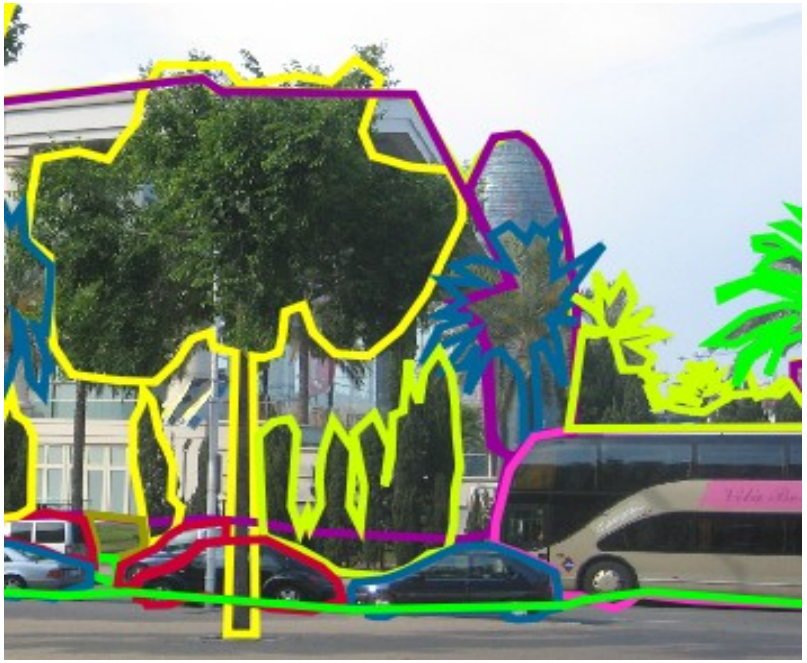
Computer Vision :
a Plea for a Constructivist View

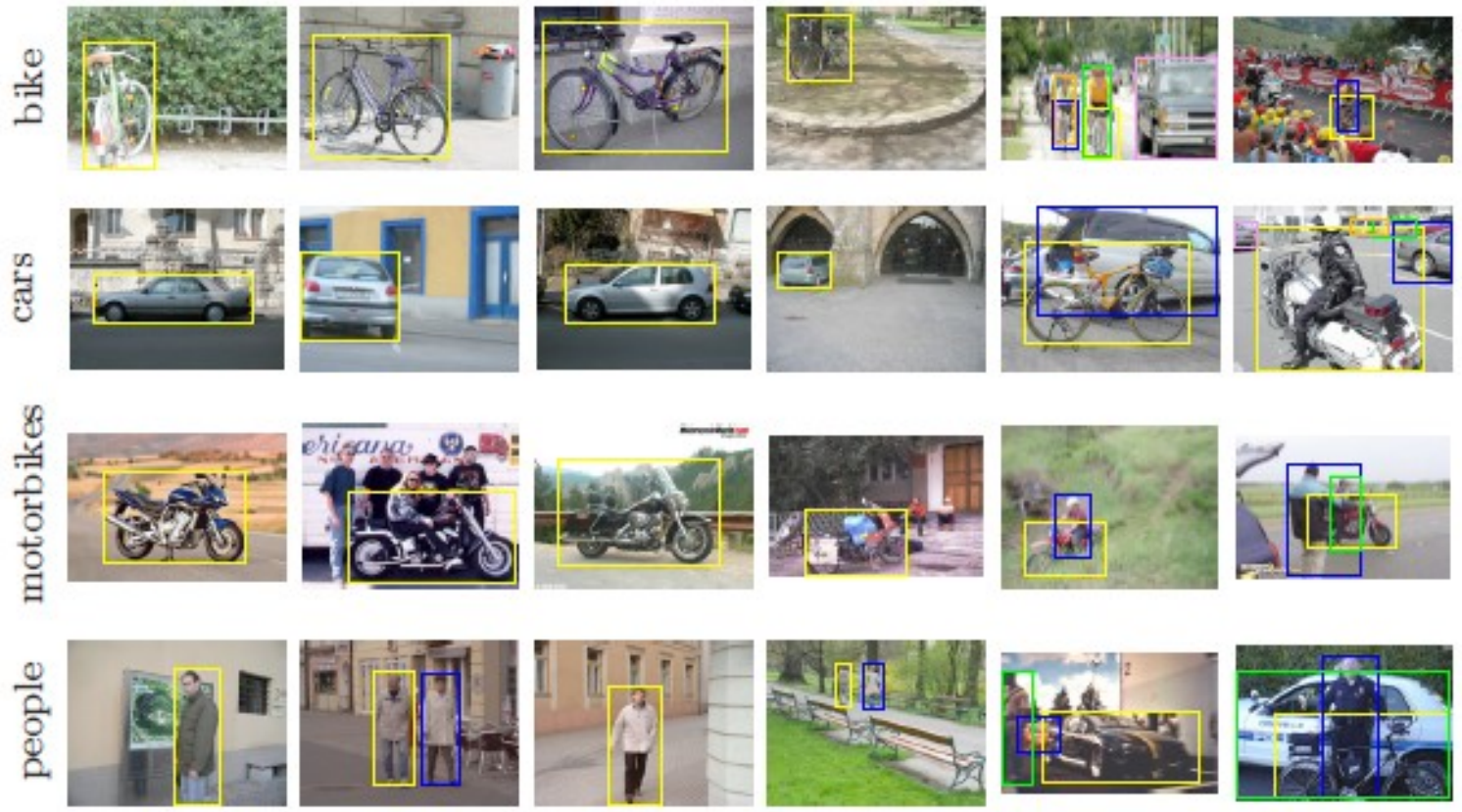
Computer vision in brief

- An ambitious goal
  - sense, process and interpret images of the outside world by means of automatic or semi-automatic means
- A variety of objectives
  - Improve the readability, enhance image quality
  - Allow fast access through natural queries
  - Extract characteristics, interest points, pattern
  - Delineate / detect / check the presence of objects, track a moving target
  - Identify a person, a monument, a situation
  - …
- Several steps and levels
  - From image sensing to high-level image interpretation, through low-level (pre)processing, 3d registration, color, texture or motion analysis, pattern recognition, classification…



http://labelme.csail.mit.edu/guidelines.html

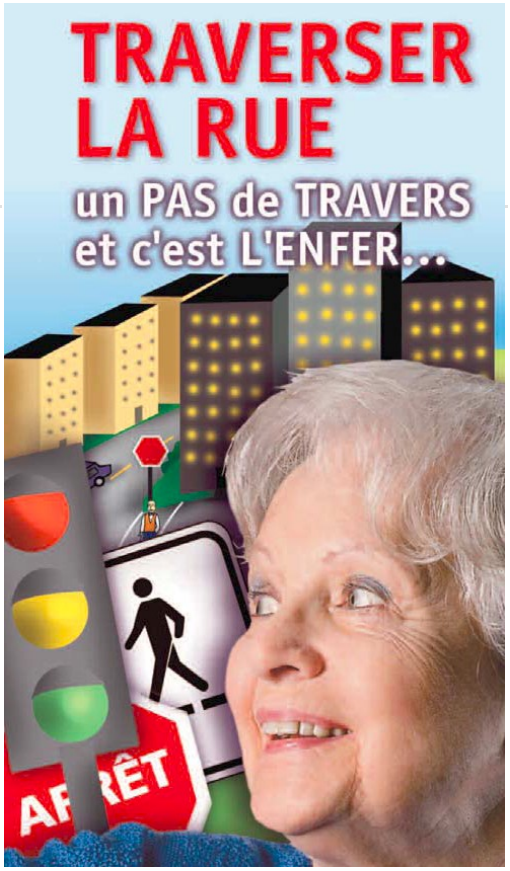
A challen-ging field of research



Dataset Issues in Object Recognition, J. Ponce et al, 2006

A stimulating relation to AI

- Bridging the gap between sensing and understanding :
  - From « neuroscience is cognition » (JP Changeux)
  - To the « embodied » intelligence (Varela)
- Viewing intelligence under its dual capacity of opening and closure
  - The brain does not « explain » intelligence
  - Intelligence does not « reduce » to solving equations but rather lies in the capacity to establish transactions with the external world
- Questionning rationality and truth
  - Vision : not a representation but a mediation to reality
  - There is no complete and consistent description of the world, even with a heavy cost
  - there is no « truth » of the world, and a rational behaviour has nothing to do with truth
- Questionning the notion of representation
  - Toward « valuable » or « true » representations?
  - The value of a representation is to neglect what is not pertinent and focus on what is related to the situation at hand.
  - *(Daniel Kayser, conf IAF, 2009)*



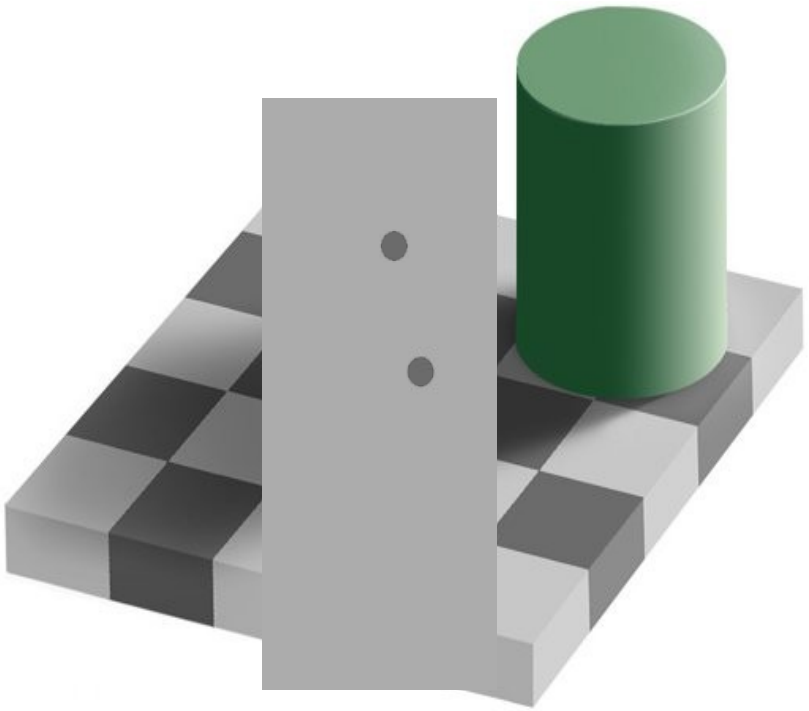
Marvin Minsky (80's) : « how can you cross a road and prove that it is secure? »

A stimulating relation to AI

- *"Whilst part of what we perceive comes through our senses from the object before us, another part (and it may be the larger part) always comes out of our own mind."*

  - W. James

- Visual illusions : not errors to avoid, nor heuristics to reproduce, but the illustration of the complexity of vision
- Vision : an ability to maintain a « viable » understanding of the world under various contexts



« Voir le monde comme je suis, non comme il est » Paul Eluard

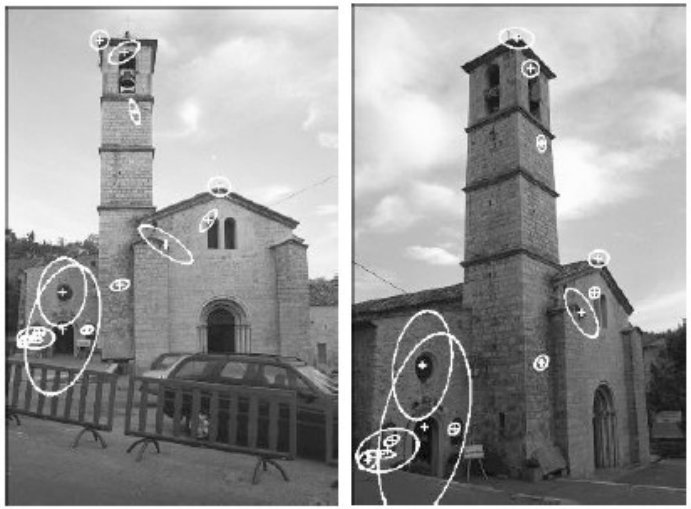
D. J. Simons 2003 - Surprising studies of visual awareness - Visual Cog Lab - http://viscog.beckman.uiuc.edu/djs_lab/

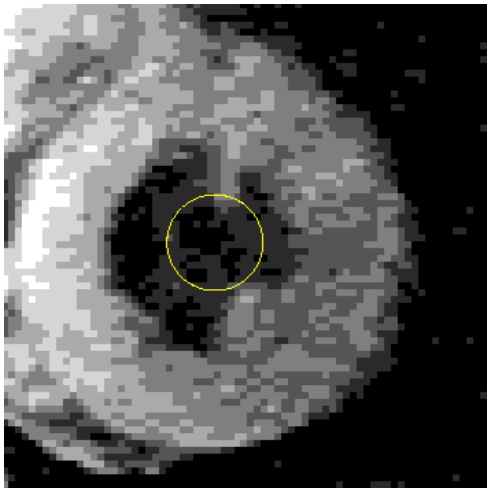
Two complementary views

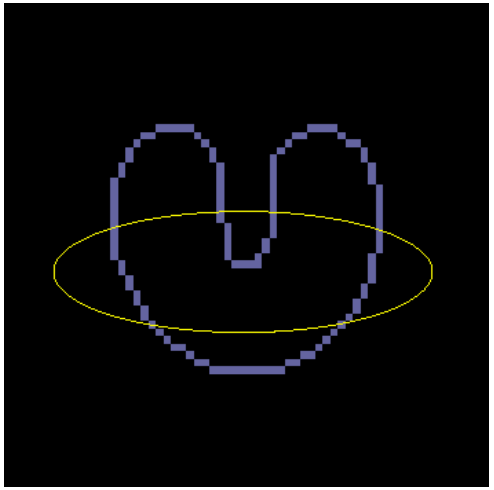
- A multidisciplinarity field of research
  - AI, robotics, signal processing, mathematical modelling, physics of image formation, perceptual and cognitive dimensions of human understanding
- A scientific domain at the crossroads of multiple influences, from mathematics to situated cognition.
- Mathematical view :
  - A positivist view, according to which vision is seen as an optimization problem.
  - A formal background under which vision is approached as a problem-solving task.
  - Rather well supported by joint work with neurophysiologist
- Constructivist view :
  - Vision as the opportunistic exploration of a realm of data, as a joint construction process, involving the mutual elaboration of goals, actions and descriptions.
  - Relies on recent trends in the field of distributed and situated cognition.

Positivism : capture variations

- Model distributions rather than means
  - Capture variations and variability rather than look for mean descriptions
  - Many difficult notions approached in extension rather than in intension
- Look for problem sensitive descriptors
  - Look for invariants (local appearance models, C. Schmid)
  - Model only the variations that are useful for the task at hand



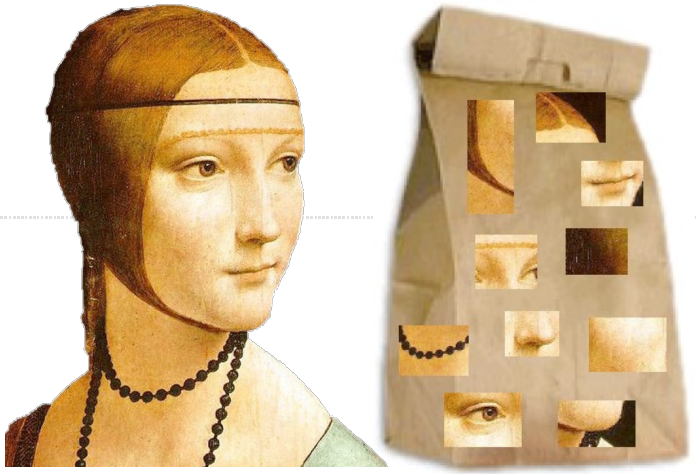




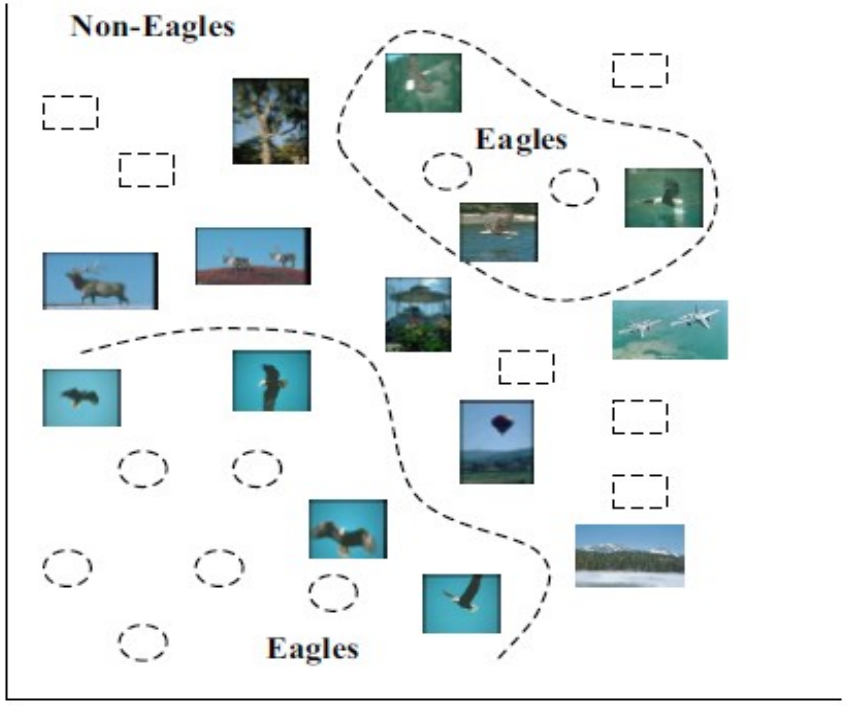
http://iacl.ece.jhu.edu/projects/gvf/heart.html

Positivism : deconstruct

- Minimize the a priori
  - minimize the *a priori* needed to recognize a scene
  - avoid the use of intuitive representations,
  - look closer to the realm of data and its internal consistency
- Deconstruct the notion of object / category
  - consider the object not as a “unity” nor as a “whole” but as a combination of patches or singular points ;
  - do not consider a concept as a being or an essence, but through its marginal elements
  - SVM classification methods



L. Fei-Fei et al. ICCV 2005 short course

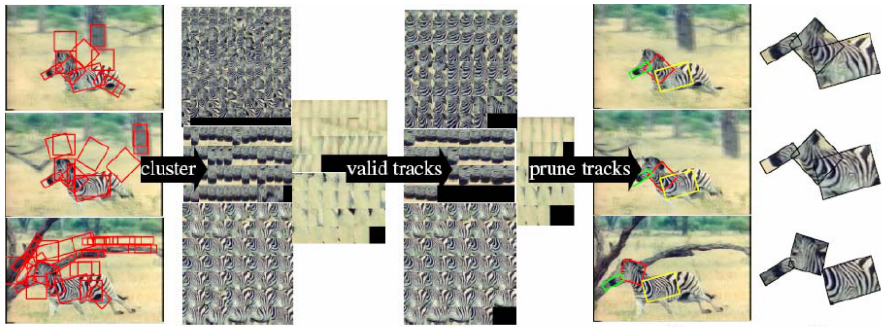


L. Zhang, F. Lin, ICIP01

Positivism : Integrate

- Integrate, model joint dependencies
  - Integrate into complex functionals heterogeneous information from different abstraction level/viewpoint
  - Model in a joint way the existence, appearance, relative position, and scale
  - Preserve contextual information

R. Fergus, ICCV 2005



Using Temporal Coherence to Build Models of Animals, D. Ramanan et al. ICCV2003

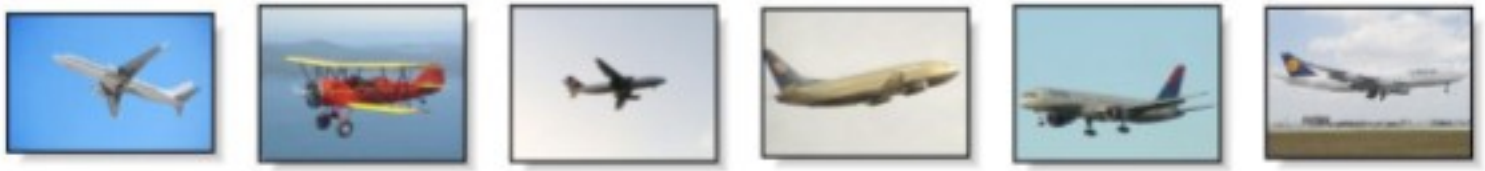
Multi-object Tracking Based on a Modular Knowledge Hierarchy - M. Spengler et al. ICVS 2003

Positivism in brief

- A focus on formal aspects, on dimensionality and scaling issues…
- A focus on how to capture variations of appearance,
- not on how to model the process of interpretation

- What has been lost in between ?

TREC Video Retrieval Evaluation - http://www-nlpir.nist.gov/projects/**trecvid**/



Pascal VOC Challenge - http://pascallin.ecs.soton.ac.uk/challenges/VOC/

Vision : what is it all about, lets try again

- **Organize affordances**
  - Interior of a room with a group of people
  - A composition involving several planes, from the back to the front
  - The viewer's eyes sees the man immediately
- **Suggest a style**
  - A construction suggestive of Degas
- **Arouse feelings**
  - Different facial expressions, captured dramatically
  - A picture full of light, a mixture between seriousness, anxiety and a feeling of joy
- **Tell a story**
  - A family surprised by an unexpected return of a political exile home
- *Il'ia Efimovich Repin: They Did Not Expect Him (1884-88)*

Not only an optimization task… but a situated activity

- [Yarbus 67]

1. No question asked ;
2. Judge economic status ;
3. Give the ages of the people
4. What were they doing before the visitor arrived ?
5. What clothes are they wearing ?
6. Remember the position of people and objects ;
7. How long is it since the visitor has seen the family ?

Images as an open universe

- The universe of images is contextually incomplete [Santini 2002] :
  - taken in isolation, images have no assertive value but rely on some external context to predicate their content.
  - A pure repository of images, disconnected from any kind of external discourse, doesn't have any meaning that can be searched, unless :
    - it is a priori inserted in restricted a domain (eg medicine)
    - It is explicitly linked to an external discourse, an intended message (eg multimedia documents)
  - The observer will endow images with meaning, depending on the particular circumstances of its observation or query.
- *« A text is an open universe where the interpret may discover an infinite range of connexions… a complex inferential mechanism »*
- *U. Ecco, The limits of interpretation, 1990*

Images as an outcome

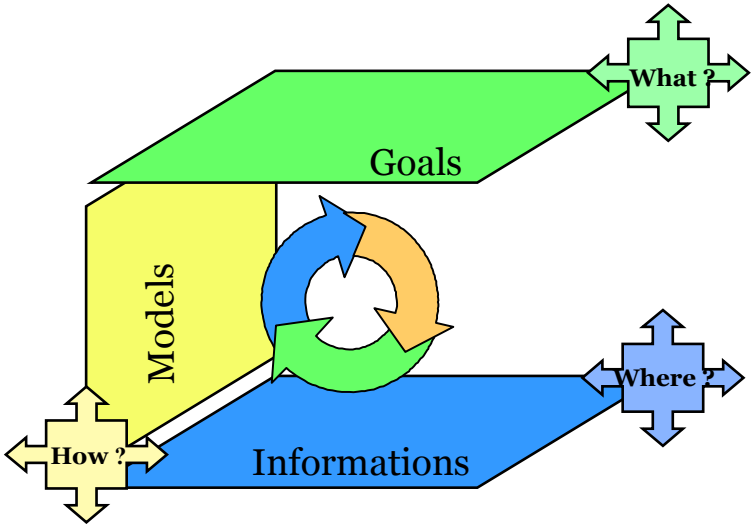
- Vision : an exploration activity
  - oriented toward the search for objects, the gathering of information, the acquisition of knowledge
- A situated process
  - A process that is context-sensitive
  - A process embodied in the action of a subject, guided by an intention, on an environment
- A constructive activity,
  - A process which do not obey any external predefined goal
  - Rather a process according to which past perceptions give rise to new intentions driving further perceptions
  - A process which operates transformations which modify the way we perceive our environment
- *Images : not a data, but a dynamical answer to a questionning process (from J. Bertin)*

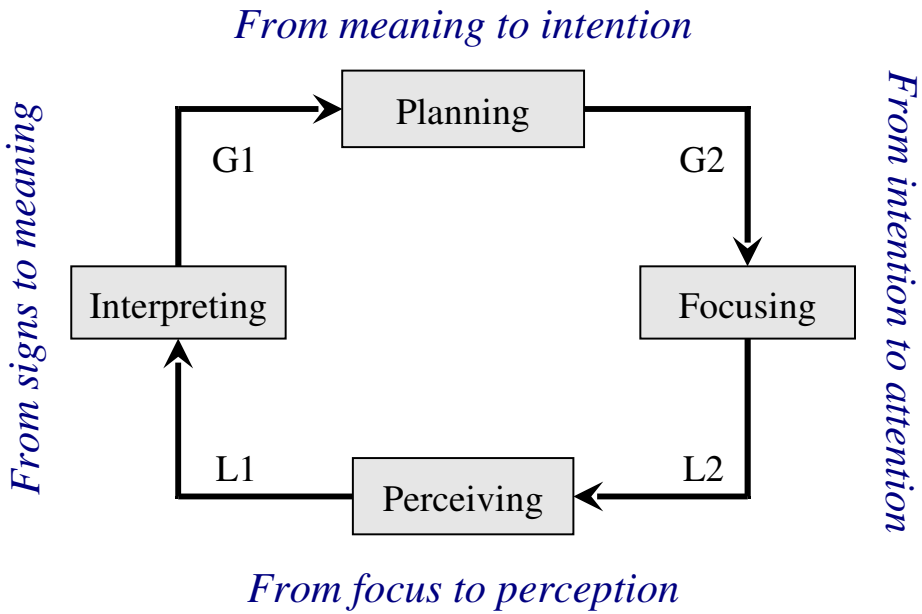
Images as a map for action

- For Bergson, there is no « pure » perception
- The human captures from objects only what appears of some « practical » interest : perception is guided primarily by the necessity of action
- Perceiving an object indicates the plan of a possible action on that object much more than it provides indications on the object itself
- Contours that we see in objects denote simply what we may reach, manipulate or modify, like ways or crossroads through which we are meant to move
- Geometrical figure recognition and memorization
  - close links between haptic exploration and vision (L. Pinet & E. Gentaz, LPNC Grenoble)

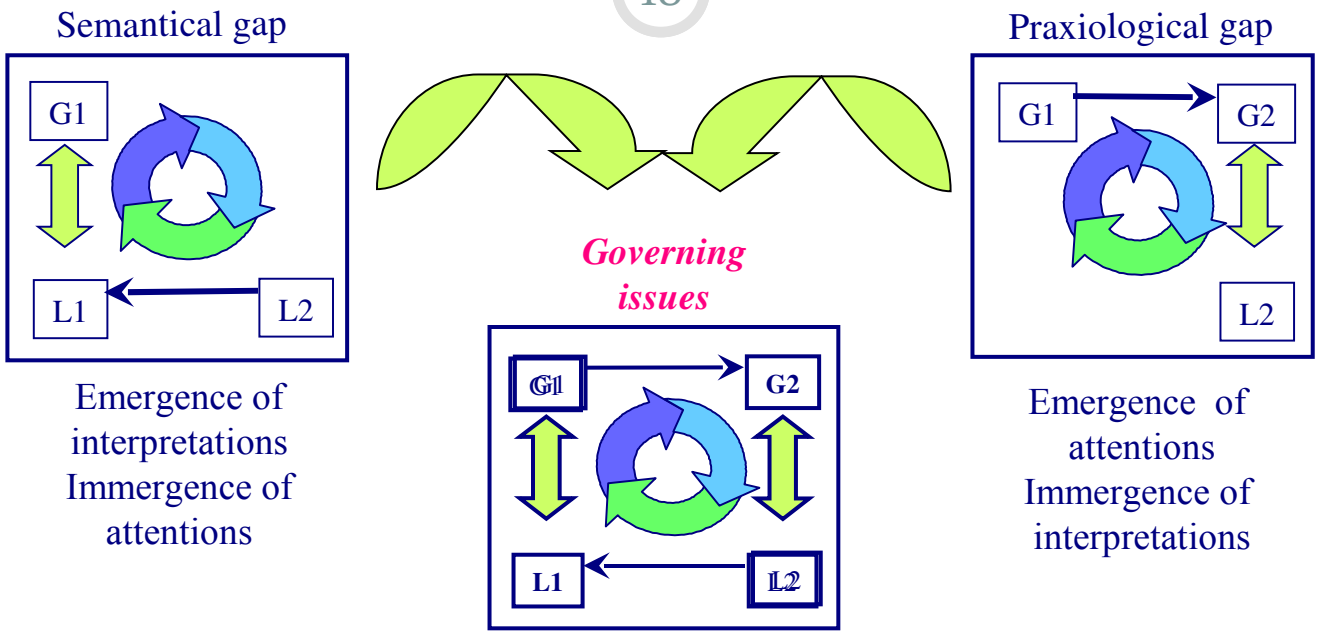
Vision : a viable coupling

- An explorative activity involving mutually dependent decisions about where to look at, what to look for, and what models to select
- Reaching a state in the decision space generates the ability to look forward
- *A process whose goal is not clearly stated in terms of a precise state to reach, but rather in terms of progressing as long as it is fruitful to do so (P. Bottoni et al., 1994)*
- *We do not just see, we look (R. Bacjsy, Active Perception, 1988)*





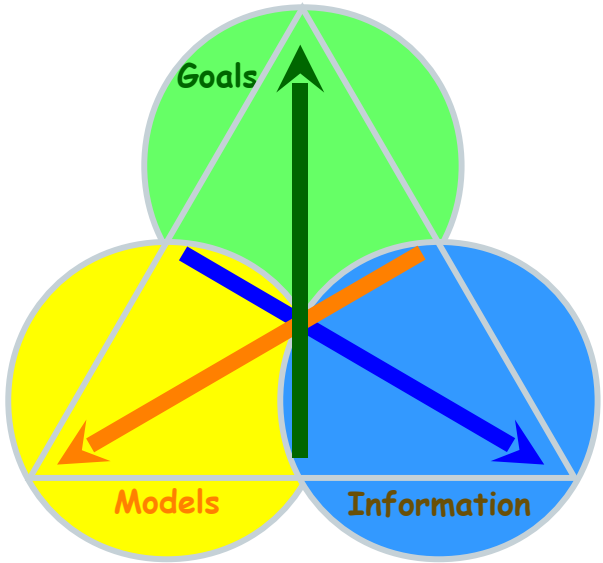
Vision : crossing gaps



- Semantic gap: how to build a global and consistent interpretation (G1) from local and inconsistent percepts (L1) acquired in the framework of given focus of attention (L2)
- Praxiological gap: how to derive local focus of attention and model selection (L2) from a global intention (G2) formulated as the result of the perceived scene understanding (G1)
- The ability to establish a viable coupling between an intentional dynamic, an attentional dynamic, and an external environment on which to act
- A constant interleaving of mutually dependent analyses occurring at different levels

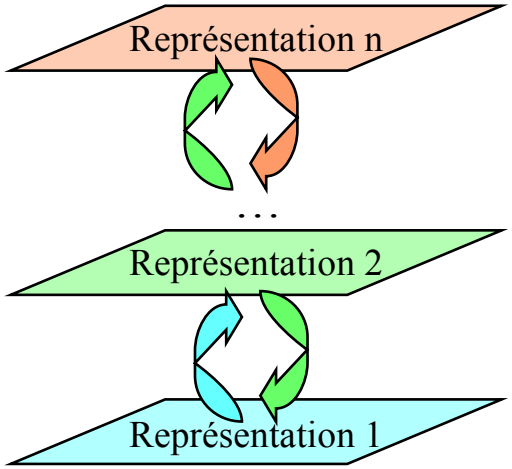
Vision : co-determination issues

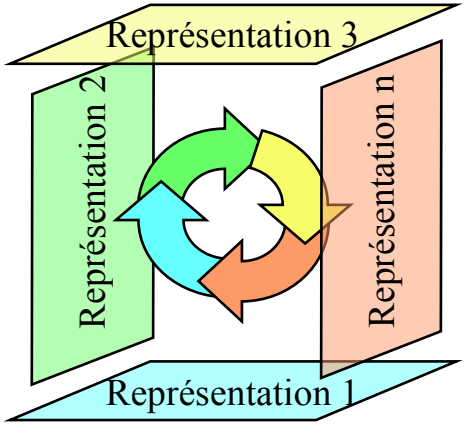
- Co-determination between goals, actions and situations :
  - I + M ➔ G
  - G + I ➔ M
  - G + M ➔ I
- A situation is built by an actor under some intention : it has no existence independently of this action
- An action may only be interpreted considering the data of the situation at hand and the possibilities for action : action exists only a posteriori
- There is no rationale for action that exists separately and independently from the action itself : a plan is a resource, not a prescription
- *The involvement in action creates circumstances that might not be predicted beforehand (Suchman, Plans and situated actions, 1987)*



Vision : back to the distribution issues

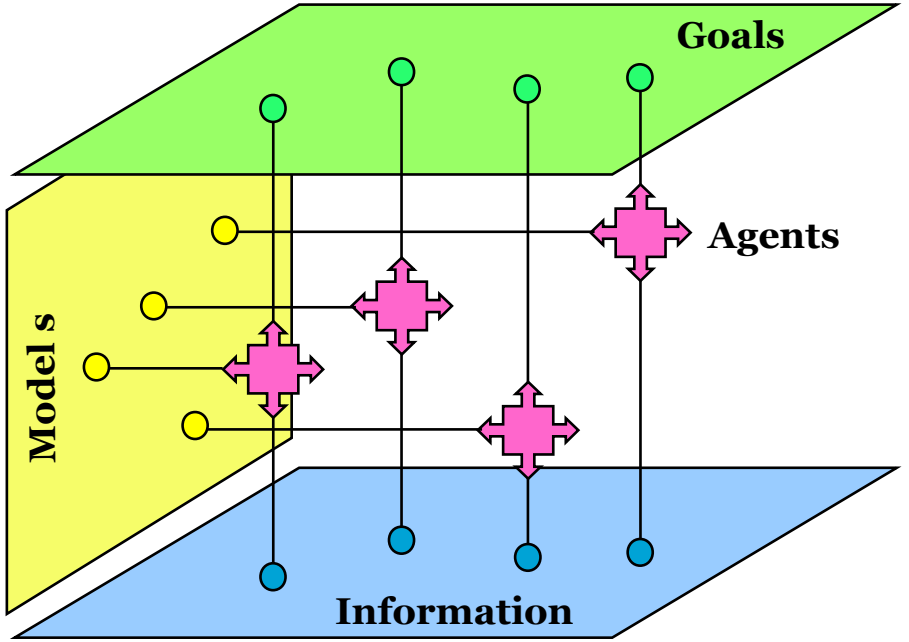
- Distribute
  - Decompose to break down the processings and cope with the semantical and praxiological gaps
  - Reduce the scope of processing, spatially and semantically
- Enrich
  - Make inferences more local, but based on richer descriptions
  - Work more slowly,but in a more robuts way : progress incrementally, in the framework of dynamically produced constraints
- Preserve the relations, cooperate
  - The principle is not to partition nor compartmentalize
  - There is no strict hierarchy in the kind of information that may be used at a given step, rather any information gained at any time, any place and any abstraction level may be used in cooperation
  - The richness of the process depends on its capacity to break down, confront, and combine information from various levels and viewpoints, providing a cooperative status to vision





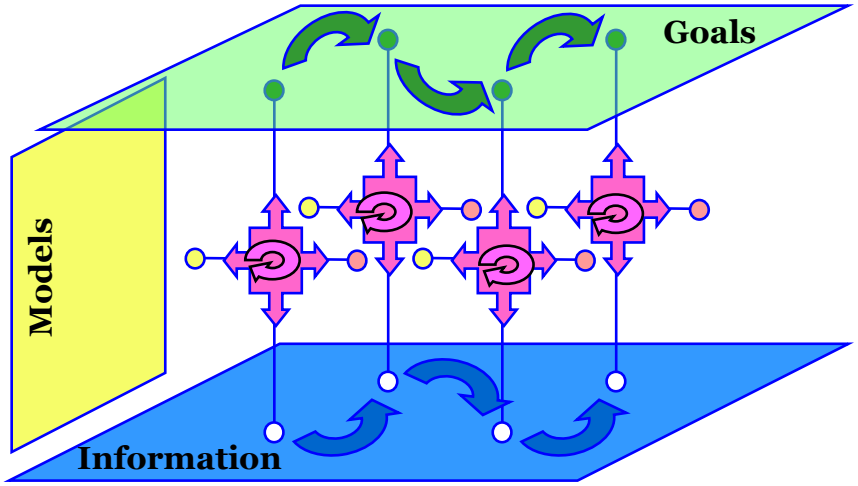
Situated agents : coupling (G, M, I)

- The agent A = f{G, M, I} is anchored
  - physically (at a given spatial or temporal location),
  - semantically (for a given goal or task) and
  - functionnally (with given models or competences) ;
- The environment E = {G, M, I} allows to share
  - Data, computed information and (partial) results
  - Models
  - Goals



Situated agents : a dual adaptation

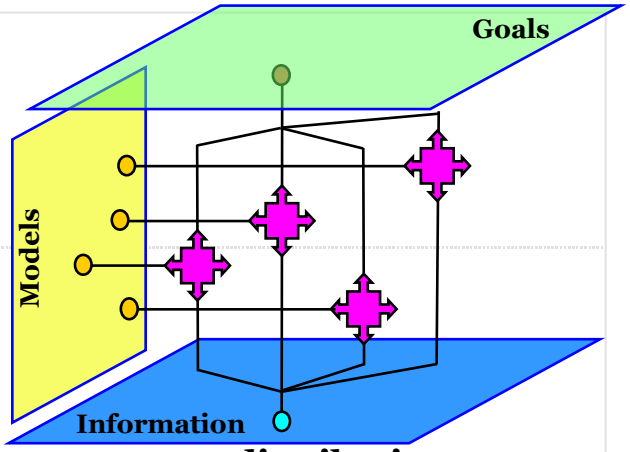
- Internal adaptation
  - Selection of adequate processing models, according to the situations to be faced and to the goals to be reached
  - Ai : Gi + Ii ➔ Mi
- External adaptation
  - Modification of the focus of attention : new situations or goals to explore
  - Creation of new agents, modifying as a consequence the organisation at the system level
  - Ai (Gi, Mi, Ii) ➔ Aj (Gj, Mj, Ij)
  - S. Giroux : Agents et systèmes, une nécessaire unité, PhD Thesis, 1993.



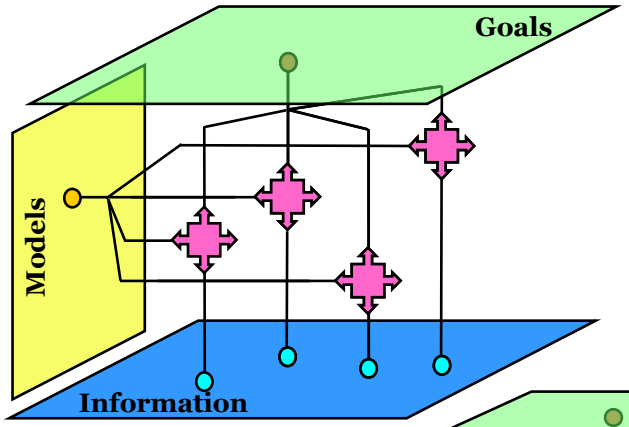
- As the system works, it :
  - completes its exploration, accumulates information, adapts and organizes according to the encoutered situations
  - A constructive approach according to which the system, its environment and goals co-evolve

Situated agents : cooperation issues

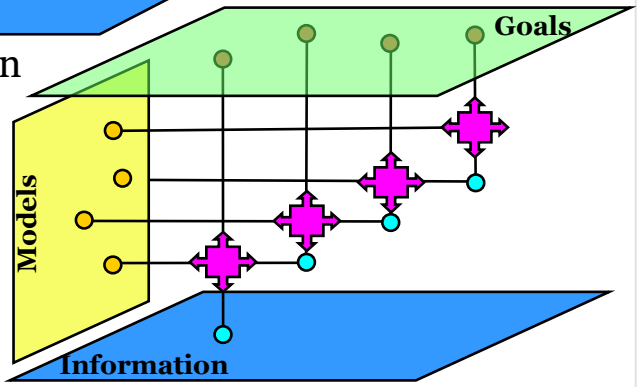
- Three cooperation styles
  - Confrontational : a task is performed by agents with competing competencies or viewpoints, operating on the same data set ; the result is obtained by fusion ;
  - Augmentative cooperation : a task is performed by agents with similar competencies or viewpoints, operating concurrently on disjoint subsets of data ; the result is obtained as a collection of partial results ;
  - Integrative cooperation : a task is decomposed into sub-tasks performed by agents operating in a coordinated way with complementary competences, ; the result is obtained upon execution completion
- J.M. Hoc, PUF, Grenoble, 1996



competence distribution



data distribution

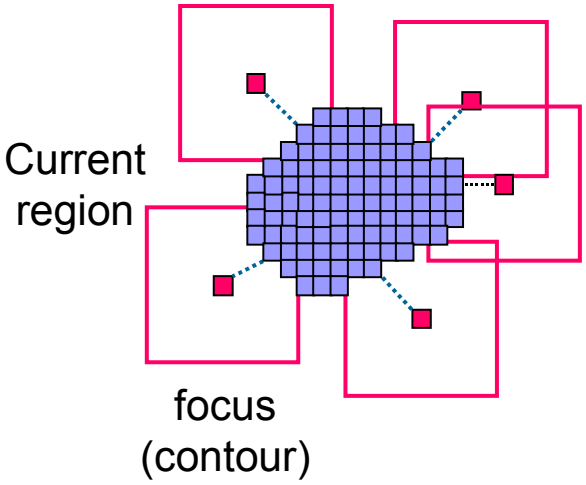


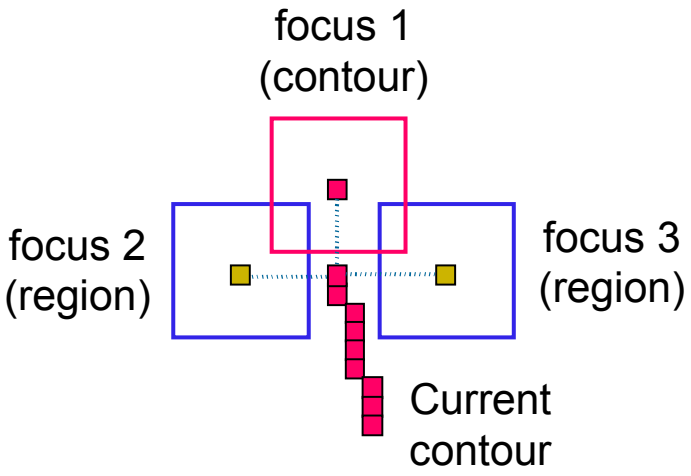
goal distribution

Two mutually dependent processes

- Two mutually dependent processes :
  - Contour following : triggered at successive steps of the region growing process ; limit their expansion
  - Region growing : triggered in case of failure of the contour following ; provide refined contextual information
  - Launching an agent expresses a lack for information
  - Each process works locally and incrementally, under dynamically and mutually elaborated constraints
- System level
  - The system of agent explores its environment in an opportunistic way
  - Under control on the system load, agent distribution (density) and agent time cycle

F. Bellet, PhD Thesis, 1998

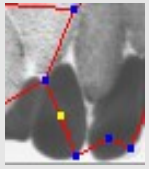




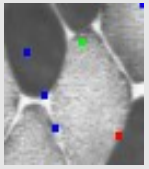
Two mutually dependent processes

- Successive focusings

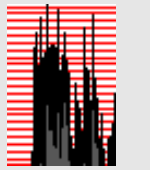
- Segmentation result



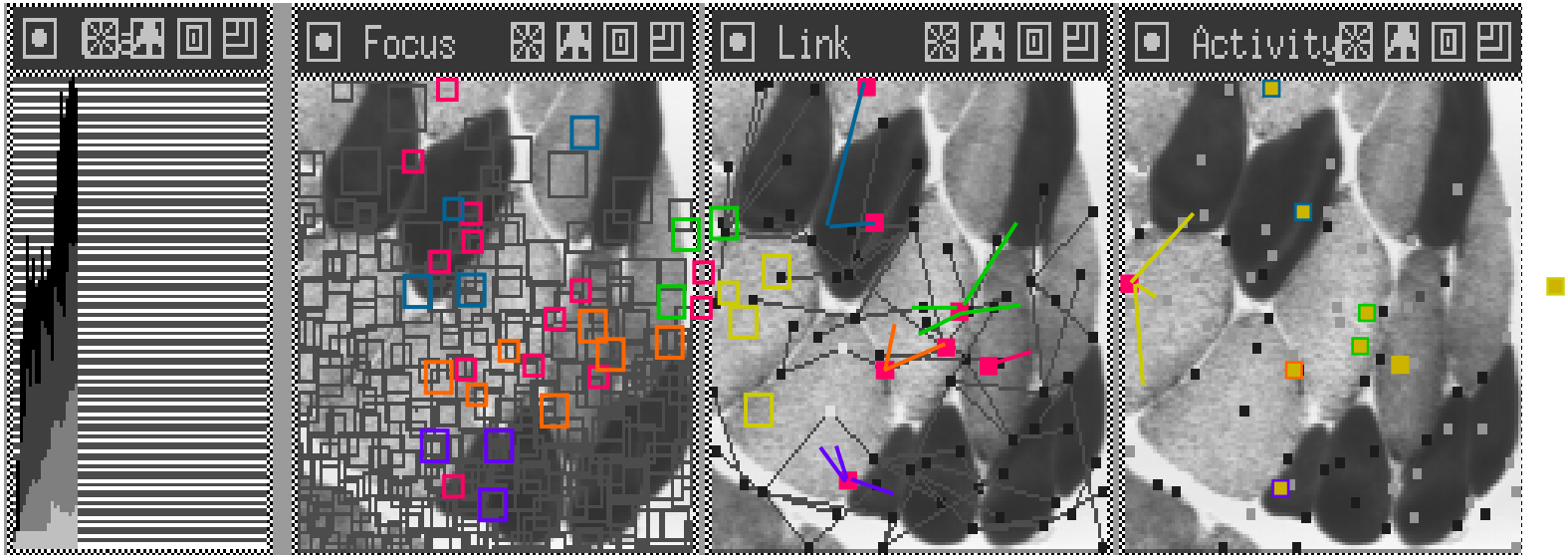
- Process linkage
  - seed process



- Process localization and state
  - executing
  - active
  - waiting

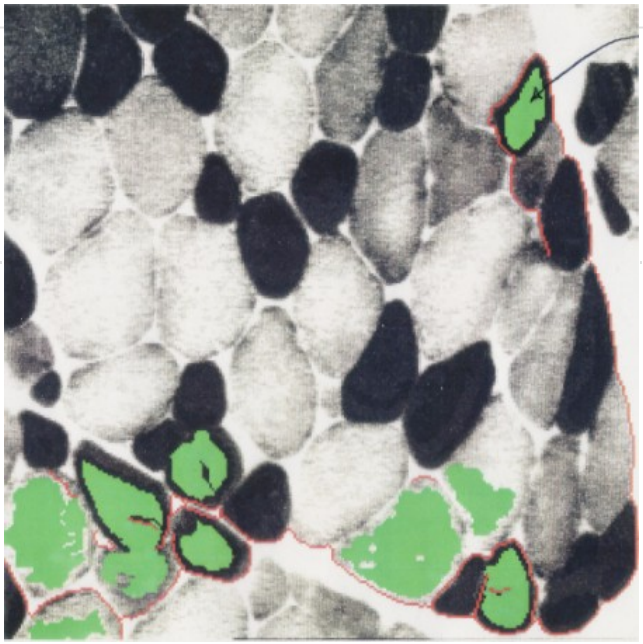


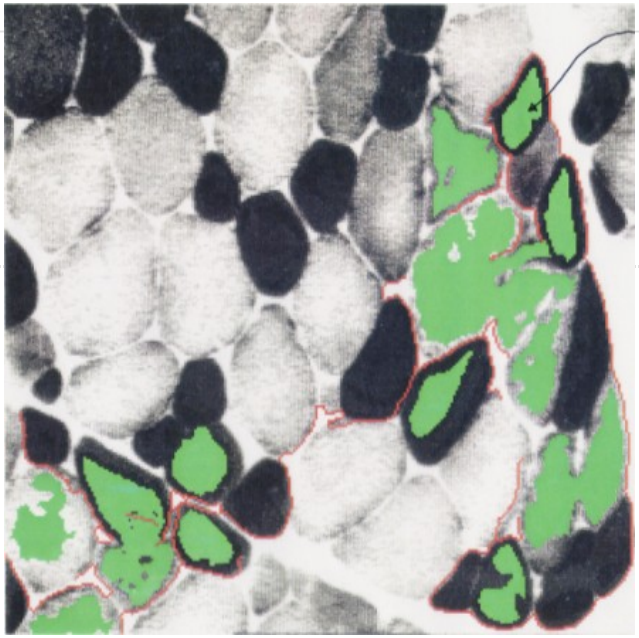
- System load

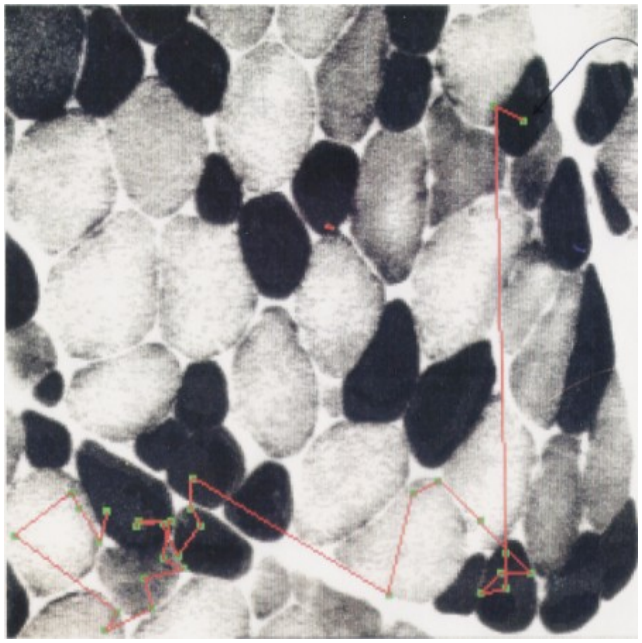


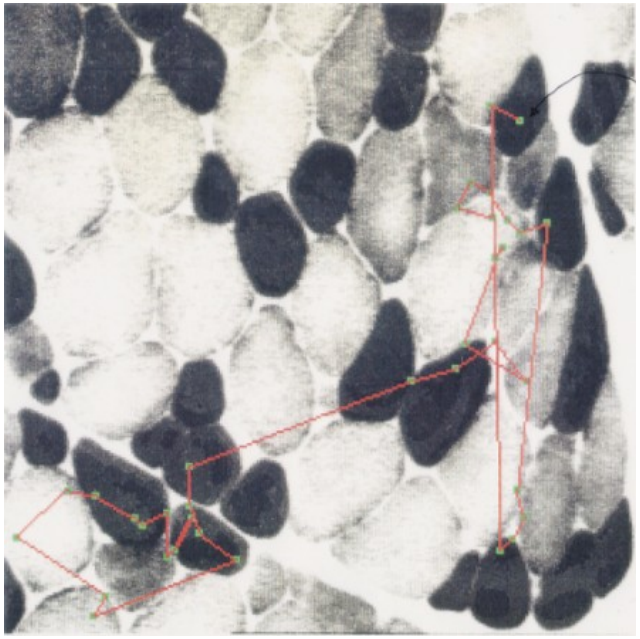
Two mutually dependent processes

- An Evolving Processing Structure
  - A coupling between :
  - A dynamically evolving processing structure ;
  - A dynamically evolving description of the initial image ;
- An Agent-Centered Design
  - A paradigm that steps back from classical procedural design ;
  - A processing approach where the time, content and partners of the interaction are not planned in advance ;
  - A problem solving approach where the solution is not sought in a global way ;

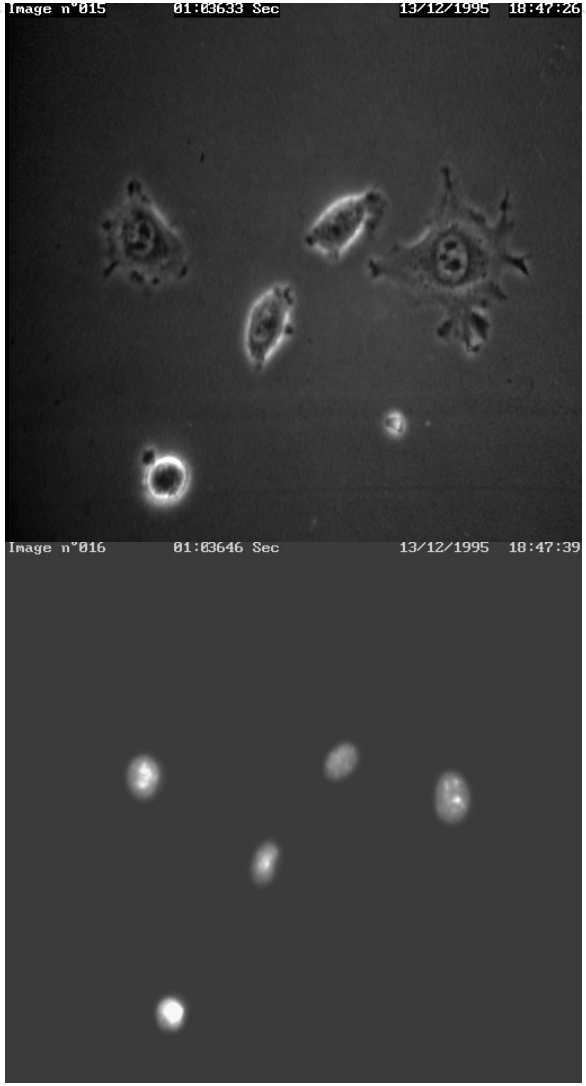


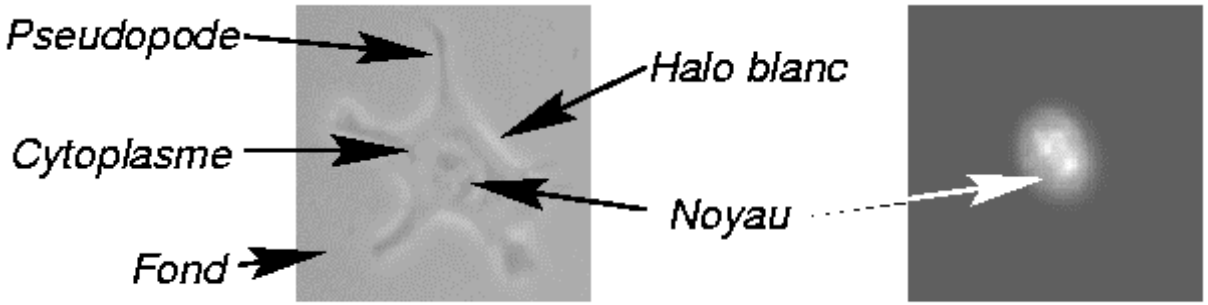


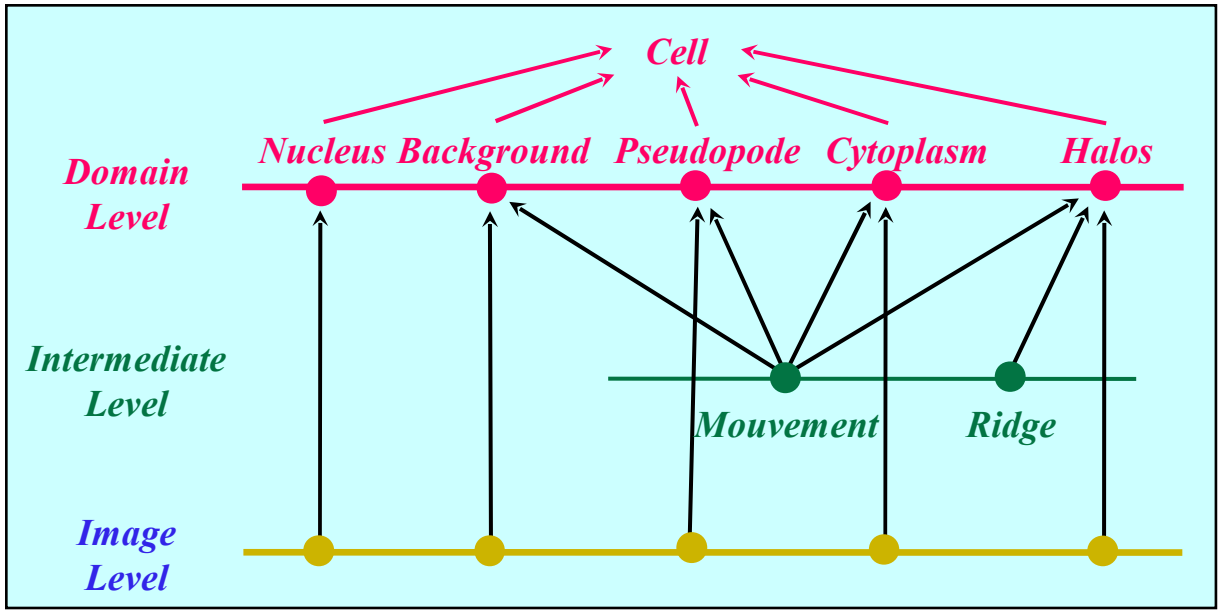




Interleaving agent behaviours



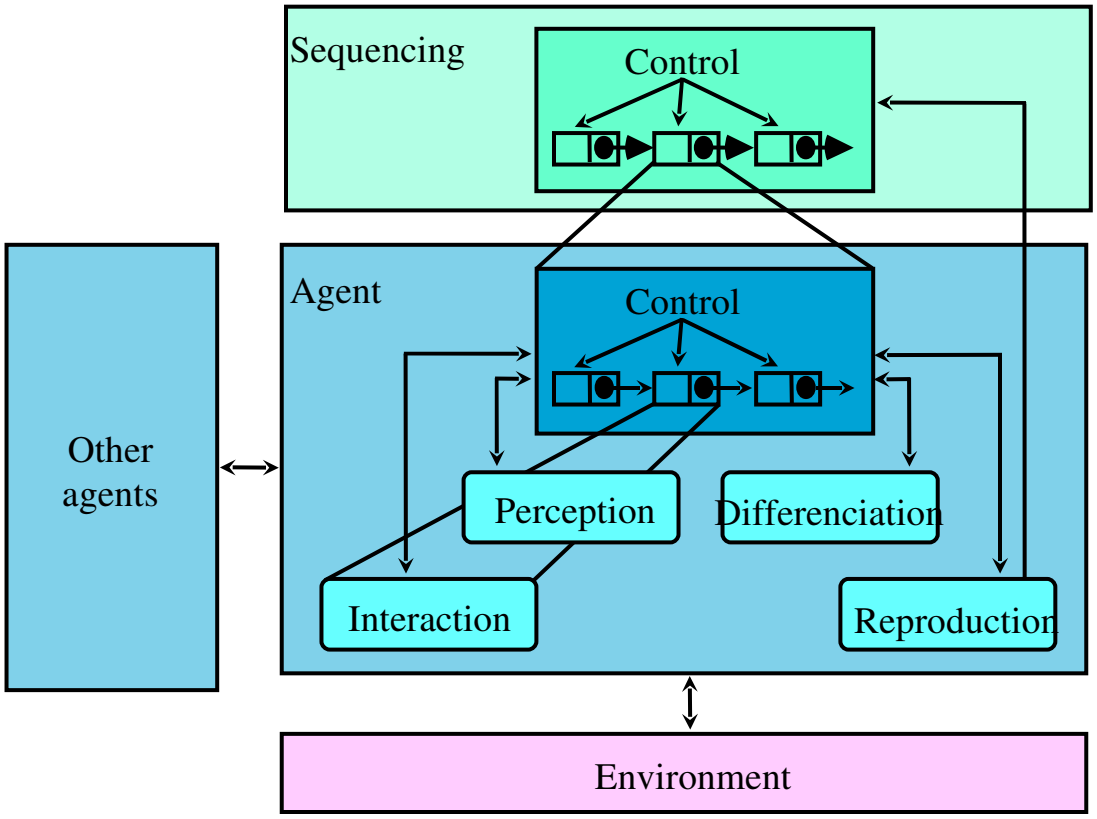




Interleaving agent behaviours

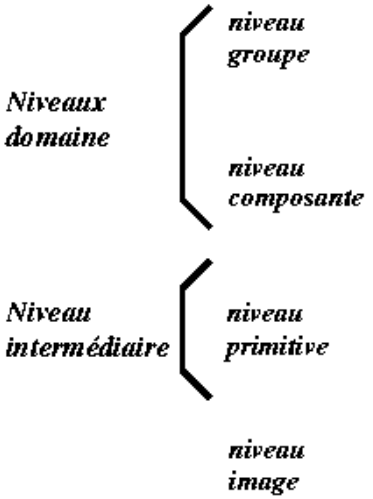
- Reactive agents
  - working asynchoronously at several representation levels and pursuing multiple goals
- Interleaving
  - perception, recognition, interaction and exploration processes

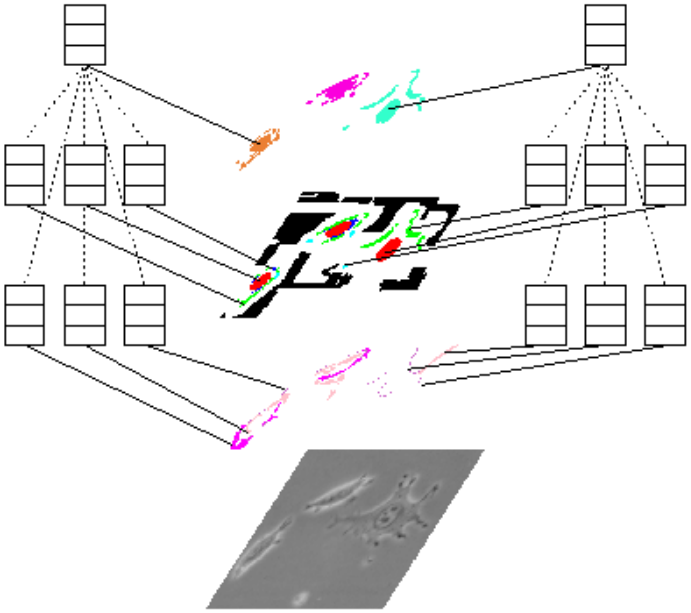
A. Boucher, PhD Thesis, 1999



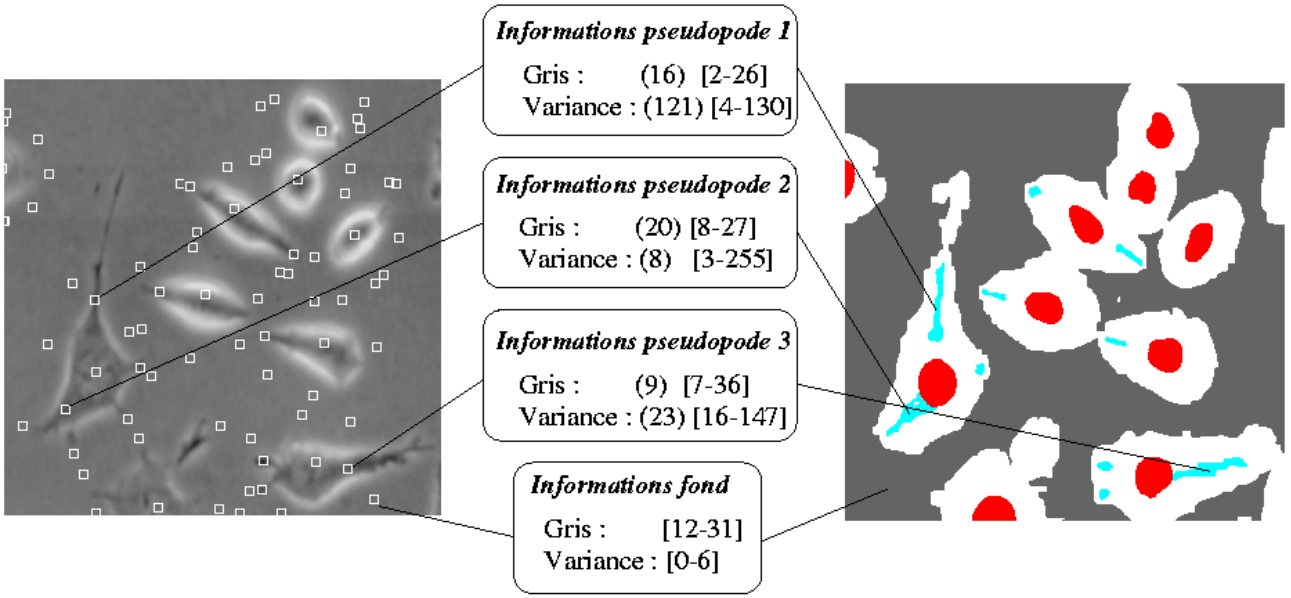
Decision making

- Multi-criteria pixel evaluation
  - Agent-specialized
  - Adapted to local contexts
  - Able to integrate heterogeneous sourc of information



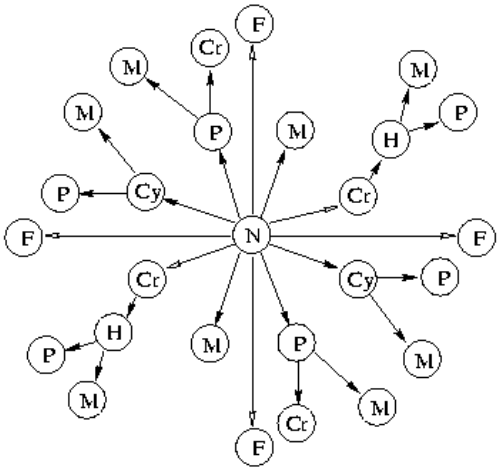


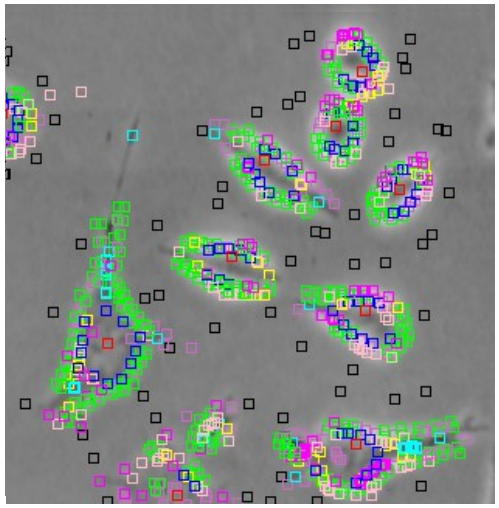
$$Evaluation_{pixel\ /\ r\hat{e}gion} = \sum_{i=1}^{n} poids_i \, crit\grave{e}re_i$$

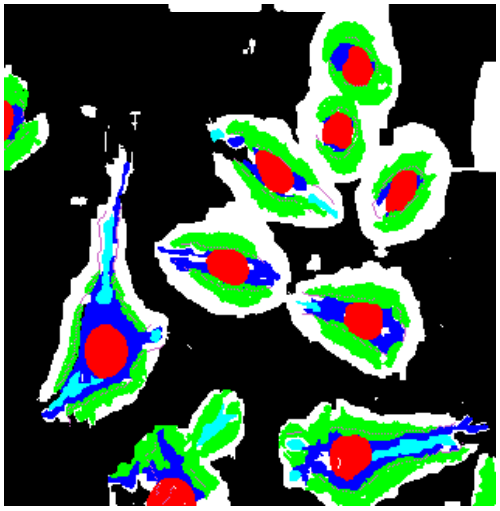


Interleaving agent behaviours

- Reproduction
  - A set of local rules specifying for each agent type
    - the type and amount of agents to be launched
    - Criteria to decide when lauching should occur
    - Criteria to detect seeds for the newly launched agents (transmitted to the created agents)
- Interaction
  - Launched in case of a « collision » between two agents of the same type
  - Ony one agent survives, depending on some criteria (eg size and confidence of the segmented zone)

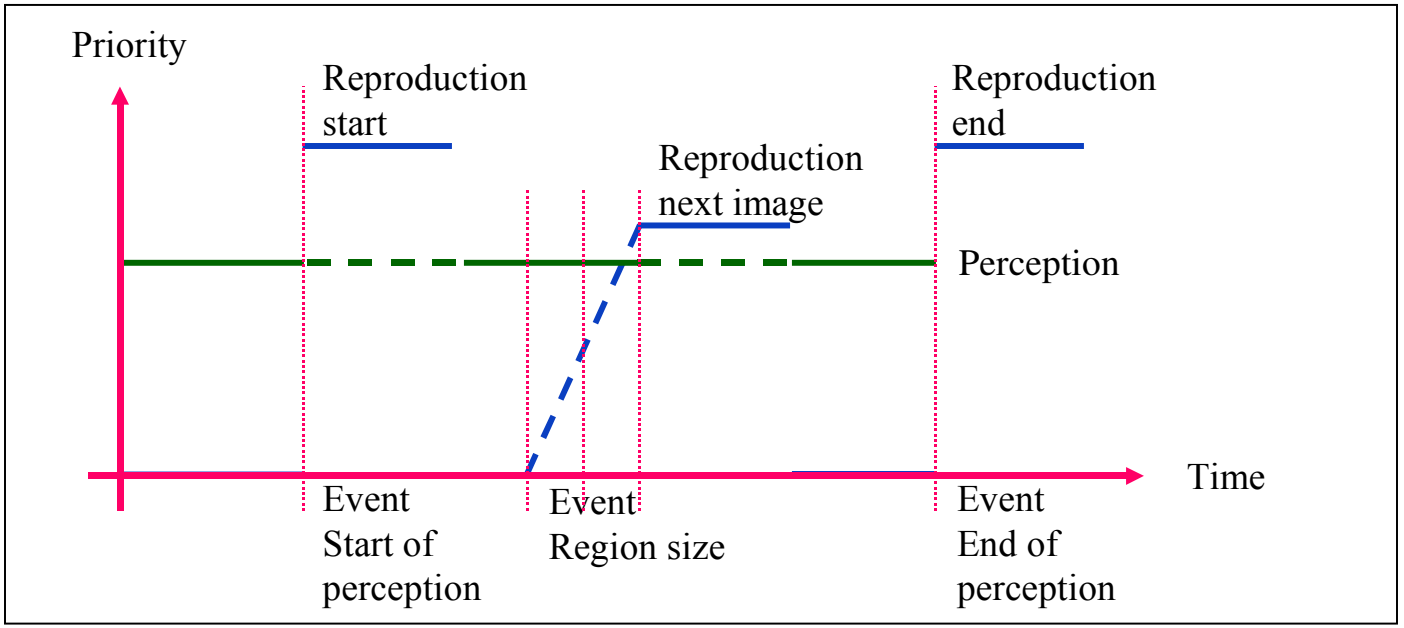






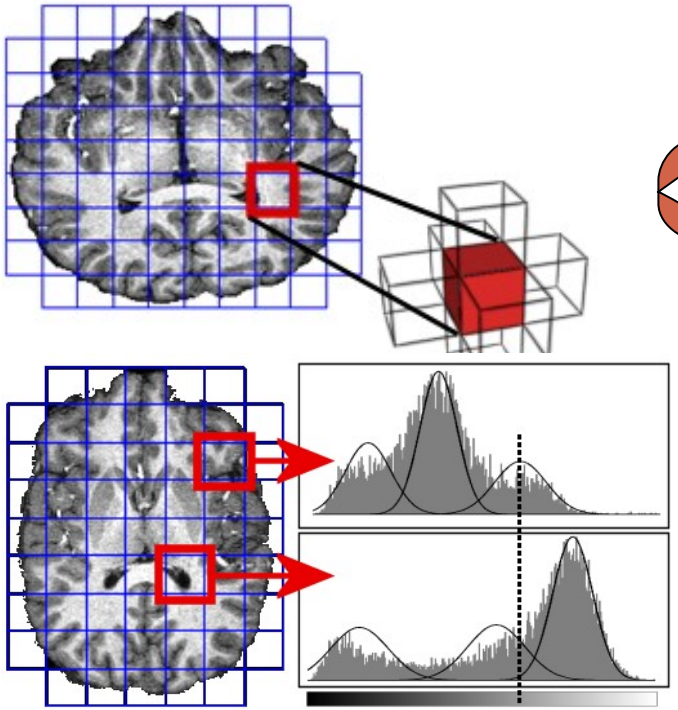
Interleaving agent behaviours

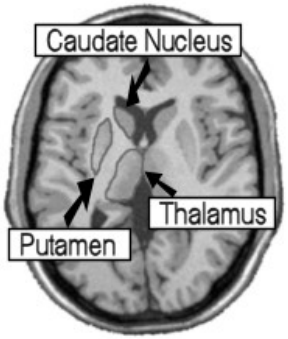
- Behaviour execution is interleaved :
  - Perception is launched first
  - Further behaviours are launched based on their priority
- Each behaviour produces events
  - The events are used to update the launching priority of behaviours

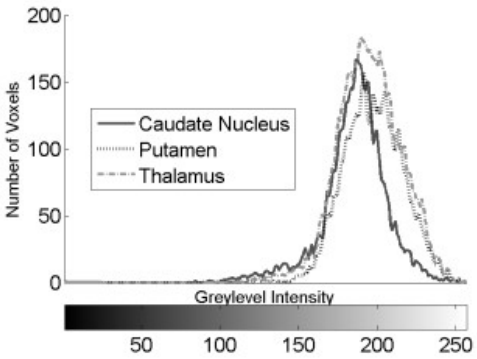


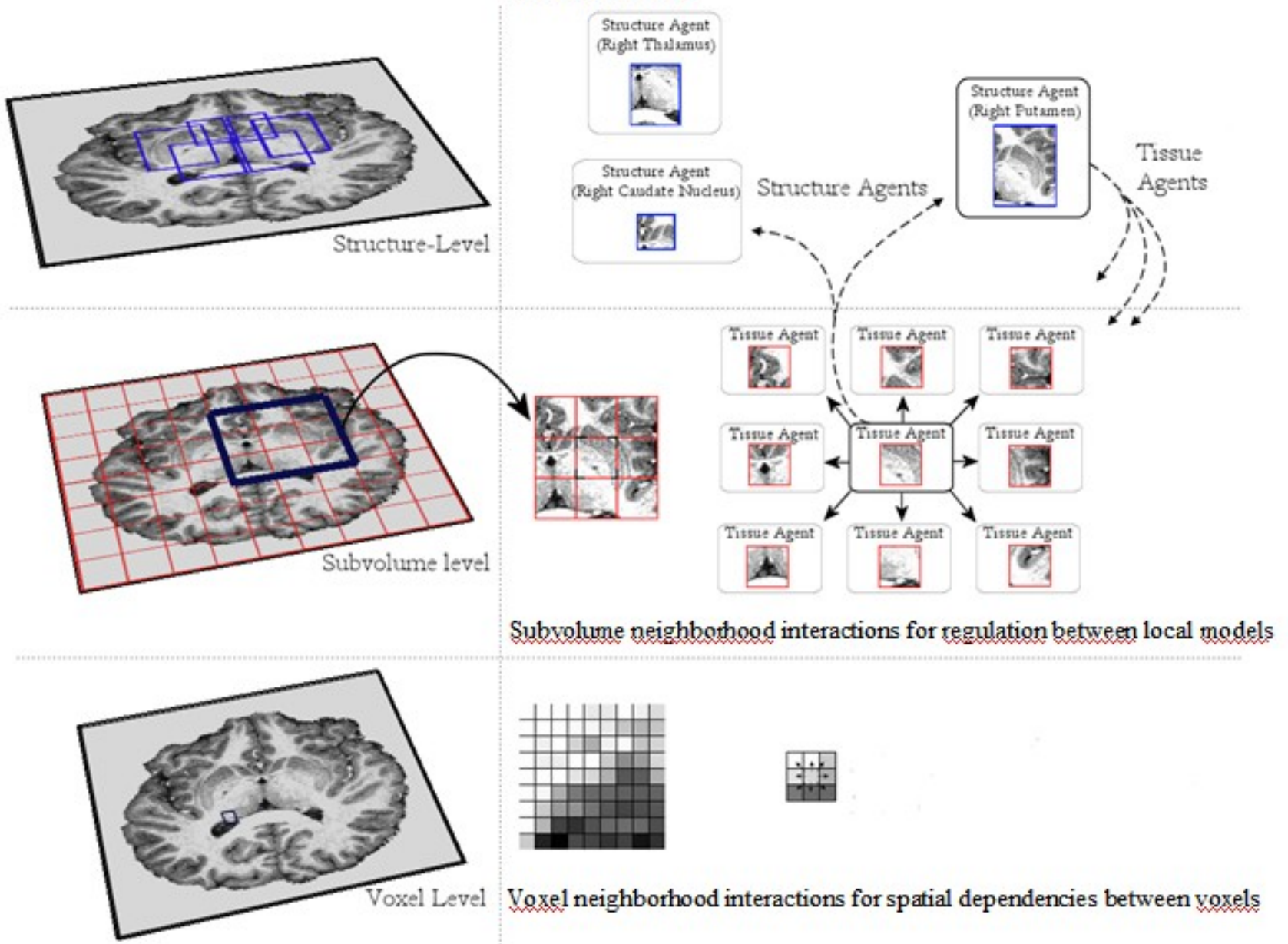
Markovian MRI Segmentation Agents

- Tissue agents (CSF, GM, WM) estimate local intensity models
- Structure agents (Frontal Horn, Caudate Nucleus…) introduce fuzzy spatial knowledge
- For each agent : a local MRF model
- B. Scherrer, PhD Thesis, 2008, with M. Dojat & F. Forbes

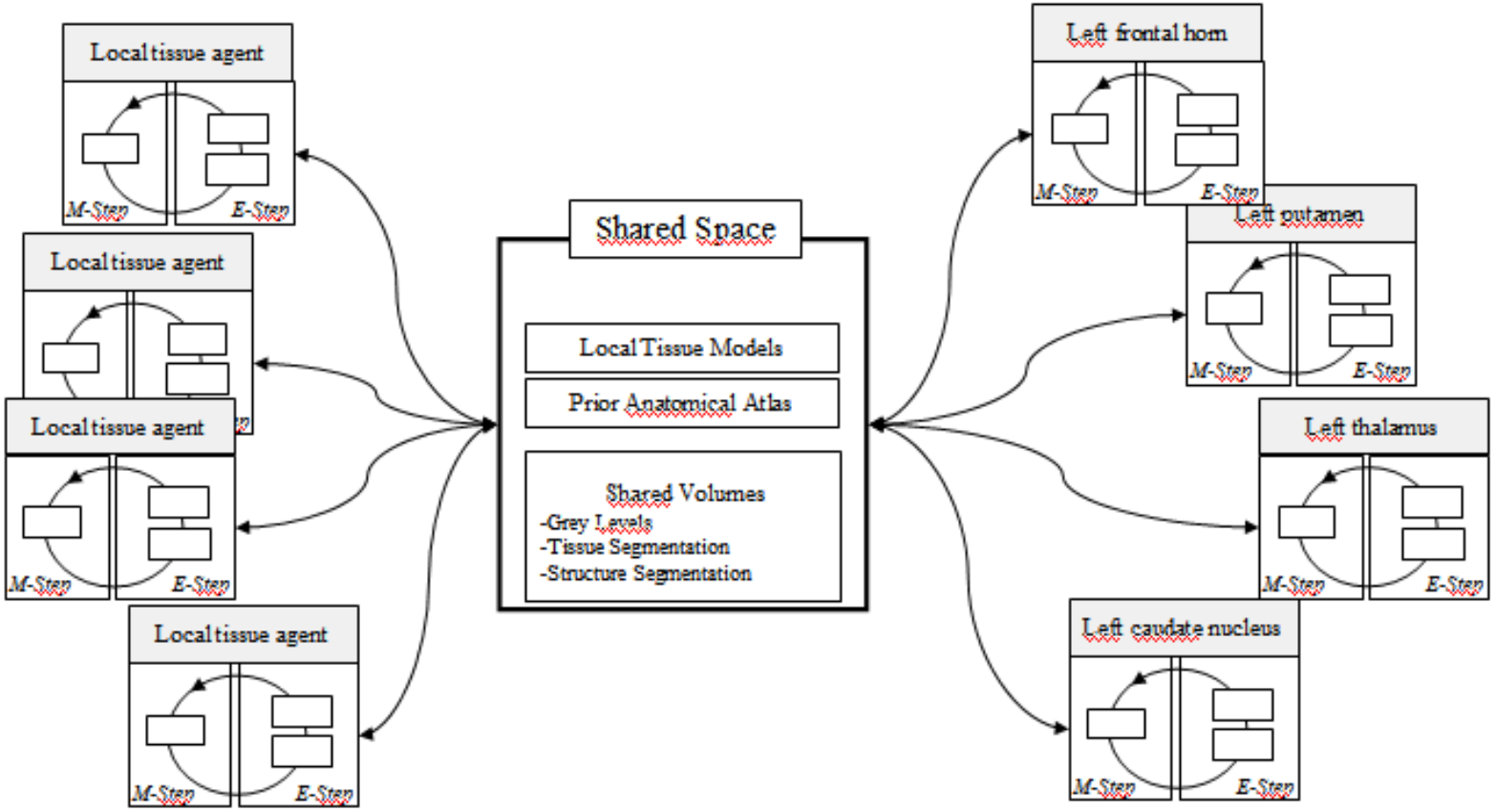







Territories based on *a priori* anatomical information for interactions with tissue models
Structure Agent (Right Thalamus)
Structure Agent (Right Putamen)
Structure Agent (Right Caudate Nucleus)
Structure Agents
Tissue Agents
Structure-Level
Tissue Agent
Subvolume level
Subvolume neighborhood interactions for regulation between local models
Voxel Level
Voxel neighborhood interactions for spatial dependencies between voxels

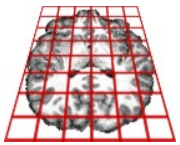
A distributed agent-based framework



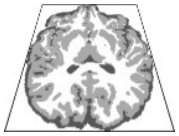
Joint Markov modelling for a situated processing

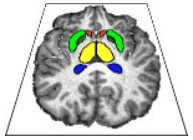
$$p(\mathbf{t}, \mathbf{s}, \theta|\mathbf{y}) \propto \exp\left(H(\mathbf{t}, \mathbf{s}, \theta|\mathbf{y})\right)$$

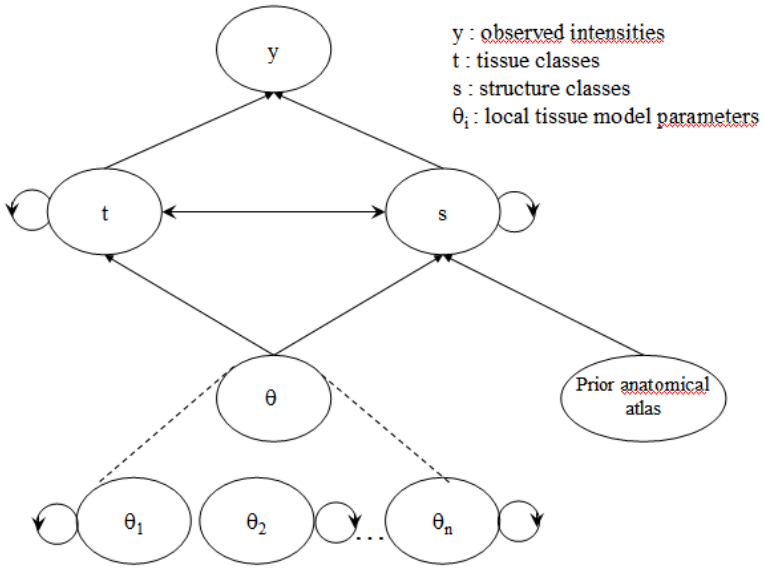
$p(\mathbf{t}, \mathbf{s}|\theta, \mathbf{y})$ $p(\theta|\mathbf{t}, \mathbf{s}, \mathbf{y})$



$p(\mathbf{t}|\mathbf{s}, \theta, \mathbf{y})$ $p(\mathbf{s}|\mathbf{t}, \theta, \mathbf{y})$







- Modelling the joint dependencies between local intensity models, and tissue and structure classifiation,
- Distributing the estimation over sub-volumes

Fully Bayesian Joint Model

- A joint probabilistic model p(t,s,θ |y)
- Three conditional Markov Random Field (MRF) models
- Optimization by means of GAM (Generalized Alternating Minimization) procedures

Structure conditional tissue model

$$H_{T|S,Y,\theta}(\mathbf{t}|\mathbf{s},\mathbf{y},\Theta) = \sum_{i\in V}\left({}^{t}t_i\gamma_i(s_i) + \sum_{j\in\mathcal{N}(i)} U_{ij}^{T}(t_i,t_j;\eta_T) + \log g_T(y_i;{}^{t}\theta_i t_i)\right)$$

Tissue model

External field : Tissue-structure interaction

Interaction between neighbouring voxels

Tissue conditional structure model

$$H_{S|T,Y,\theta}(\mathbf{s}|\mathbf{t},\mathbf{y},\Theta) = \sum_{i\in V}\left({}^{t}s_i \log f_i + \sum_{j\in\mathcal{N}(i)} U_{ij}^{S}(s_i,s_j;\eta_S) + \log g_S(y_i|t_i,s_i,\theta_i)\right)$$

Tissue-structure interaction

A priori knowledge on structure

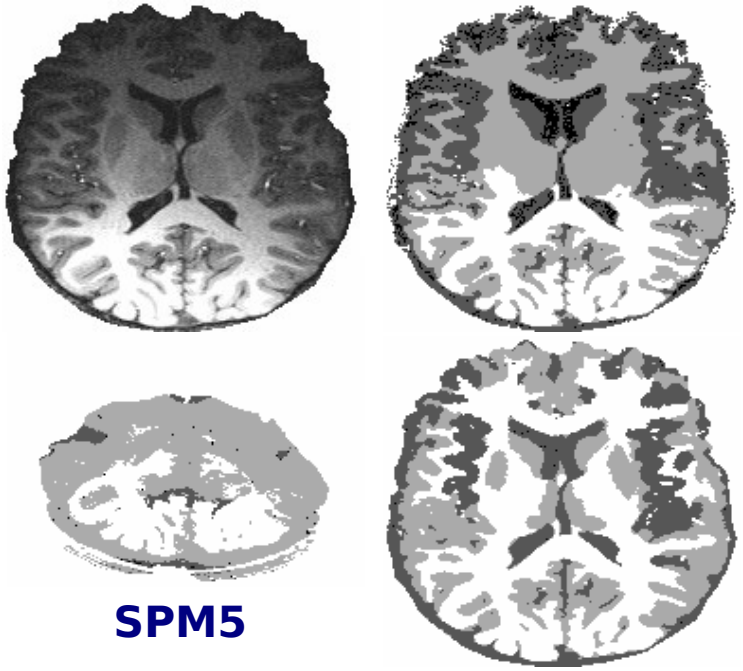
Tissue/structure conditional parameter model

Dependency between neighbouring sub-volumes

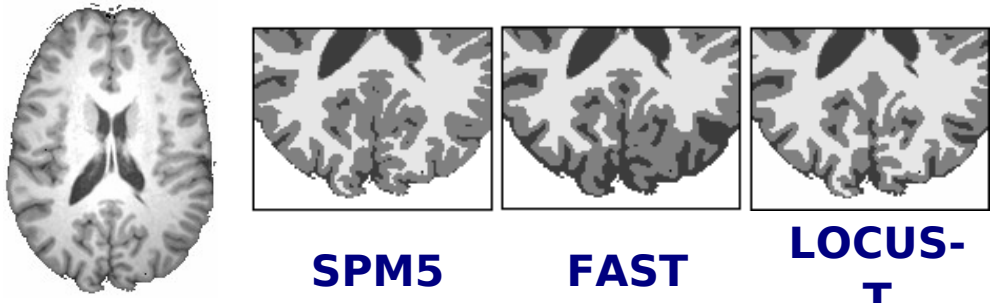
$$H_{\theta|Y,T,S}(\Theta|\mathbf{y},\mathbf{t},\mathbf{s}) = H_{\Theta}^{\mathcal{C}}(\theta) + \sum_{c\in\mathcal{C}} \log \prod_{i\in V_c} g_S(y_i|t_i,s_i,\theta_c)$$

Model constancy over a sub-volume

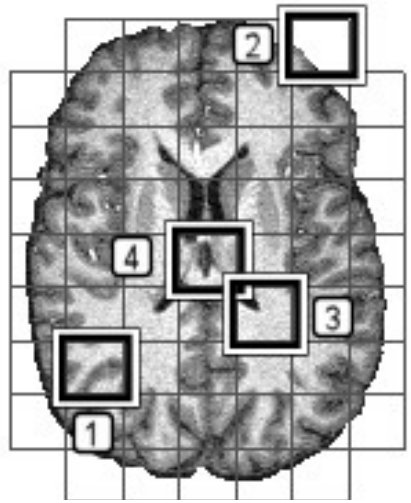
High inhomogeneity (surface antenna)



Real 3T Image

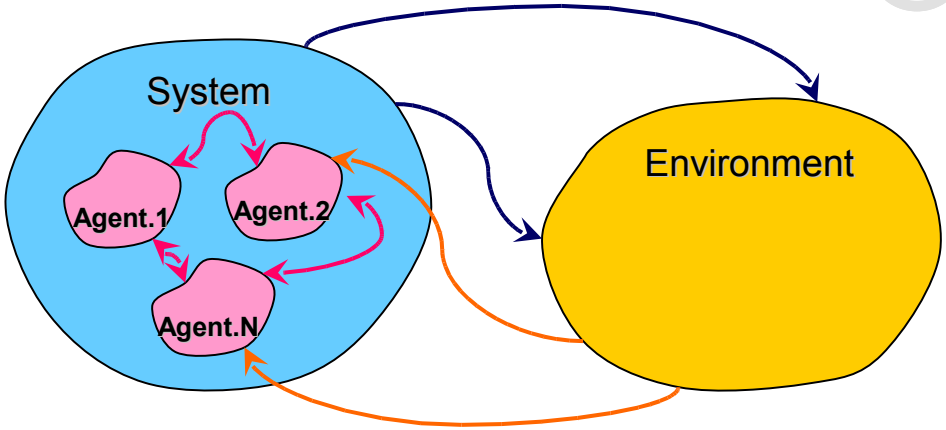


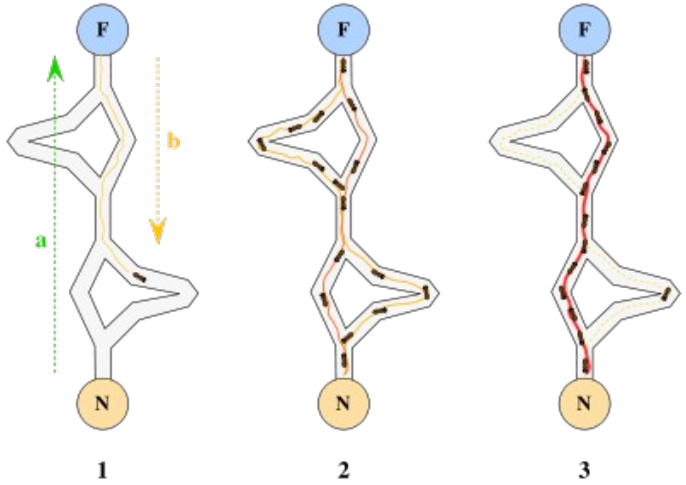
Adaptation to local image complexity



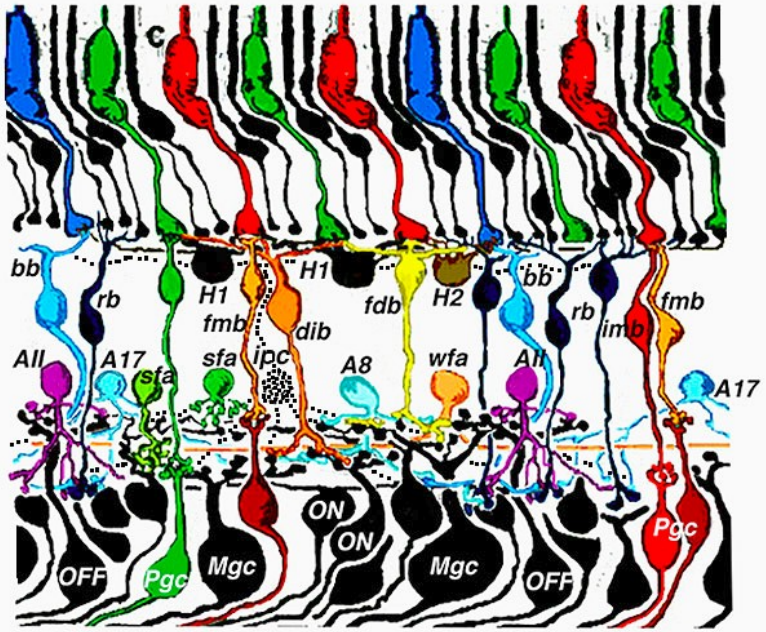
Iteration number per agent

Why is this an important question ?





- Rationality under two different viewpoints
- Bounded rationality :
  - The agent rationality is « limited » when its cognitive abilities do not allow him to reach an optimal behaviour or when the complexity of the environment is beyond the capacities of the agent
  - The environment is a constraint to which the agents must adapt
- Situated rationality
  - Rationality as a property of the interaction between the agent, its environment, the other agents and the system as a whole
  - The environment provides resources which complement the agents own resources and support their action : « a digital housing environment »
  - Problem solving as a co-construction resulting from the agent (inter)actions and the resources in their environment
  - *F. Laville, 2000 « La cognition située, une nouvelle approche de la rationnalité limitée »*

➔ Swarm intelligence, social cognition…

Mobilize all the heterogeneous styles of computational design to build tomorrow's AI