

# Apprentissage Machine

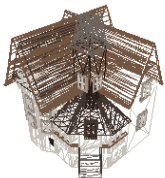
## *“Machine Learning”*

***Jean-Gabriel Ganascia***

*LIP6 – CNRS*

*Université Pierre et Marie Curie (Paris VI)*

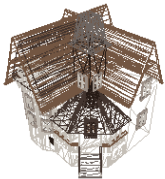
*Jean-Gabriel @Ganascia.name*



ACASA - Agents Cognitifs et Apprentissage  
Symbolique Automatique

J-G Ganascia – Apprentissage artificiel

# Centralité de l'apprentissage en Intelligence Artificielle



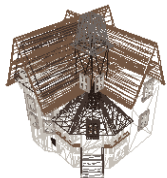
ACASA - Agents Cognitifs et Apprentissage  
Symbolique Automatique

J-G Ganascia – Apprentissage artificiel

# *Une machine peut-elle penser ?*

## Turing 1947, 1950

### Test de Turing: jeu de l'imitation



ACASA - Agents Cognitifs et Apprentissage  
Symbolique Automatique

J-G Ganascia – Apprentissage artificiel

# Un exemple de dialogue

C: Quelle est la longueur de vos cheveux?

A: Je suis coiffée à la garçonne et les mèches les plus longues ont à peu près 20 centimètres

C: S'il vous plait, faites-moi un sonnet sur la Tour Eiffel

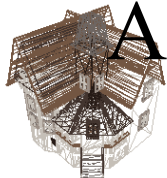
A : Ne comptez pas sur moi pour ça. Je n'ai jamais été capable d'écrire un poème.

C: Ajouter 34957 à 70764.

A: (Pause d'environ 30 secondes avant de donner la réponse) 105621.

Q: Savez-vous jouer aux échecs?

A: Oui.



# Nécessité de l'apprentissage pour Turing



VOL. LIX. No. 236.]

[October, 1950

## MIND

A QUARTERLY REVIEW

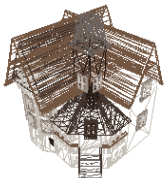
OF

PSYCHOLOGY AND PHILOSOPHY

— 506 —

I.—COMPUTING MACHINERY AND  
INTELLIGENCE

By A. M. TURING



Épistémologie de l'apprentissage machine

IN THIS BUILDING DURING THE SUMMER OF 1956

JOHN McCARTHY (DARTMOUTH COLLEGE), MARVIN L. MINSKY (MIT)  
NATHANIEL ROCHESTER (IBM), AND CLAUDE SHANNON (BELL LABORATORIES)  
CONDUCTED

# THE DARTMOUTH SUMMER RESEARCH PROJECT ON ARTIFICIAL INTELLIGENCE

FIRST USE OF THE TERM "ARTIFICIAL INTELLIGENCE"

FOUNDING OF ARTIFICIAL INTELLIGENCE AS A RESEARCH DISCIPLINE

"To proceed on the basis of the conjecture  
that every aspect of learning or any other feature of intelligence  
can in principle be so precisely described that a machine can be made to simulate it."

IN COMMEMORATION OF THE PROJECT'S 50th ANNIVERSARY  
JULY 13, 2006

# Suite de la proposition...

*The following are some aspects of the artificial intelligence problem:*

- 1. Automatic Computers*
- 2. How Can a Computer be Programmed to Use a Language*
- 3. Neuron Nets*
- 4. Theory of the Size of a Calculation*
- 5. Self-Improvement*
- 6. Abstractions*
- 7. Randomness and Creativity*

# Modélisation – 5 fonctions cognitives

1. *Les fonctions réceptives* : elles autorisent l'acquisition, le traitement, la classification et l'intégration de l'information.
2. La *mémoire* et l'*apprentissage* permettant le stockage et le rappel de l'information.
3. Le *raisonnement*, la *pensée*. Cela concerne aussi l'organisation et la réorganisation mentale de l'information ainsi que son utilisation.
4. Les *fonctions expressives* qui rendent possible la communication.
5. Les *fonctions exécutives* de prise de décision et d'action.



# Panorama général (« IA at large »)

## Modélisation – 5 classes de fonctions cognitives

- i. *Les fonctions réceptives* : elles autorisent l'acquisition, le traitement, la classification et l'intégration de l'information.
- ii. La *mémoire* et l'*apprentissage* permettant le stockage et le rappel de l'information.
- iii. Le *raisonnement*, la *pensée*. Cela concerne aussi bien l'organisation et la réorganisation mentale de l'information que son utilisation.
- iv. Les *fonctions exécutives* de prise de décision et d'action.
- v. Les *fonctions expressives* qui rendent possible la communication.

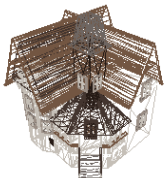
# Apprentissage

=

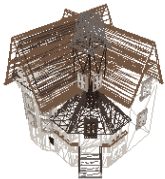
## Acquisition des connaissances

*Qu'est ce que  
l'Apprentissage Artificiel ?*

*Comment faire une machine qui tire  
parti de ses propres expériences pour  
accroître ses connaissances?*



# Y a-t-il une science de l'apprentissage artificiel ?

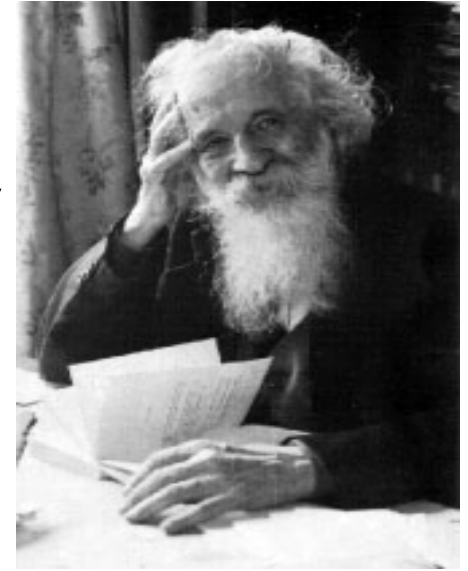


ACASA - Agents Cognitifs et Apprentissage  
Symbolique Automatique

J-G Ganascia – Apprentissage artificiel

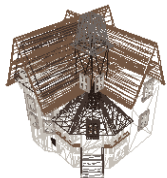


# Du concret à l'abstrait



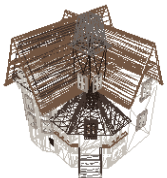
## *Plan*

- *Les images*
- *Les algorithmes*
- *Les généralisations*
- *Les théories formelles*



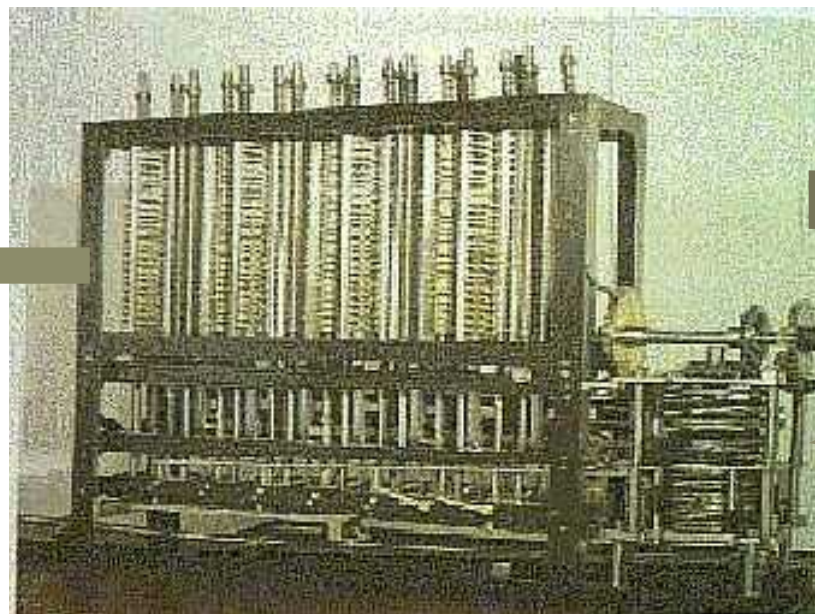
# Les images – *sources d'inspiration*

## *L'homme: théories classique de l'apprentissage*

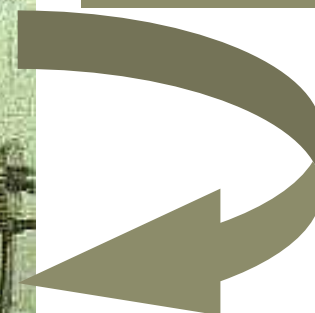


# Différentes approches de l'apprentissage artificiel

Observation

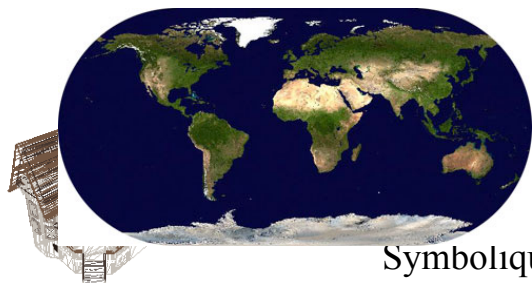


Reflection



TDIT  
Neural Network  
Genetic algorithm...

Learning by doing,  
Explanation based-learning,  
Case Based Reasoning  
Reinforcement learning...



Symbolique

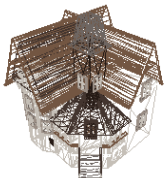
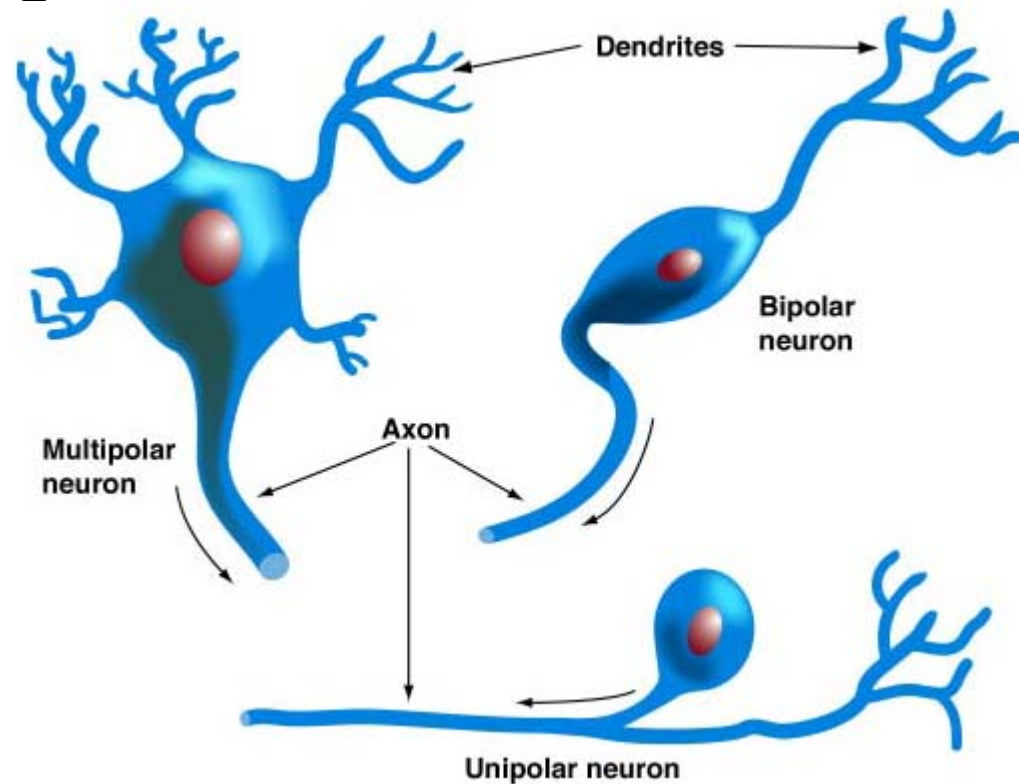
## Techniques d'apprentissage machine



# Les images – *sources d'inspiration*

*L'homme: théories classique de l'apprentissage*

*La physiologie: le cerveau*

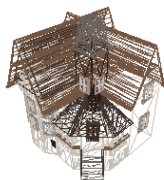
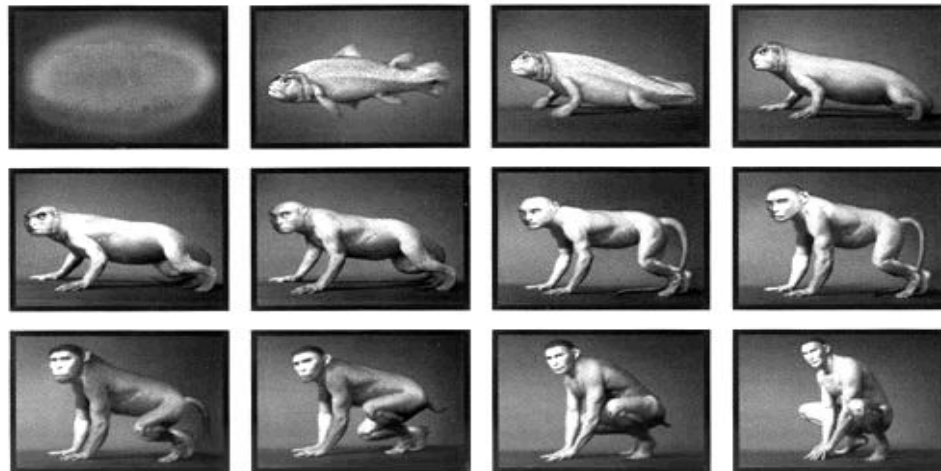


# Les images – *sources d'inspiration*

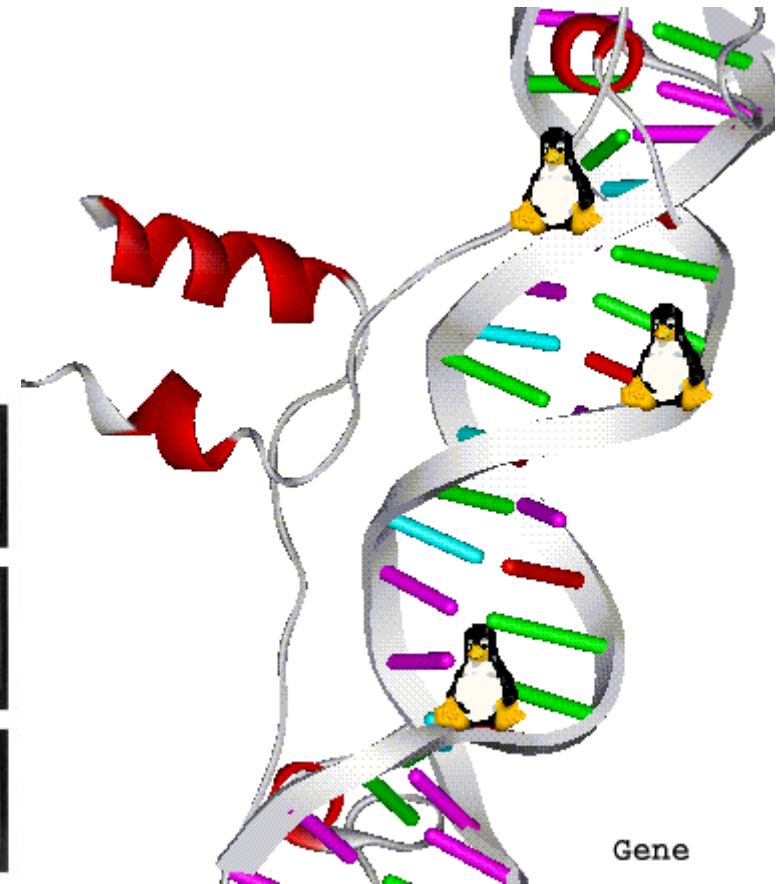
*L'homme: théories classique de l'apprentissage*

*La physiologie: le cerveau*

*La nature: l'évolution*



ACASA - Agents Cognitifs et Apprentissage  
Symbolique Automatique



J-G Ganascia – Apprentissage artificiel



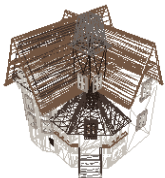
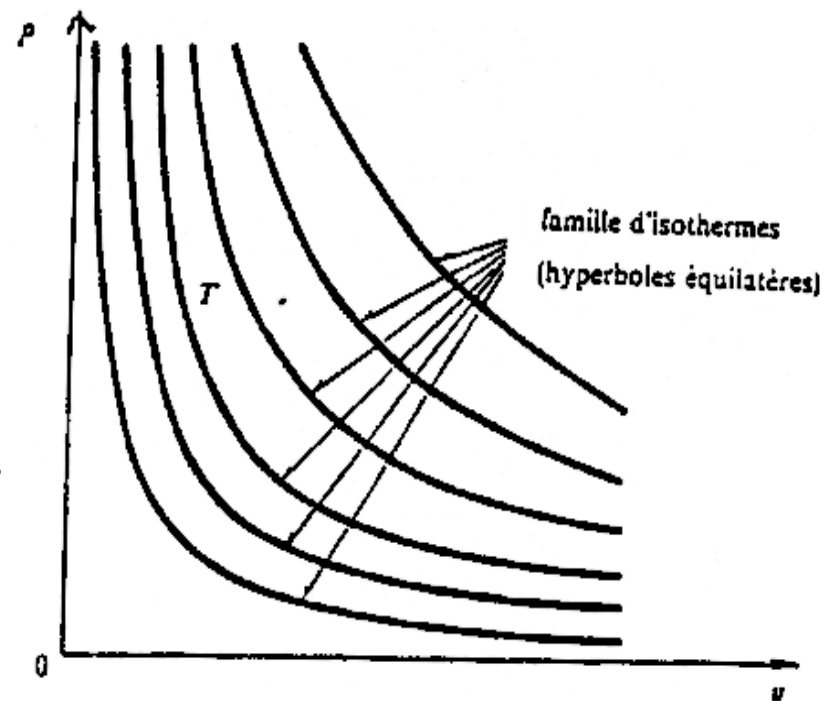
# Les images – *sources d'inspiration*

*L'homme: théories classique de l'apprentissage*

*La physiologie: le cerveau*

*La nature: l'évolution*

*La culture: l'épistémologie*



# Les images – *sources d'inspiration*

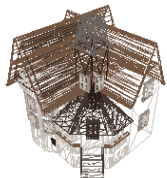
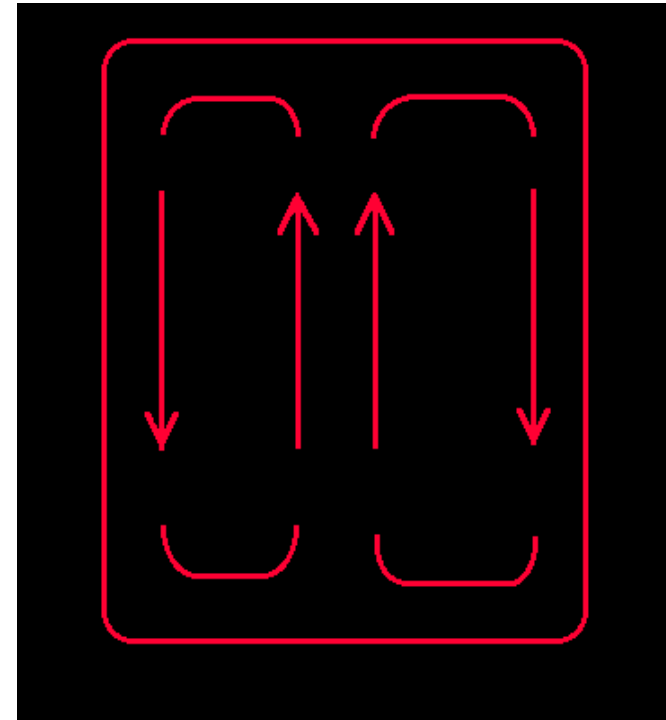
*L'homme: théories classique de l'apprentissage*

*La physiologie: le cerveau*

*La nature: l'évolution*

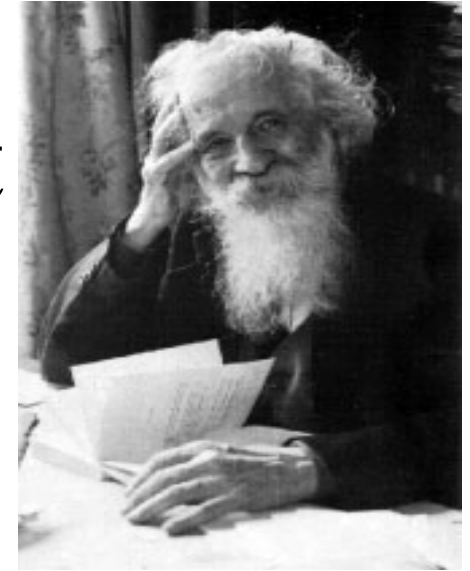
*La culture: l'épistémologie*

*La physique: les systèmes dynamiques*



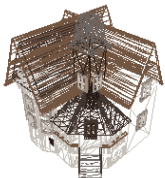


# Du concret à l'abstrait



## *Plan*

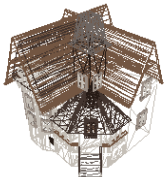
- *Les images*
- *Les algorithmes*
- *Les généralisations*
- *Les théories formelles*



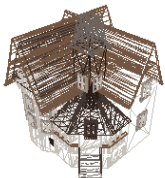
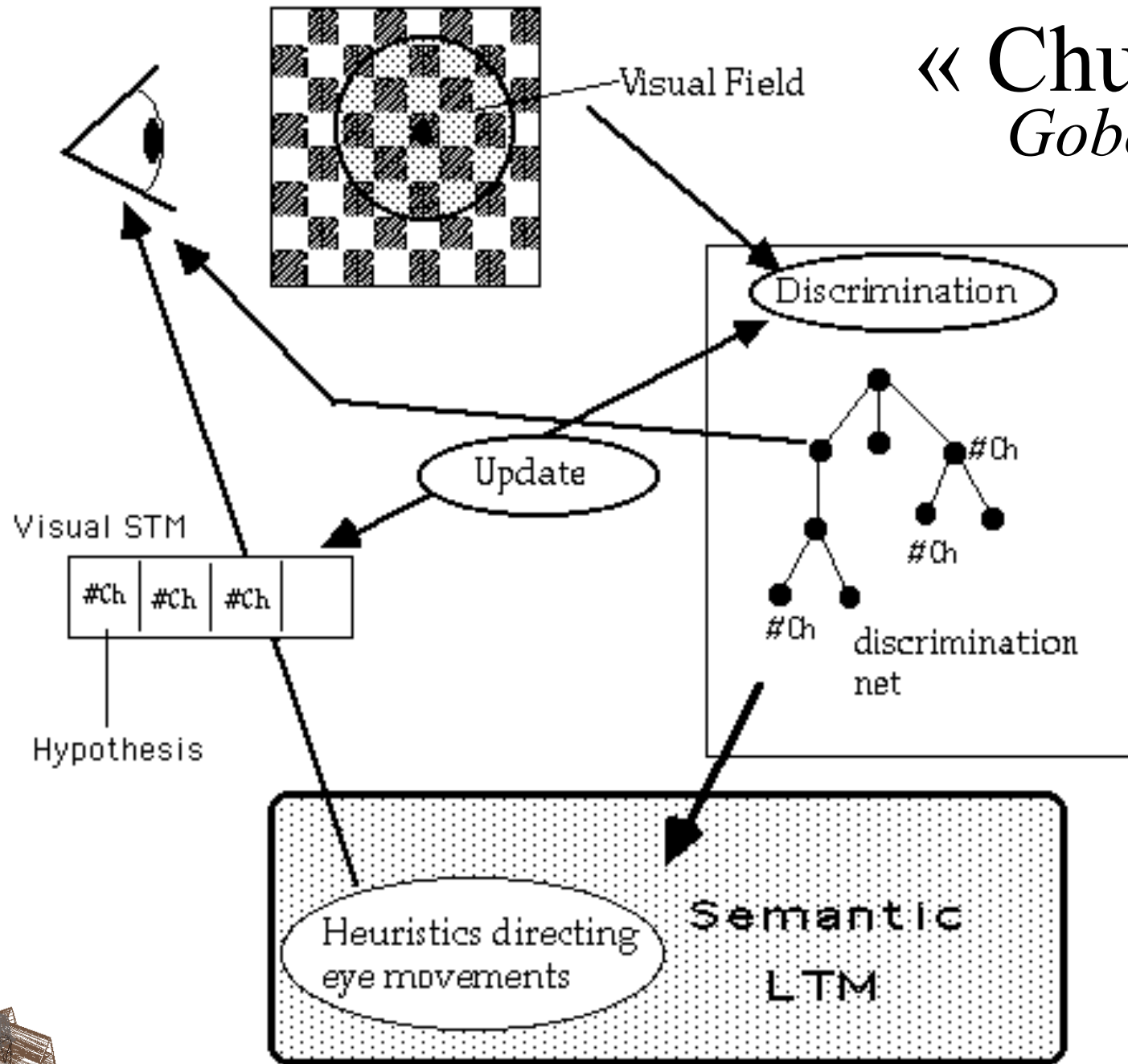
# Les algorithmes

## *L'homme: théories classique de l'apprentissage*

- *Apprentissage par renforcement (Samuel)*
- *EPAM – apprentissage non supervisé (Feigenbaum)*
- *Apprentissage par l'action - “chunking” (Newell et al.)*



# « Chunking » *Gobet & Al.*



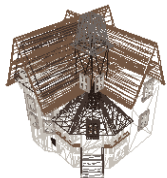
# Les algorithmes

## *L'homme: théories classique de l'apprentissage*

- *Apprentissage par renforcement (Samuel)*
- *EPAM – apprentissage non supervisé (Feigenbaum)*
- *Apprentissage par l'action - “chunking” (Newell et al.)*

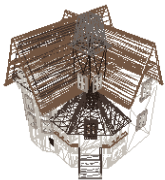
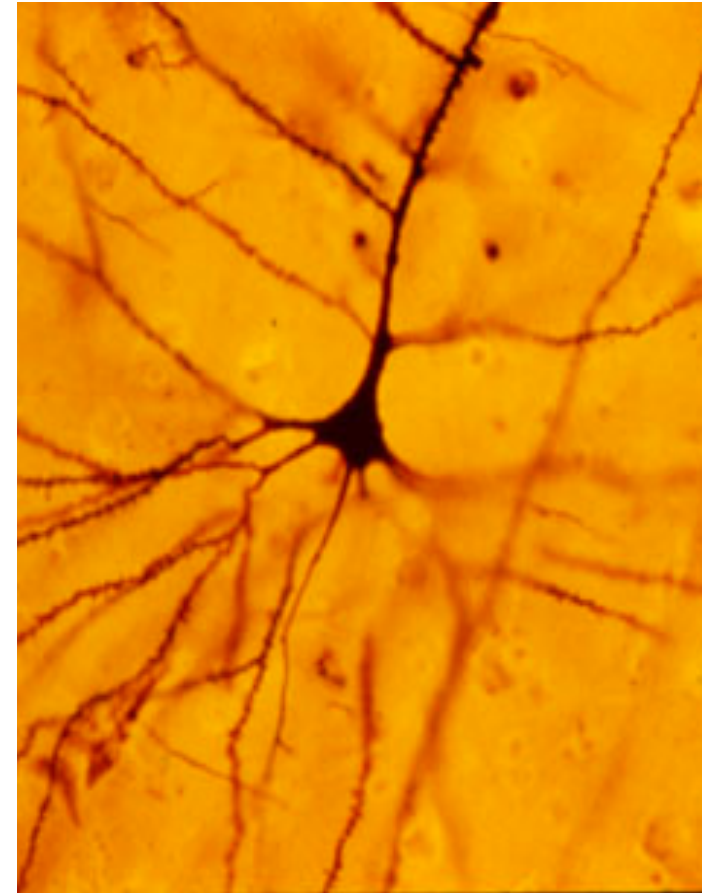
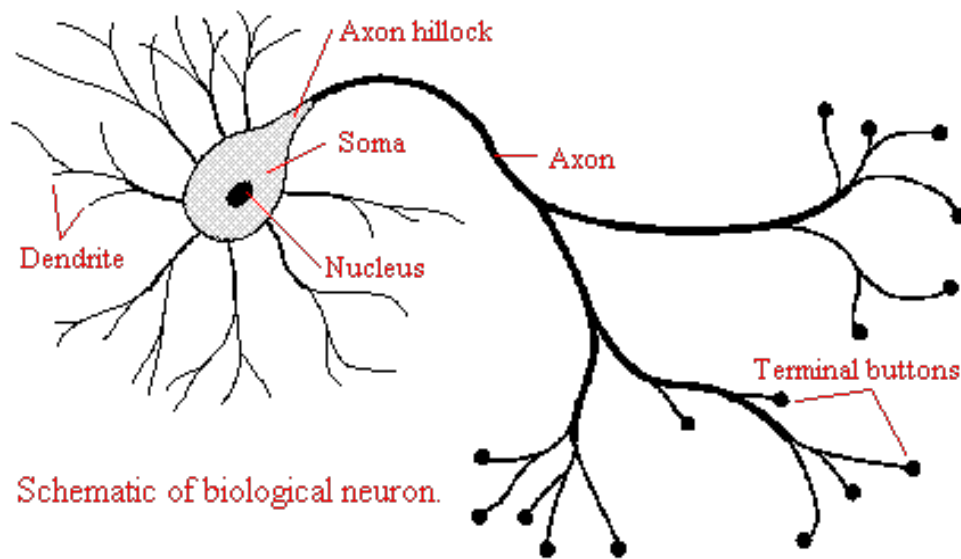
## *La physiologie: le cerveau*

- *Perceptron (Rosenblatt)*



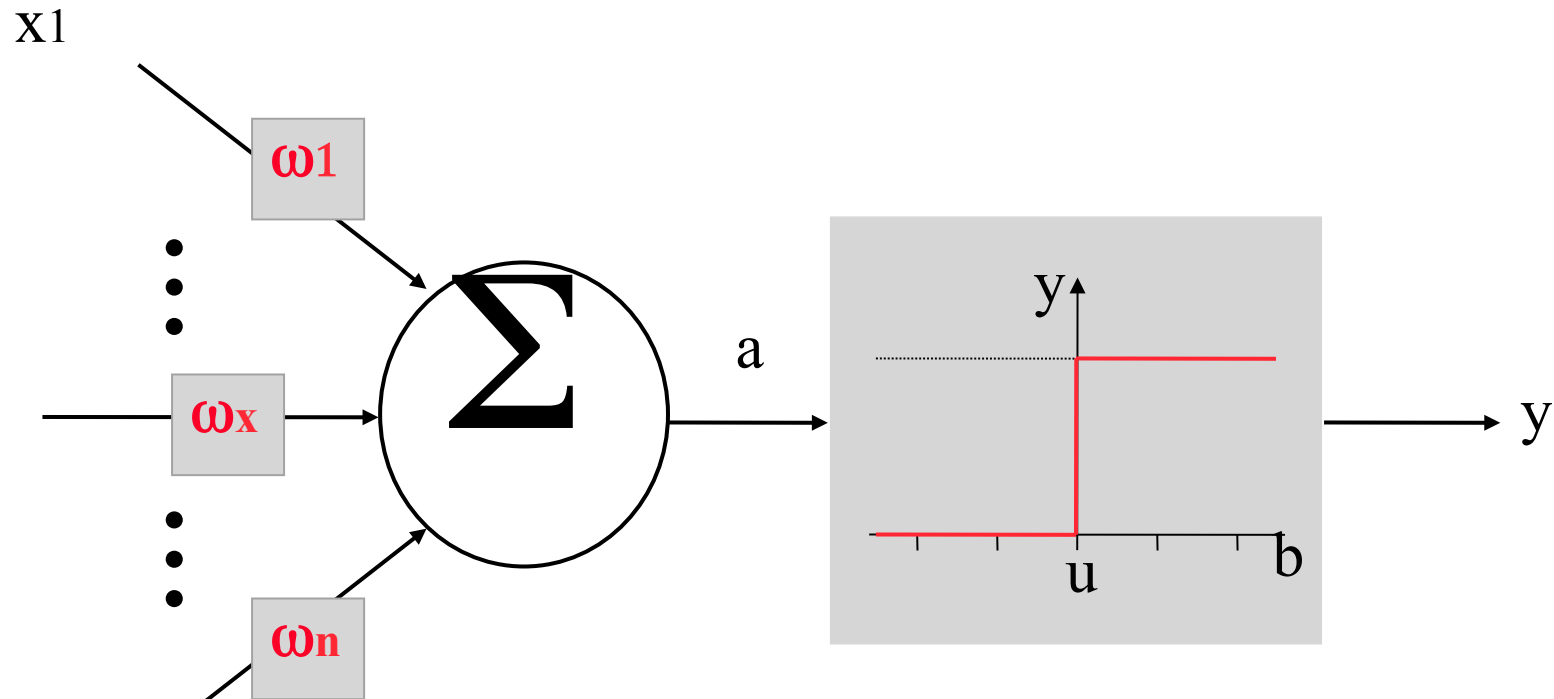
# Réseaux de neurones

- Prix Nobel 1906
  - Camilo Golgi
  - Santiago Ramón Y Cajal

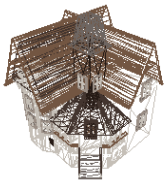


# Réseaux de neurones formels

(Mc Culloch and Pitt, 1943)



$$y(t+1) = f\left(\sum_{k=1}^n w_k x_k(t) - u\right) = f\left(\sum_{k=0}^n w_k x_k(t)\right)$$

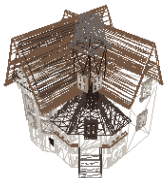
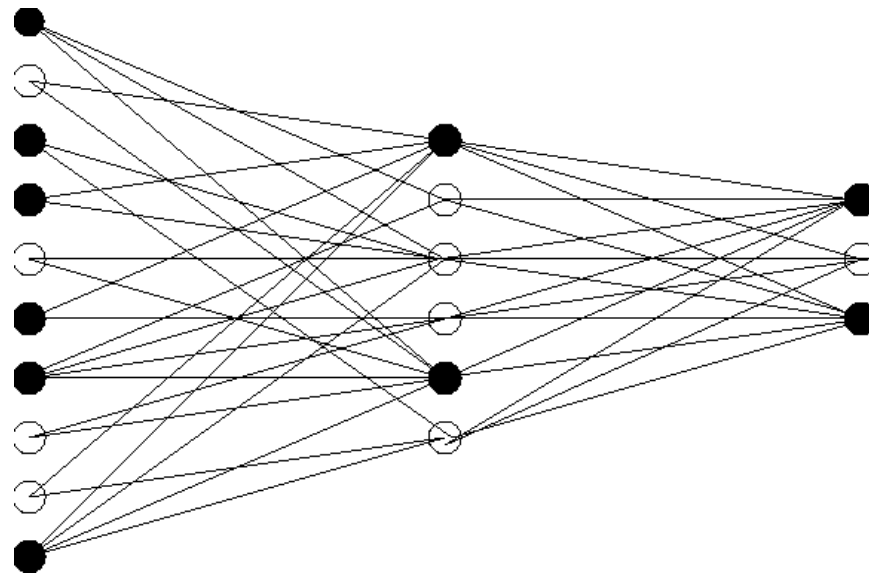




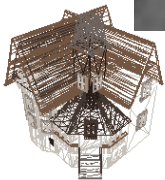
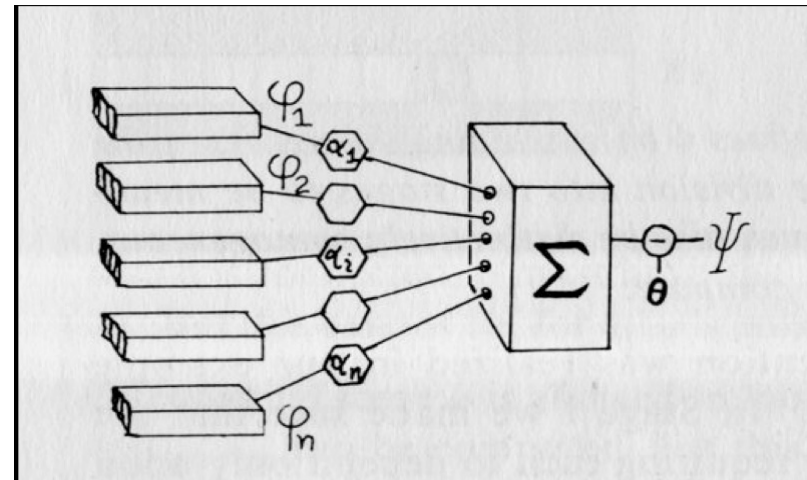
# Fonction logique

## Propriété d'universalité

- Un réseau de neurones à deux couches avec une fonction d'activation en escalier peut réaliser **n'importe quelle fonction booléenne**, à condition que le nombre de neurones de la couche cachée  $H$  soit suffisamment grand (*Mc Culloch and Pitts, 1943*) .

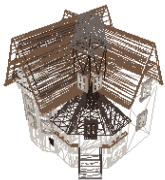
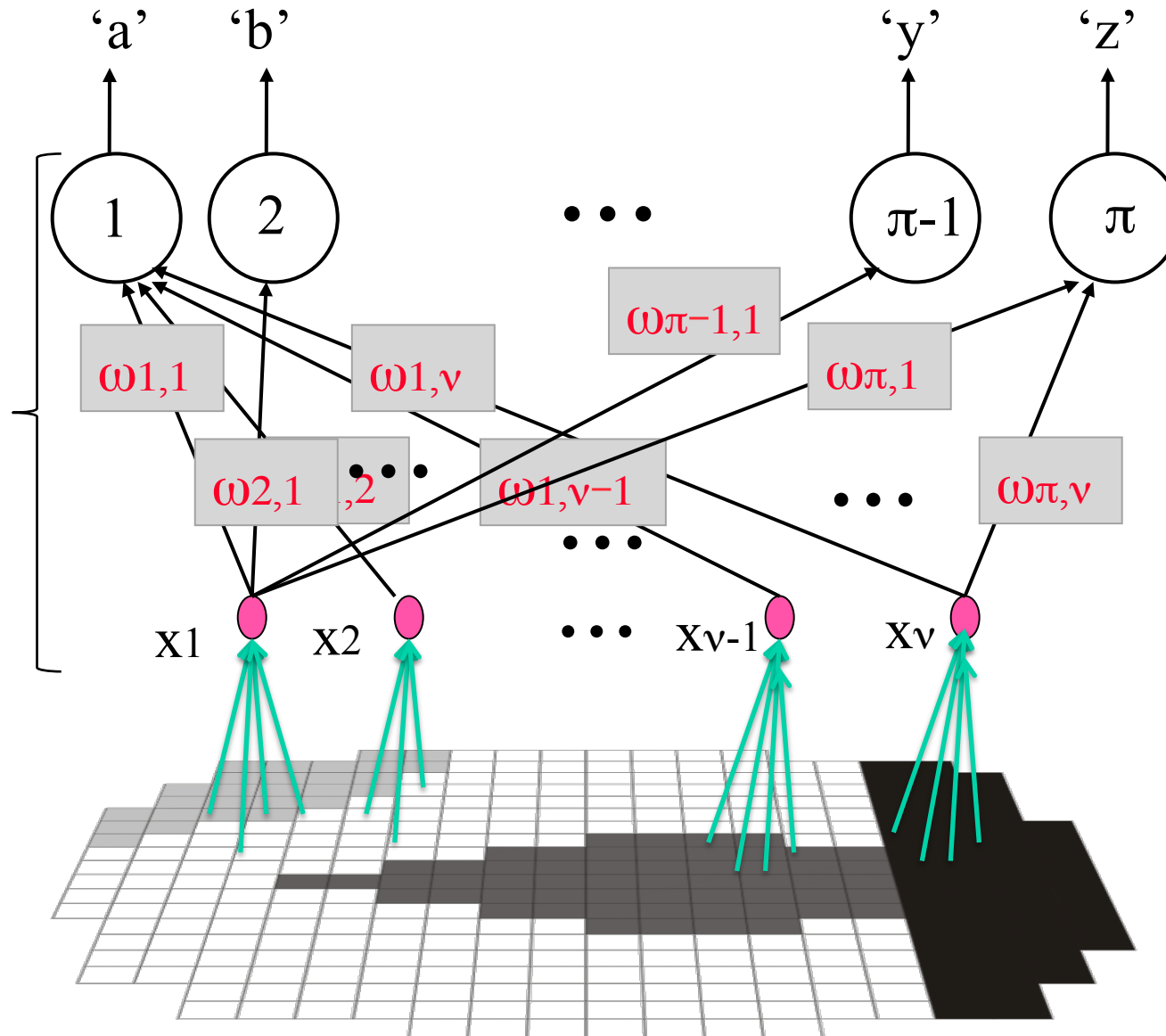


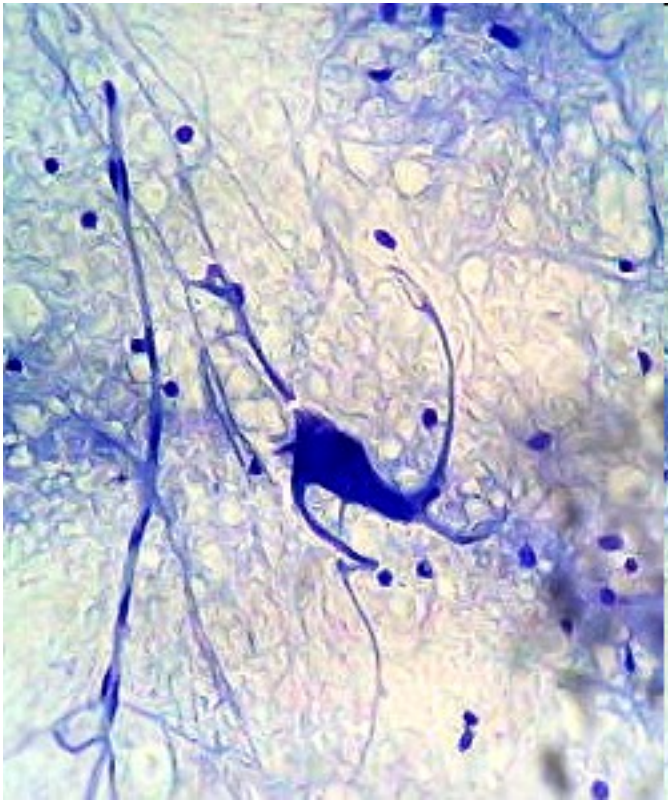
# Marvin Minsky



Épistémologie de l'apprentissage machine

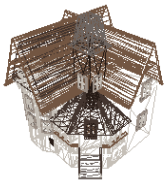
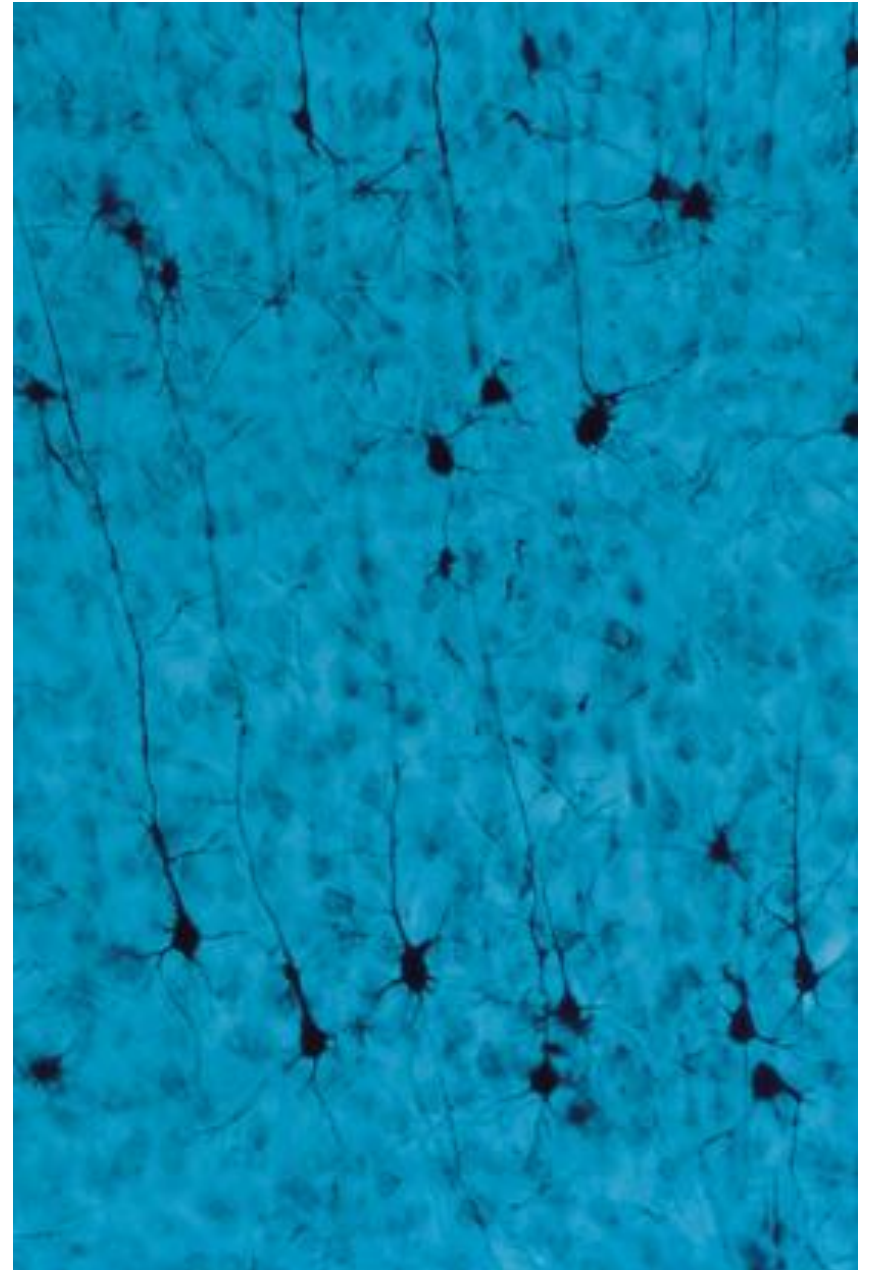
# Perceptron – Rosenblatt 1958



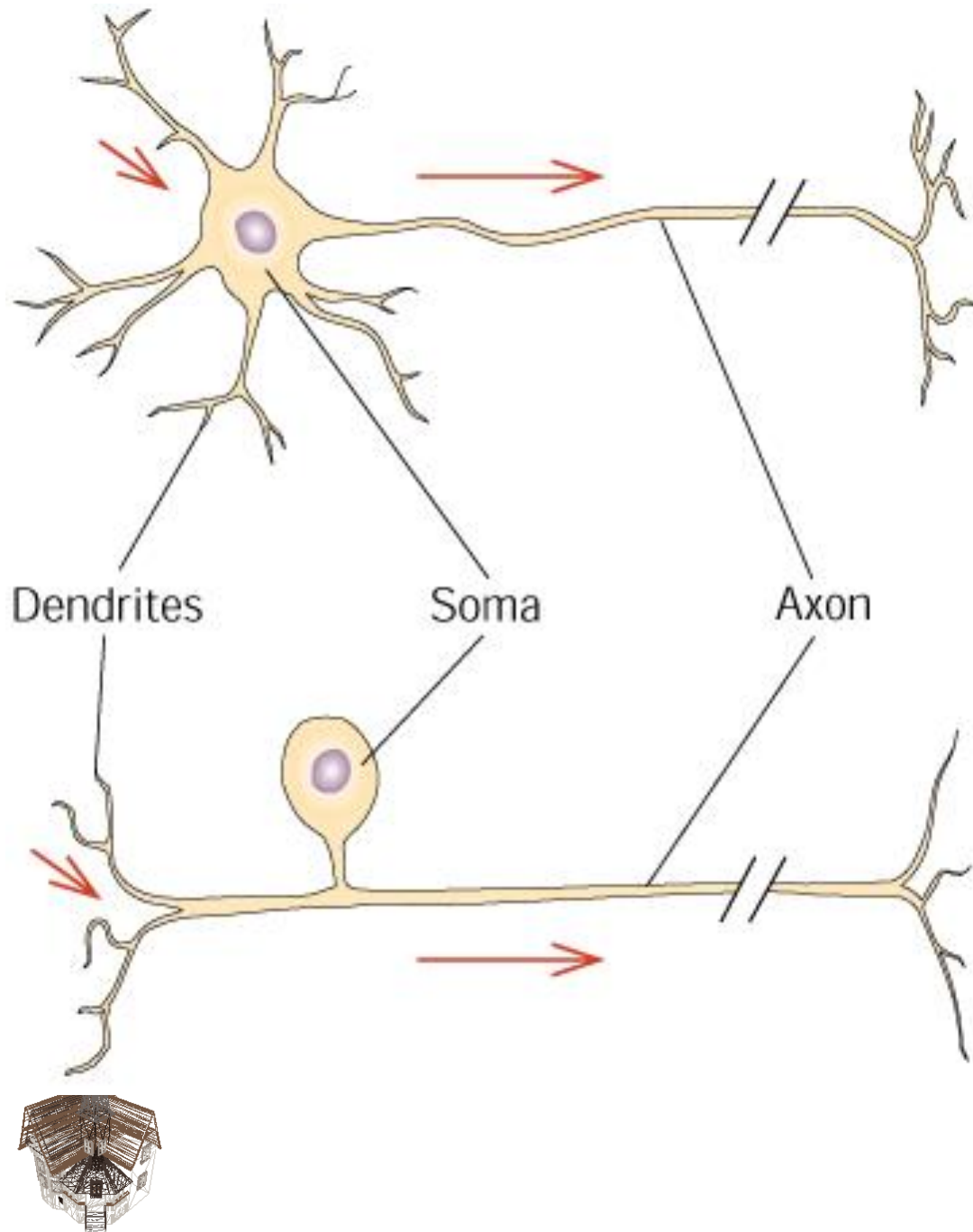


# Neurones

Ordres de grandeur



# Éléments du neurone



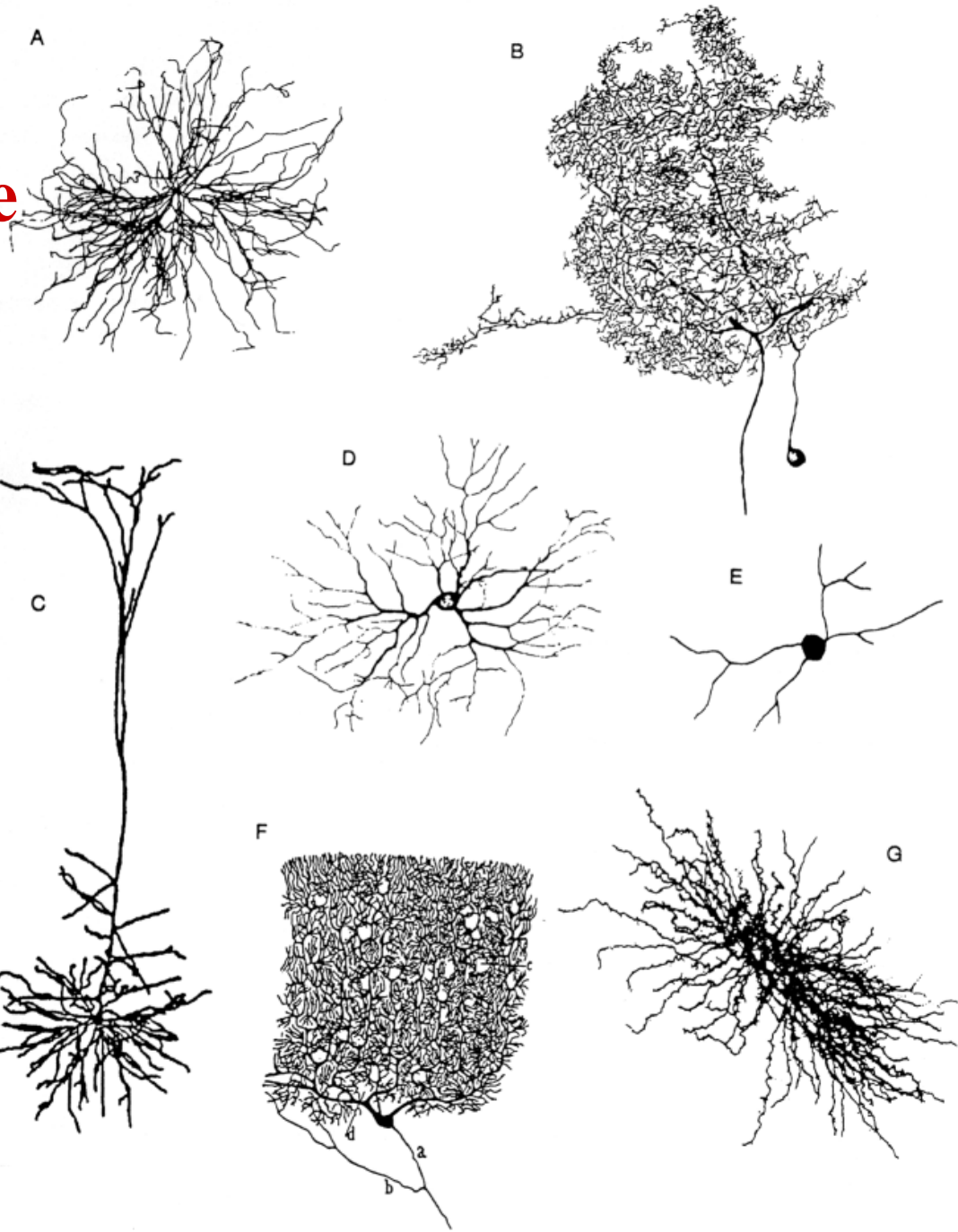
- **Corps de la cellule** (ou soma): contient un noyau
- **Dendrites**: transporte l'information vers le corps de la cellule
- **Axones**: transportent l'information loin de la cellule (muscles, ...)
- **Synapses**: points de communication entre neurones



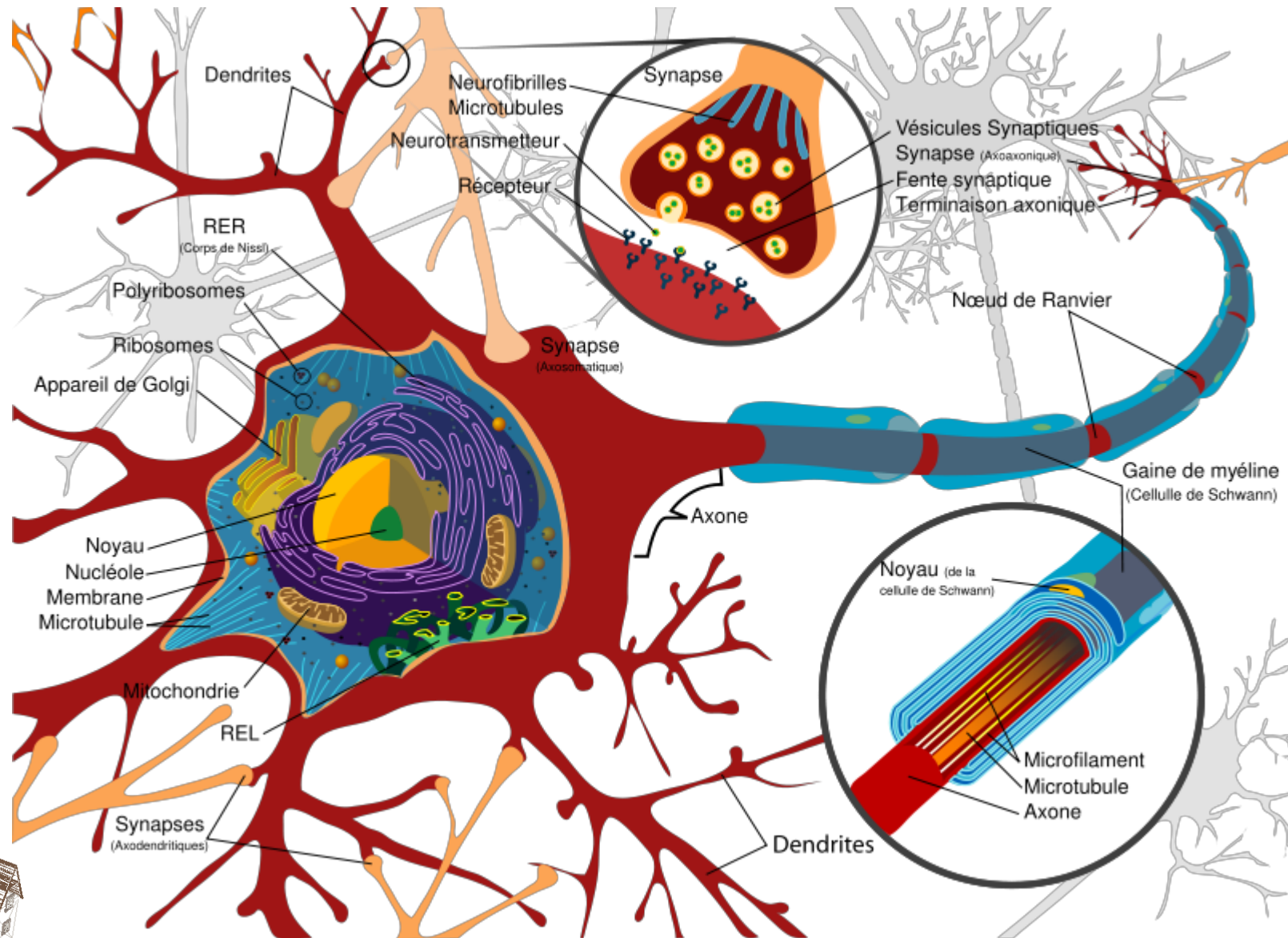
# Diversité des neurones

## structure anatomique unique

- A) motor neuron in spinal cord of cat,
- B) interneuron from a ganglion in the thorax of a locust,
- C) pyramidal cell from the cerebral cortex of a rat,
- D) retinal ganglion cell from a kitten,
- E) amicroine cell from the retina of a tiger salamander,
- F) Purkinje cell from the cerebellum of a human,
- G) relay neuron from the thalamus of a rat



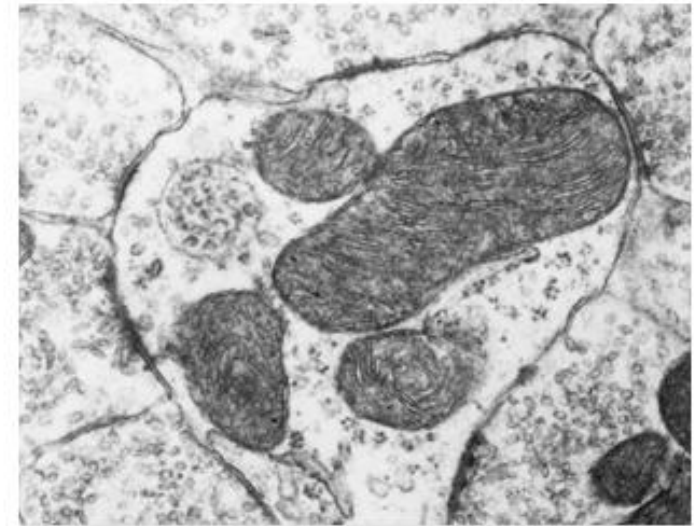
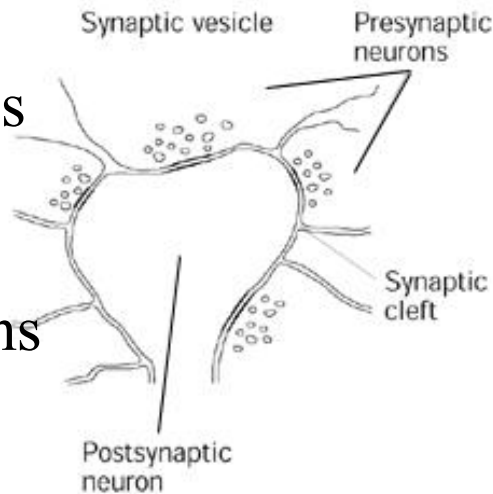
# Connexions entre neurones



# Synapses: transmission entre cellules

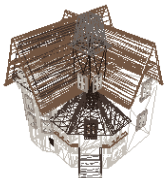
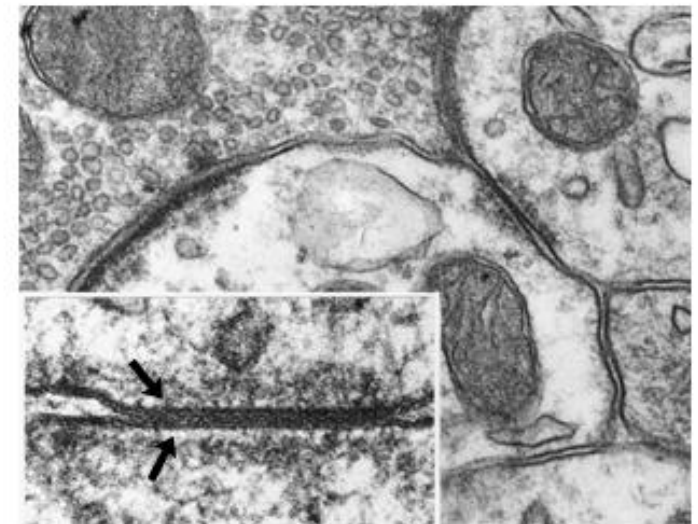
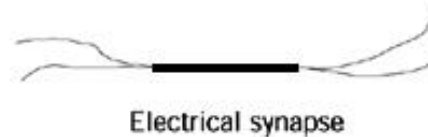
- Synapses chimiques <sup>A</sup>

- Neurotransmetteurs
- Vésicules synaptiques
- Polarisation: la communication intervient dans un sens



- Synapses électriques <sup>B</sup>

- Courant électrique
- Pas de vésicules synaptiques
- Pas de polarisation

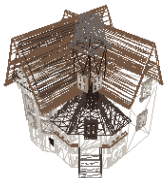






# Quelques données

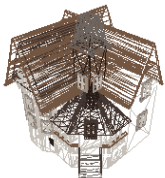
- Poids du cerveau
  - Homme adulte: 1,3 – 1,4 kg
  - Eléphant: 6kg
  - Homme du Pithécanthrope: 0,85 – 1kg
  - Girafe: 0,68 kg
  - Rat: 0,002 kg
  - Chimpanzé: 0,420 kg
  - Chat: 0,030 kg
- Nombre moyen de neurones
  - 100 milliards (homme)
  - 300 millions (pieuvre)
  - Quelques centaines (vers nématodes)





## Quelques données (suite)

- Poids d'un neurone:  $10^{-6}$  grammes
- Nombre de synapses: 1000 – 10000
- Diamètre d'un neurone: 4 – 100  $\mu\text{m}$
- Vitesse de conduction de l'influx nerveux: 0,6 – 120m/s



# Les algorithmes

## *L'homme: théories classique de l'apprentissage*

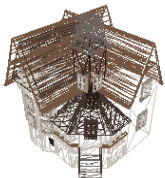
- *Apprentissage par renforcement (Samuel)*
- *EPAM – apprentissage non supervisé (Feigenbaum)*
- *Apprentissage par l'action - “chunking” (Newell et al.)*

## *La physiologie: le cerveau*

- *Perceptron (Rosenblatt)*

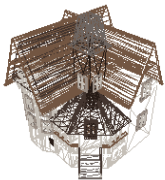
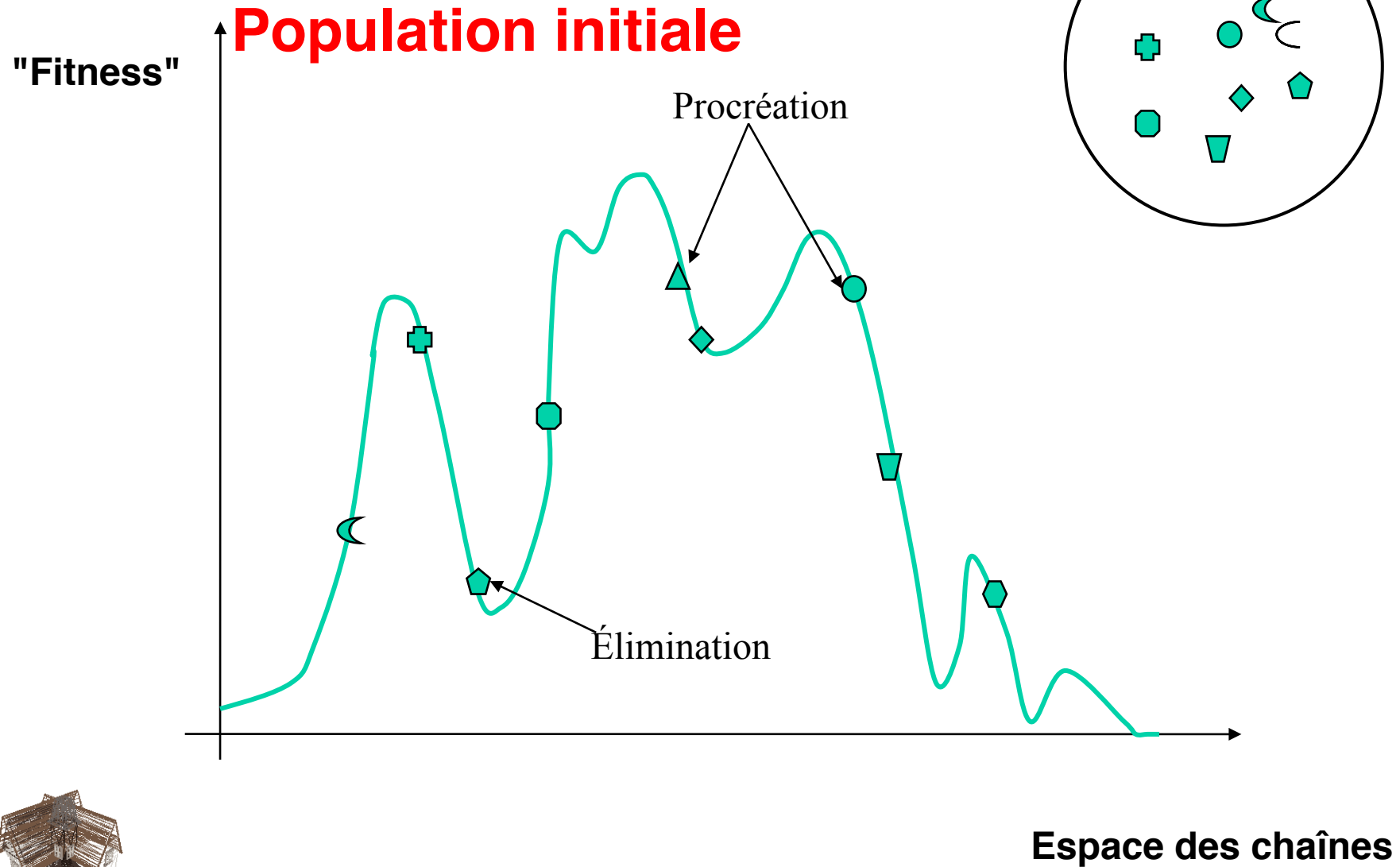
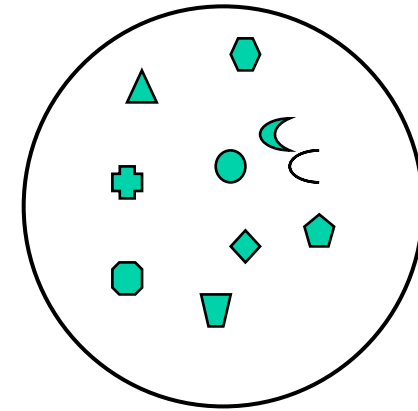
## *La nature: l'évolution*

- *Algorithmes génétiques (Holland)*

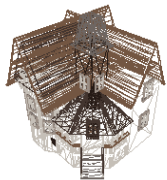
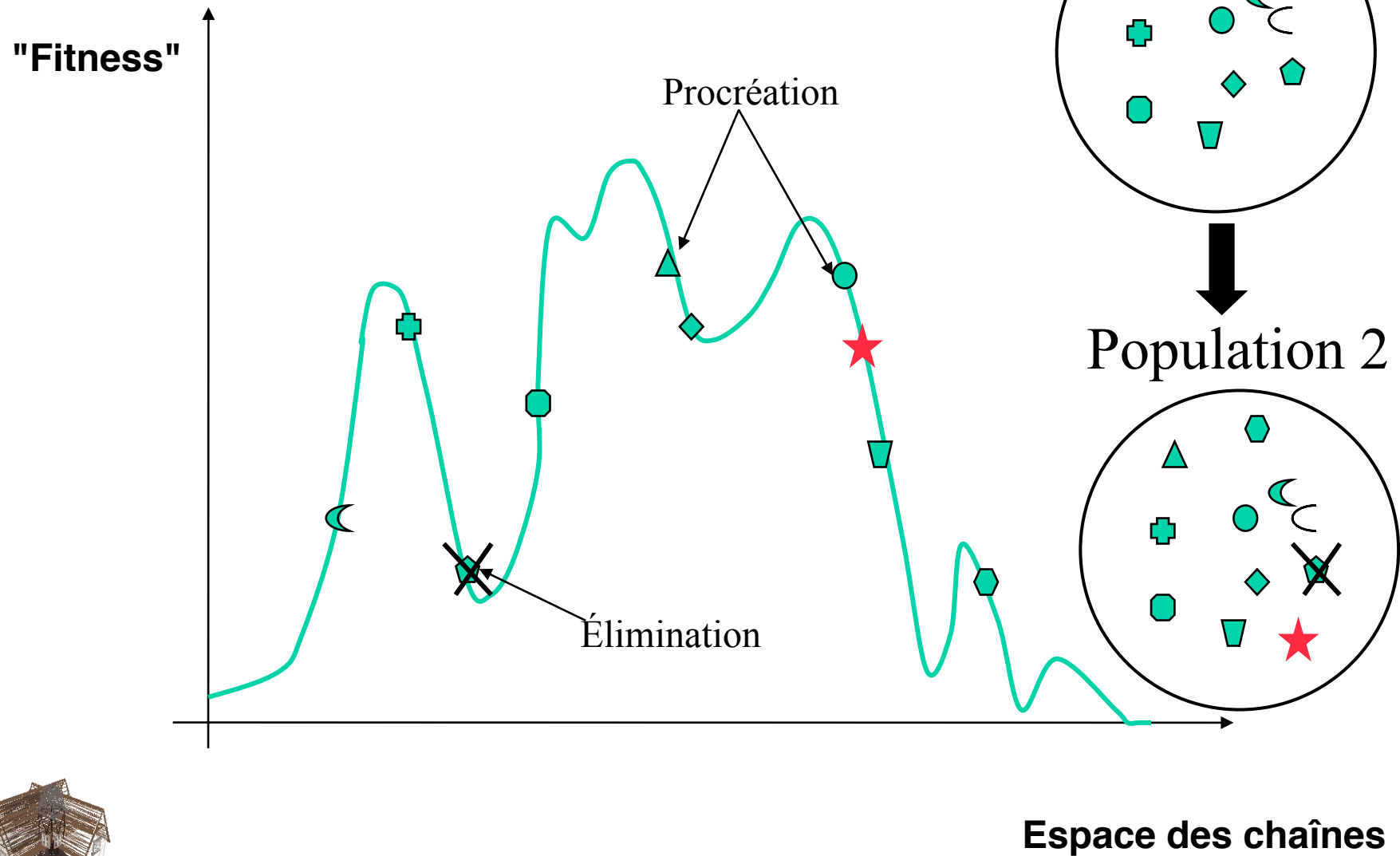


# Adaptation

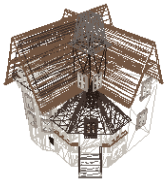
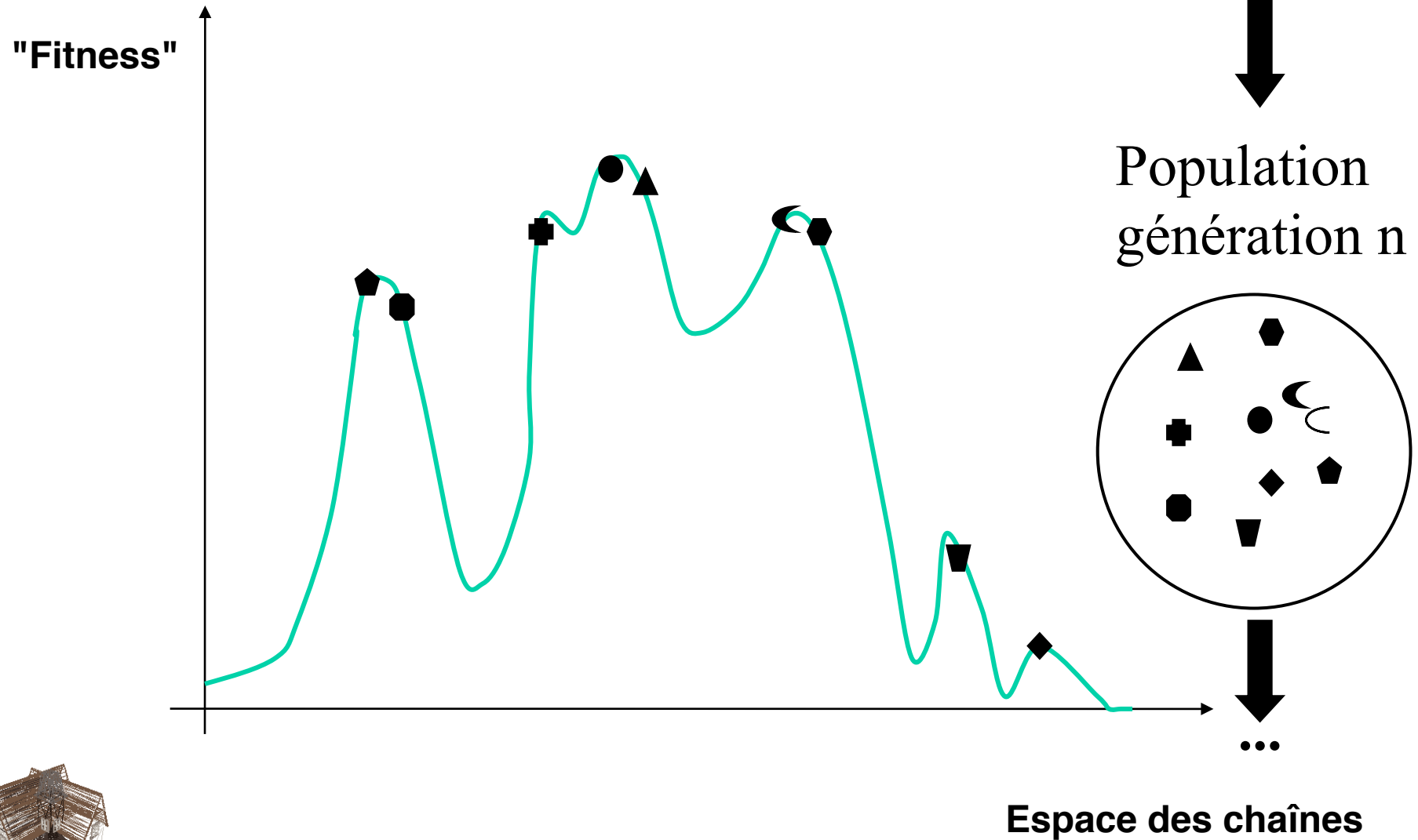
Population 1



# Adaptation



# Adaptation



# Les algorithmes

## *L'homme: théories classique de l'apprentissage*

- *Apprentissage par renforcement (Samuel)*
- *EPAM – apprentissage non supervisé (Feigenbaum)*
- *Apprentissage par l'action - “chunking” (Newell et al.)*

## *La physiologie: le cerveau*

- *Perceptron (Rosenblatt)*

## *La nature: l'évolution*

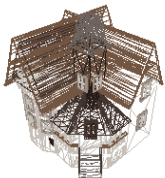
- *Algorithmes génétiques (Holland)*

## *La culture: l'épistémologie*

- *Découverte scientifique (Simon, Langley, Zytkow...)*

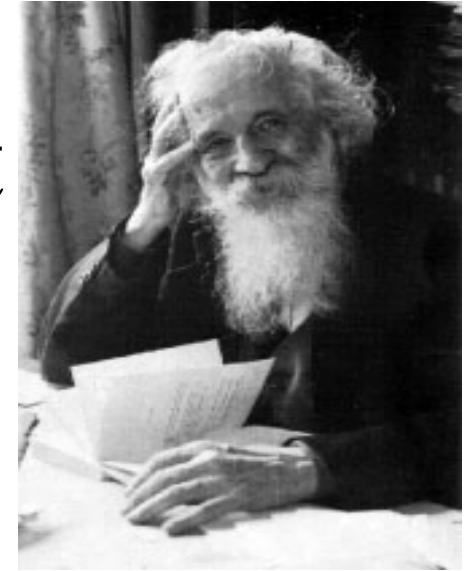
## *La physique:*

- *Arbres de décision – théorie de l'information (Quinlan)*



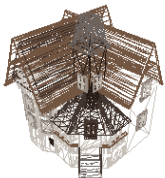


# Du concret à l'abstrait



## *Plan*

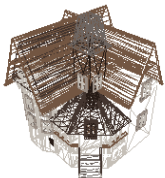
- *Les images*
- *Les algorithmes*
- *Les généralisations*
- *Les théories formelles*





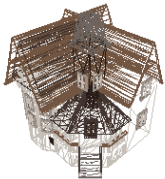
# Les généralisations

- Apprentissage supervisé/non supervisé
- Codage des exemples (extraction de traits) et connaissance
- Symbolique vs. numérique
- Evaluation des résultats
- Minimisation du risque / optimisation de la récompense
- Fourniture des exemples: méthodes d'ensemble (“bagging”, “boosting”)



# Les généralisations

- **Apprentissage supervisé/non supervisé**
- Codage des exemples (extraction de traits) et connaissance
- Symbolique vs. numérique
- Evaluation des résultats
- Minimisation du risque / optimisation de la récompense
- Fourniture des exemples: méthodes d'ensemble (“bagging”, “boosting”)



# Apprentissage supervisé/non supervisé

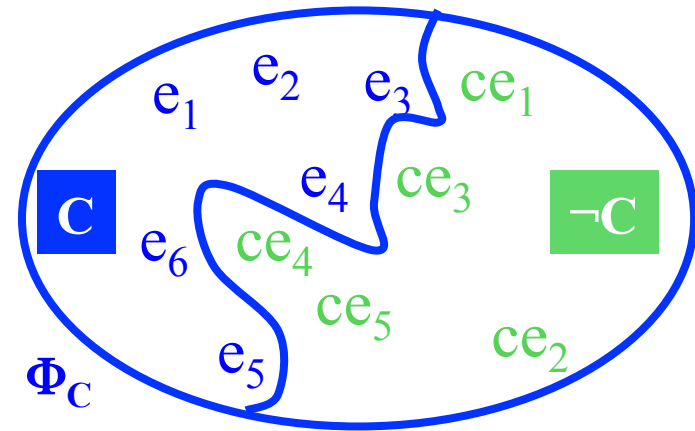
## Apprentissage supervisé

Étant donné:

- Un ensemble d'exemples de  $C$
- Un ensemble de contre exemples de  $C$

construire:

Une fonction  $\Phi_C$  compatible avec les exemples et les contre exemples



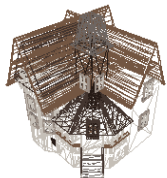
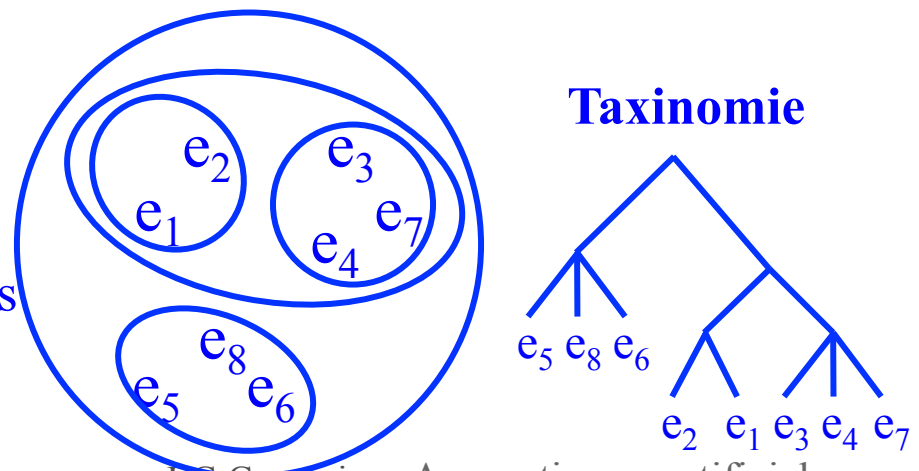
## Apprentissage non supervisé

Étant donné:

- Un ensemble d'exemples

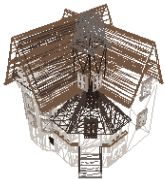
Construire:

Une structure  $S$  organisant les exemples eu égard à une distance  $D$

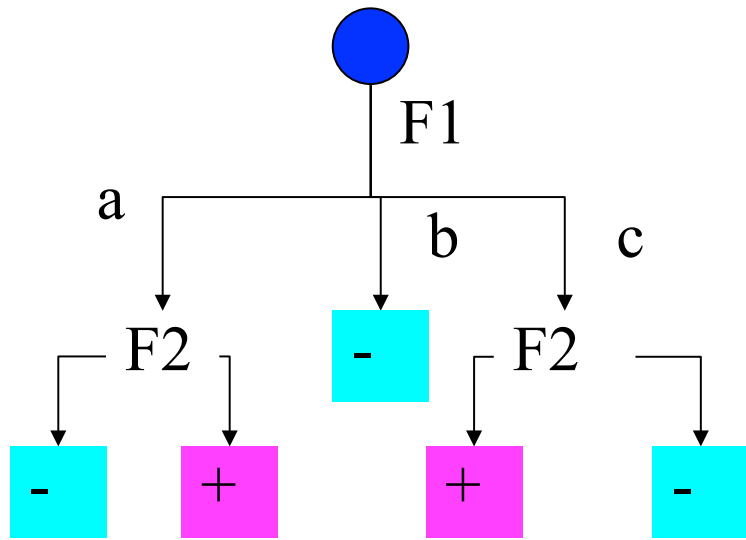


# Les généralisations

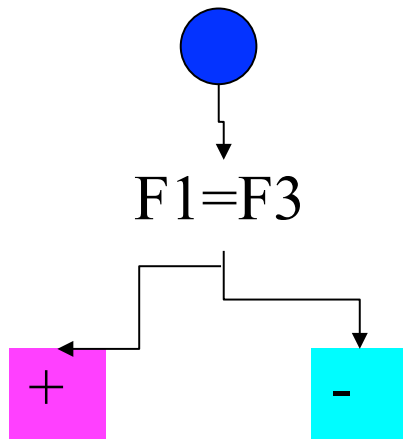
- Apprentissage supervisé/non supervisé
- **Codage des exemples (extraction de traits) et connaissance**
- Symbolique vs. numérique
- Evaluation des résultats
- Minimisation du risque / optimisation de la récompense
- Fourniture des exemples: méthodes d'ensemble (“bagging”, “boosting”)



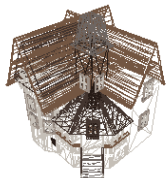
# Illustration



	E1	E2	E3	E4	E5
F <sub>1</sub>	a	c	a	b	c
F <sub>2</sub>	c	a	a	c	c
F <sub>3</sub>	a	c	c	a	b
C	+	+	-	-	-

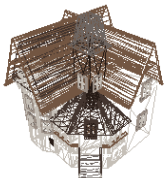


	E1	E2	E3	E4	E5
F <sub>1</sub>	a	c	a	b	c
F <sub>2</sub>	c	a	a	c	c
F <sub>3</sub>	a	c	c	a	b
F <sub>1</sub> =F <sub>3</sub>	Vrai	Vrai	Faux	Faux	Faux
C	+	+	-	-	-



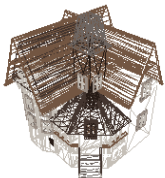
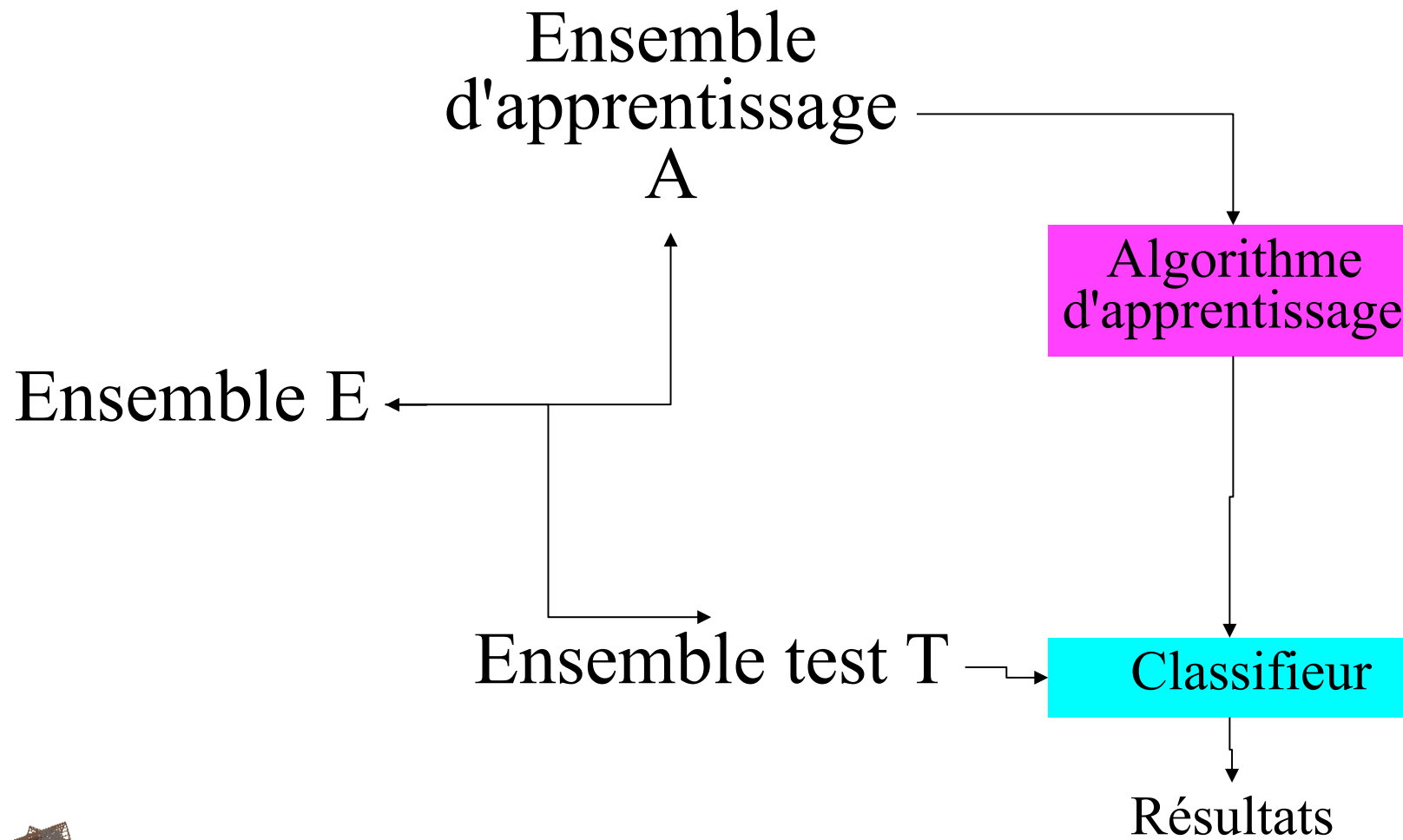
# Les généralisations

- Apprentissage supervisé/non supervisé
- Codage des exemples (extraction de traits) et connaissance
- Symbolique vs. numérique
- **Evaluation des résultats**
- Minimisation du risque / optimisation de la récompense
- Fourniture des exemples: méthodes d'ensemble (“bagging”, “boosting”)

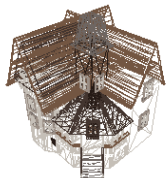
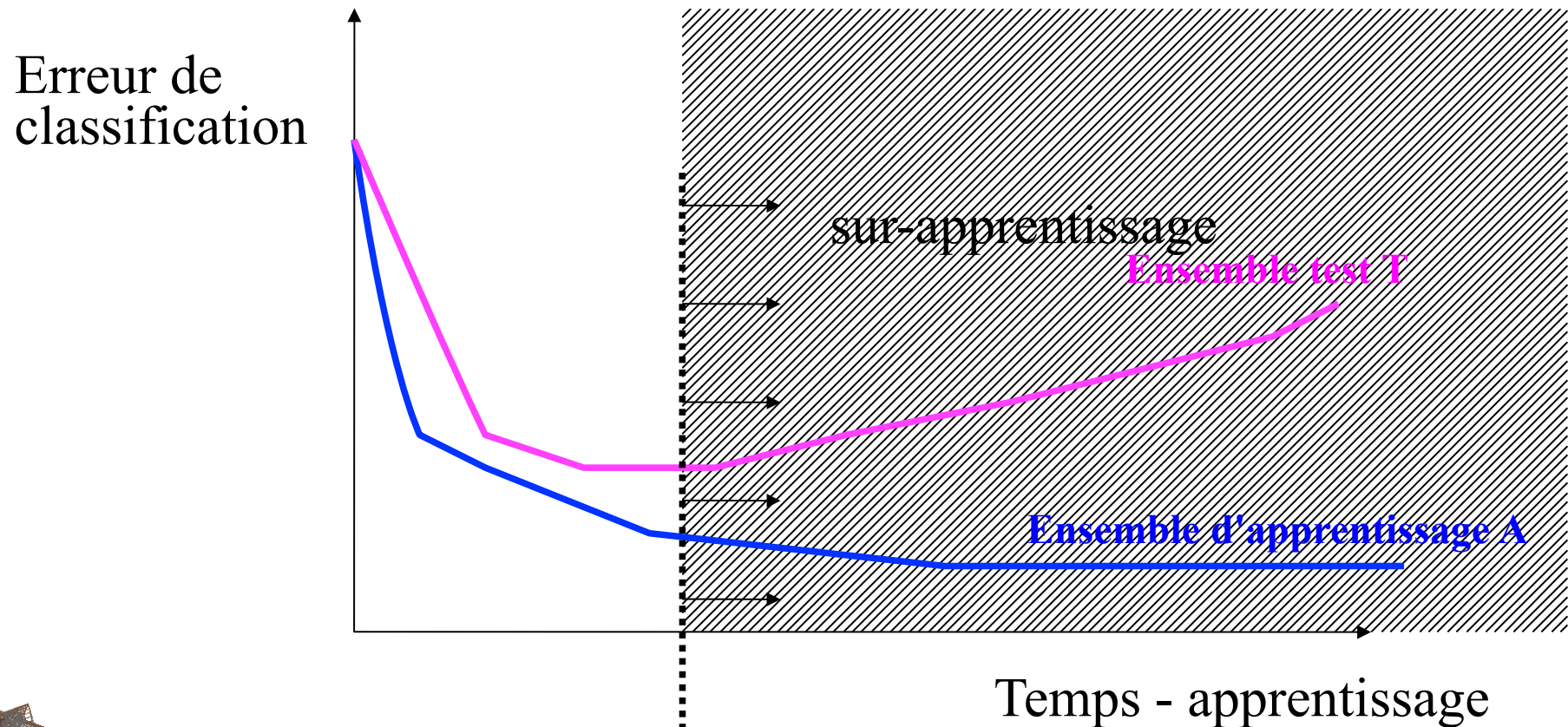




# Evaluation

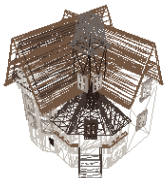


# Sur-apprentissage (“overfitting”)



# Les généralisations

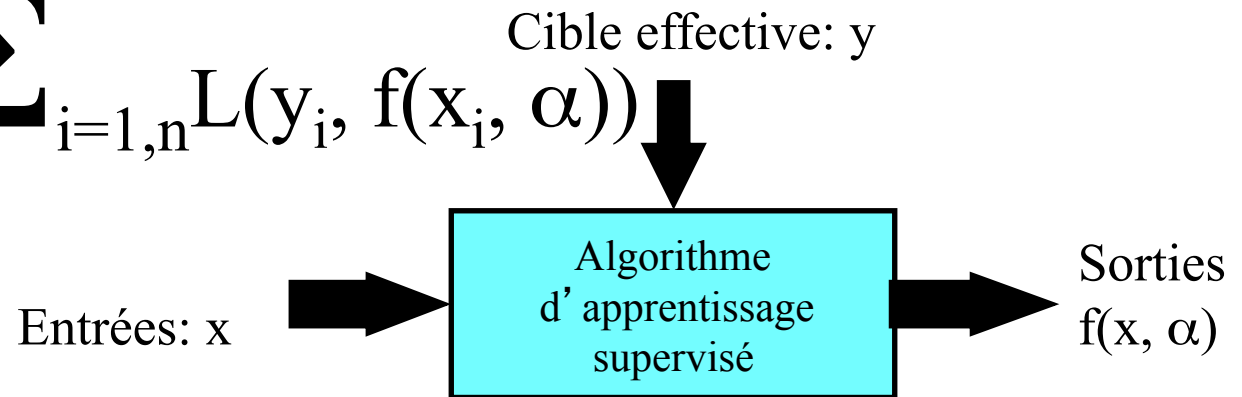
- Apprentissage supervisé/non supervisé
- Codage des exemples (extraction de traits) et connaissance
- Symbolique vs. numérique
- Evaluation des résultats
- **Minimisation du risque / optimisation de la récompense**
- Fourniture des exemples: méthodes d'ensemble (“bagging”, “boosting”)



# Objectifs...

- Minimisation du risque empirique:  
*fonction de coût*  $L(y, f(x, \alpha))$

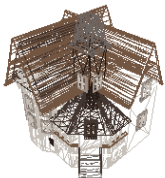
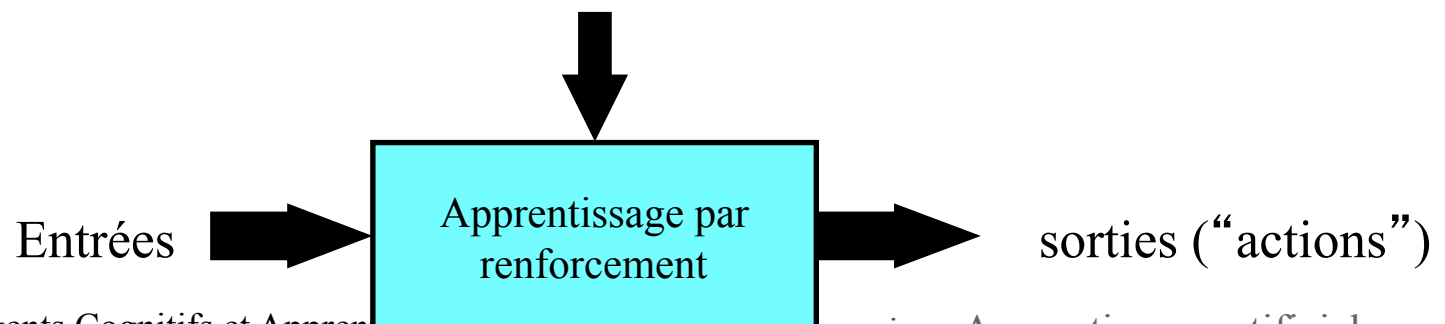
$$R_{\text{emp}}(\alpha) = 1/n \sum_{i=1, n} L(y_i, f(x_i, \alpha))$$



- Optimisation des récompenses  
*récompenses*  $r$

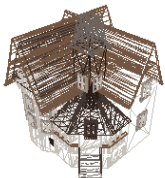
$$E(\sum_{t=0.. \infty} \gamma^t r_t)$$

évaluation ("récompenses" / "punitions")



# Les généralisations

- Apprentissage supervisé/non supervisé
- Codage des exemples (extraction de traits) et connaissance
- Symbolique vs. numérique
- Evaluation des résultats
- Minimisation du risque / optimisation de la récompense
- **Fourniture des exemples: méthodes d'ensemble (“bagging”, “boosting”)**



# Bagging: “Bootstrap Aggregating”

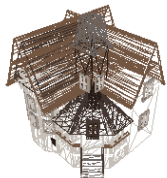
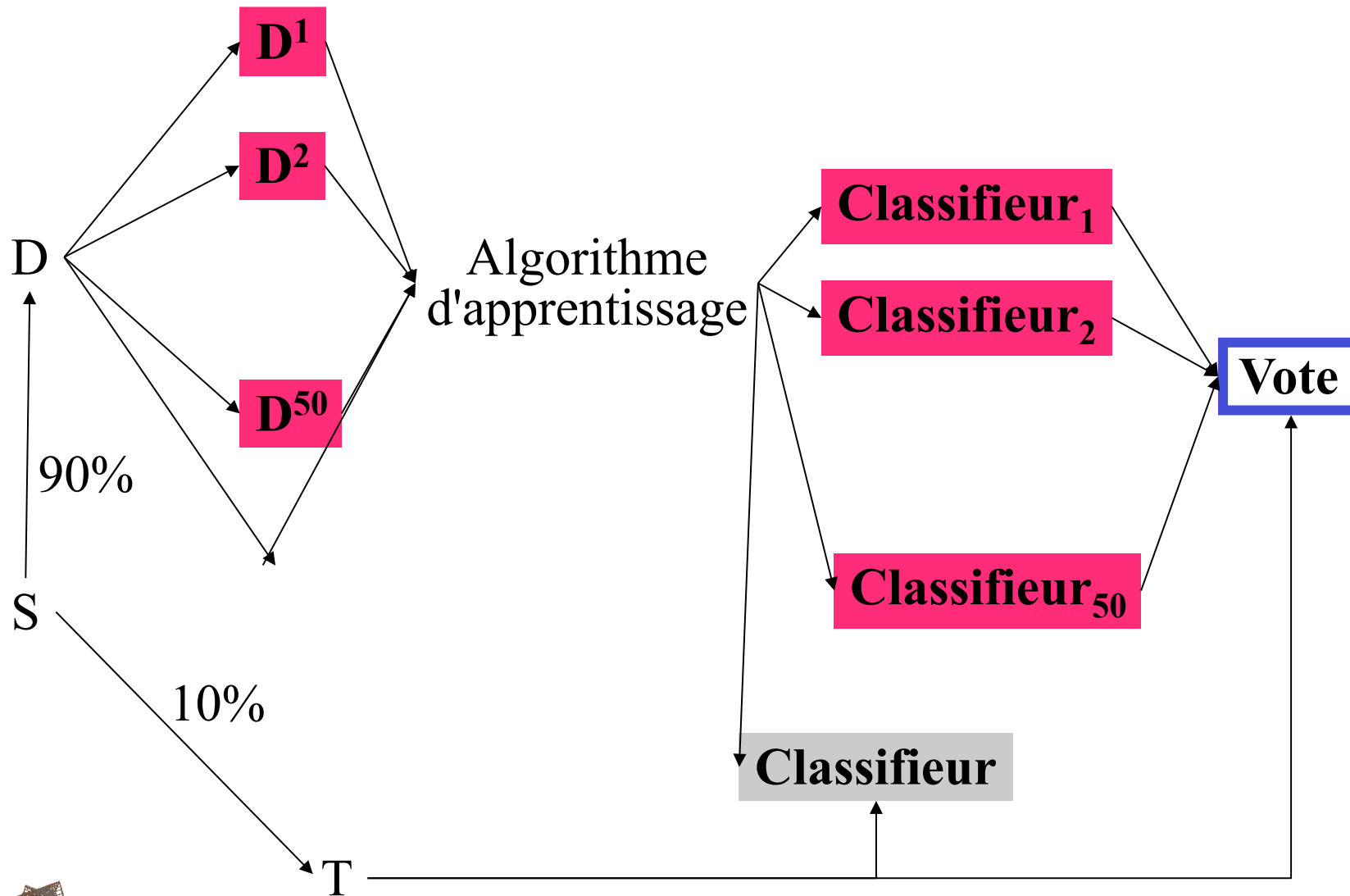
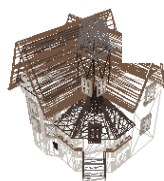


Table 5 Test Set Error (%)

<u>Data Set</u>	arc-fs	arc-x4	bagging	CART
heart	1.1	1.0	2.8	4.9
breast cancer	3.2	3.3	3.7	5.9
ionosphere	6.4	6.3	7.9	11.2
diabetes	26.6	25.0	23.9	25.3
glass	22.0	21.6	23.2	30.4
soybean	5.8	5.7	6.8	8.6
-----				
letters	3.4	4.0	6.4	12.4
satellite	8.8	9.0	10.3	14.8
shuttle	.007	.021	.014	.062
DNA	4.2	4.8	5.0	6.2
digit	6.2	7.5	10.5	27.1

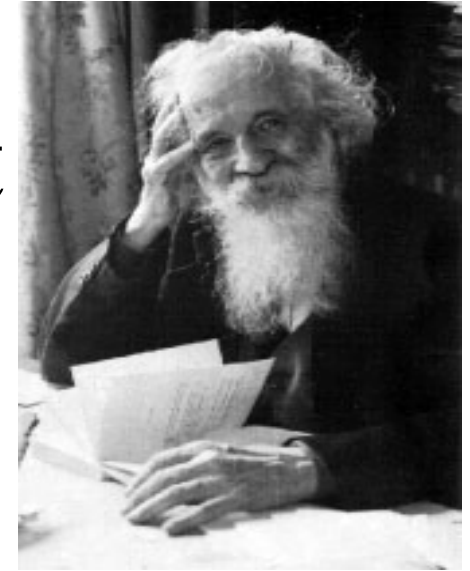
# Résultats expérimentaux





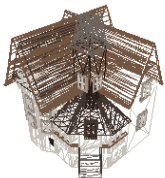


# Du concret à l'abstrait



## *Plan*

- *Les images*
- *Les algorithmes*
- *Les généralisations*
- *Les théories formelles*



# Théories formelles de l'apprentissage



A

C

Que peut-on apprendre?  
Combien faut-il d'exemples?  
A Quel temps faut-il?



# Théories formelles de l'apprentissage

Théorie statistique de l'apprentissage (*Vapnik*)

« PAC » – *Probably Approximately Correct*

(*Valiant*)

- Apprentissage d'une bonne approximation  $q$  de la fonction  $f$ :

$$\Pr\{\Pr(f \Delta q) \leq \varepsilon\} \geq 1 - \delta$$

# Théories formelles de l'apprentissage

Théorie statistique de l'apprentissage (*Vapnik*)

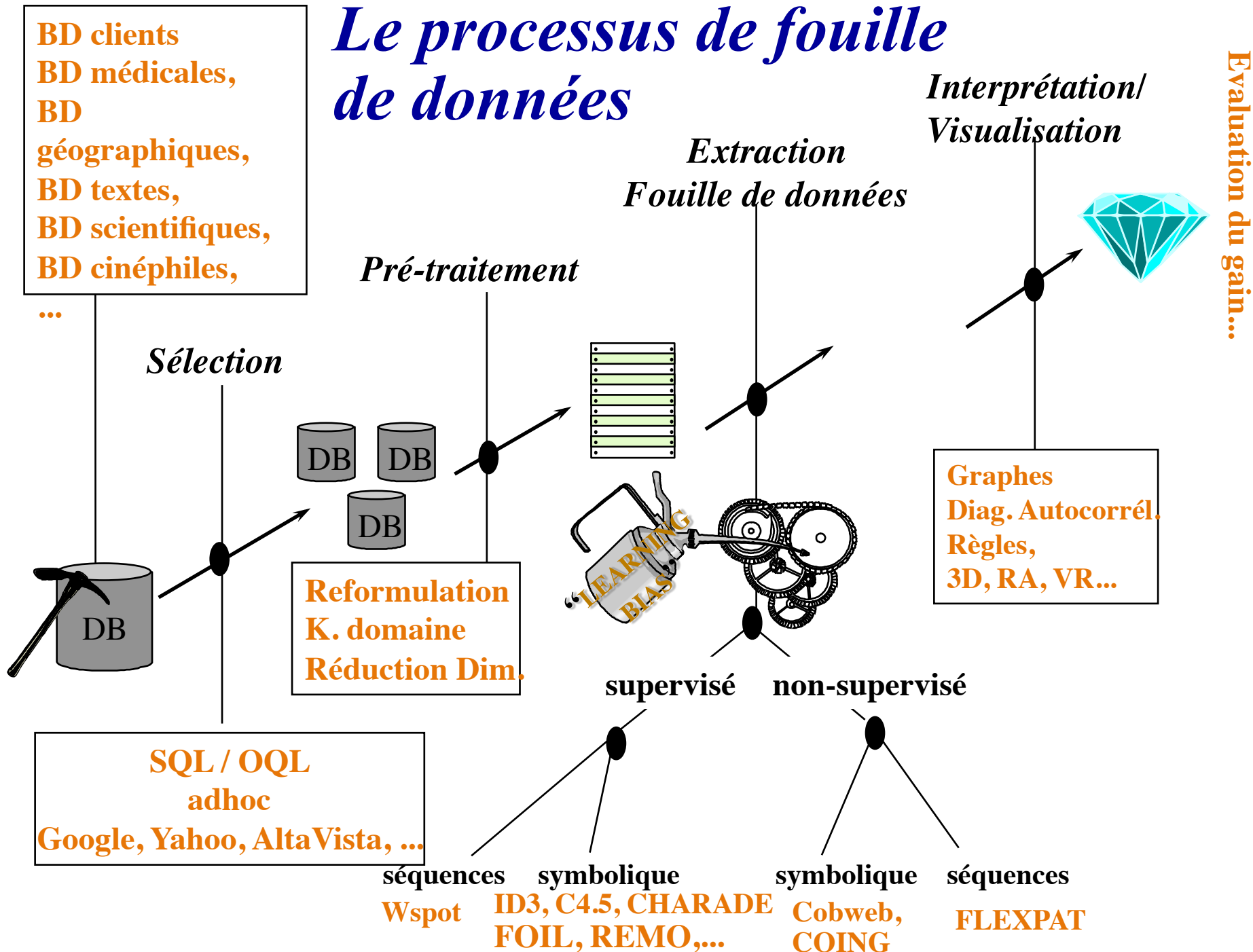
« PAC » – *Probably Approximately Correct* (*Valiant*)

- Apprentissage d'une bonne approximation  $q$  de la fonction  $f$ :

$$\Pr\{\Pr(f \Delta q) \leq \varepsilon\} \geq 1 - \delta$$

- Complexité d'échantillonnage
  - Nombre d'exemples en fonction de  $n$ ,  $1/\varepsilon$  et  $1/\delta$
  - Introduction de la dimension dite de *Vapnik-Chervonenkis* d'une classe de fonctions:  $D_{vc}(F^{[n]})$
- Complexité algorithmique
  - Nombre de pas de l'algorithme d'apprentissage
  - Vitesse de convergence...

# Le processus de fouille de données



Invention Imagination

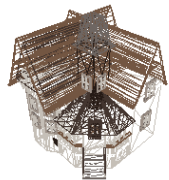
Découverte **CRÉATIVITÉ**

Évolution

Apprentissage...

Adaptation

Acquisition

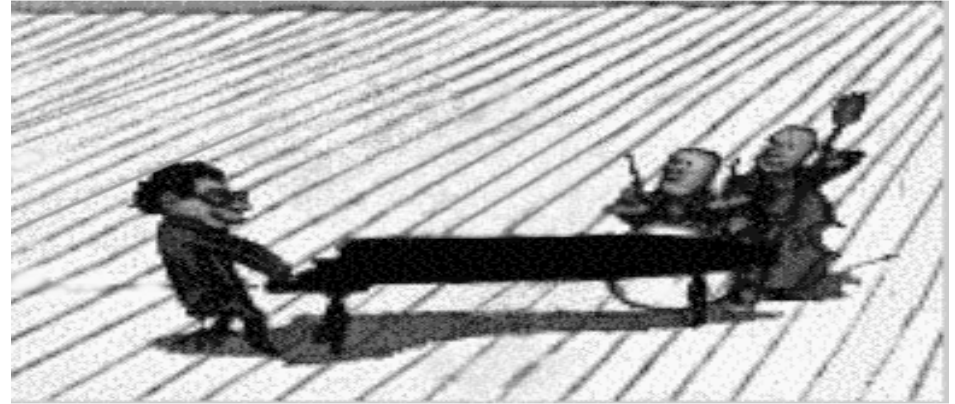


**Renforcement**





# Perspective de l'IA sur la créativité



- **Schank: Explanation patterns**
  - La création n'est pas une affaire magique : tous les hommes créent dans leurs activités quotidiennes, tandis qu'ils ne font pas tous des maths...

**Imagination**

=

**recombinaison d'éléments mémorisés**

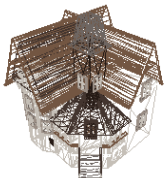




# Labex OBVIL

Sorbonne Universités (*UPMC – Paris Sorbonne*)  
« Digital Humanities » – *Humanités numériques*

- Stylistique électronique:
  - Extraction de motifs syntaxiques récurrents
- Détection de réemplois sur de gros corpus (« big data »)
- Comparaisons de versions
- Cartographie de contenus
- Indexation sémantique
- ...



# Apprentissage symbolique

D

- Arbres de décision
- Règles d'association , analyse formelle de concepts
- Induction structurelle.

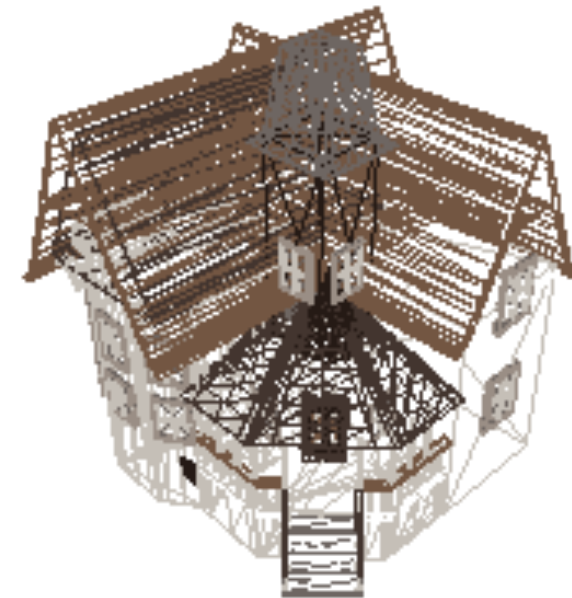
A

- Règles de résolution
- Apprentissage relationnel & ILP

- Acquisition des connaissances
  - Modélisation et apprentissage

C

- Apprentissage non supervisé
  - Construction de taxinomies
  - Extraction de motifs dans les séquences, les arbres et les graphes



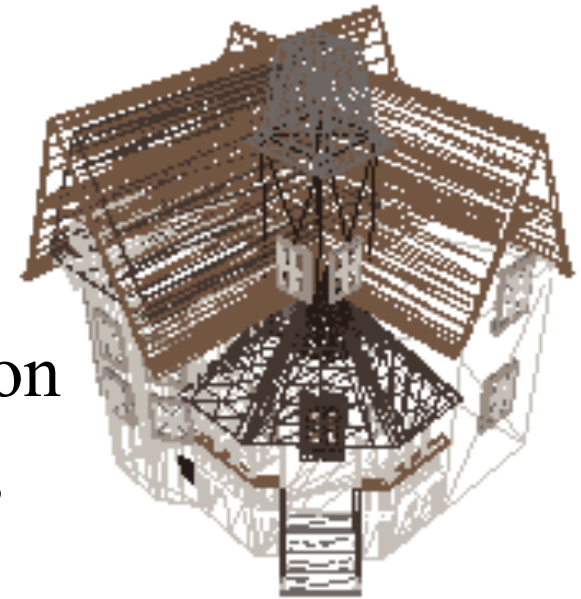
# Apprentissage symbolique

D

- Découverte scientifique
- Apprentissage et traitement automatique des langues

A

- Étude de la créativité – humanités numériques
  - MEDITE: comparaison de version
  - PHOEBUS + Détection galaxies
  - DeSeRT (moteur sémantique)
  - Constellations conceptuelles
  - Reconnaissance entités nommées, etc.



# Questions

D

- Connaissances:
  - Ontologies
  - Construction arbres de décision
  - Règles d'association
  - Règle de résolution

A

C

