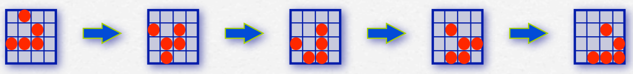
Apprentissage artificiel et complexité

Andrés Pérez-Uribe
HEIG-VD

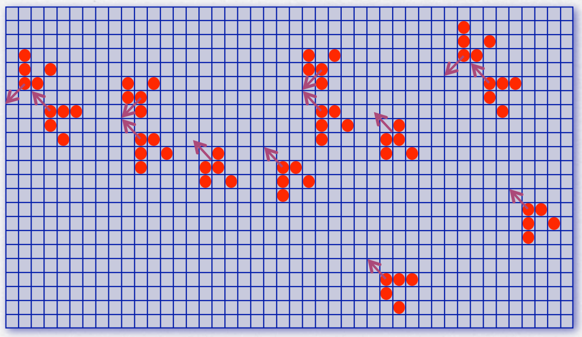
Jeu de la vie (1)

un glider

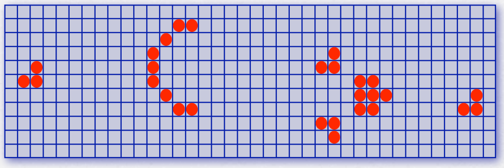


après 4 pas de temps, il se déplace en diagonale

treize gliders:
après 67 pas de
temps, un pistolet
à glider se forme

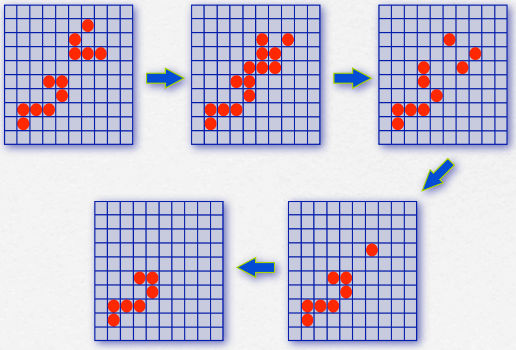


Jeu de la vie (2)

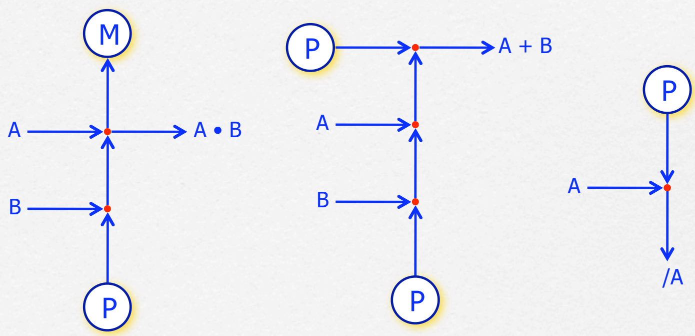


tous les 15 pas de temps, ce pistolet tire un glider

mangeur à glider:



Jeu de la vie (3)



Jeu de la vie (4)

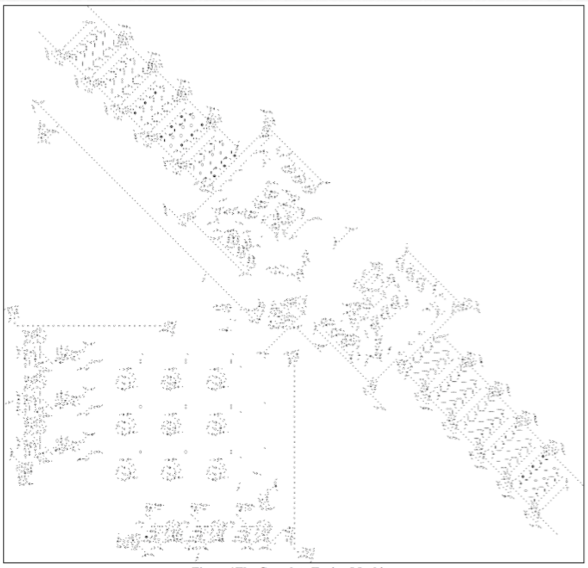
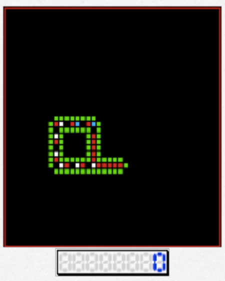


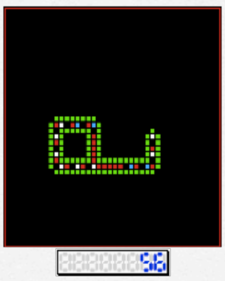
Figure 1The Complete Turing Machine

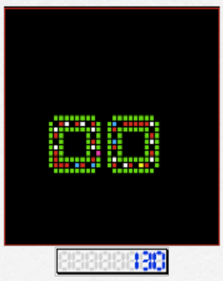
L’automate de Langton

C. Langton

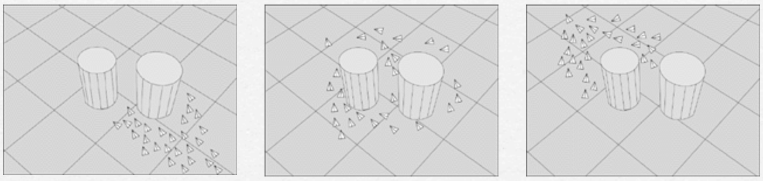
- Huit états par cellule
- Quatre voisines
- Tableau de description avec 219 transitions



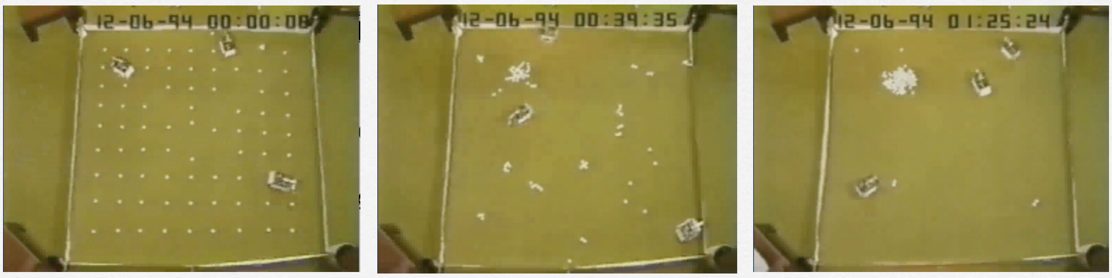




□ Boids (Craig Reynolds)



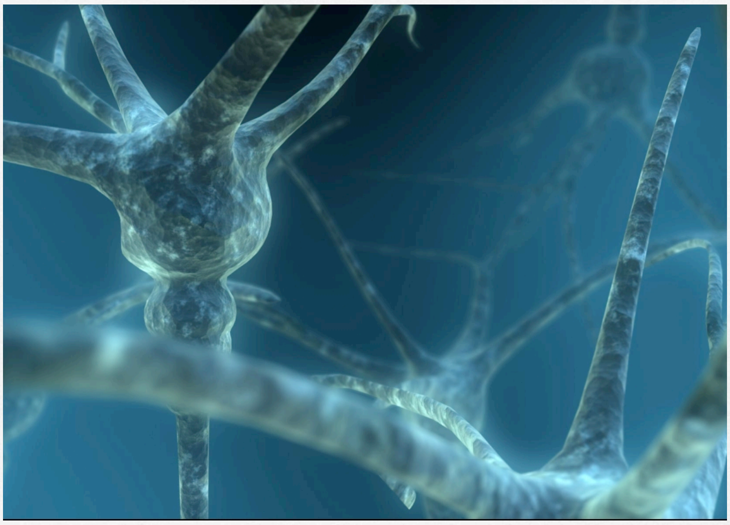
□ Object clustering (O. Holland)



□ Comment construire des systèmes tels que leur comportement soit complexe ?

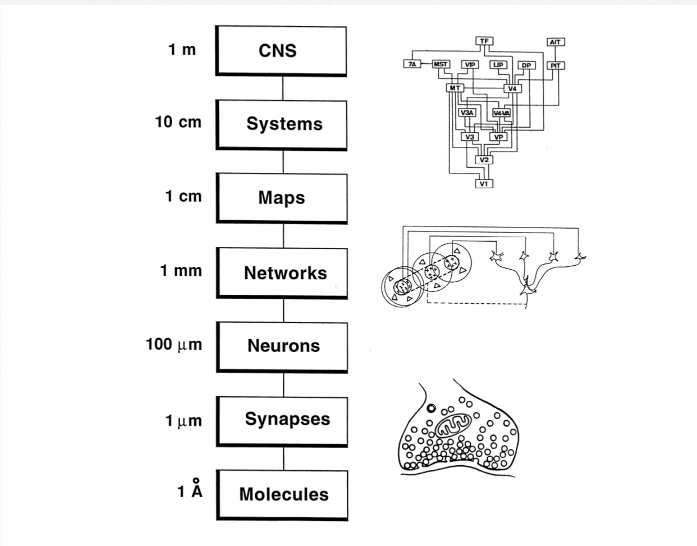
On peut programmer le comportement des unités individuelles, mais comment garantir qu'un comportement complexe va émerger ?

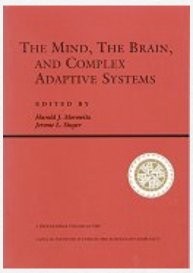
par essai-erreur ?



Connectionism: modeling of mental or behavioral phenomena as emergent processes of interconnected networks of simple units

Système nerveux

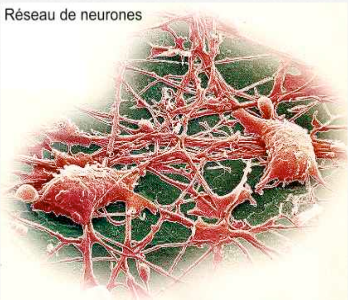


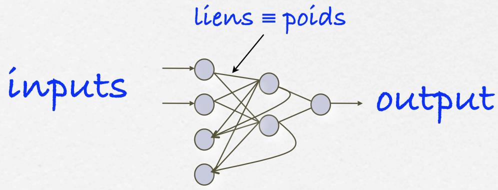


Santa Fe Institute proceedings XXII

Systèmes connexionistes

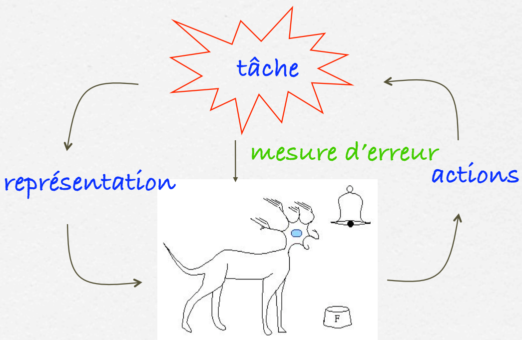
On fait (idéalement) appel à une architecture massivement parallèle où chaque élément (neurone) réalise une sorte de corrélation entre les entrées et des valeurs stockées (poids synaptiques).





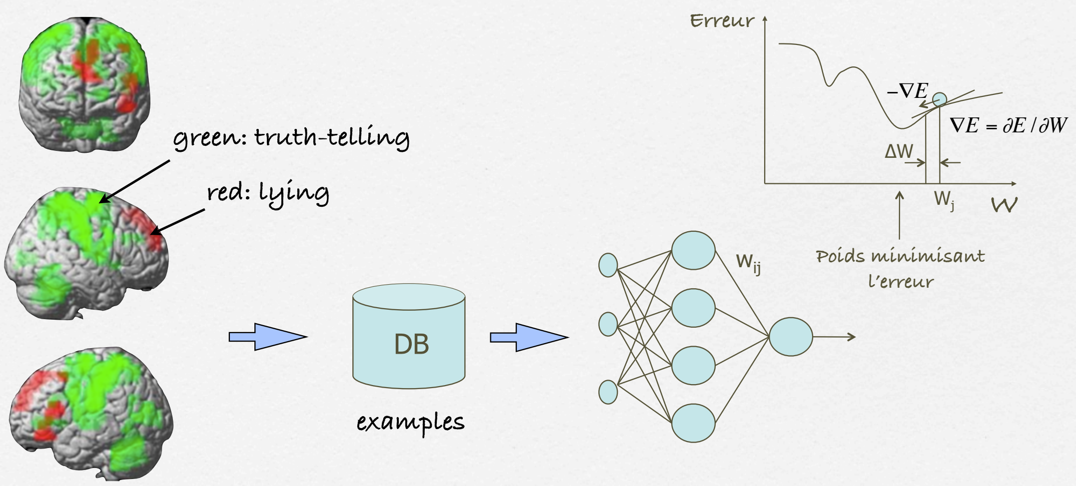
Apprentissage artificiel

- On remplace la programmation par l'apprentissage
- L'apprentissage se fait par l'exemple ou par essai-erreur liée a des récompenses/punitions



- Apprentissage supervisé
- Apprentissage non supervisé
- Apprentissage par renforcement
- Apprentissage par évolution

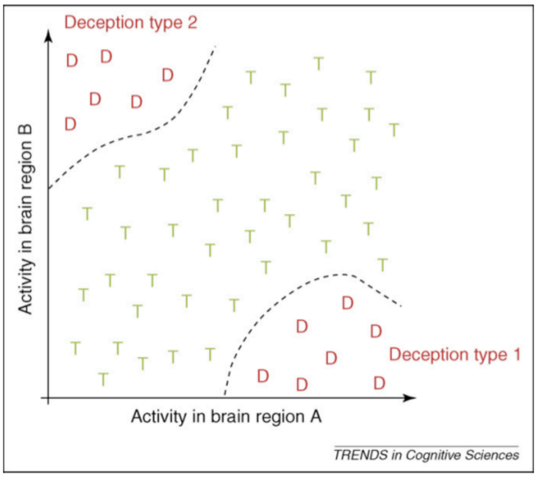
Apprentissage supervisé



NeuroImage 28 (2005) 663 – 668

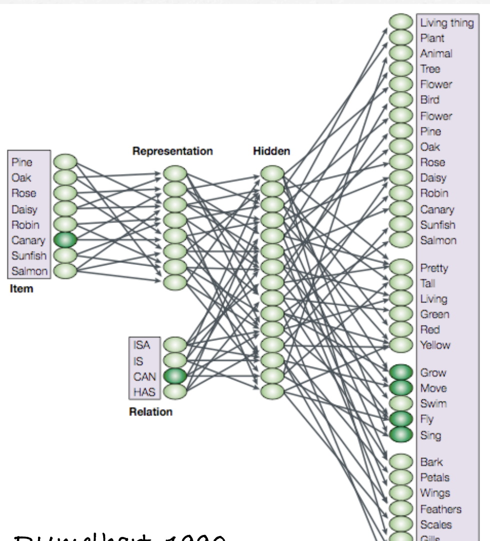
Paul Werbos (1974)
Rumelhart, Hinton, Williams (1986)

Apprentissage supervisé (2)

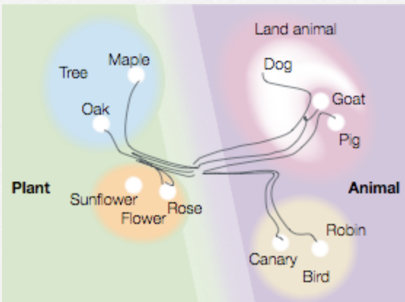


Trends in Cog. Sci, vol 12, No 4, March 2008

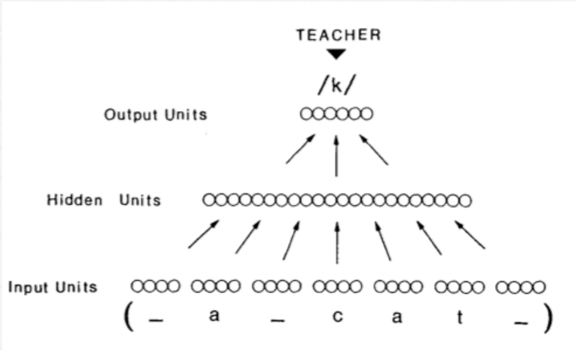
Différentiation de concepts



Rumelhart, 1990
Rogers, McClelland, 2005

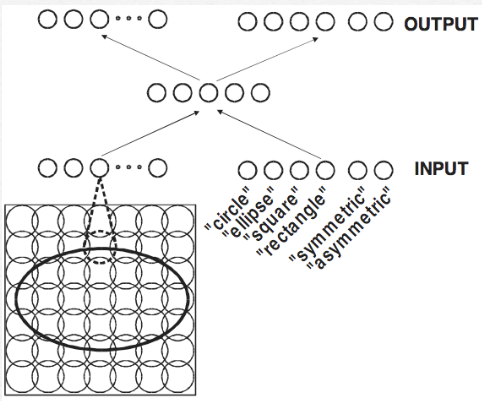


NETtalk



Sejnowski, Rosenberg (Complex Systems 1, 1987)

Symbol grounding

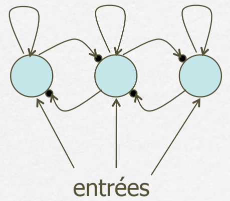


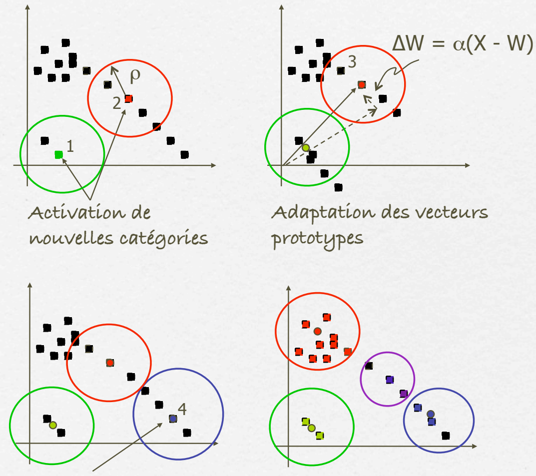
prototype sorting → entry-level naming → higher-level naming → grounding test

Cangelosi, Greco, Harnad, 2000

Apprentissage non supervisé

- Réseau à compétition
- "The winner-takes-all"





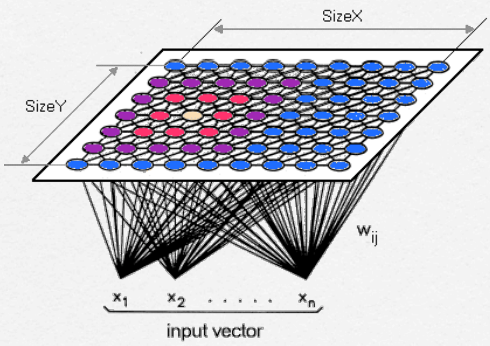
Activation de nouvelles catégories

Adaptation des vecteurs prototypes

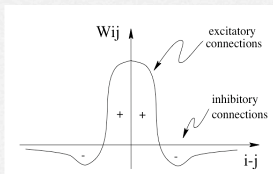
Activation d'une nouvelle catégorie

ART - Adaptive Resonance Theory (S. Grossberg, 1987)

Cartes auto-organisatrices



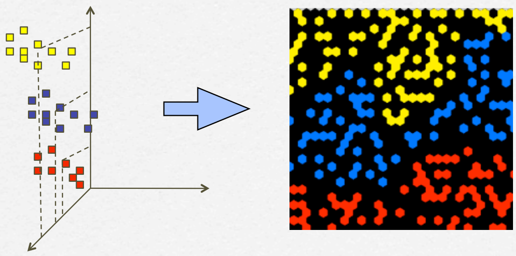
- Matrice de neurones
- Voisinage
- The "Winner takes tbe most"



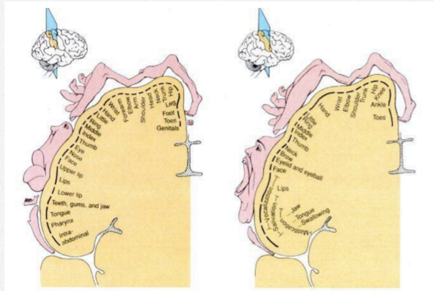
Christof von der Malsburg (1970s)
Teuvo Kohonen (1980s)

Cartes auto-organisatrices (2)

mapping



- cartes somatotopiques



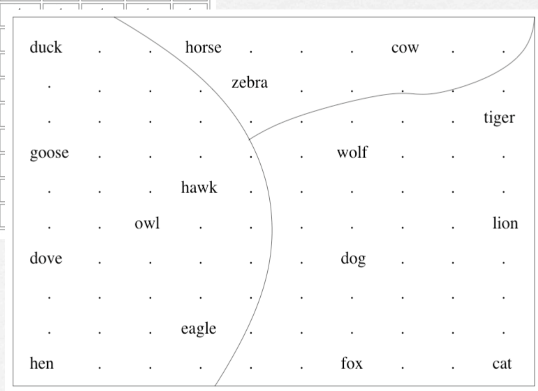
Cortex somatosensoriel | Cortex moteur

- carte tonotopiques
- carte rétinotopiques

Cartes auto-organisatrices (3)

Table I: Input Data Set: Animals

	Attribute	Dove	Hen	Duck	Goose	Owl	Hawk	Eagle	Fox	Dog	Wolf	Cat	Tiger	Lion	Horse	Zebra	Cow
is	small	1	1	1	1	1	1	0	0	0	0	1	0	0	0	0	0
	medium	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0
	big	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1
has	2 legs	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0
	4 legs	0	0	0	0	0	0	0	1	1	1	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	hair	0	0	0	0	0	0	0	1	1	1	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	hooves	0	0	0	0	0	0	0	0	0	0	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	mane	0	0	0	0	0	0	0	0	0	1	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	feathers	1	1	1	1	1	1	1	0	0	0	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
likes	hunt	0	0	0	0	1	1	1	1	0	1	1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
to	run	0	0	0	0	0	0	0	0	1	1	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	fly	1	0	0	1	1	1	1	0	0	0	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	swim	0	0	1	1	0	0	0	0	0	0	0	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]



Apprentissage par renforcement

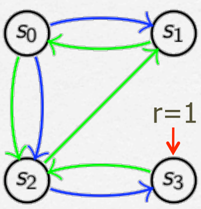
Problèmes de décision séquentielle

Temporal credit assignment problem

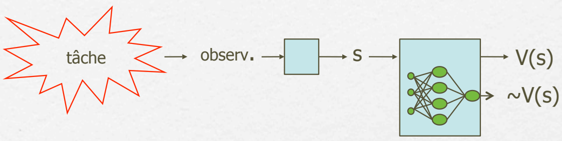
- Pour résoudre des problèmes qui évoluent au cours du temps, l'agent apprend un mapping perceptions -actions en maximisant le cumul des récompenses
- Le renforcement (récompense ou punition) est souvent retardé : on peut ne voir l'intérêt des actions que dans le futur.

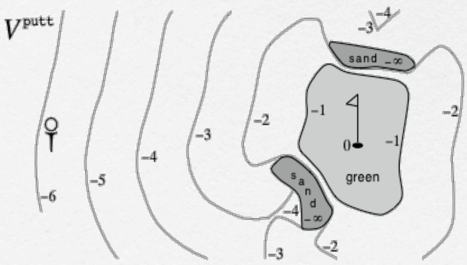
Apprentissage par renforcement (2)

Markov Decision Process (MDP)



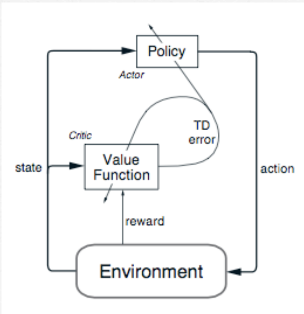
Fonction de valeurs





Sutton & Barto, 1998

Apprentissage par la méthode de différences temporelles



TD - error :

$\delta_t = r_{t+1} + \gamma V(s_{t+1}) - V(s_t)$

δ>0 bonne surprise
δ<0 mauvaise surprise

Actions are determined by preferences :

$$\pi_t(s,a) = \Pr\{a_t = a | s_t = s\} = \frac{e^{p(s,a)}}{\sum_b e^{p(s,b)}},$$

Update the preferences :

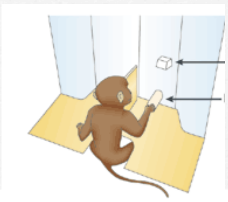
$p(s_t, a_t) \leftarrow p(s_t, a_t) + \beta\delta_t$

The value function update :

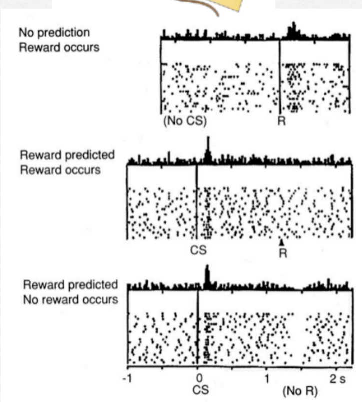
$V(s_t) \leftarrow V(s_t) + \alpha\delta_t$

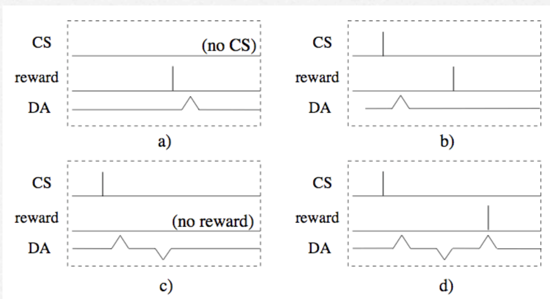
Sutton & Barto, 1988

Les différences temporelles révélées



food box
resting key





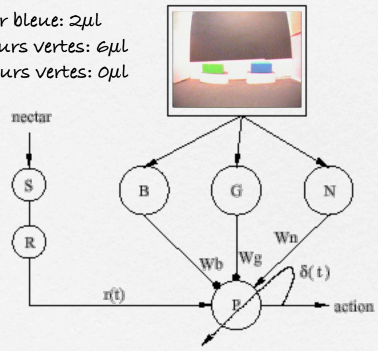
Dopamine neurons encode an error in the temporal prediction of reward.

Schultz, Dayan, Montague, 1997

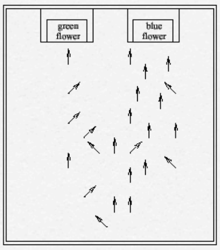
Signaux dopaminergiques dans un "robot abeille"

D'après les expériences de L. Real, 1991

fleur bleue: 2μl
1/3 fleurs vertes: 6μl
2/3 fleurs vertes: 0μl



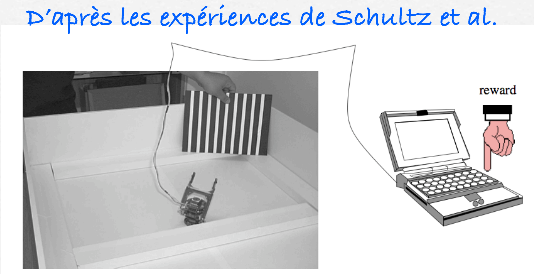
comportement après apprentissage

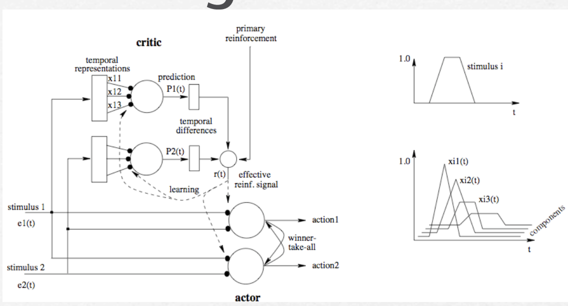


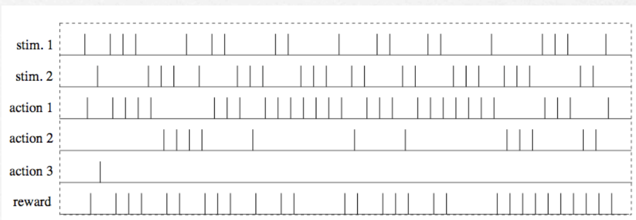
Perez-Uribe and Hirsbrunner, 2000a

Signaux dopaminergiques dans un "robot singe"

D'après les expériences de Schultz et al.



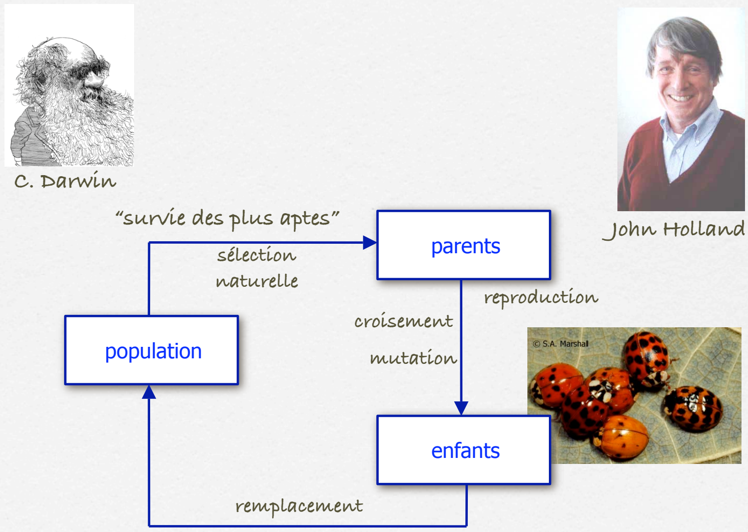




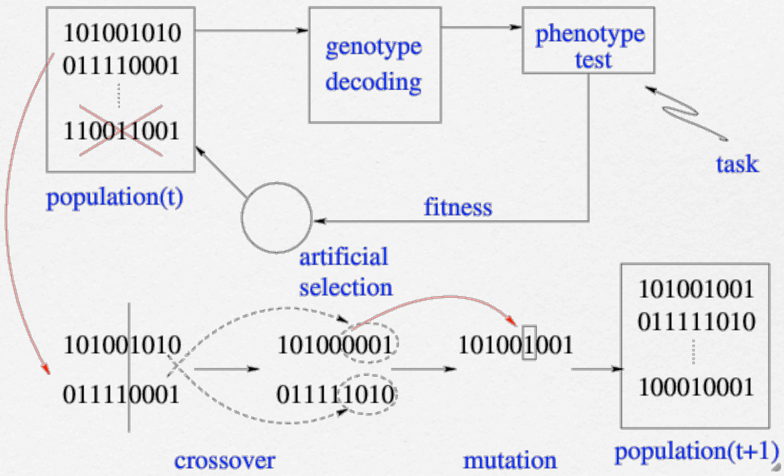
variable
action-reward delay
500 - 3000 ms

Perez-Uribe, 2000b

Evolution artificielle

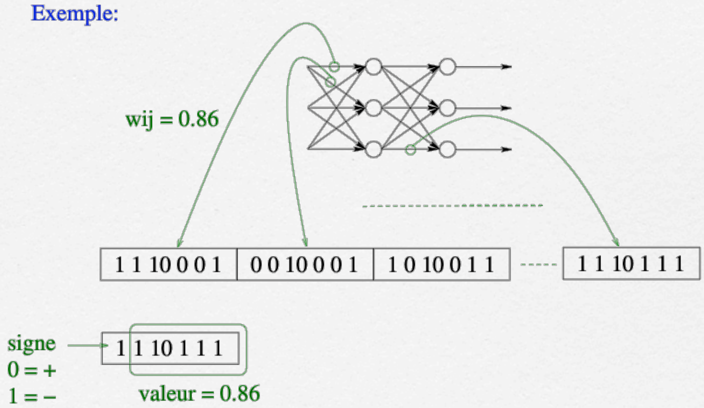


Evolution artificielle (2)



Apprentissage par évolution

Exemple:

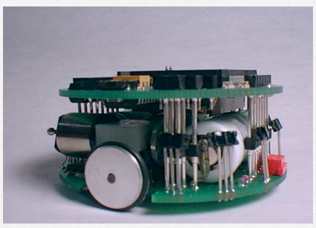


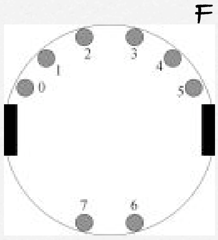
Fitness = 1 – erreur

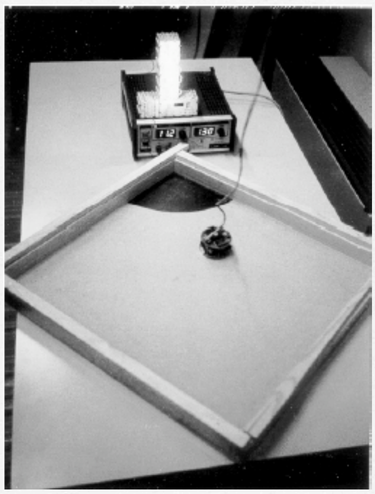
Evolutionary robotics

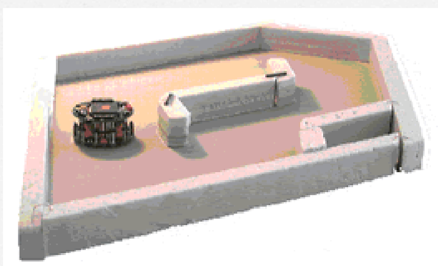
(D. Floreano, LIS, EPFL)

F = f(vitesse, activation capteurs IR)

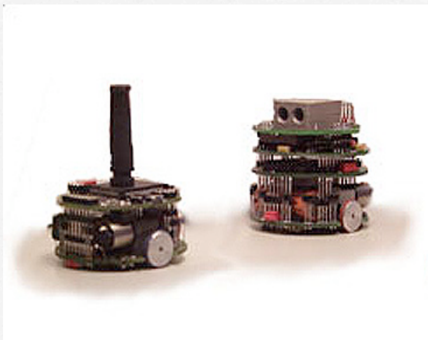


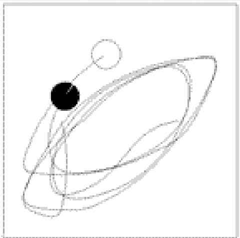
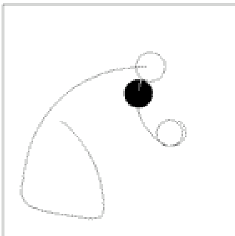
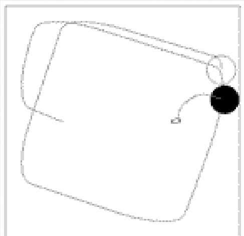




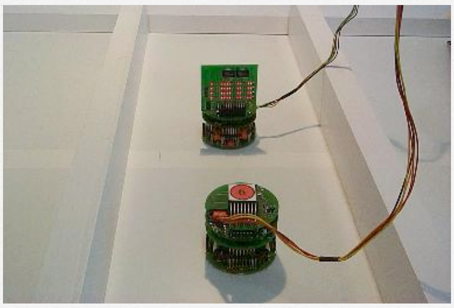


Co-évolution: proie - prédateur
(D. Floreano, LIS, EPFL)



Co-évolution d'un système de communication



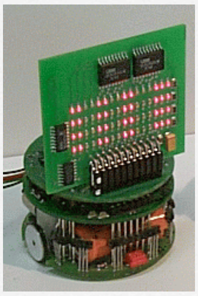
Perez-Uribe and Courant, 2001

Red Queen effect

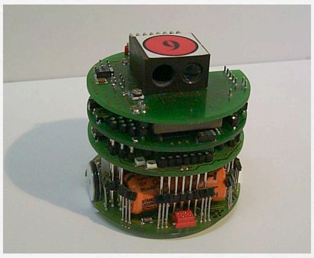
"Well, in our country," said Alice, "you'd generally get to somewhere else. ... "A slow sort of country!" said the Queen. "Now, here, you see, it takes all the running you can do, to keep in the same place."

Alice in Wonderland

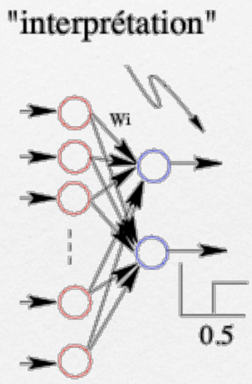
Co-évolution d'un système de communication (2)



Signaler



Receiver



réseau de neurones

poids synaptiques

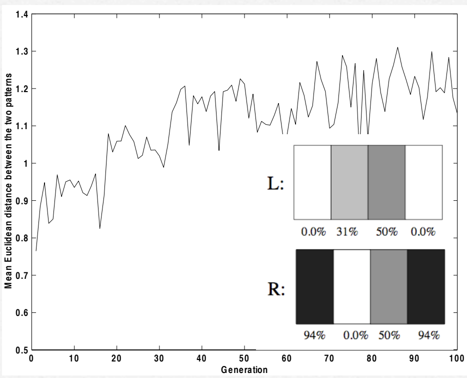
LEDs à allumer

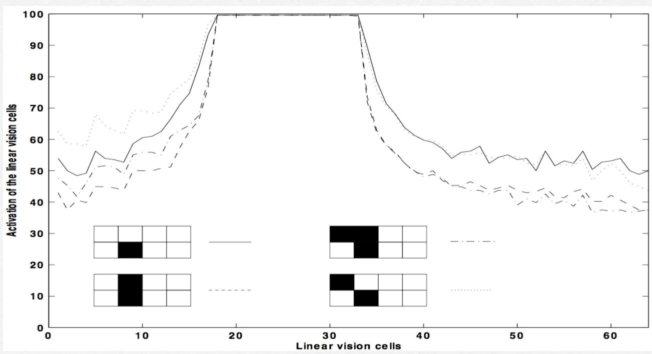
Wi = +0.75

1 1 1 1 -----

1 1 1 0 0 -----

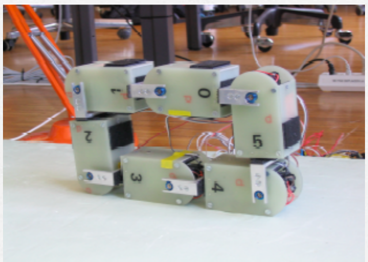
Co-évolution d'un système de communication (3)

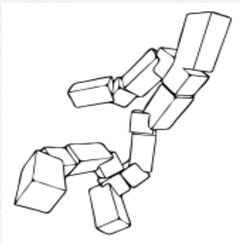


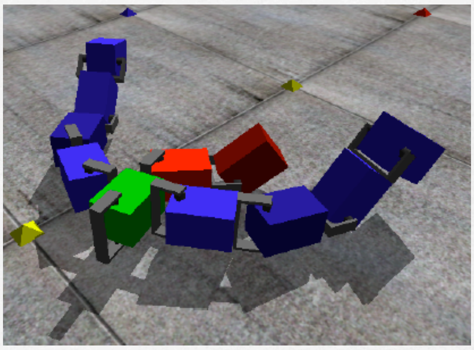


Robots modulaires

(Auke Ijspeert, BIRG, EPFL)







"What you test is what you get"

D. Floreano