(Neuromimetic) Action Selection

Benoît Girard
benoit.girard@isir.upmc.fr

ISIR - CNRS / UPMC
http://www.isir.upmc.fr

Sélection de l'Action en Robotique Comportementale

Back to Brooks...

- Braitenberg vehicles do not allow for much behavioral complexity.
- How can we go beyond that ?

- **Specifications** (Brooks, 1991) :
  - A robot must react appropriately and rapidly to changes of the environment
  - A robot must be robust wrt. its environment (minor changes in the environment ➔ no major behavioral changes)
  - A robot must be able to simultaneously maintain multiple goals (ex. walk and find an object)

Brooks' Method

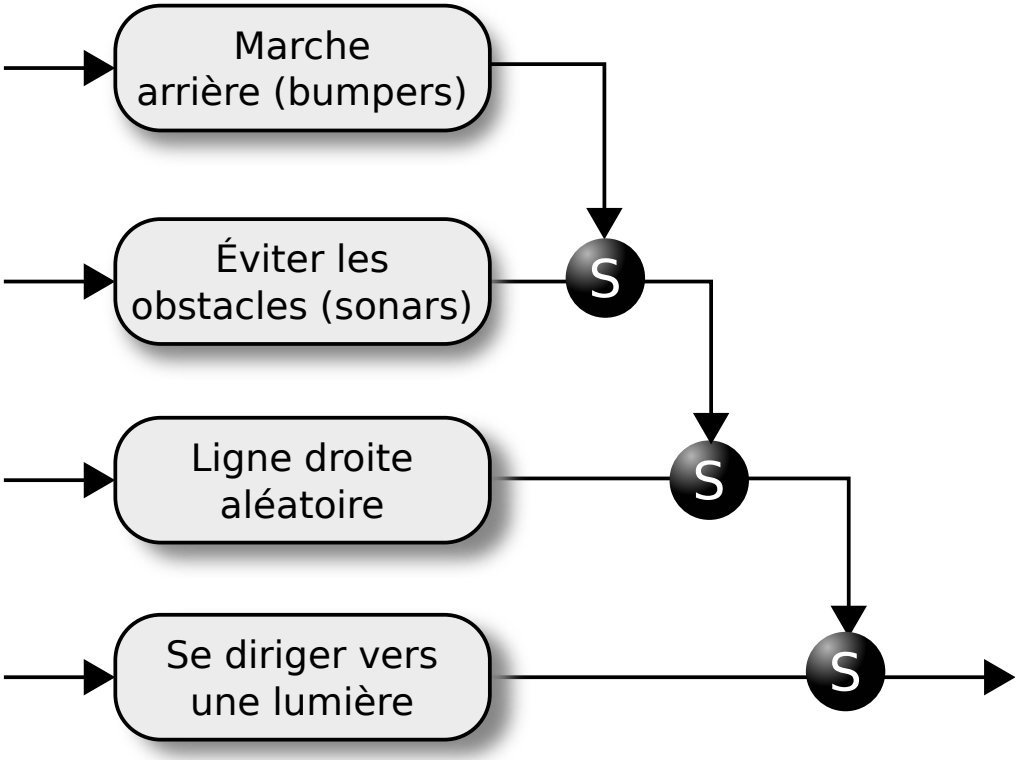
- Incrementally build the capabilities of the robot, with a complete operational system at each step
- At each step, the robot has to be tested in real environments with real sensors

  ➔ go beyond simulations and fully controlled evts.

"When we examine very simple level of intelligence, we find that explicit representation and models of the world simply get in the way. It turns out to be better to use the world as its own model." – Brooks, 1991

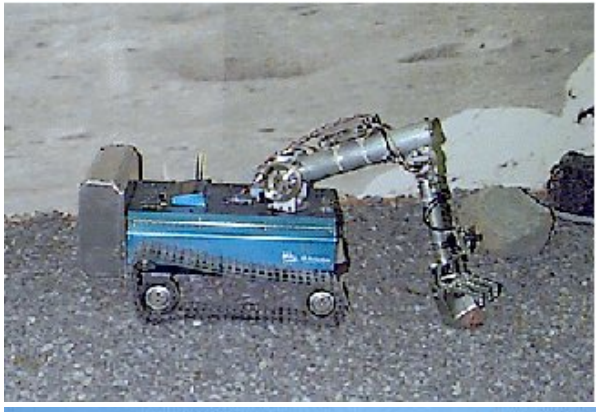
➔ the question of representations is not central for robotic AI.

Subsumption architecture

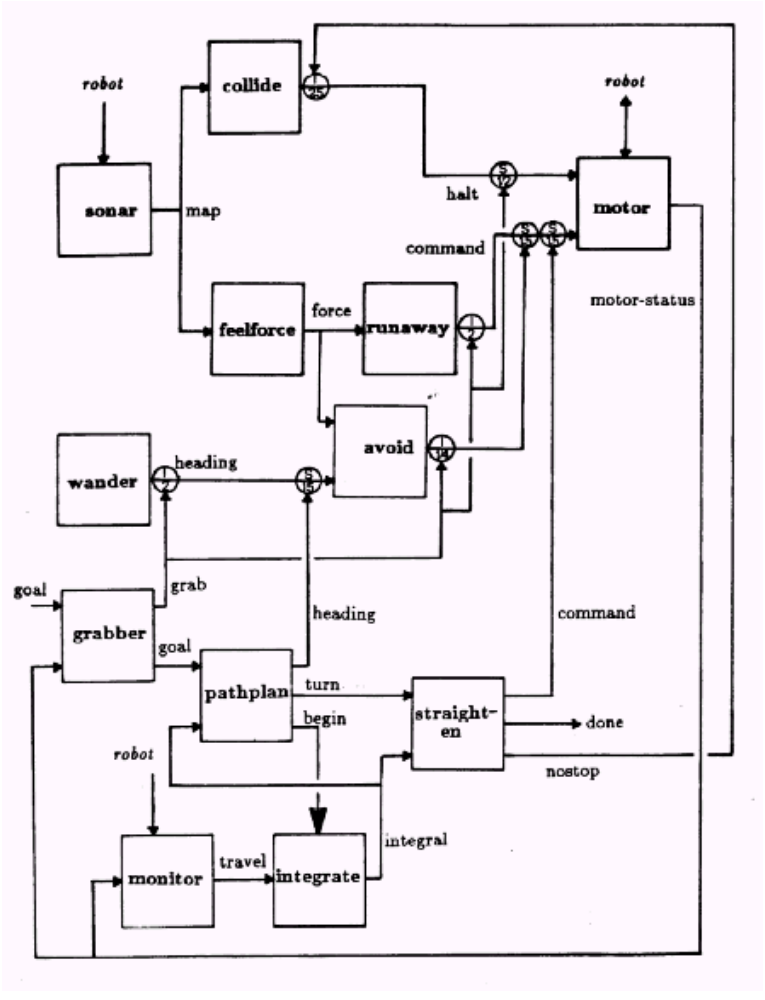


- Each nehavior operates in parallel
- Behaviors of higher priority inhibit the lower ones
- The global behavior **emerges** from the interactions between the behaviors and the environment.

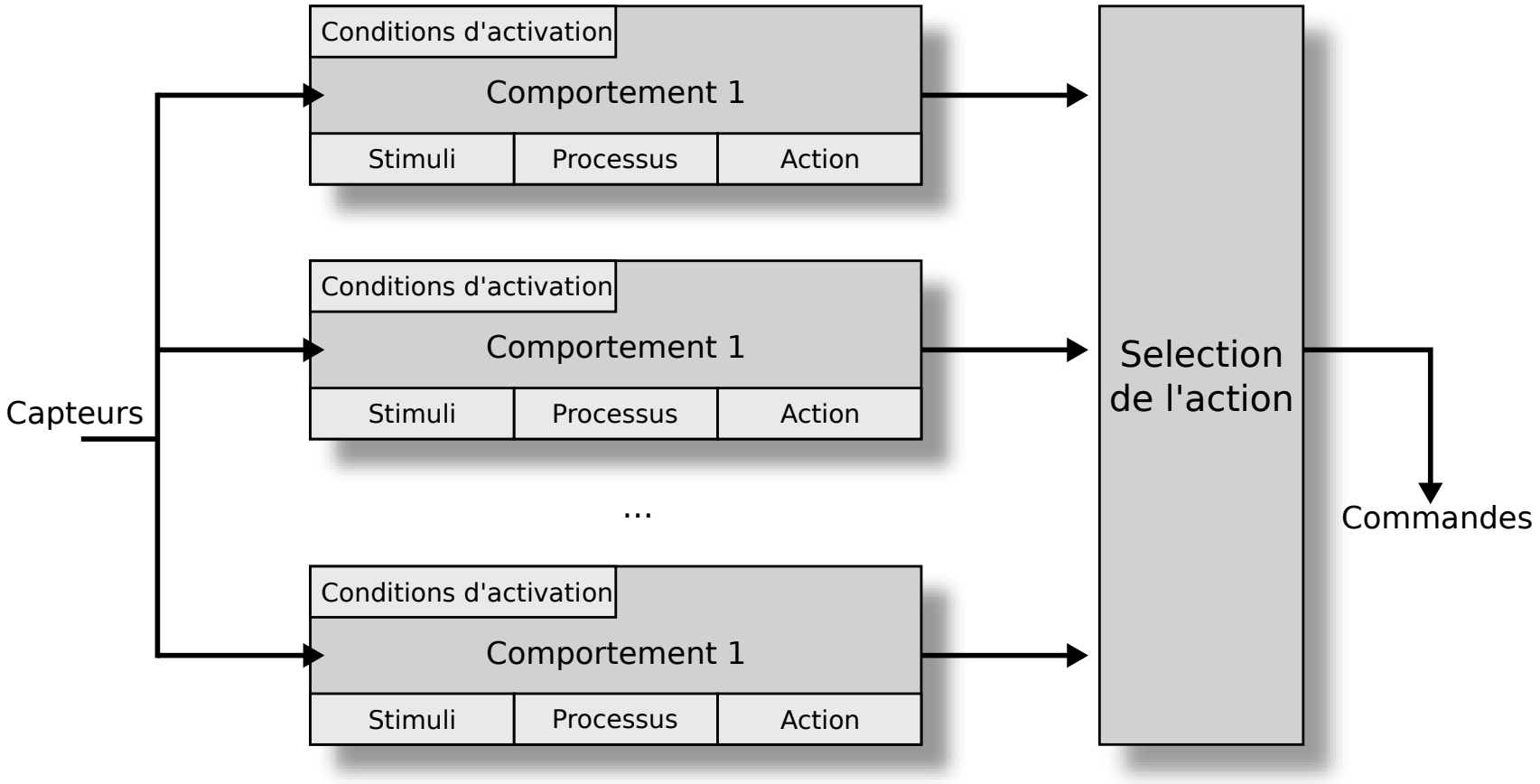
Implementations



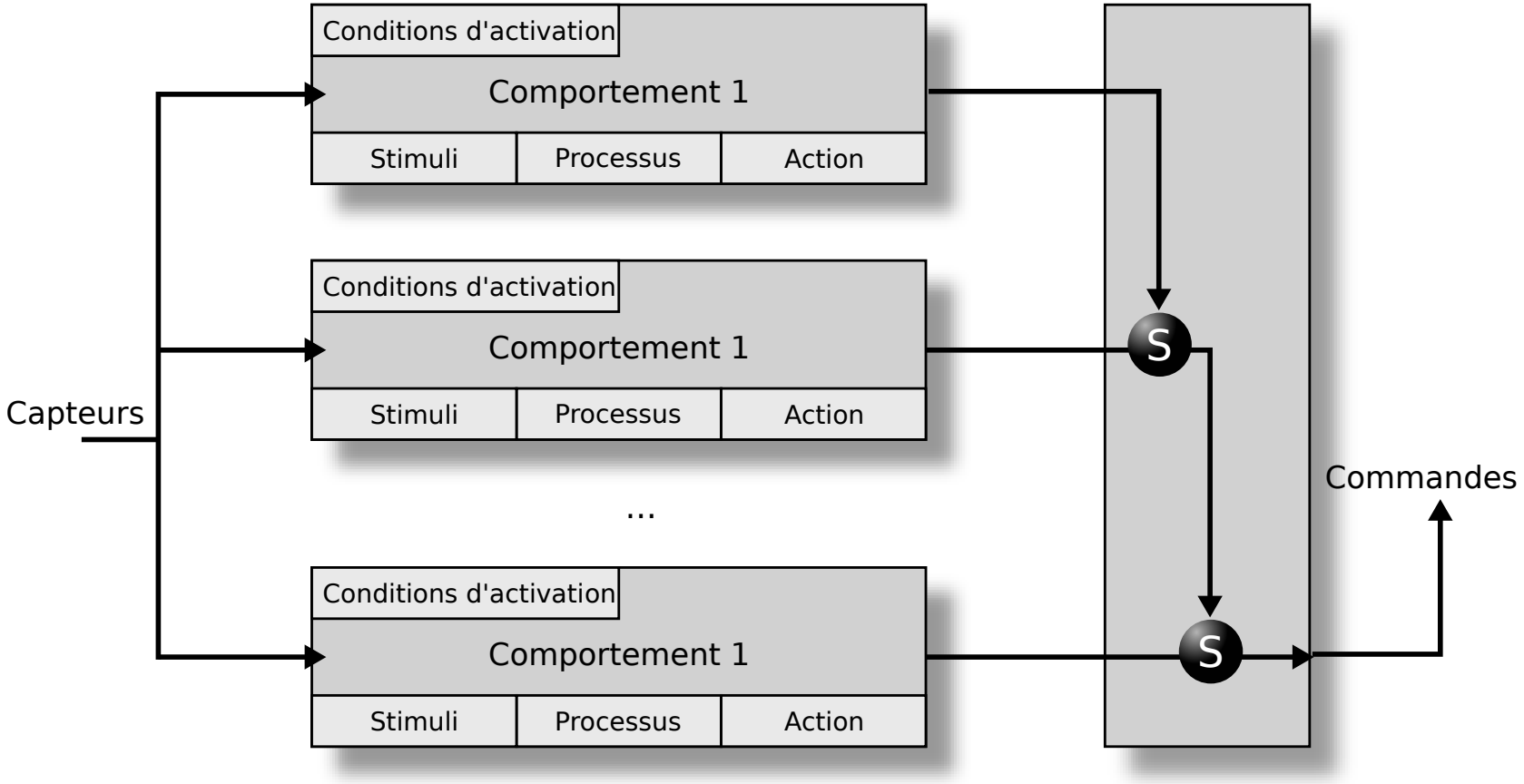
But it rapidly becomes messy...



Generalization

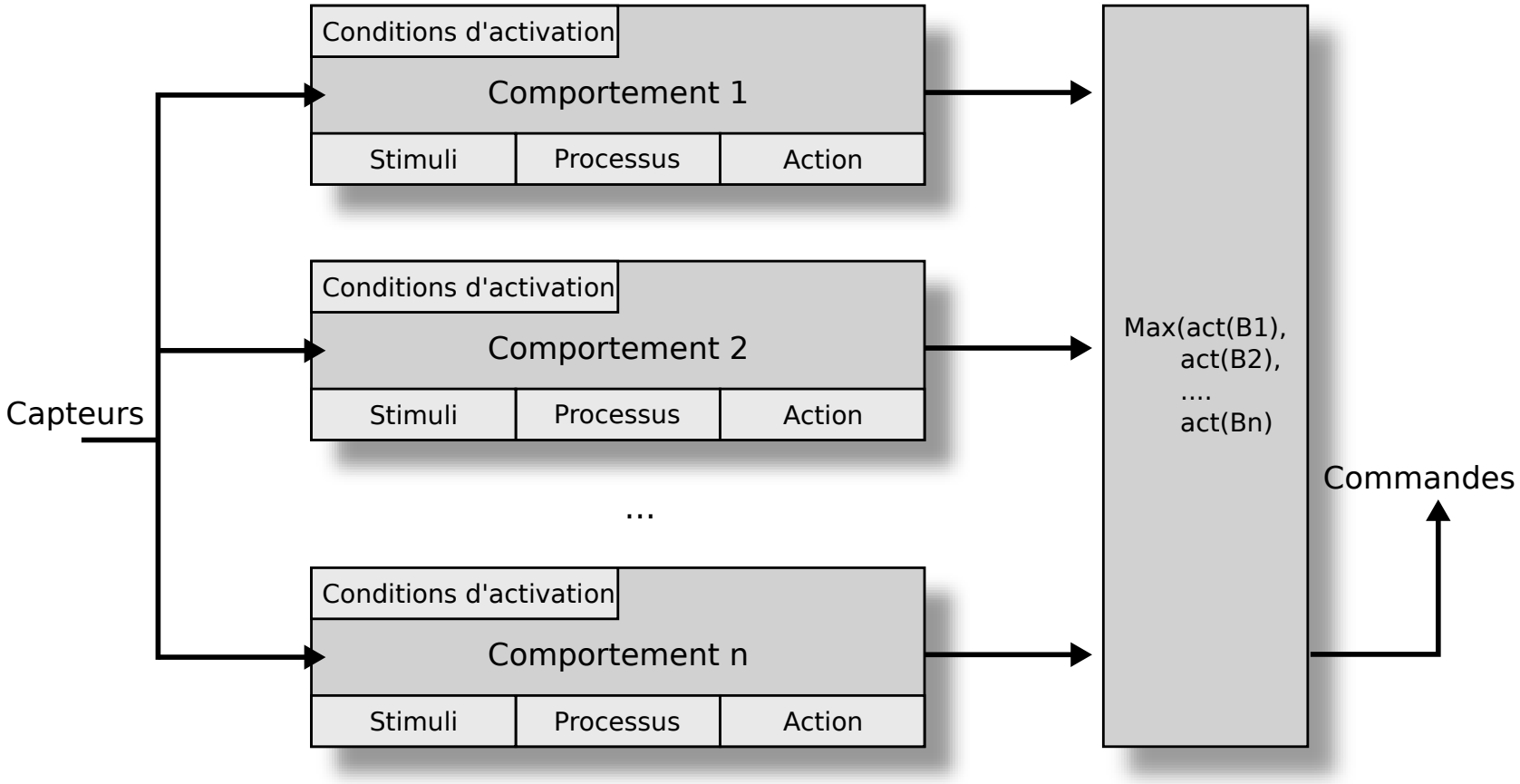


Subsumption



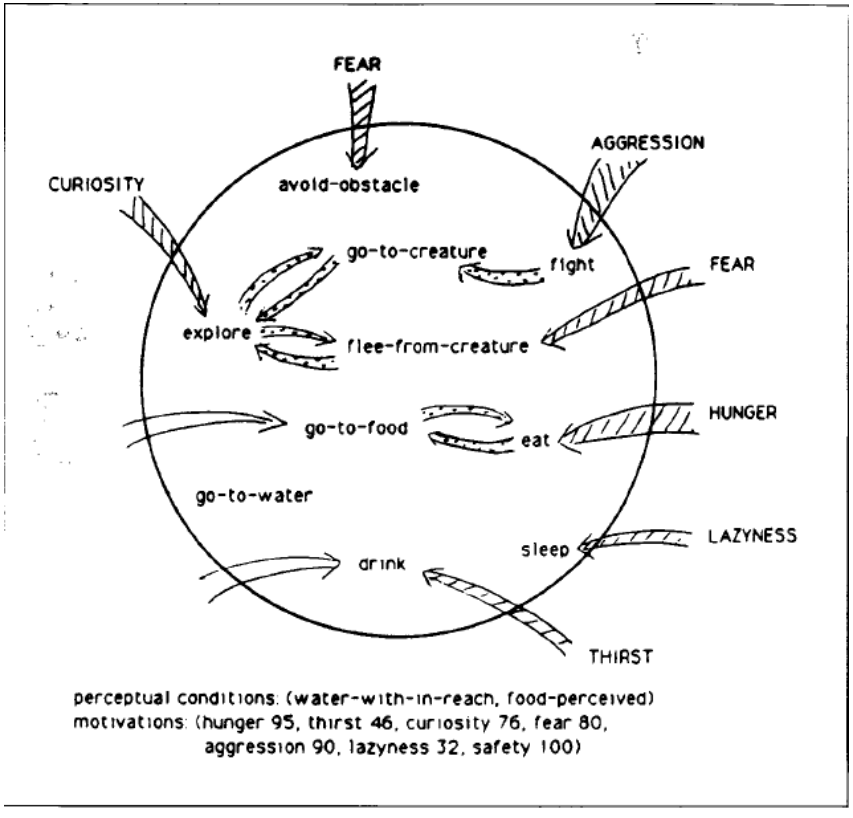
Selected behavior : highest priority

Action Selection



Selected behavior : highest activation

Système non-hiérarchique (Maes, 1991)

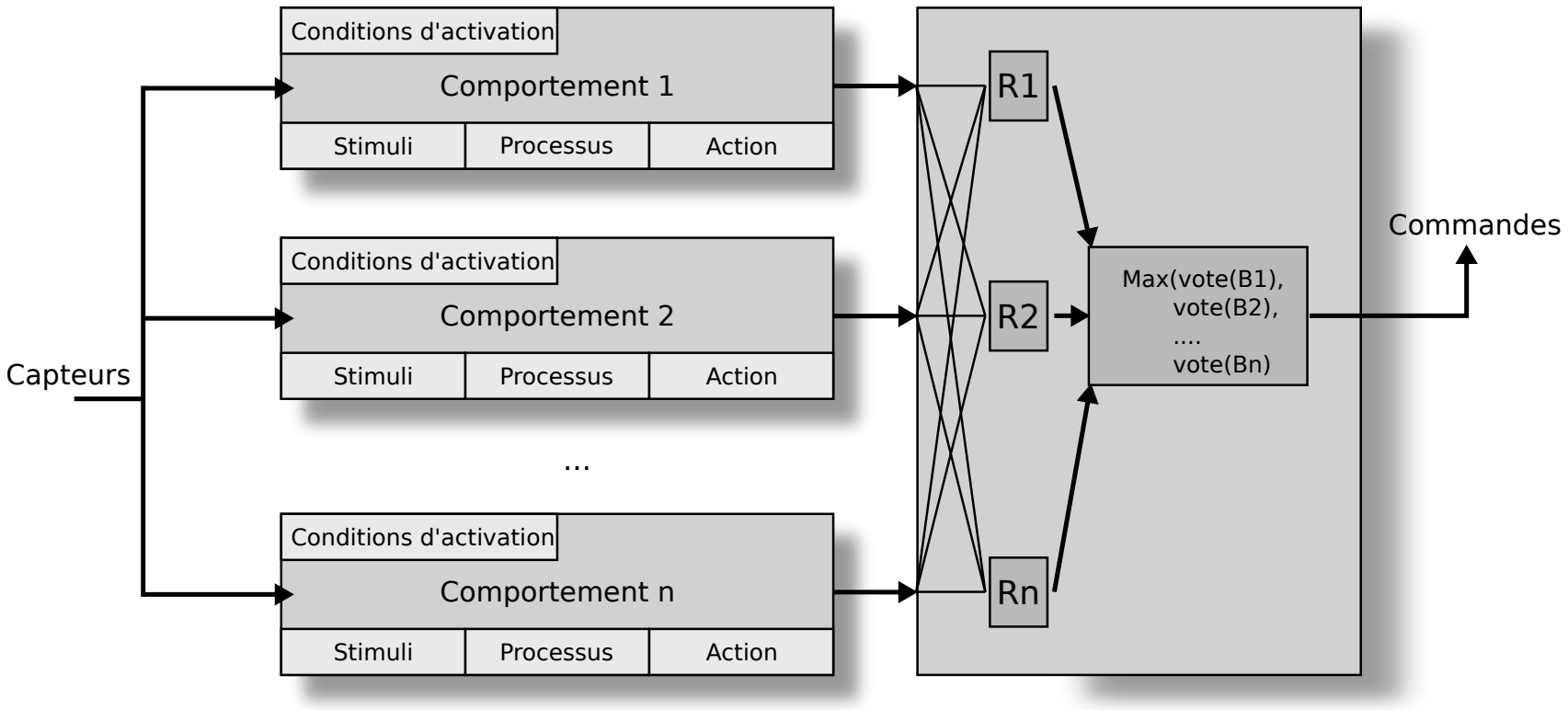


Liens de type :

- prédécesseur,
- successeur,
- antagoniste.

Flux d'énergie dans le réseau, similaire à un réseau de neurones. Conditions logiques d'activation.

Vote



Selected : highest number of votes

DAMN (Distributed Architecture for Mobile Navigation)

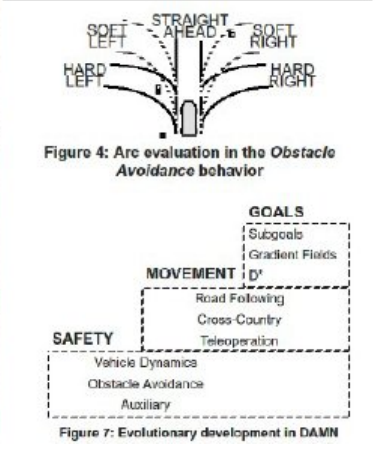


Figure 4: Arc evaluation in the *Obstacle Avoidance* behavior

Figure 7: Evolutionary development in DAMN

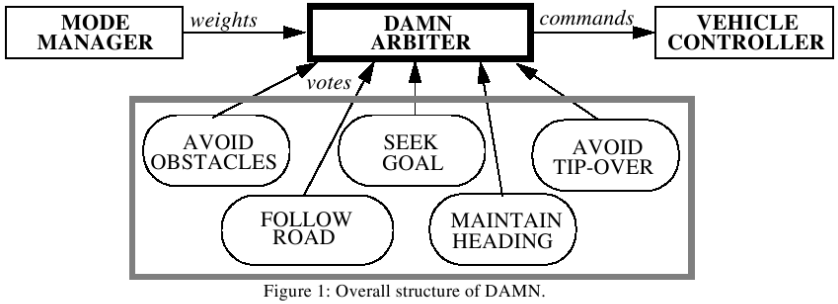


Figure 1: Overall structure of DAMN.

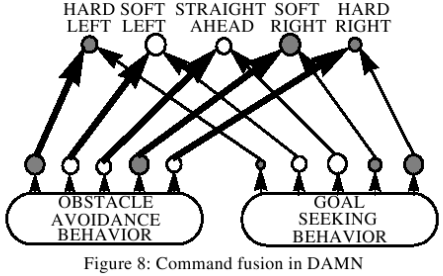


Figure 8: Command fusion in DAMN

Rosenblatt & Payton (1989–1996)

Hiérarchie à libre flux (Tyrrell, 1995)

Travail fondamental de formalisation du problème de la sélection de l'action.

- Génération de séquences d'actions contiguës,
- Priorités dans le choix des comportements,
- Choix opportunistes,
- Persistance,
- Interruptabilité des comportements,
- Décomposition des problèmes en sous-problèmes,
- Préférence des actions consummatoires *vs.* appétitives,
- Utiliser toute l'information disponible,
- Être extensible,
- Variation des objets d'attention,
- Permettre des actions parallèles,
- Maximiser la survie/les reproductions.

Hiérarchie à libre flux (Tyrrell, 1995)

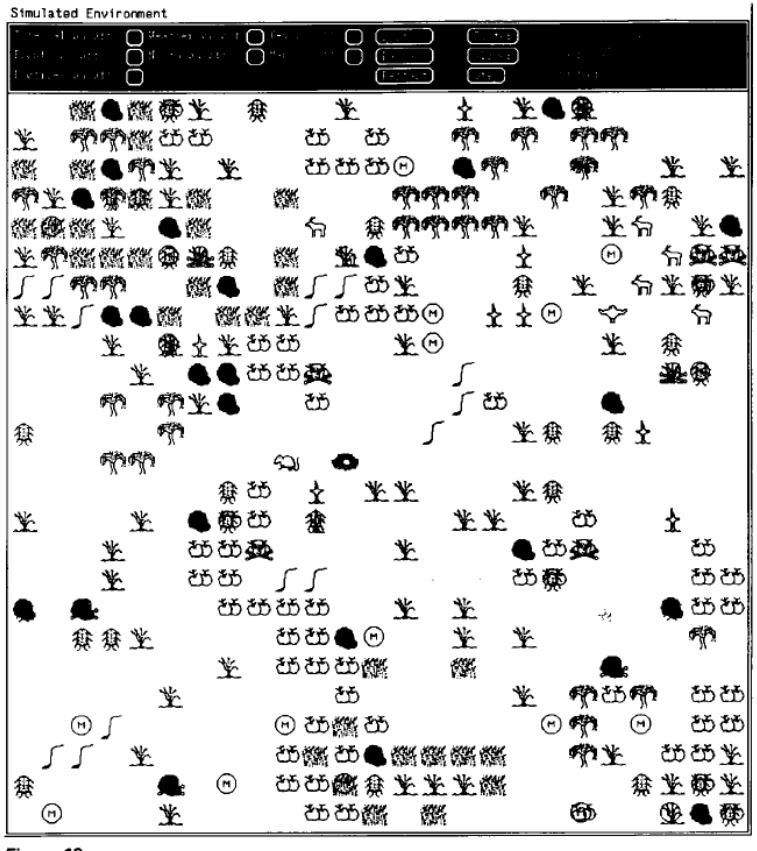


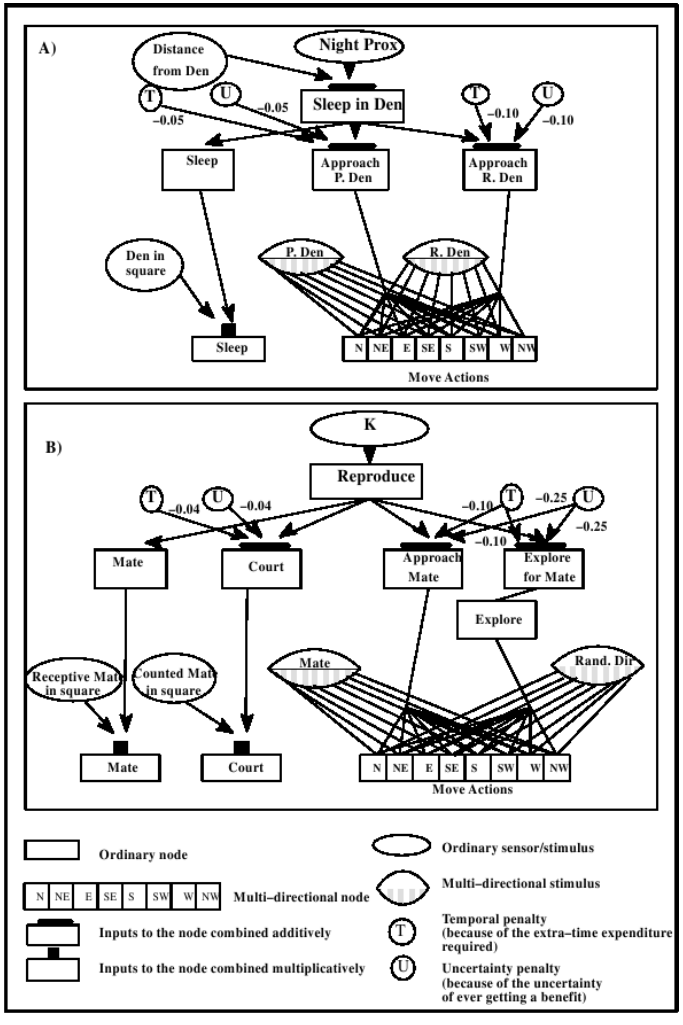
Figure 13
A randomly generated instance of the simulated environment showing the positions of features, the position of the animal, the current action of the animal, the time, the day, and the current part of the day.

Benchmark :

- 9 actions consummatoires,
- 20 actions appétitives,
- 17 motivations,
- 27 stimuli.

Une complexité excessive compte-tenu de l'objectif.

Hiérarchie à libre flux (Tyrrell, 1995)



Hiérarchie à libre flux.
Dérivé de DAMN.
Pb de granularité des actions.
Montre les limitations intrinsèques de Maes.

Hiérarchie à libre flux (Tyrrell, 1995)

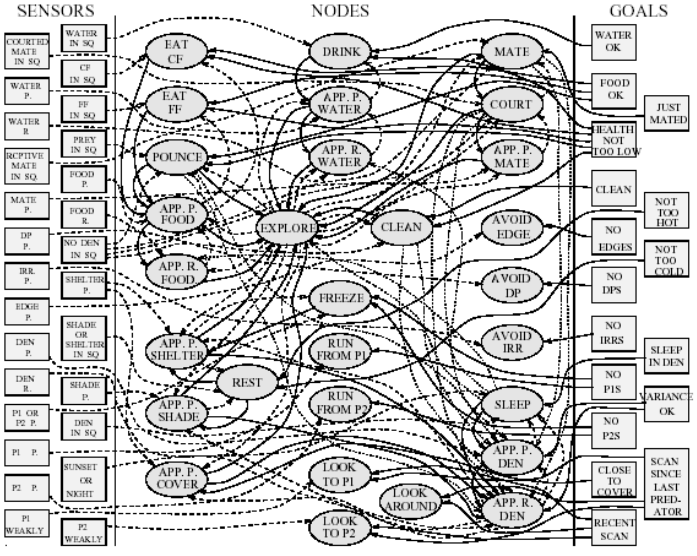


Figure 9.4: The implementation of Maes' ASM to solve the action selection problem posed by the SE. Solid lines denote goal or predecessor connections, dashed lines denote sensor or successor connections, and dotted lines denote protected goal or conflictor connections. 'P' stands for 'perceived' in sensor or node names, and 'R' for 'remembered'. 'APP' stands for 'approach' and 'SQ' for 'square'.

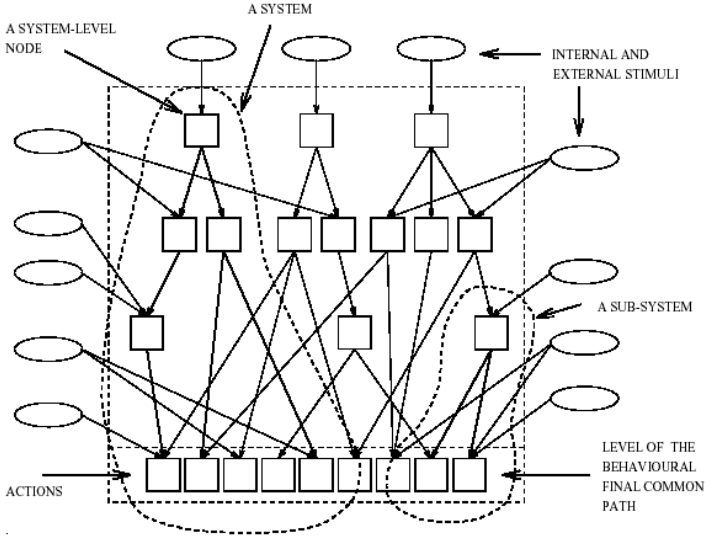
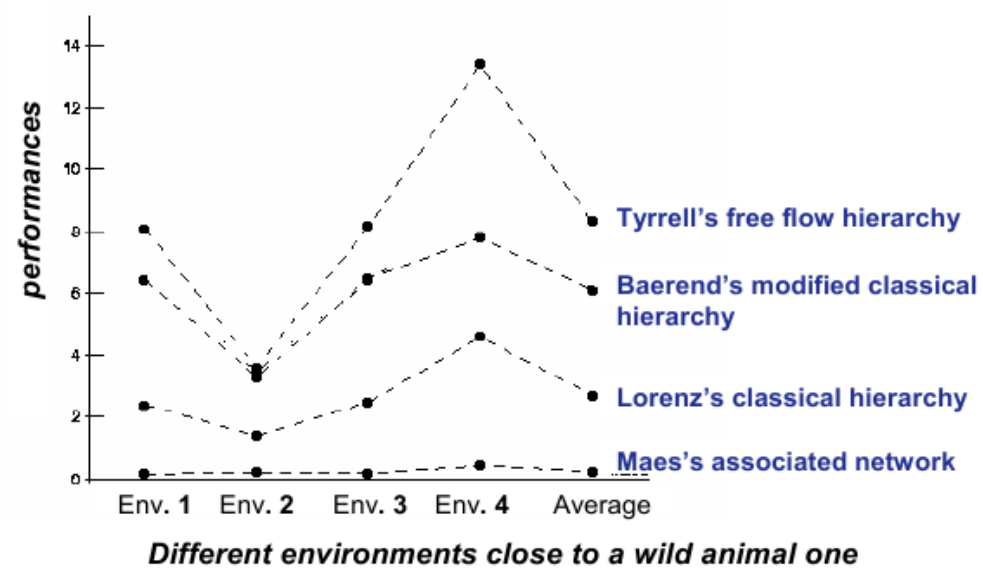


Figure 8.16: An example of the structure of a Rosenblatt & Payton-like ASM. A single system is shown outlined by the curved dashed line. Stimuli are also shown impinging on the mechanism.

Comparaison de la résolution de la tâche de Tyrrell par son système et celui de Maes.

Hiérarchie à libre flux (Tyrrell, 1995)



Comparaison de la résolution de la tâche de Tyrrell par son système et celui de Maes.

Modèle sans sélection ni actions... (?) (Seth, 1998)

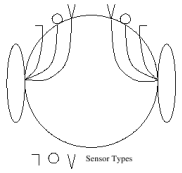


Figure 1 *Architecture of the animat: there are three sensor types, with each sensor connected directly to the wheel on the same side. Each illustrated connection actually comprises three genetically specified links, so the animat consists of 18 concurrent and direct transformations of sensory input into motor output.*

- 3 type de senseurs (nourriture, eau, pièges),
- 3 connexions par senseurs vers les moteurs,
- fonctions linéaires par morceaux dont les paramètres résultent d'un algo. génétique.

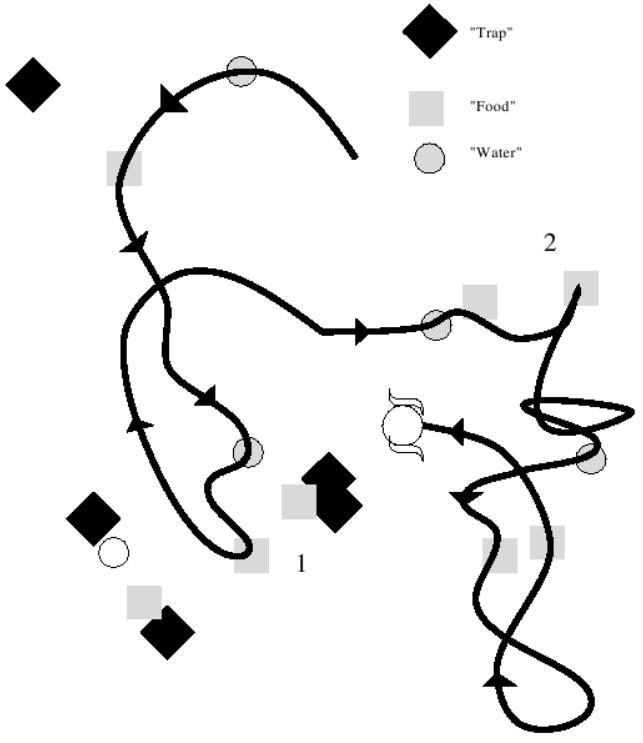


Figure 3 *A sample trajectory; animat passes through a series of food and water sources, backtracking to avoid traps at point 1, and displaying opportunistic behaviour at point 2.*

Modèle sans sélection ni actions... (?) (Seth, 1998)

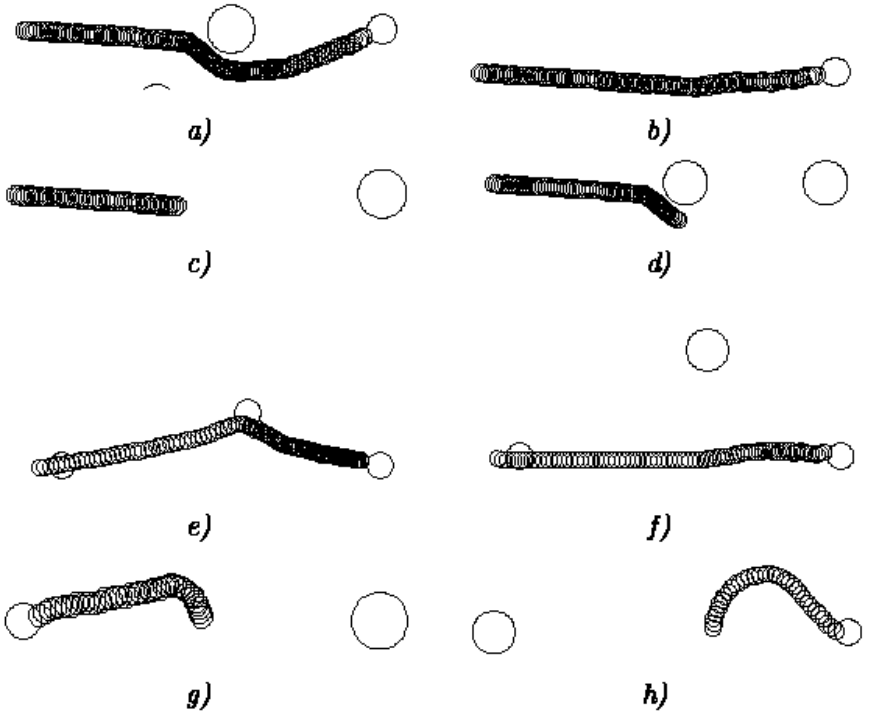
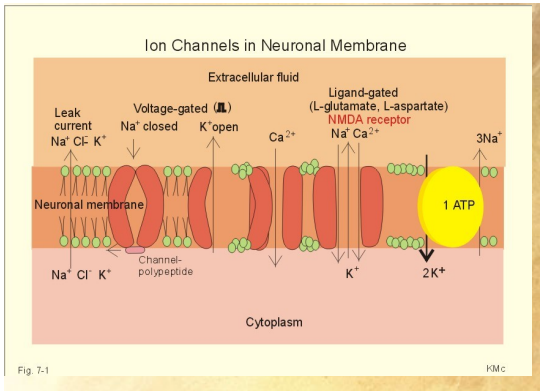
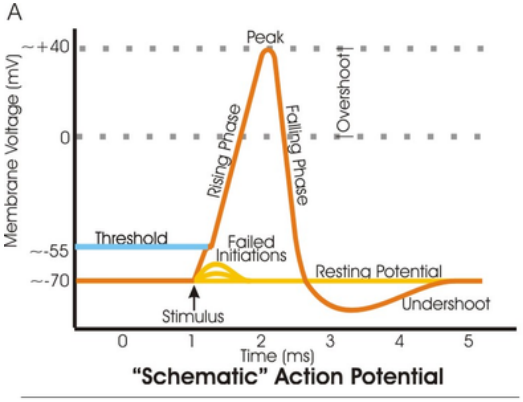


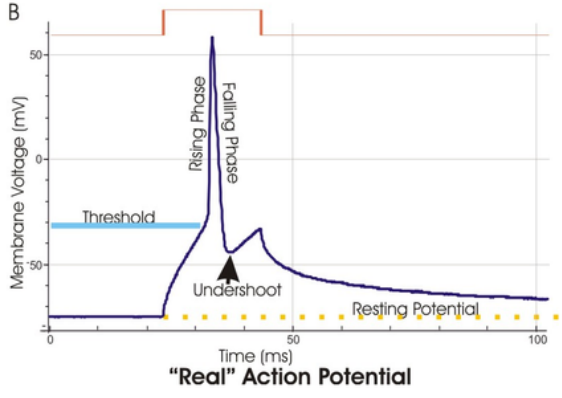
Figure 6 *Action selection tests; consumed sources appear as smaller circles; a) a trap suddenly appears, and the animat interrupts its progress towards the food in order to avoid it. b) with no trap, the animat progresses directly to the food source. c) trajectory of the animat at the point just before the trap appears... d)..and just after the trap appears. e) animat passes through a first food source, and will divert to a second food source (en route to water) if the diversion is not too far. f) if the second food source is too far away, the animat will proceed from the first food source to the water source. g) animat starts between two equidistant sources, and will head towards one needed most. h) if the 'most needed source' is too far away, then the animat will proceed to the other source.*

Approche Neuromimétique

A short foreword about neurons & neuron models

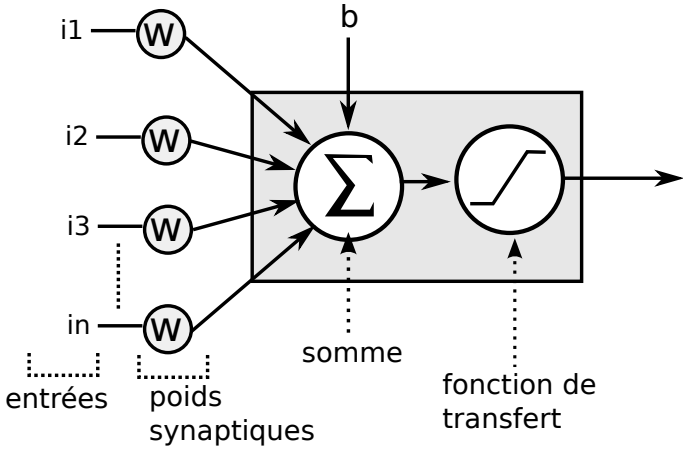


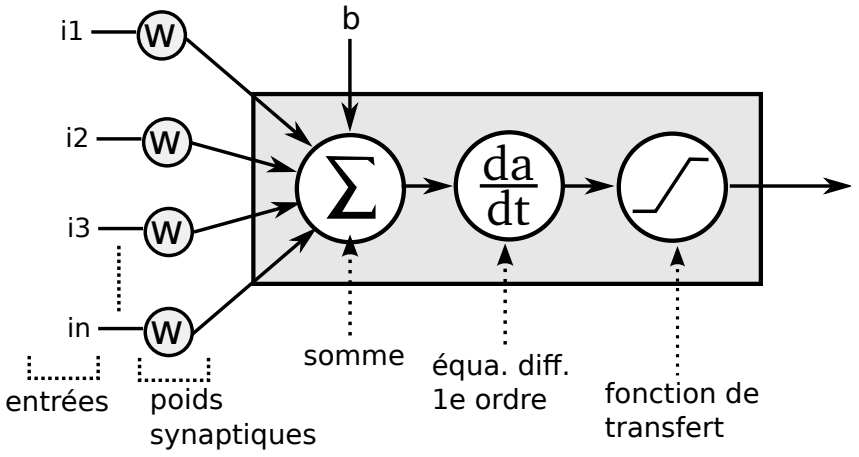




A short foreword about neurons & neuron models

Firing rate models :





- McCulloch et Pitts (1943) model (& variations)

$$y = \varphi(\sum_{j=1}^{n} w_j i_j + b)$$

$$\varphi(x) = \frac{2}{1 + \exp(-\alpha x)} - 1$$

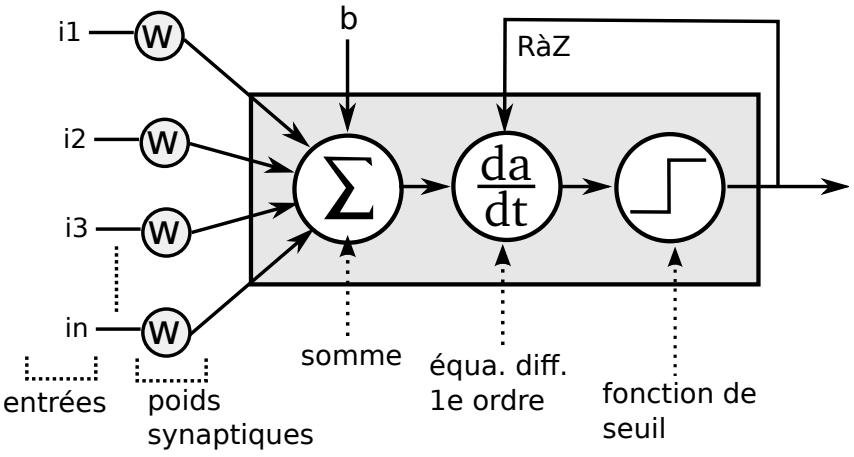
- Leaky-integrator

$$\tau \frac{da}{dt} = \left(\sum_{j=1}^{n} w_j i_j + b - a \right)$$

$$y = \varphi(a)$$

A short foreword about neurons & neuron models

Spiking models :

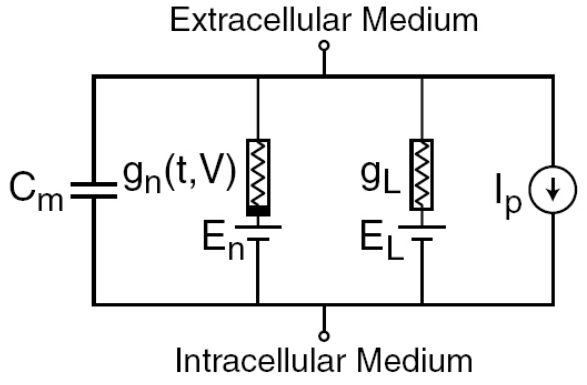


- Leaky integrate-and-fire (LIF)

$$\tau \frac{da}{dt} = \left(\sum_{j=1}^{n} w_j i_j + b - a \right)$$

if $a(t) > \theta \rightarrow a(t) = 0, y(t) = 1$

otherwise $y(t) = 0$



- Hodgkin & Huxley (1952)

$$\dot{V_m} = -\frac{1}{C_m}\left(\sum_{j=1}^{n} I_j\right)$$

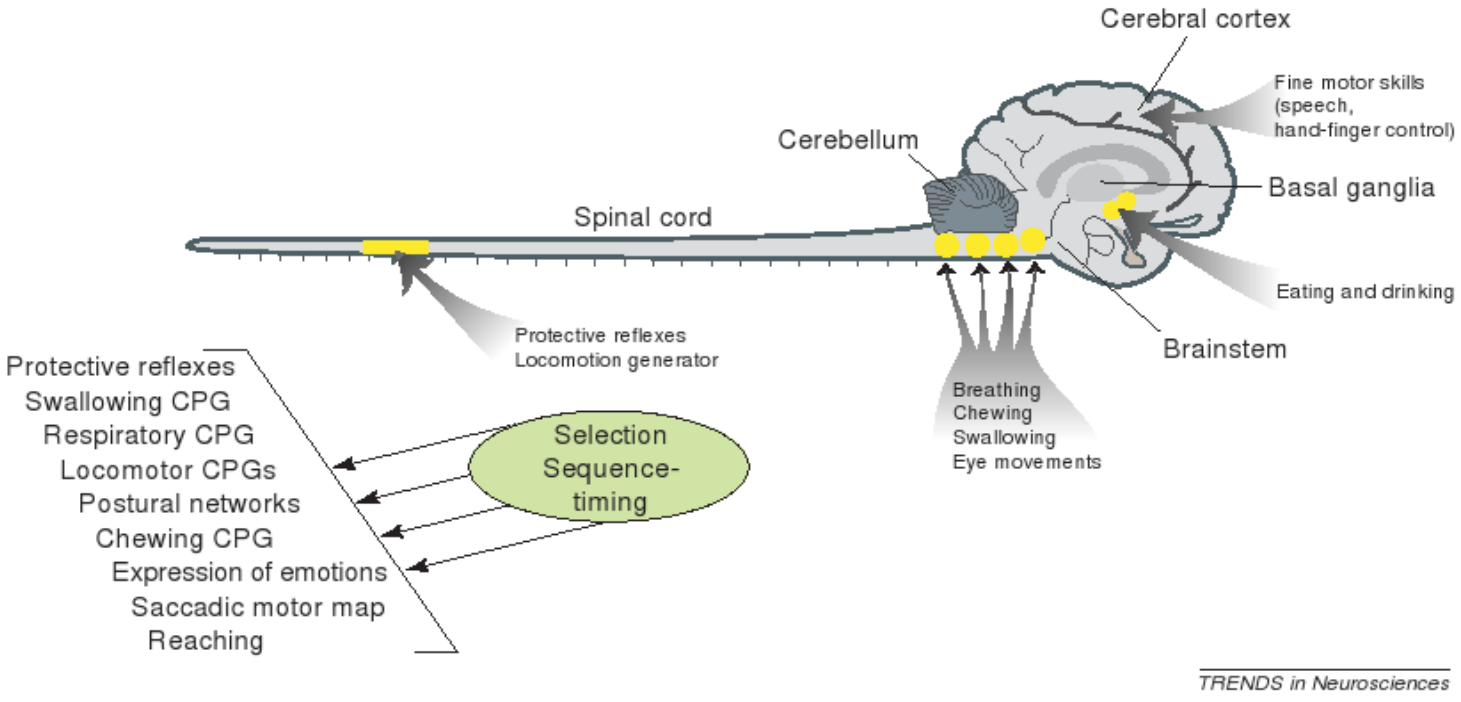
$$I_i(V_m,t) = (V_m - E_i) g_i$$

$$g_n(V_m) = \bar{g_n} \phi^{\alpha} \chi^{\beta}$$

$$\dot{\phi}(V_m) = \frac{(\phi_{inf} - \phi)}{\tau_\phi}$$

$$\dot{\chi}(V_m) = \frac{(\chi_{inf} - \chi)}{\tau_\chi}$$

Coordination du répertoire moteur des vertébrés

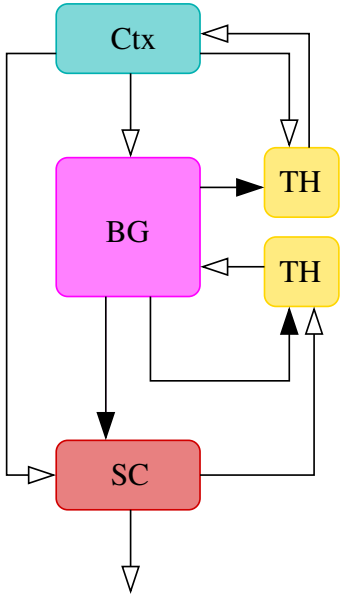


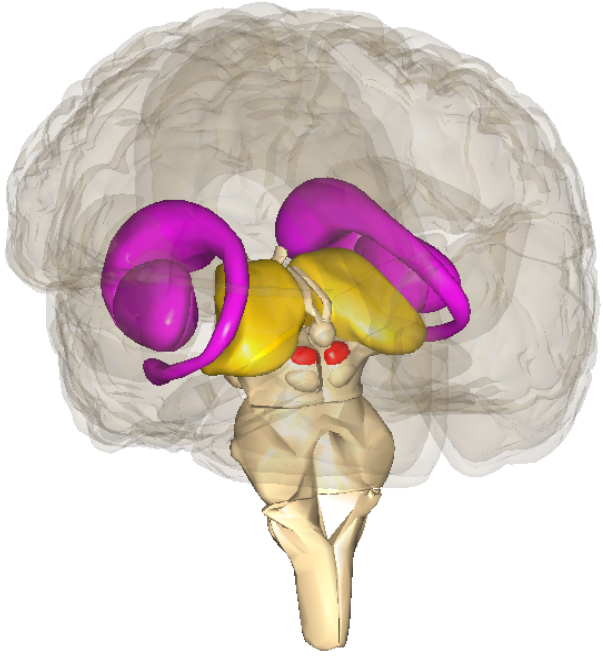
Répertoire moteur des vertébrés.

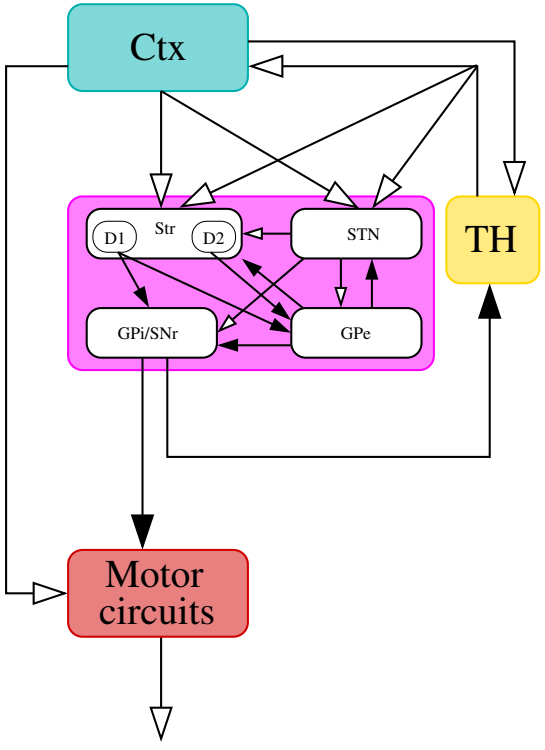
- Le cortex n'est pas indispensable pour générer des comportements moteurs variés et efficaces (chats corticaux).
- Les ganglions de la base semblent jouer le rôle de coordinateurs.

(Grillner et al, 2005)

Action Selection in the brain : the basal ganglia

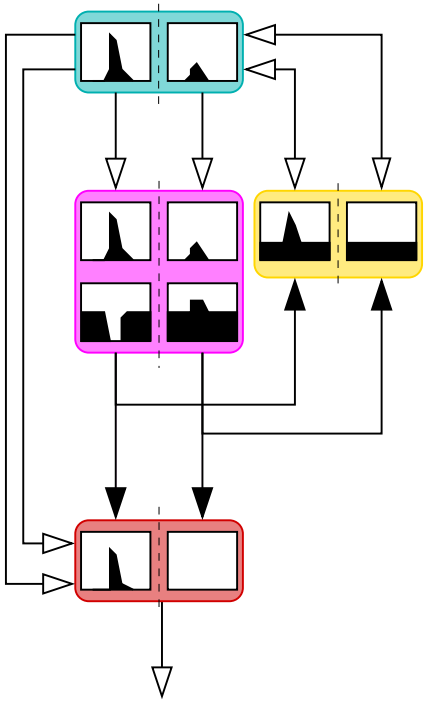


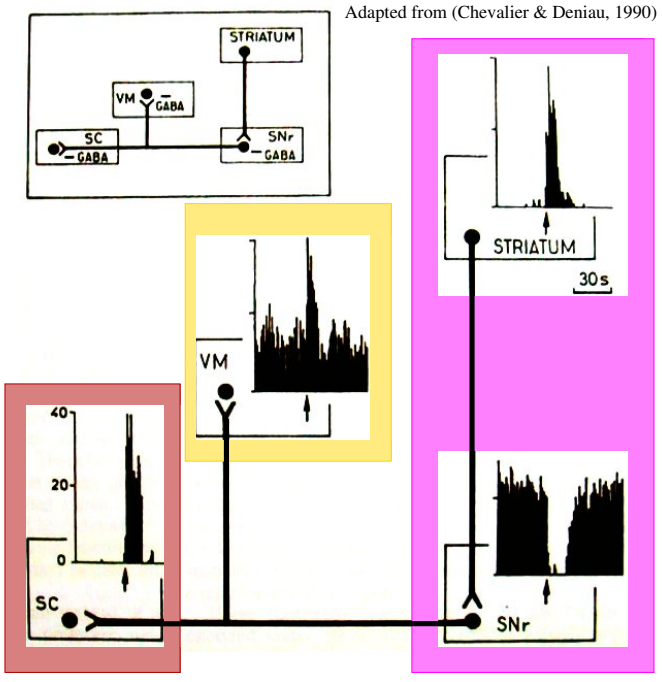




- interconnected subcortical nuclei, common to all vertebrates,
- forming multiple loops with the cortex,
- but also subcortical targets (superior colliculus, cerebellum, etc.)

The basal ganglia – Selection

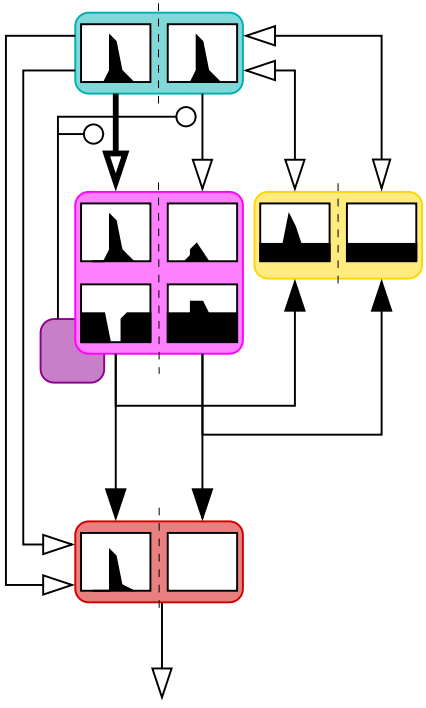


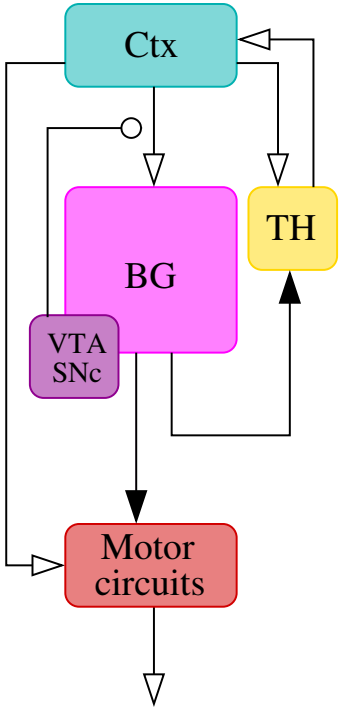


Generic selection role

- Selection by disinhibition,

The basal ganglia – Selection

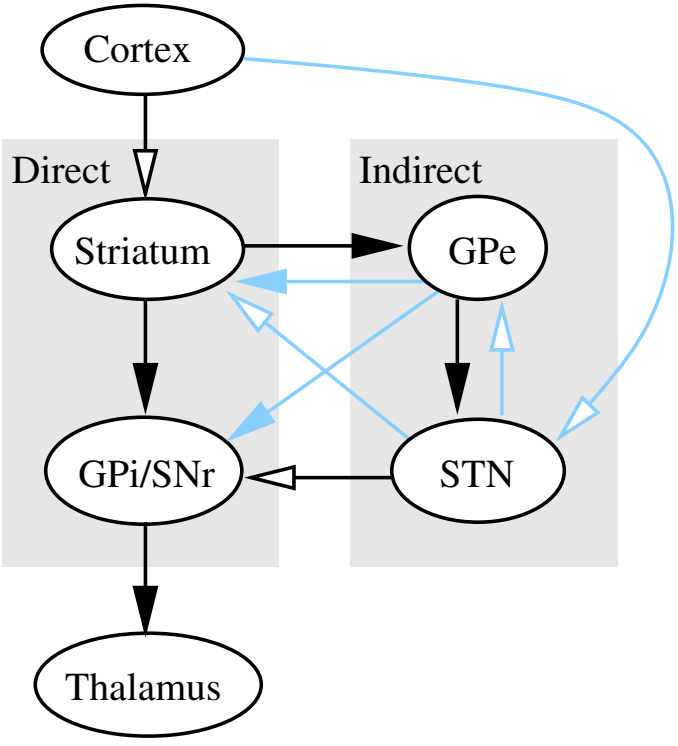


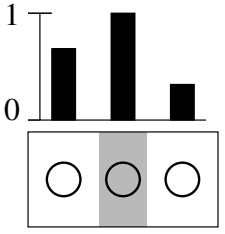


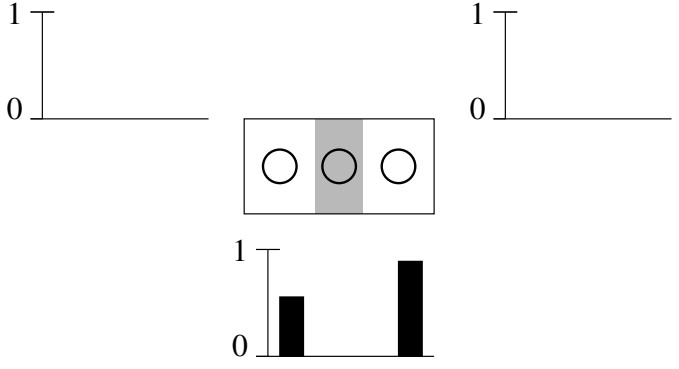
Generic selection role

- Selection by disinhibition,
- Reinforcement Learning
  $\Rightarrow$ Cortico-striatal plasticity, under dopaminergic control

Direct / indirect pathways

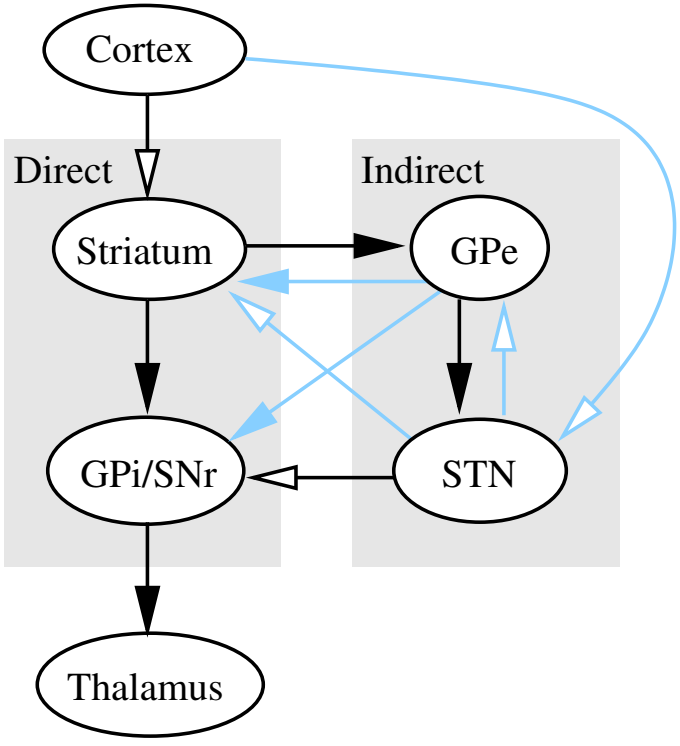


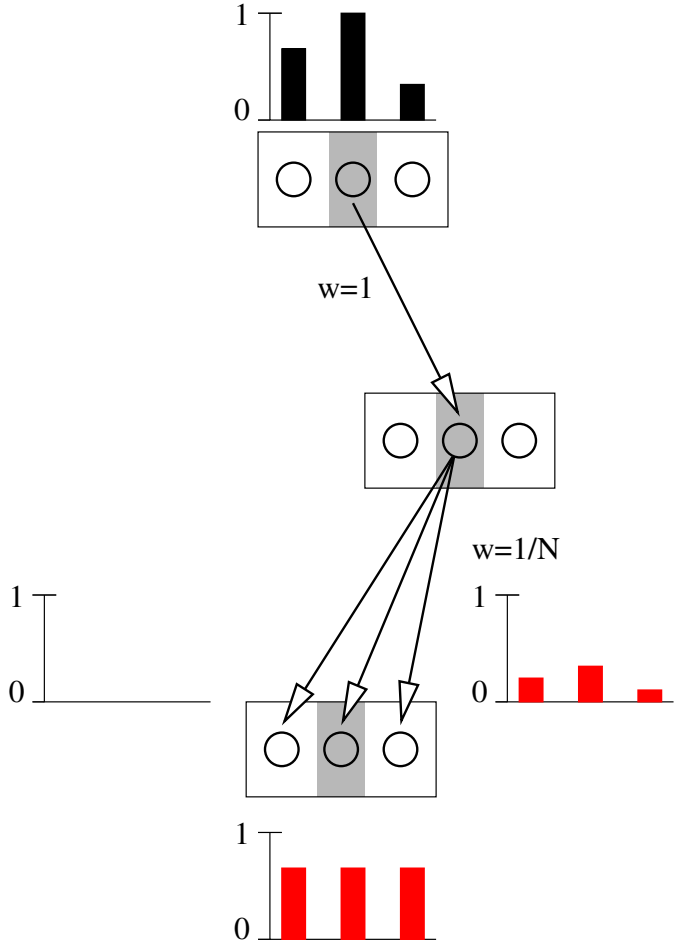




(Albin et al. 1989)
(Berns & Sejnowski, 1996)

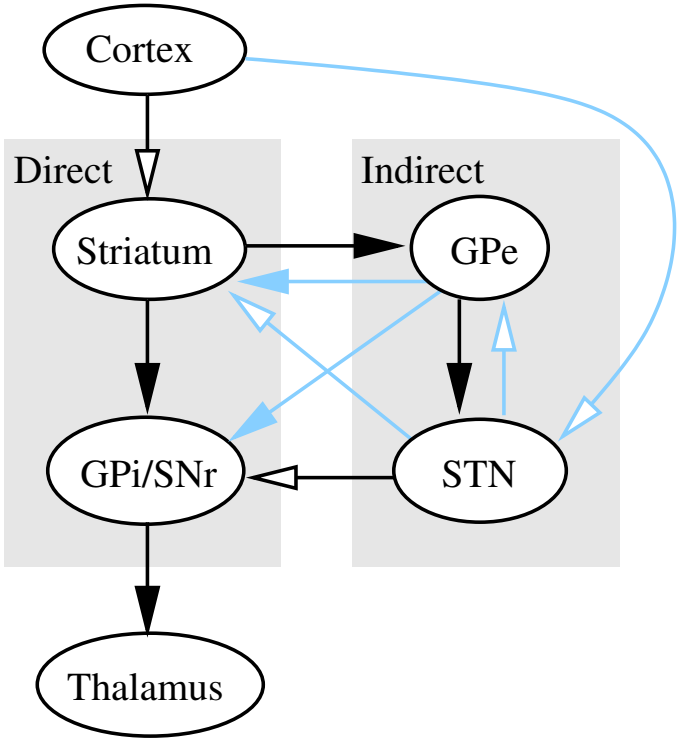
Direct / indirect pathways



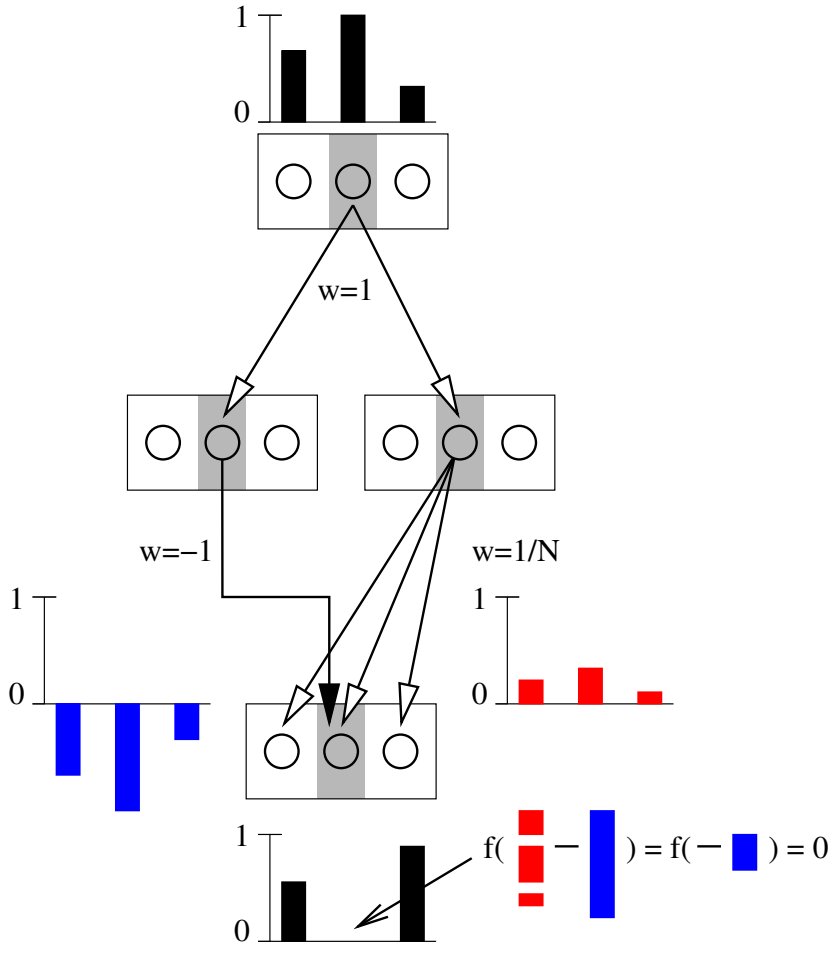


(Albin et al. 1989)
(Berns & Sejnowski, 1996)

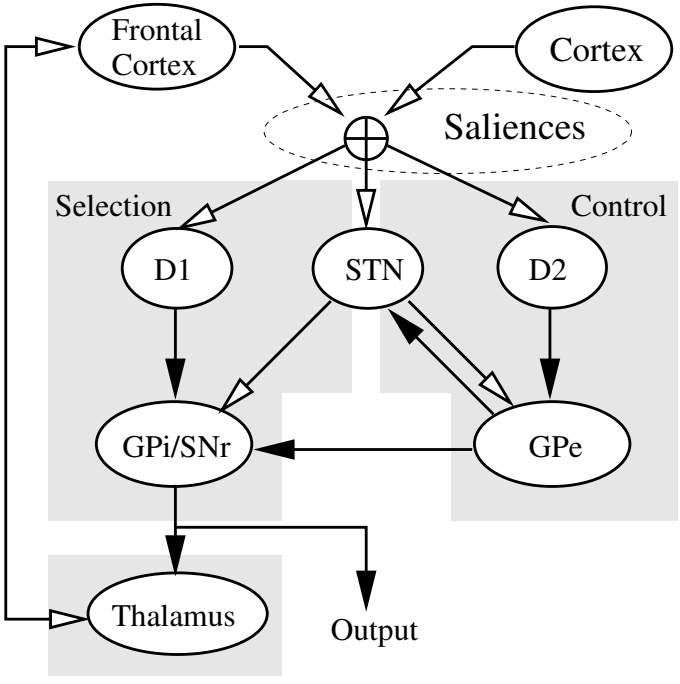
Direct / indirect pathways

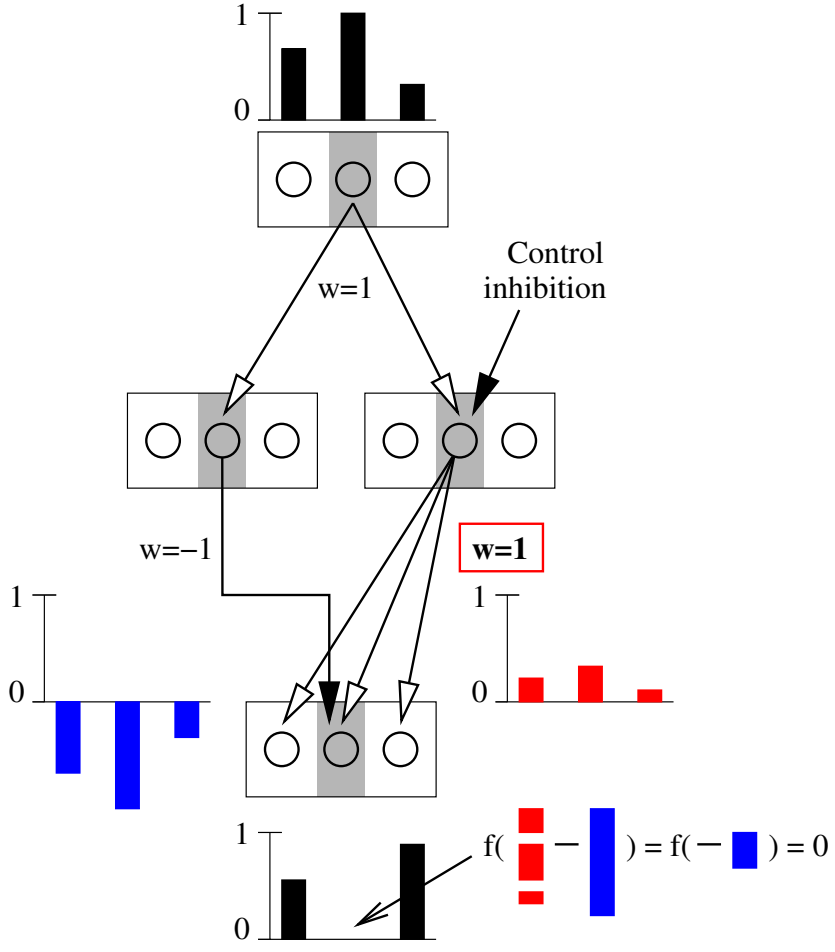


(Albin et al. 1989)
(Berns & Sejnowski, 1996)

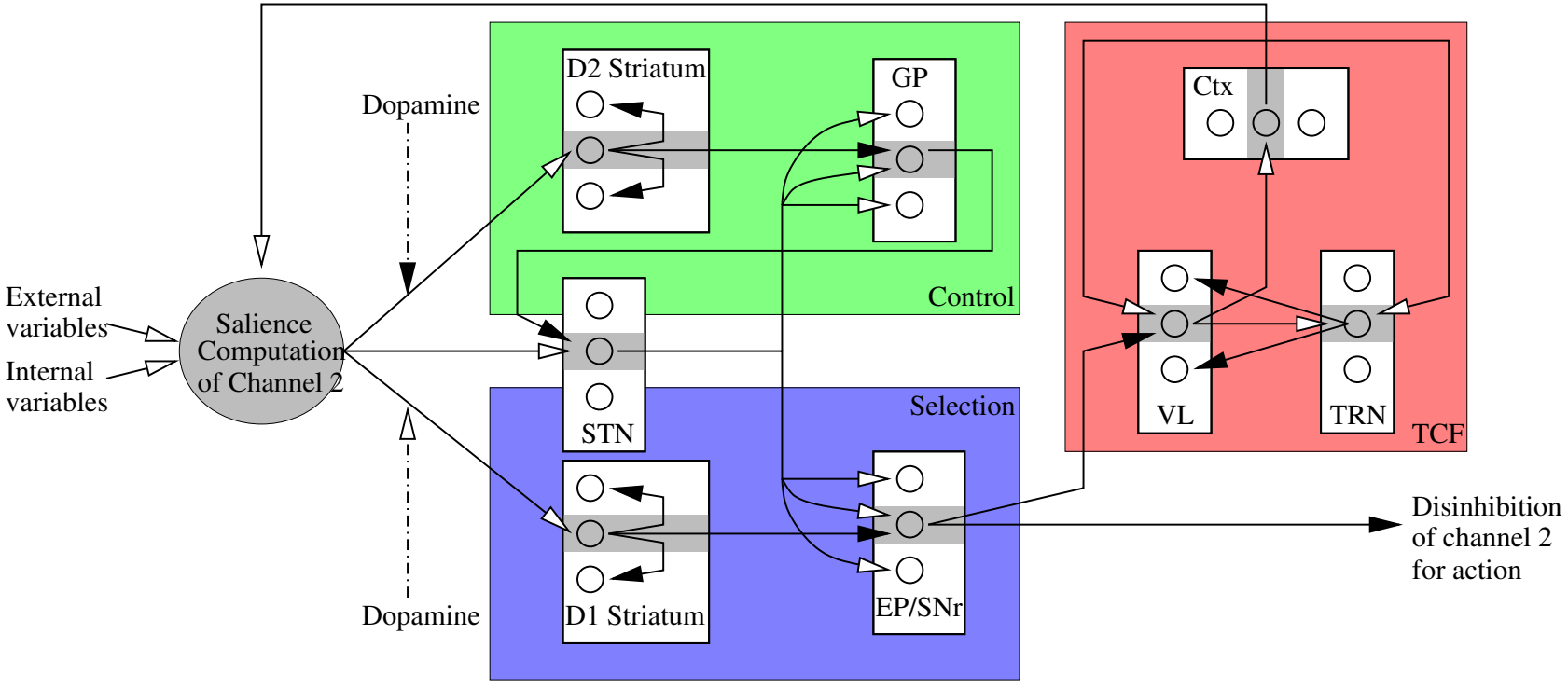


GPR model (Gurney et al., 2001)

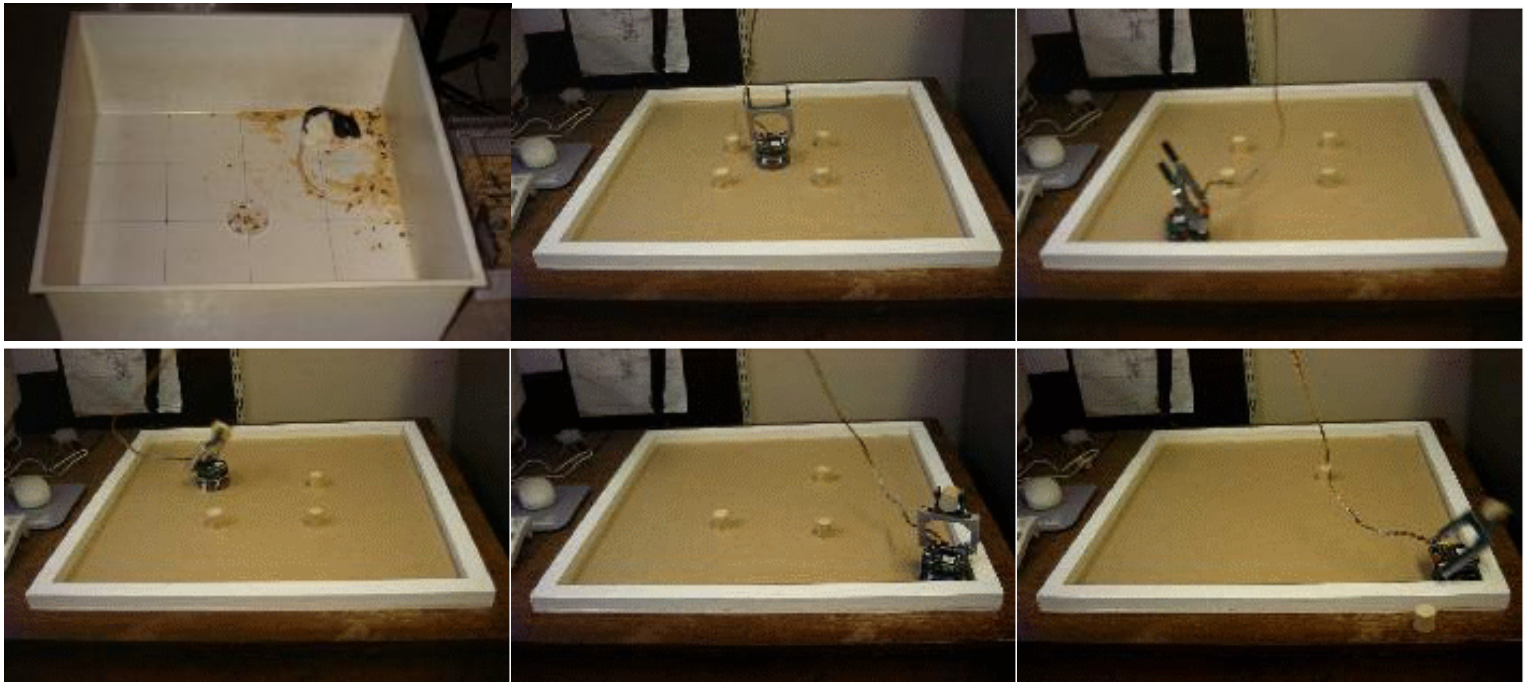




Selection, control and feedback circuits



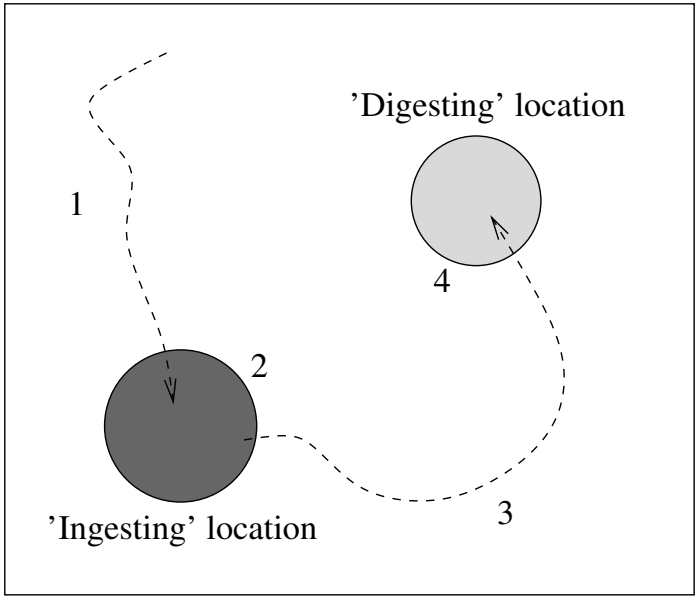
Robotic evaluation of the GPR

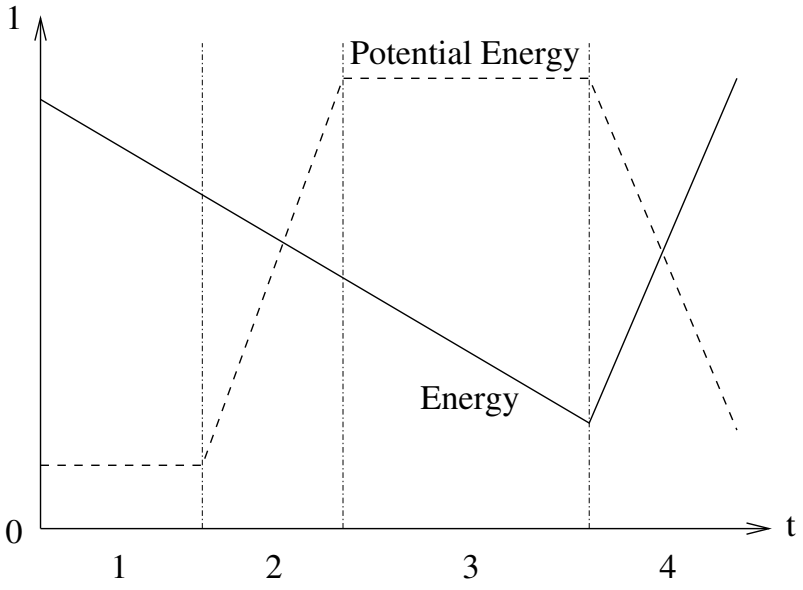


Qualitative imitation of the rat behavior.

(Montes-Gonzalez et al., 2000 ; Prescott et al., 2006)

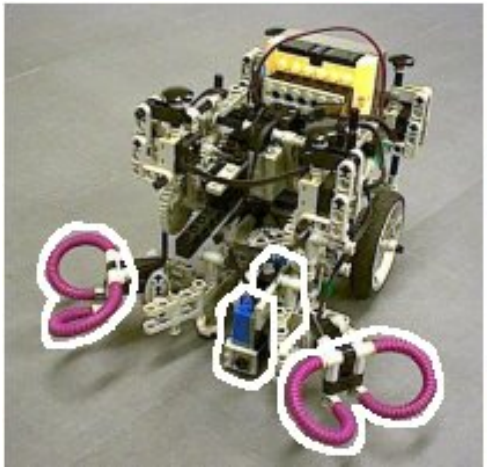
Minimal survival task (Girard et al., 2003)

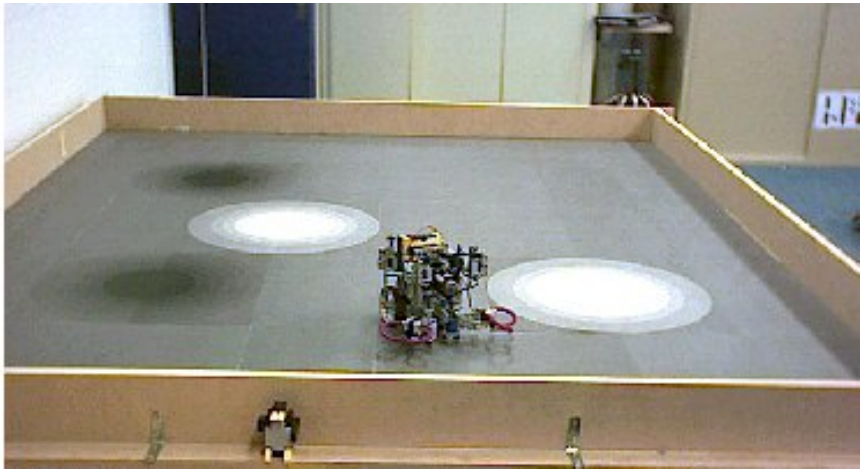




- 2 internal variables : Energy (E) et Potential Energy (E_p),
- Continuous energy consumption : $E = 0 \rightarrow$ end of the trial.
- Ingestion zones : reload E_p,
- Digestion zones : transforms the stored E_p into E.

Robot & environment





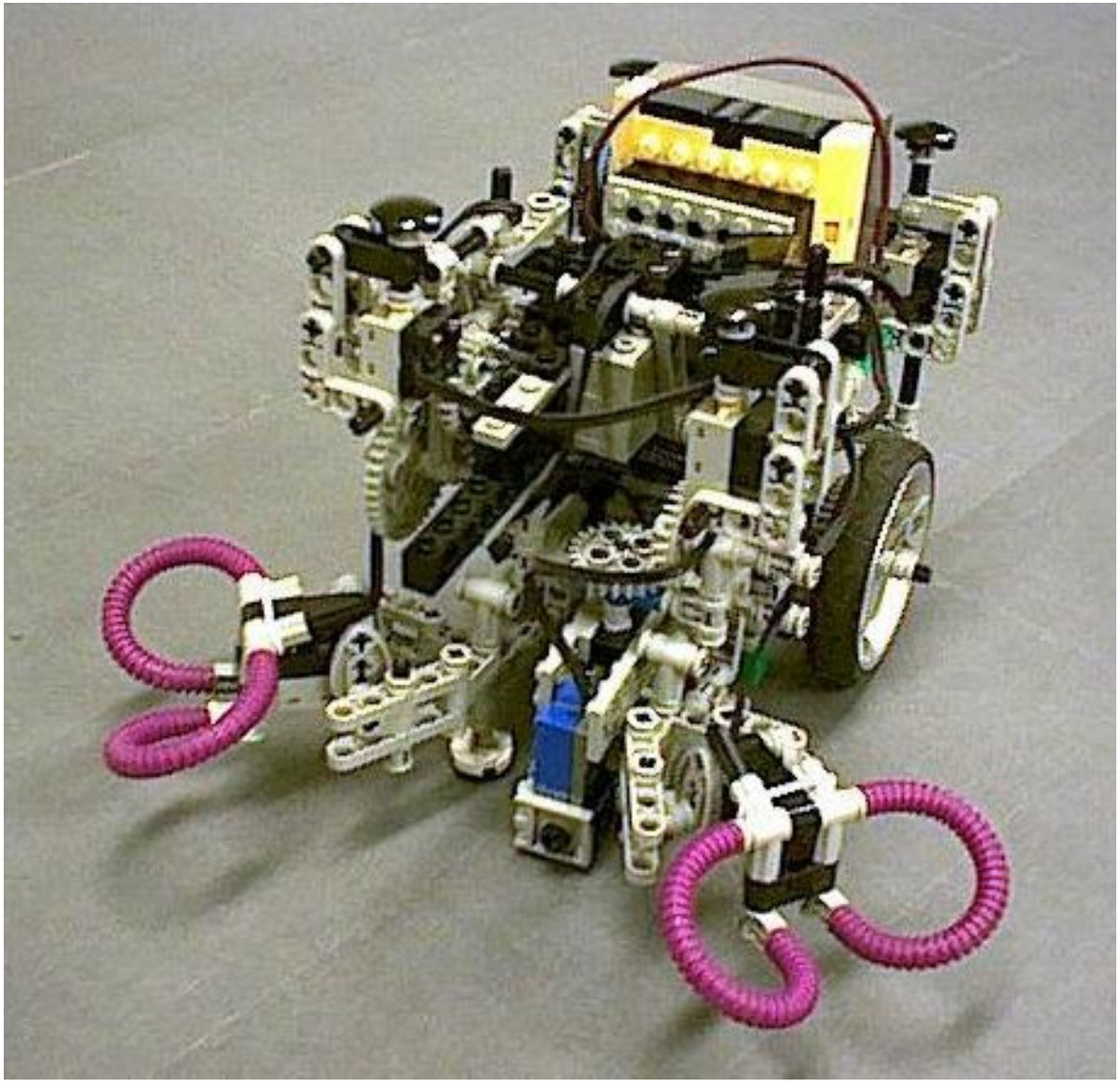
Behaviors

- Exploration
- ReloadOnBlack
- ReloadOnWhite
- AvoidObstacle
- (Rest)

Variables

- Darkness
- Brightness
- Left / right contact
- E and E_p

Minimal survival task (Girard et al., 2003)



Evaluation en survie du GPR (Girard et al., 2003)

Comparaison avec un algorithme "Winner-takes-all" élémentaire

- calcul des **saliences** sans rétroaction (senseurs, E_p et E),
- choix du comportement de plus forte **salience**

Evaluation en survie du GPR (Girard et al., 2003)

Comparaison avec un algorithme "Winner-takes-all" élémentaire

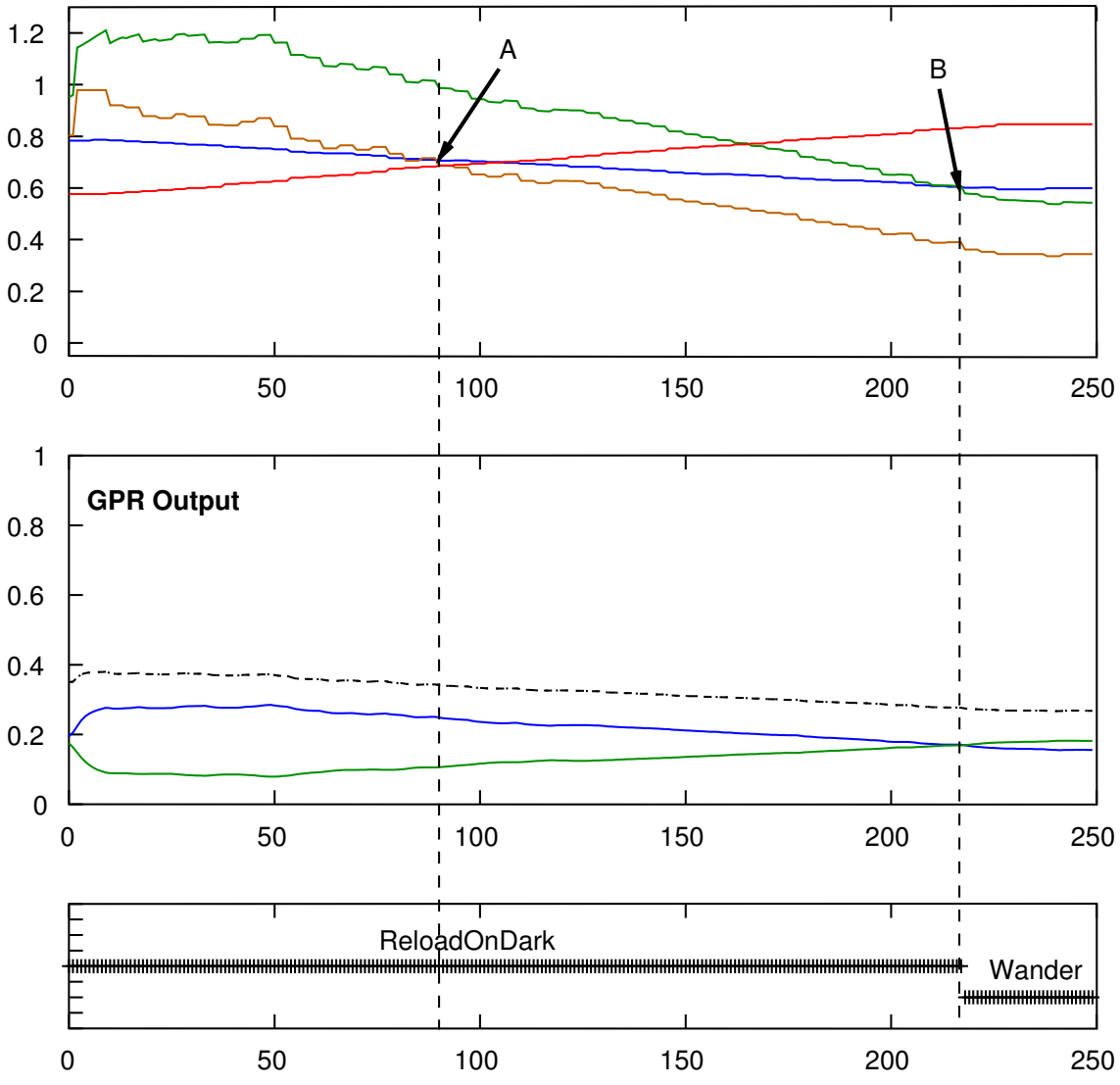
- calcul des **saliences** sans rétroaction (senseurs, E_p et E),
- choix du comportement de plus forte **salience**

Résultats

- WTA et GPR résolvent la tâche,
- Oscillation comportementales du WTA,
- Recharge du WTA significativement moins longues et plus fréquentes,
- GPR qualitativement dans une meilleure zone de confort de sa variable E_p.

Interprétation : exploitation de la persistance des comportements de recharge grâce à la boucle de rétroaction.

Effet de la rétroaction positive

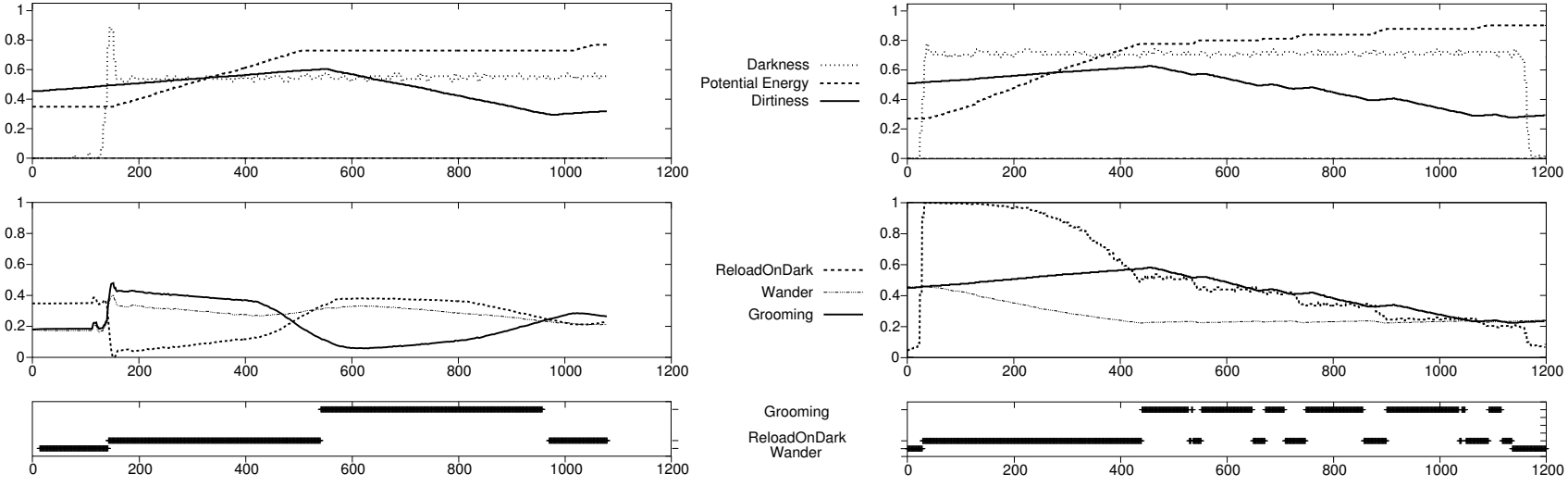


Saliences :
- Wander
- ReloadOnDark
- ROD no persistence

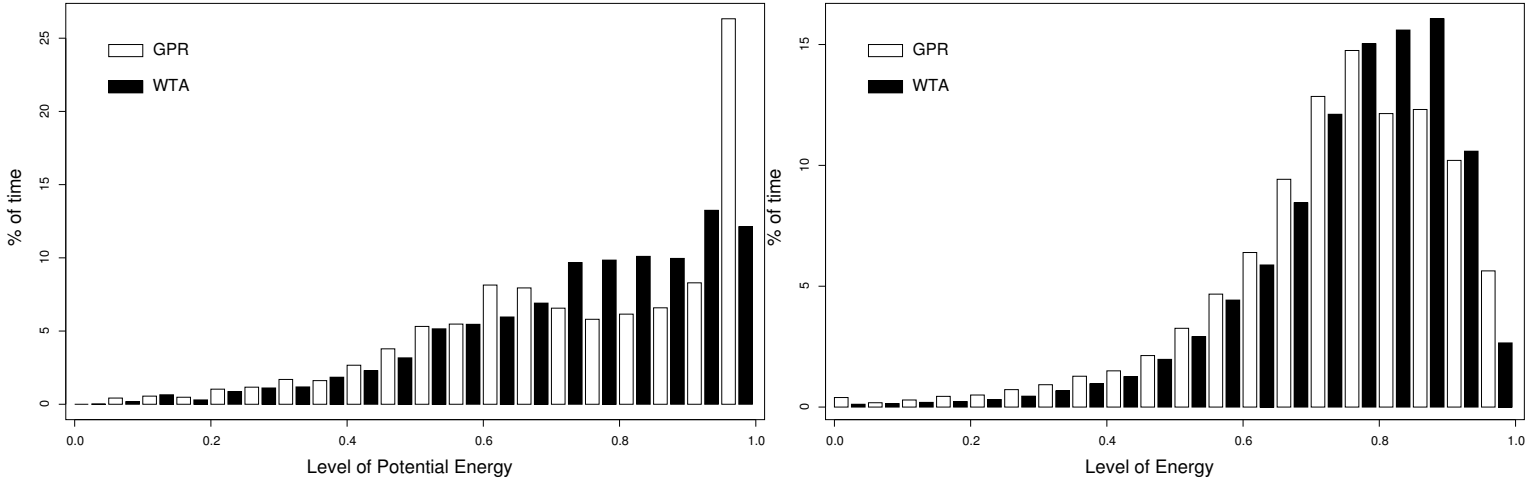
Potential Energy

- Wander
- ReloadOnDark
- ReloadOnBright
- AvoidObstacle

Oscillations comportementales



Zone de confort E_p



Exploitation des zones de confort

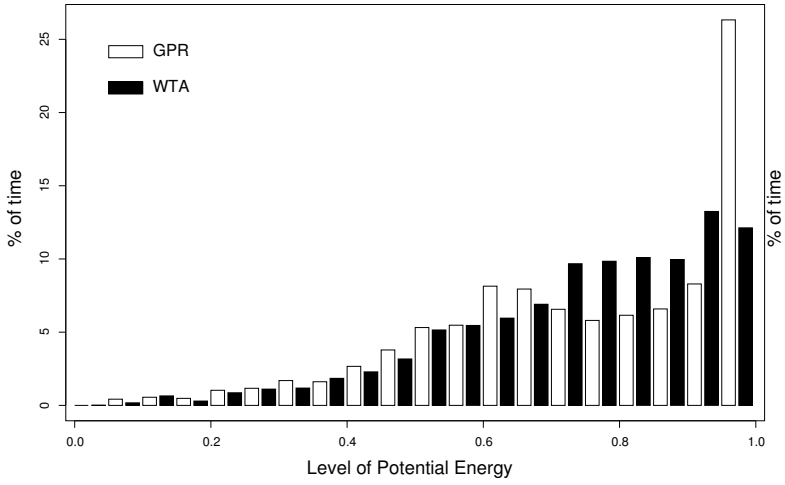
Ajout comportement repos :

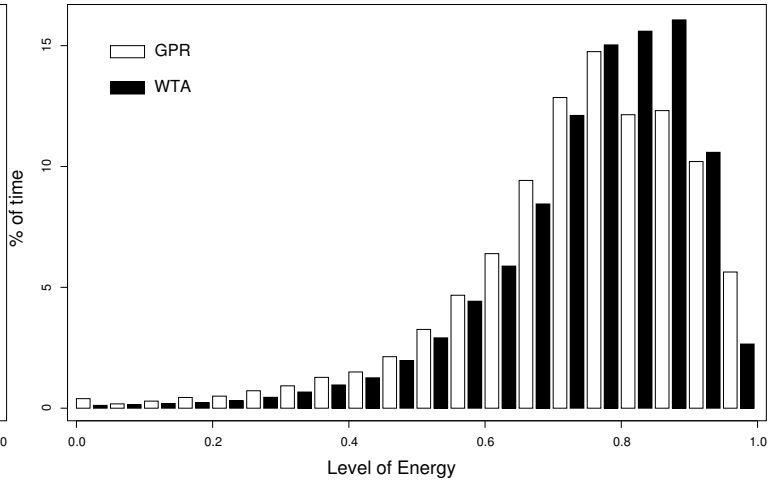
- Robot immobile,
- Consommation d'Energie divisée par deux.

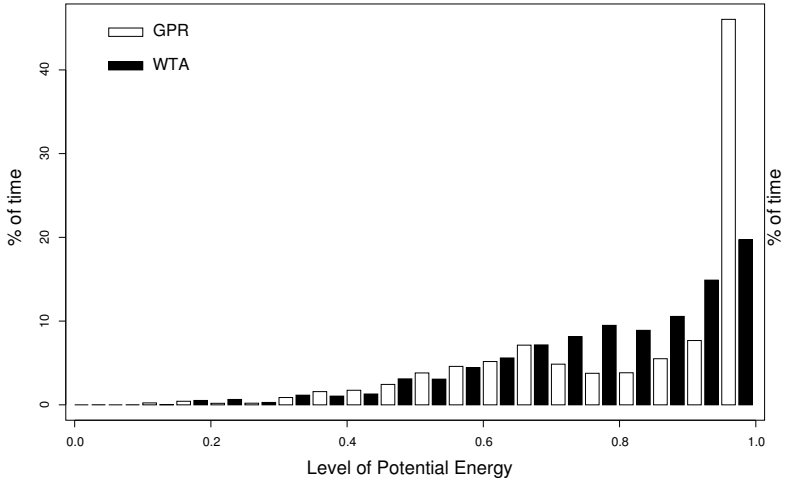
Résultats :

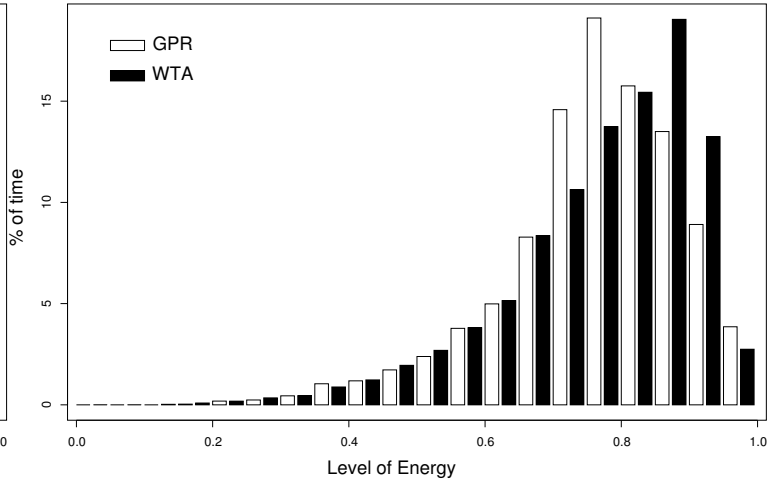
- Activations de Repos significativement plus longues pour le GPR,
- GPR extrait moins d'énergie de l'evt (20%).

Zone de confort E_p

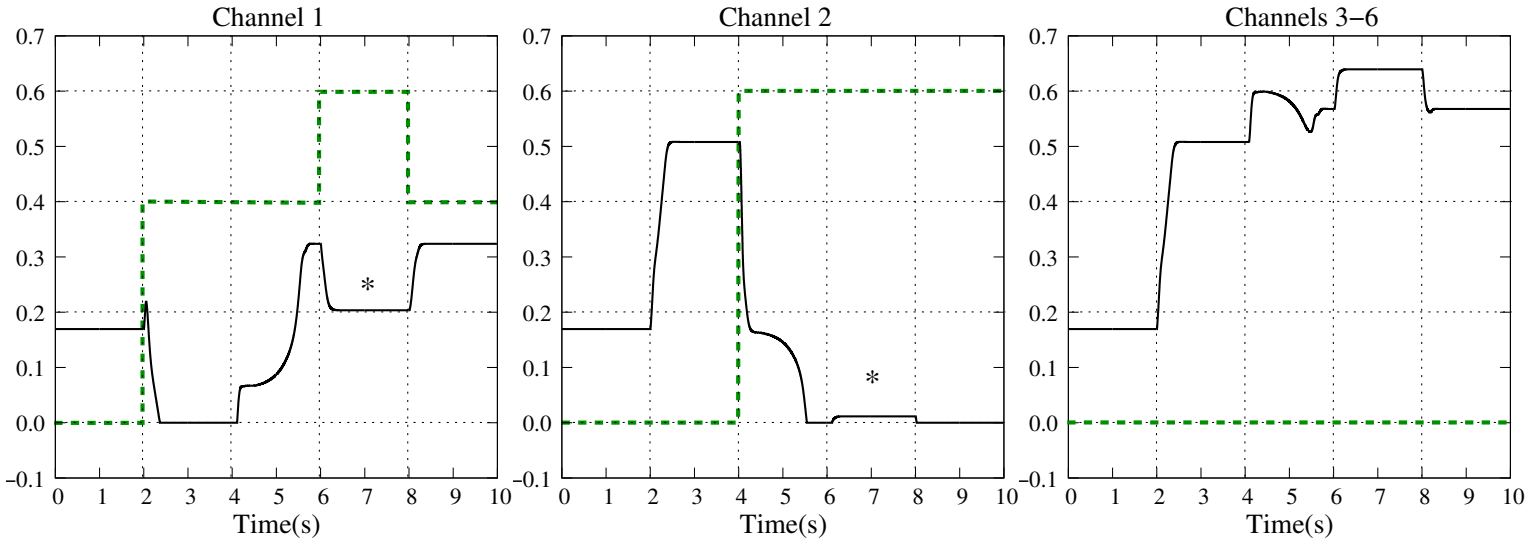






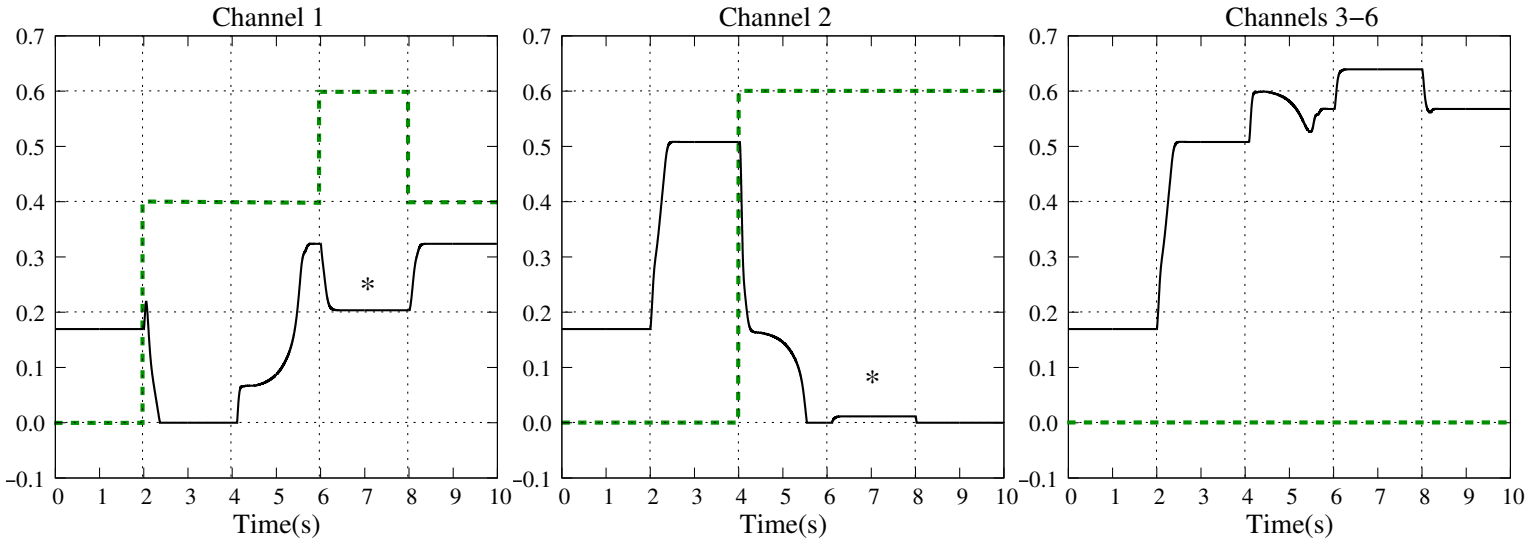


GPR Test



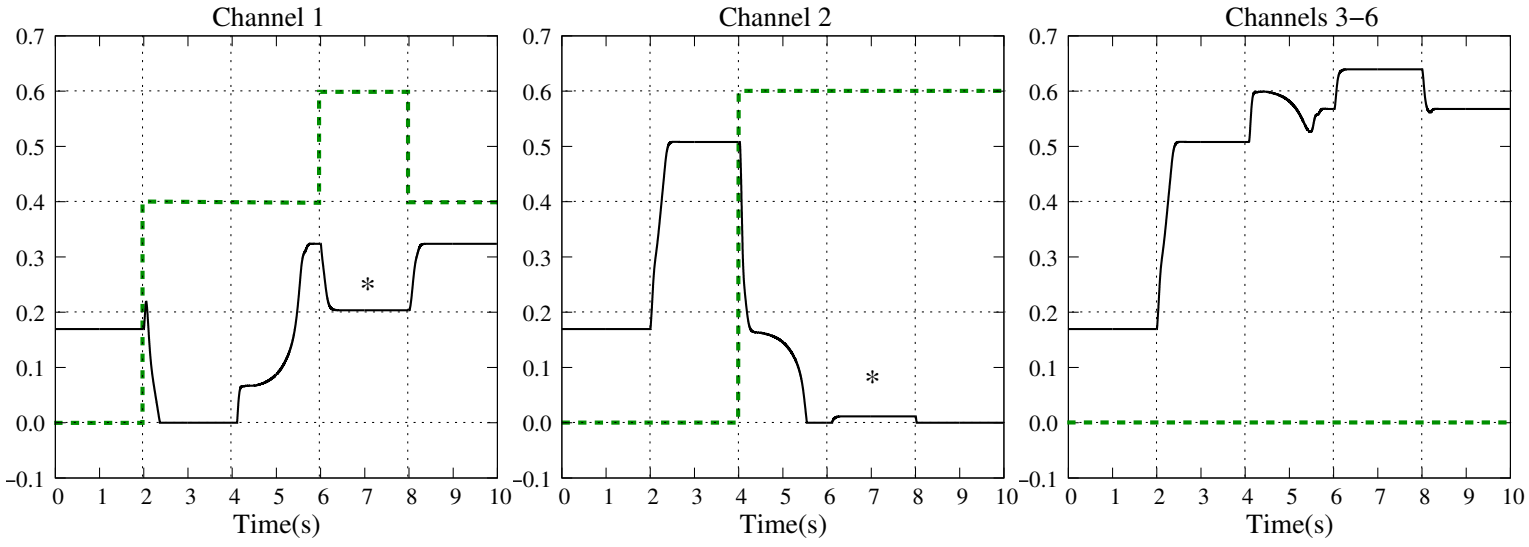
With random input saliences :
the « selected » is more inhibited than at rest in 30% of the cases.
From an engineering point of view, the model does not operate well.

GPR Test



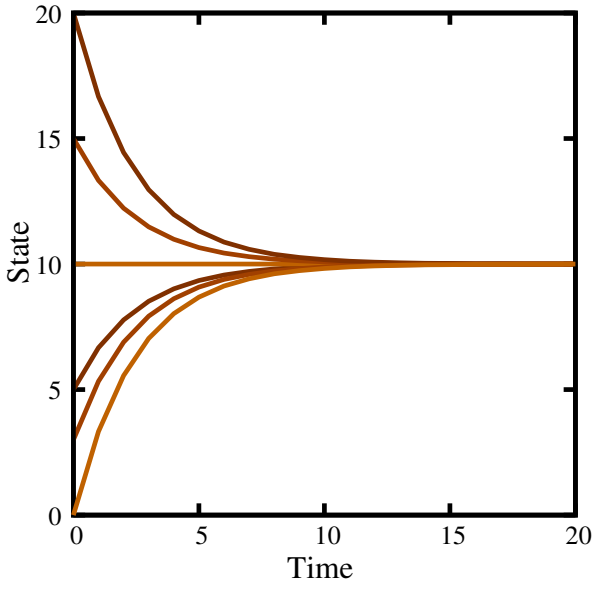
With random input saliences :
the « selected » is more inhibited than at rest in 30% of the cases.
From an engineering point of view, the model does not operate well.

GPR Test



With random input saliences :
the « selected » is more inhibited than at rest in 30% of the cases.
From an engineering point of view, the model does not operate well.

Contraction



Contraction

Consider a system of the following form :

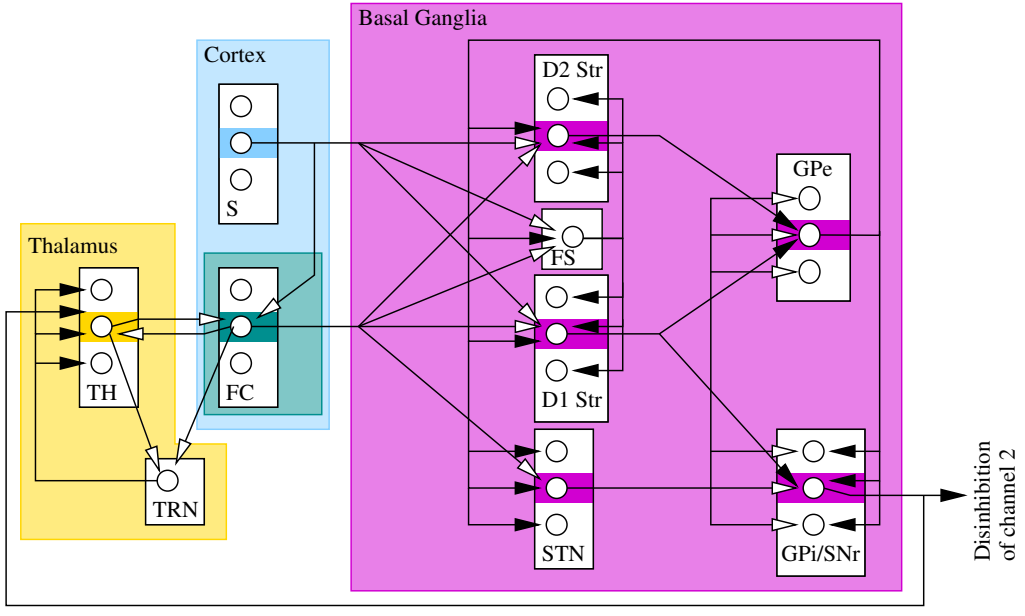
$$\dot{x}(t) = f(x(t), t)$$

with $x \in \mathbb{R}^n$ et $t \in \mathbb{R}_+$.
It forgets its initial conditions exponentially fast if and only if the Jacobian of f is uniformly definite negative in a given metric.
Such a system is **contracting**.

(Lohmiller et Slotine, 1998)

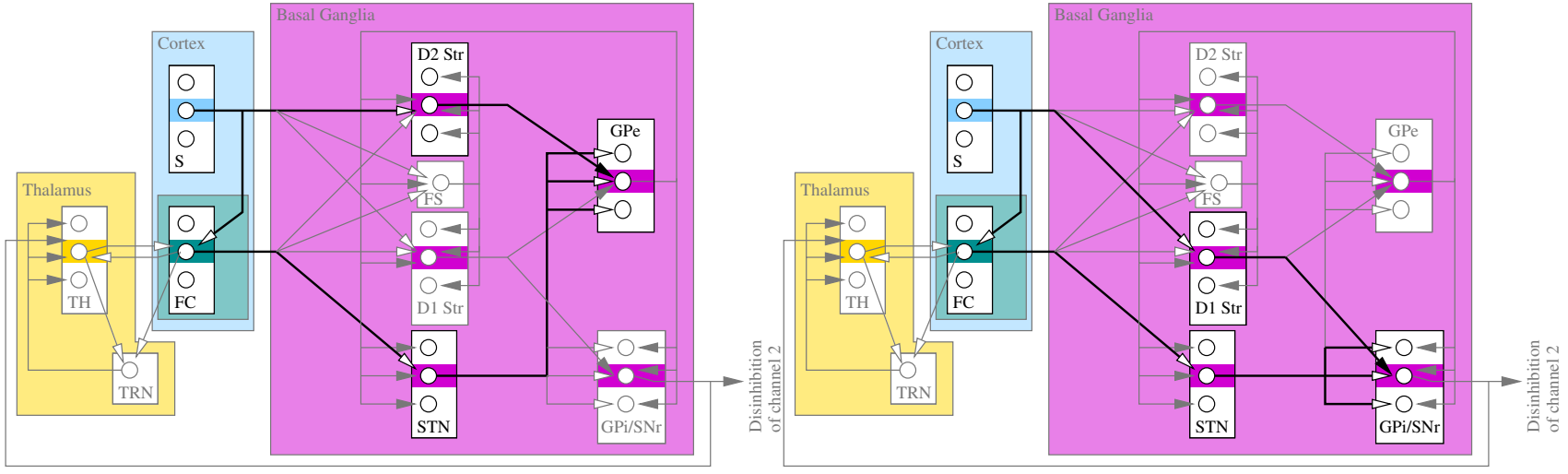
The combination of contracting systems preserves contraction (as long as a number of constraints are verified)
$\Rightarrow$ Contracting rate neuron model based on lPDS (**locally Projected Dynamical Systems**).

Contracting basal ganglia model (Girard et al., 2008)



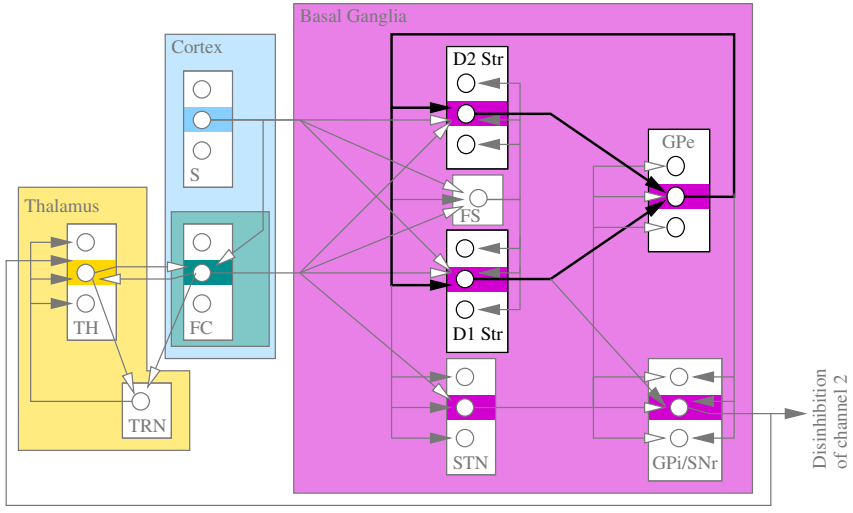
- Integrates connexions ignored by previous models.
- Amplificatory thalamo-cortical loop.
- Selective amplification in the frontal cortex, selection by strict disinhibition in the subcortical targets,
- Numerous internal loops : is it stable ?

Structure



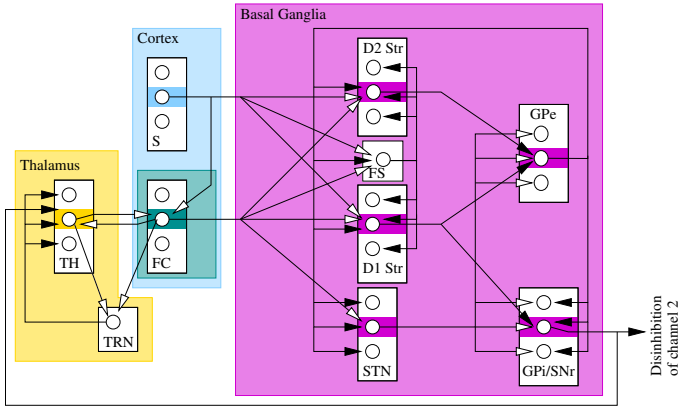
- « Off-center on-surround » circuits,
- Negative « feedback » amplifying the contrast,
- Contraction conditions defined analytically (BG and TC),
- Numeric check of the contraction of the whole system.

Structure



- « Off-center on-surround » circuits,
- Negative « feedback » amplifying the contrast,
- Contraction conditions defined analytically (BG and TC),
- Numeric check of the contraction of the whole system.

Contraction



BG module :

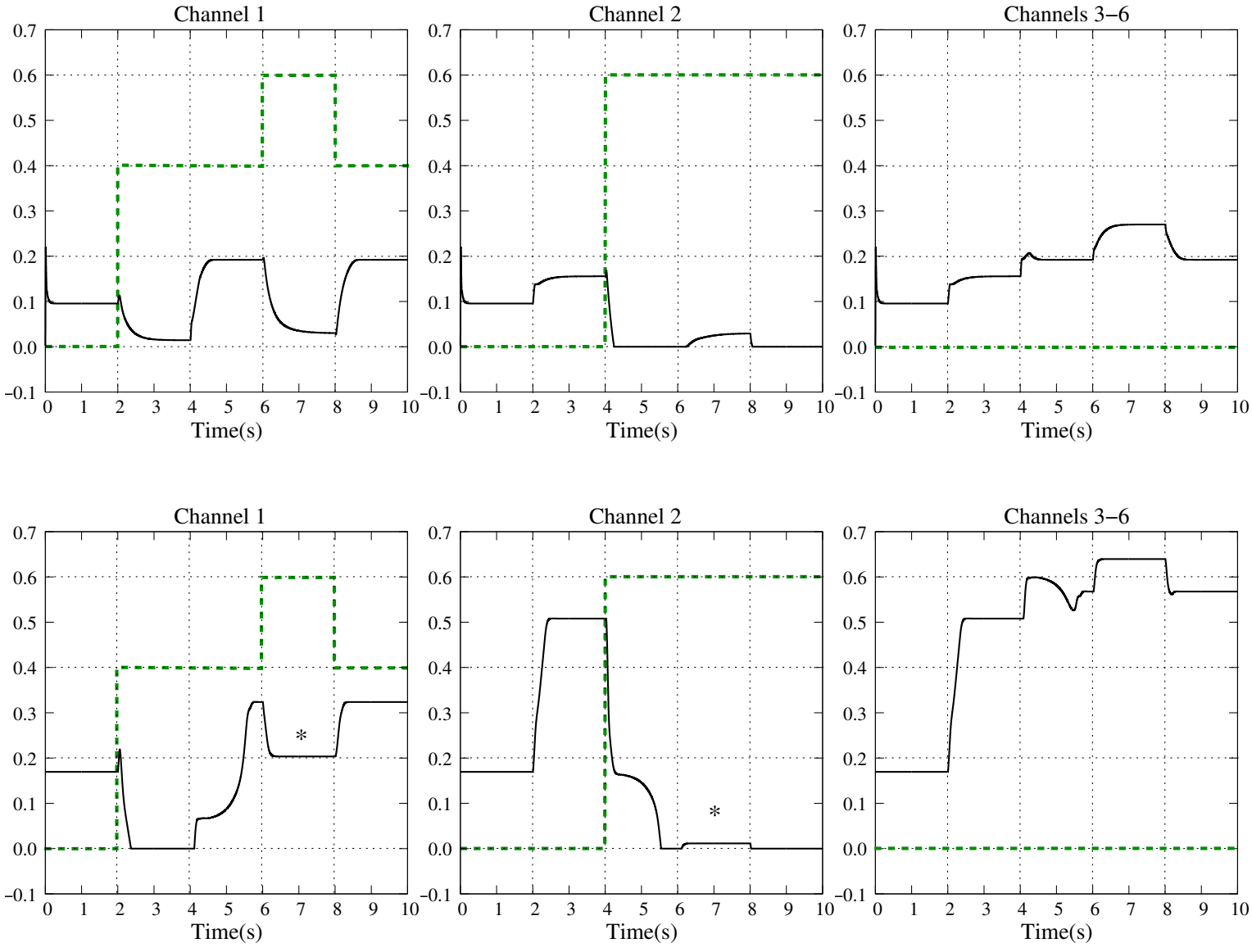
$$\left((1+\gamma)w_{D1}^{GPe}w_{GPe}^{D1}\right)^2+\left((1-\gamma)w_{D2}^{GPe}w_{GPe}^{D2}\right)^2<1$$

TH-FC module, if $w_{TH}^{TRN}=_{TRN}^{TH}$:

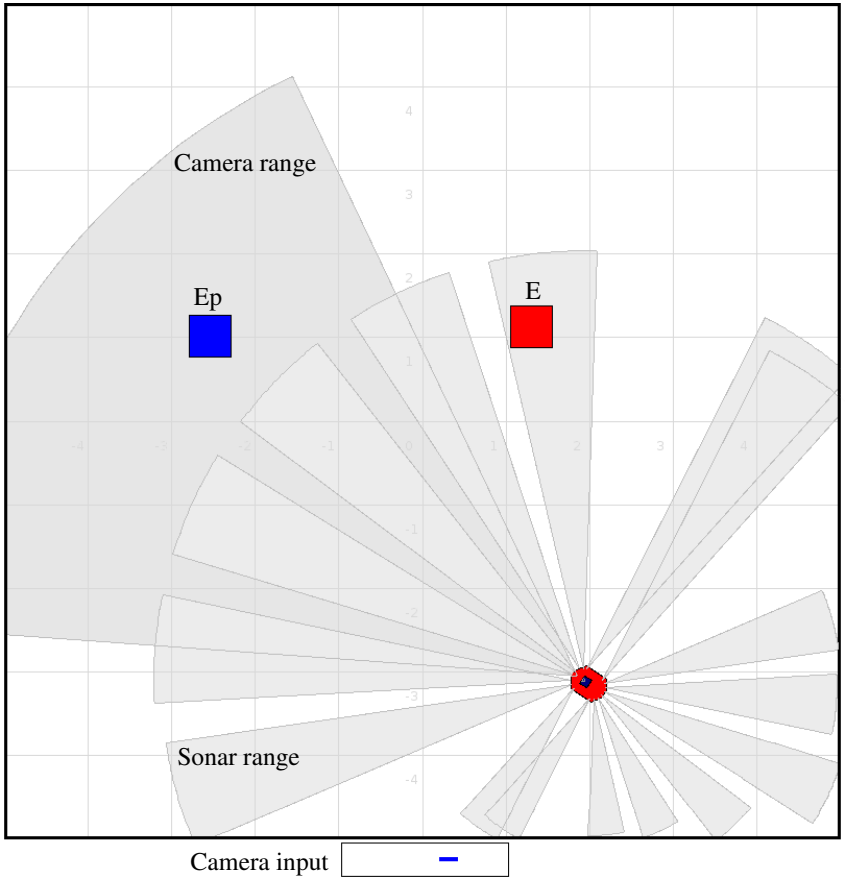
$$w_{TH}^{FCtx}(_{FCtx}^{TH}+\sqrt{{w_{FCtx}^{TH}}^2+N w_{FCtx}^{TRN^2}})<1$$

- « Off-center on-surround » circuits,
- Negative « feedback » amplifying the contrast,
- Contraction conditions defined analytically (BG and TC),
- Numeric check of the contraction of the whole system.

Offline model evaluation



Minimal survival task

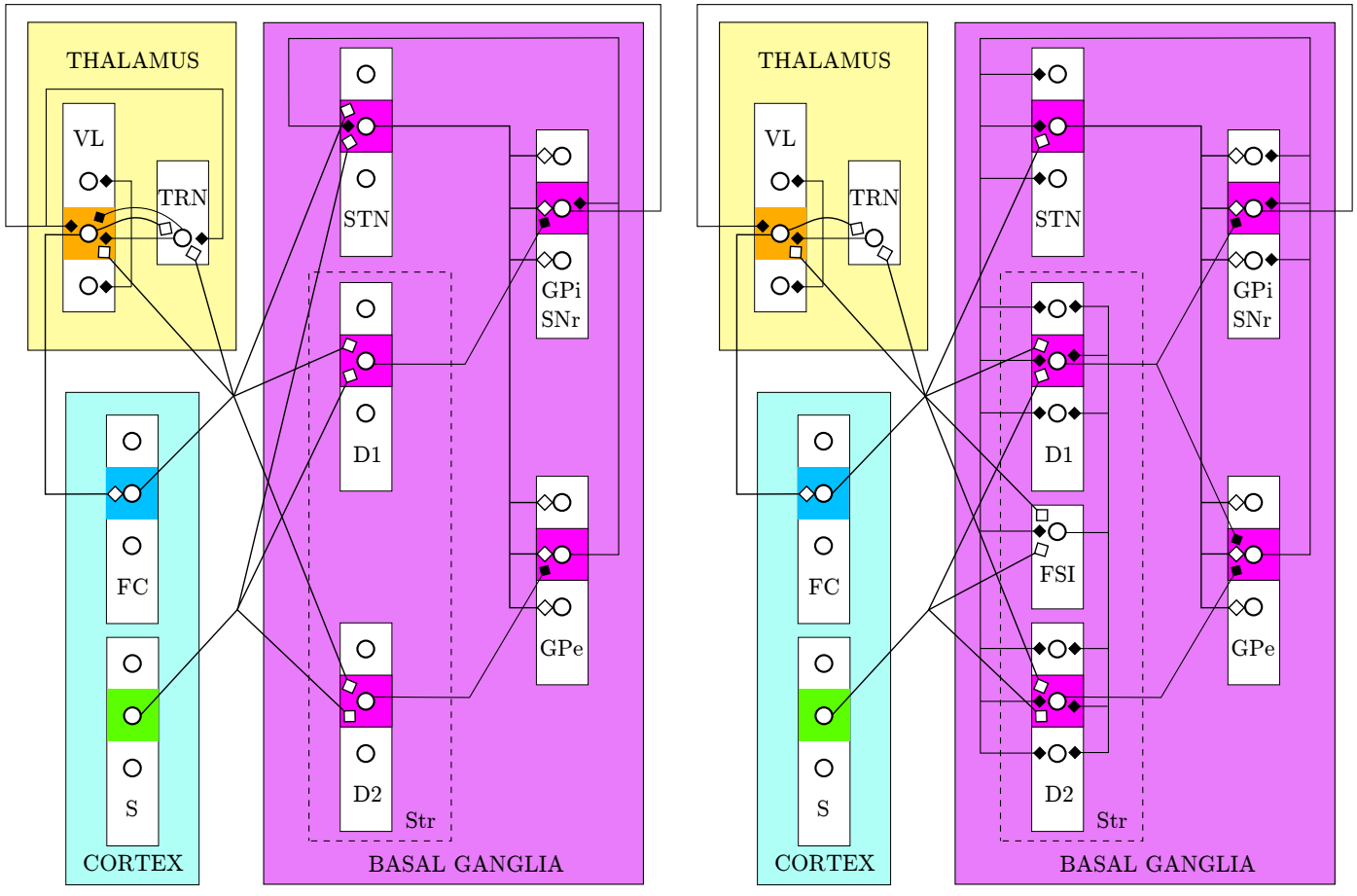


Comparison if-then-else

- limitation of the behavioral oscillations,
- higher energy stocks,
- reduced energy consumption.

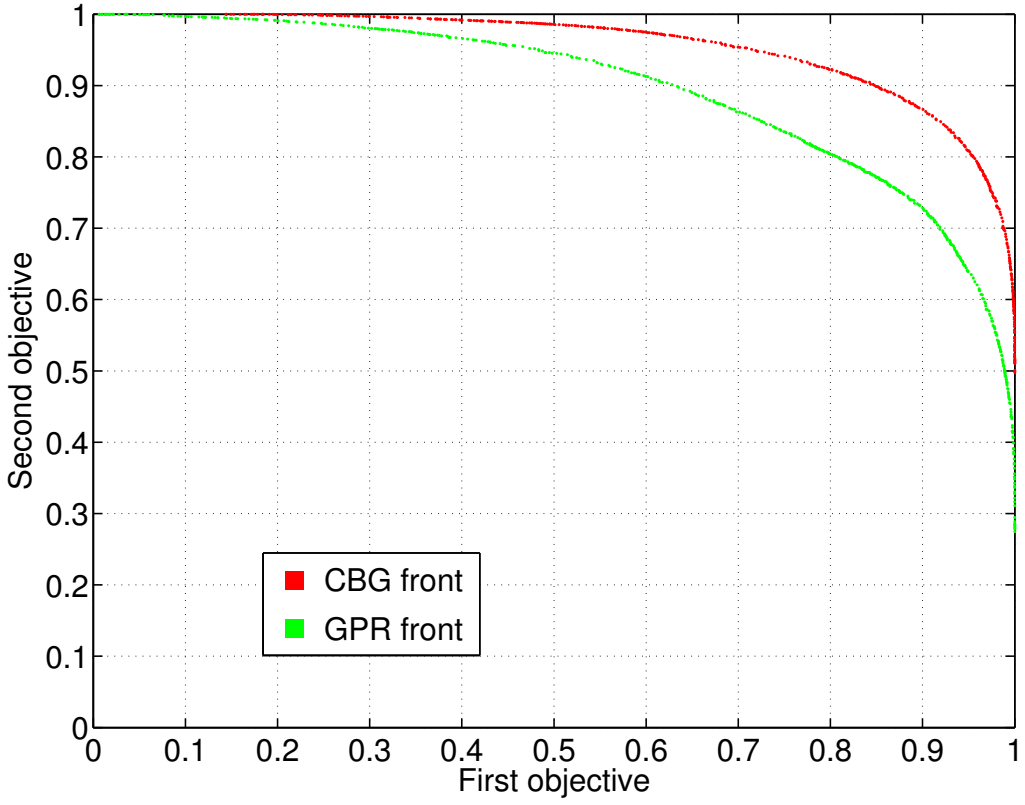
⇒ The "memory" of a GPR is not necessary to keep the good properties.

Comparaison GPR-CBG par évolution artificielle



Evolution multiobjectif des paramètres des modèles.

Comparaison GPR-CBG par évolution artificielle

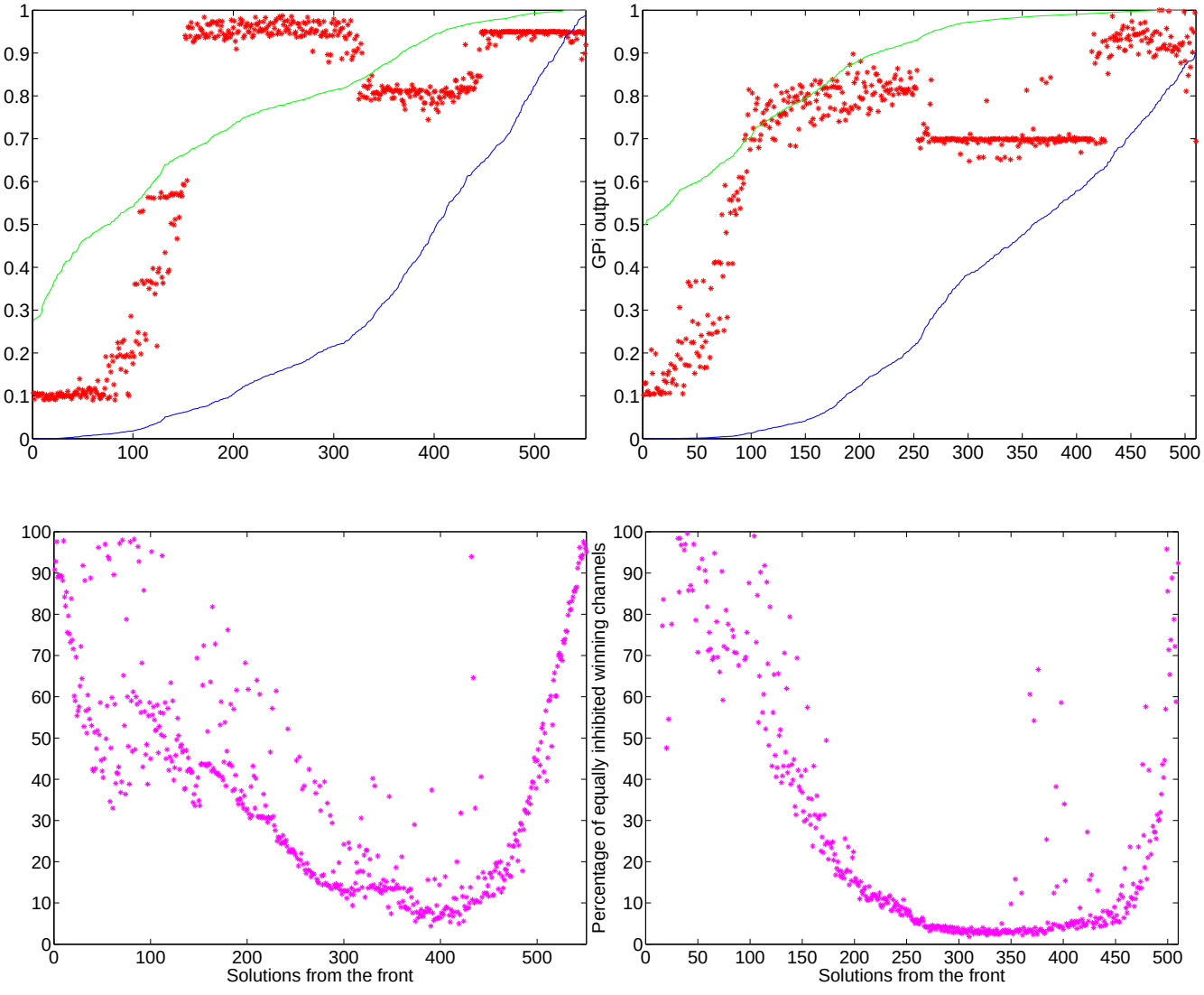


f1 : desinhibition du plus fort

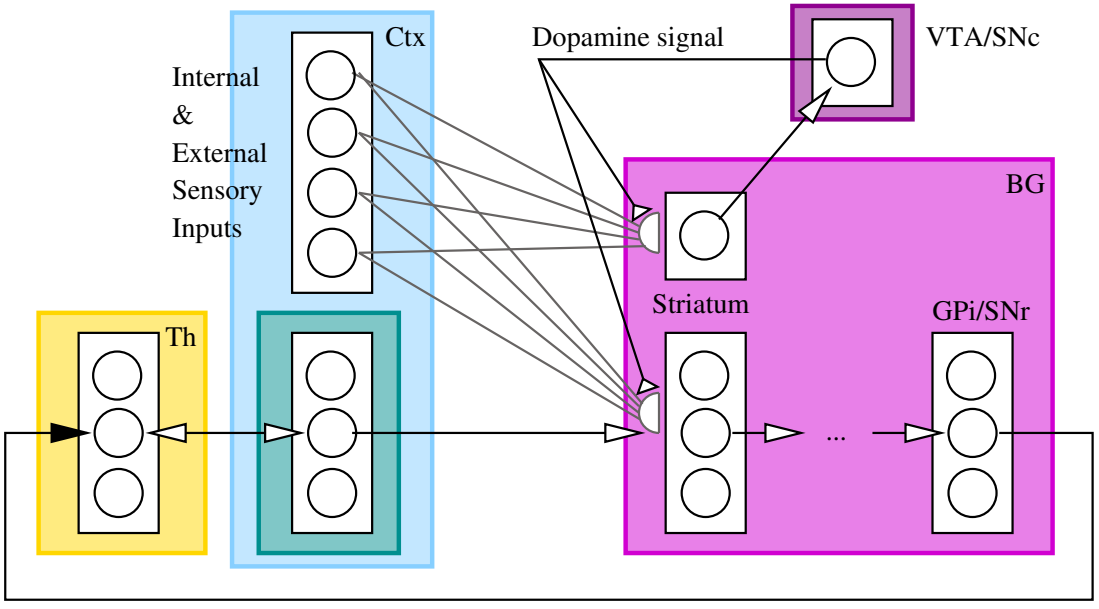
f2 : inhibition moyenne des autres

Domination CBG sur GPR sur l'ensemble du front $\Rightarrow$ connexions additionnelles CBG exploitables pour améliorer la sélection.

Comparaison GPR-CBG par évolution artificielle

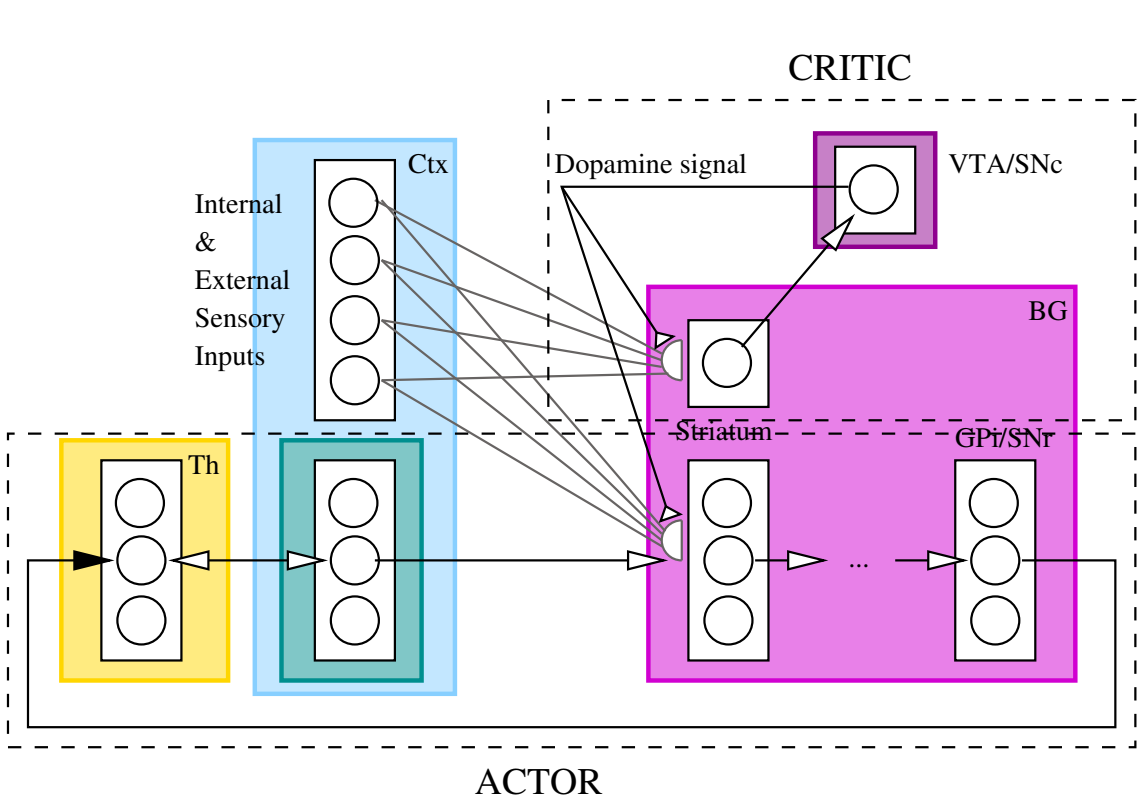


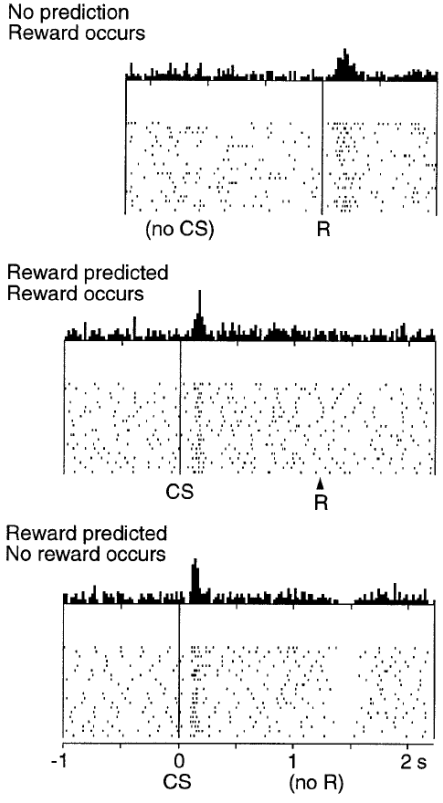
Apprentissage des saliences



- Voir modèles TD de l'apprentissage par renforcement (cours O. Sigaud).

Apprentissage des saliences





- Voir modèles TD de l'apprentissage par renforcement (cours O. Sigaud).

Conclusion

Pour la sélection de l'action :

- Robotique & Neurosciences se retrouvent autour de plusieurs concepts
  - Répertoires de comportements opérant en parallèle,
  - Centralisation de la fonction de sélection d'action,
  - Structuration hiérarchique (strictes ou non).