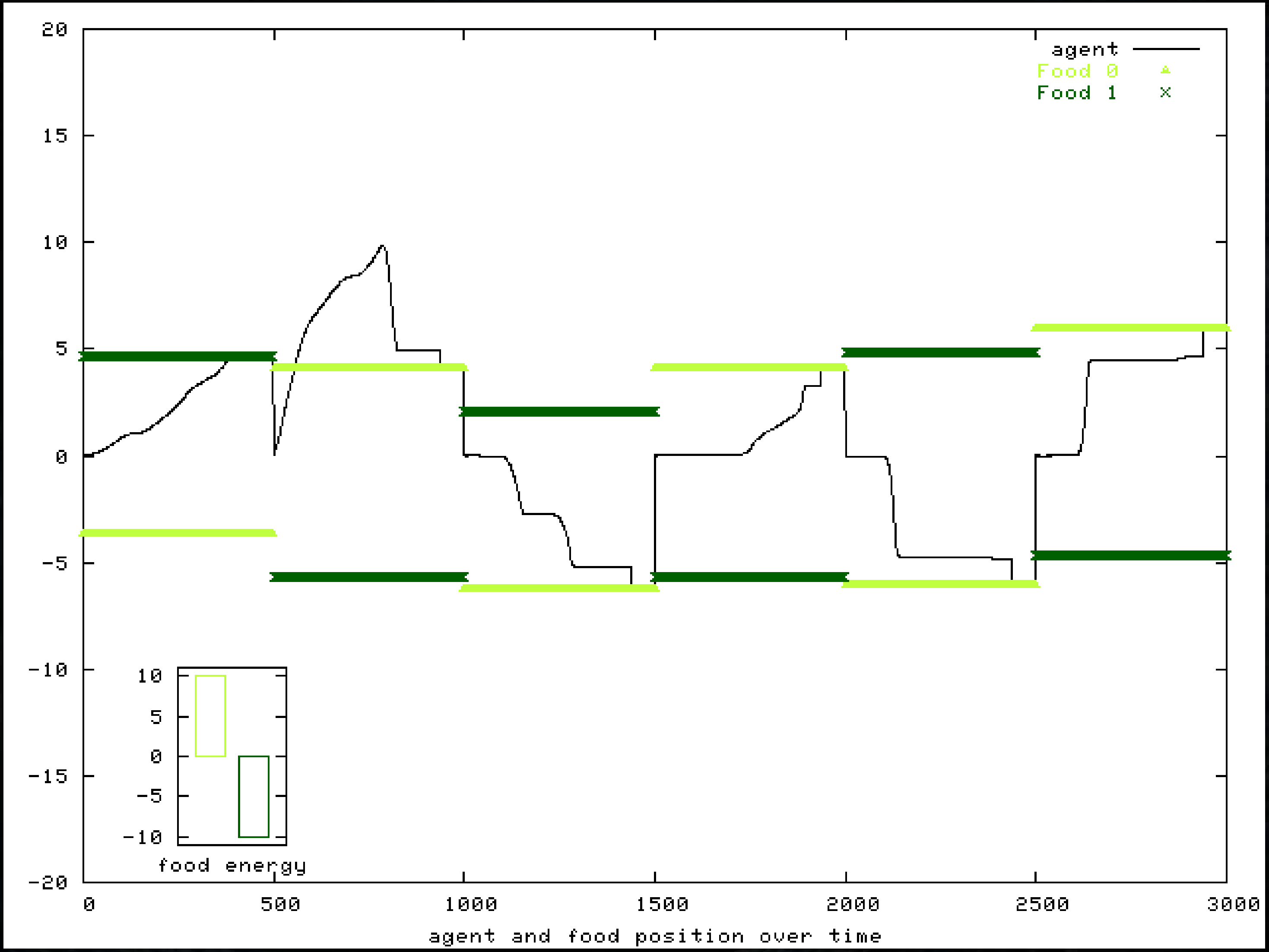
Learning with time constants (condition 0)



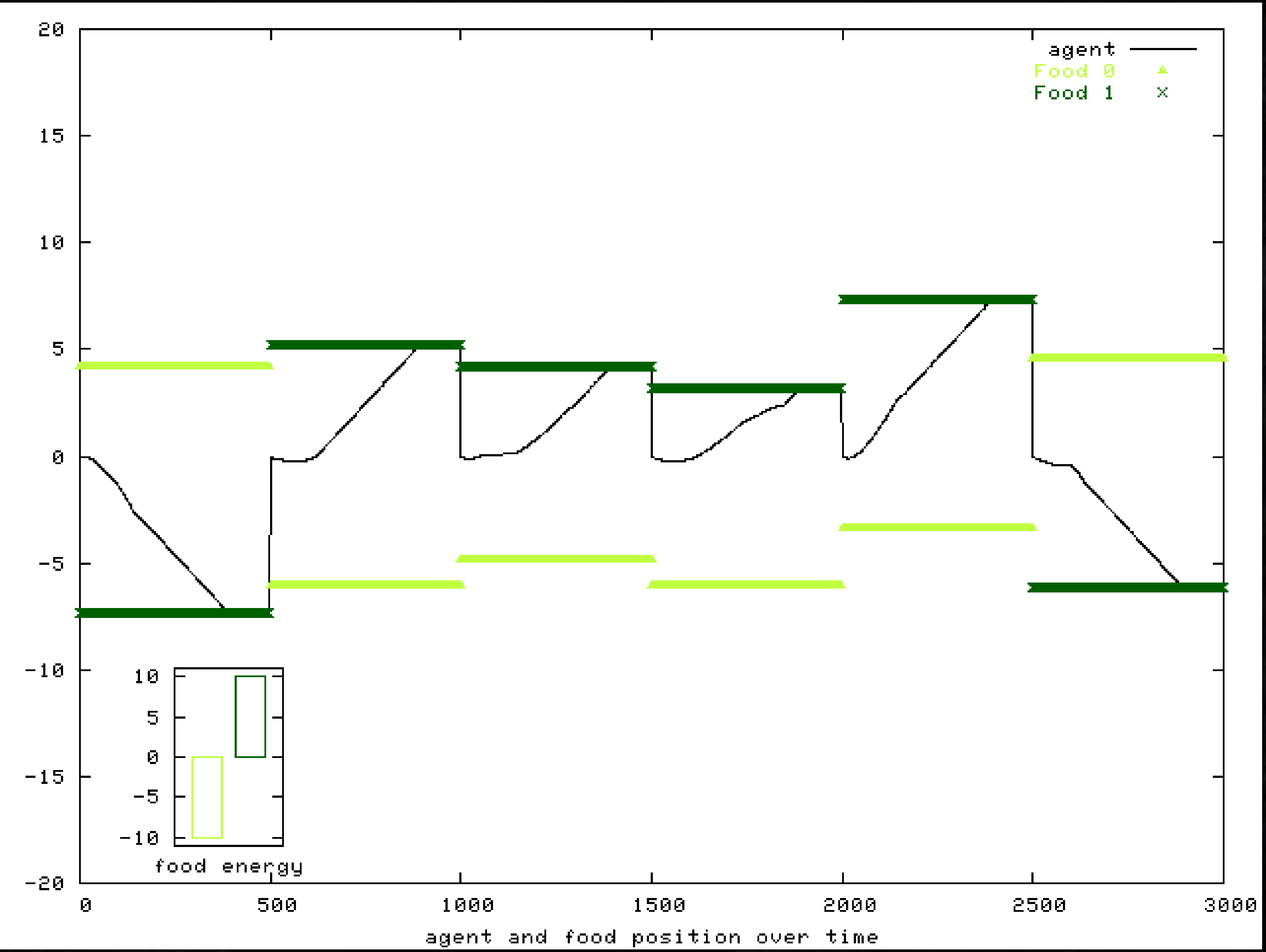
Learning with time constants (condition 1)



Learning with synaptic plasticity (condition 0)



Learning with synaptic plasticity (condition 1)



THANK YOU !!!

(so... *are plants cognitive?*)