

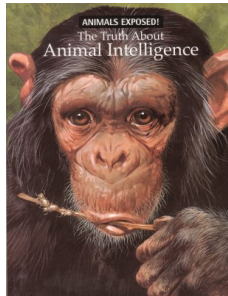
# Algorithmes d'apprentissage et intelligence artificielle

**Yoshua Bengio**

**3 septembre 2008**

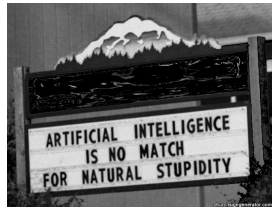
# Intelligence Naturelle

- S'adapter à de nouvelles situations pour survivre / se reproduire.
- Nécessite l'acquisition de schémas comportementaux variés (évolution + individu)
- Permettent de généraliser à de nouvelles situations



# Intelligence Artificielle

- Plus d'un demi-siècle de recherches, et le but semble encore très loin
- Pourquoi ?
- Trop pressés d'obtenir des résultats, plutôt que de comprendre ?



# Où prendre les connaissances ?

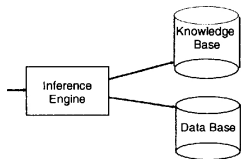
- L'IA doit capter un grand nombre de connaissances sur notre monde.
- L'approche explicite-symbolique :  
**Cyc** = collection de règles et faits écrits par des humains



Titanesque



Incohérent/Incomplet



Pas robuste

# A apprendre les connaissances ?

- Animaux et humains : connaissances innées et acquises.
- Apprennent des tâches non prévues par l'évolution.
- Plusieurs méthodes statistiques = apprentissage
- Apprentissage statistique maintenant dans la plupart des applications de l'IA : vision, reconnaissance de la parole, traduction, moteurs de recherche, robotique, bio-informatique, chemo-informatique, intelligence d'affaires...
- Pari de l'existence de stratégies/principes assez généraux.

# Apprentissage supervisé



© 1991 CAMUSO - SYRACUSE-HERALD JOURNAL

# Apprentissage supervisé

réponse

Spider Lizard Insect Other

apprenant

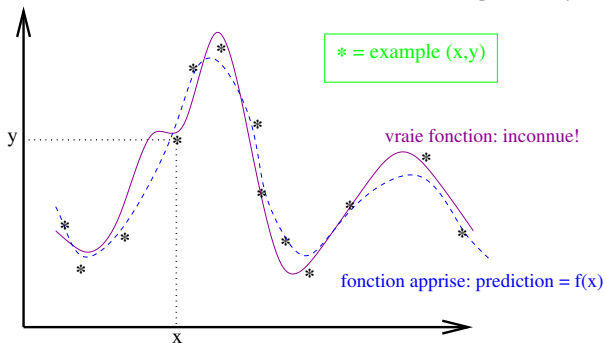
entrée



correction du professeur

# Apprentissage supervisé

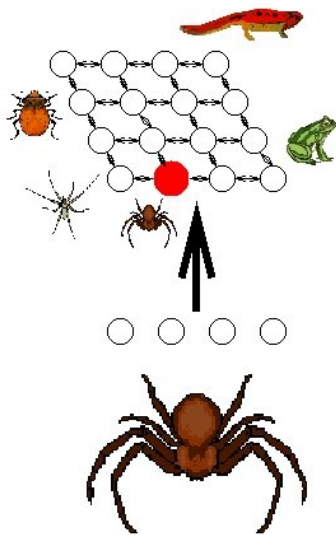
- La forme la plus simple d'apprentissage
- Ensemble d'apprentissage : paires ( $x$  =entrée,  $y$  =sortie désirée)
- Tirées d'un processus **inconnu**  $\mathcal{P}$
- Critère d'erreur( $y$ ,prédiction)
- On cherche une fonction  $f$  qui minimise  $E[\text{erreur}(y, f(x))]$  sous  $\mathcal{P}$



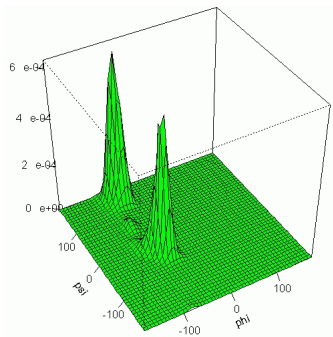


# Apprentissage non-supervisé

Découvrir des catégories et autres facteurs naturels

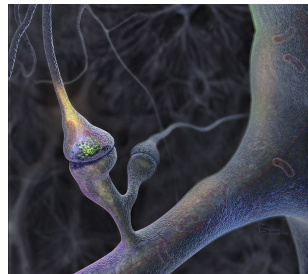


- La plupart des exemples disponibles aux humains et aux machines ne sont pas sémantiquement étiquetés
- **Objectif** : capter les régularités statistiques dominantes
- = découvrir où la vraie densité se concentre

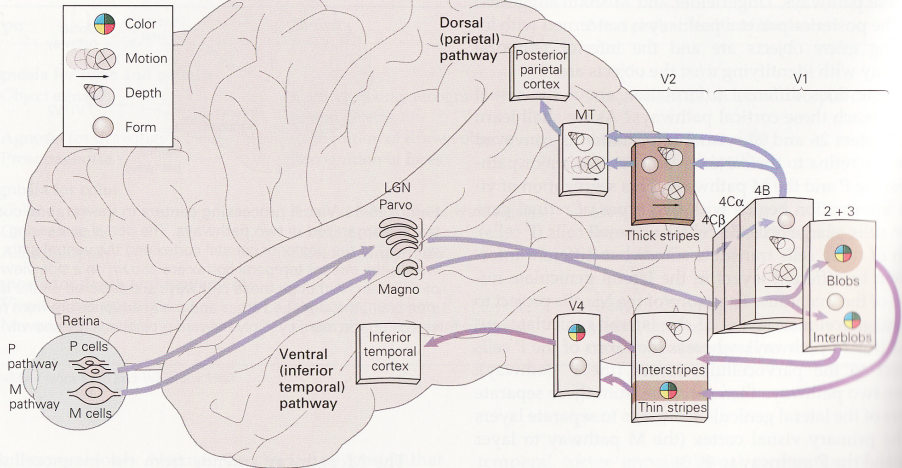


# Apprentissage dans le cerveau

- $10^{11}$  neurones et  $10^{14}$  synapses (connexions)
- Apprentissage = modification / destruction / création de synapses
- Le cortex est la partie la plus grande et la plus moderne
- Fraction significative des neurones actifs à la fois
- Apprentissage au fur et à mesure (“online”)



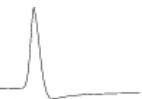
# Système visuel des primates



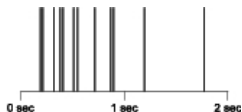
Système visuel : séquence de transformations/niveaux d'abstraction

# Apprentissage dans le cerveau

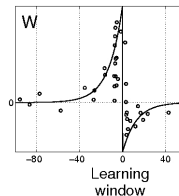
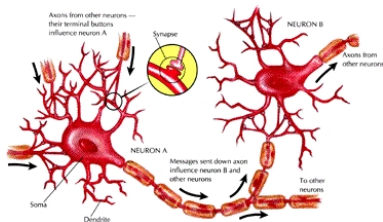
- Sortie du neurone = train d'impulsions ( $\sim 1$  par 5ms à 100ms ) avec aléas apparents
- Neurone accumule signaux d'autres neurones (1000 en moy.)
- Fréquences in/out : non-linéarité monotone bornée du neurone
- L'effet dominant apprentissage : STDP (à la Hebb)



(a)



(b)



# Généraliser

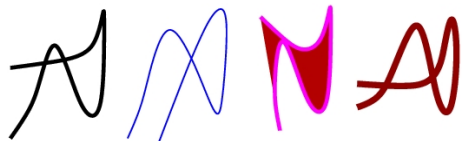
Voici un exemple de la lettre Zuf en Klingon :



Voici d'autres exemples :



Lesquels parmi ces caractères Klingons sont des Zuf ?



Pas de réponse universelle  $\Rightarrow$  problème mal posé !

# Généraliser $\neq$ Apprendre par cœur

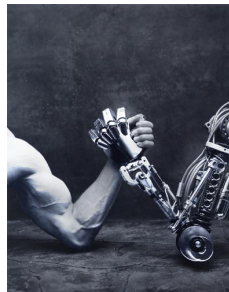
Facile pour l'ordi d'apprendre par cœur



Pas pour nous !



$\Rightarrow$  Mais nous, nous généralisons naturellement !



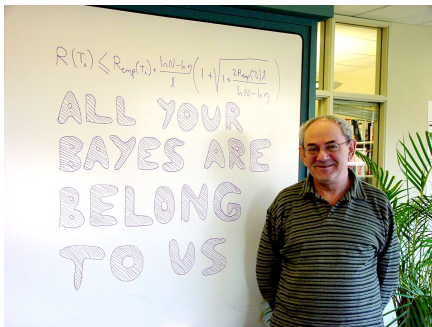
# Compresser pour généraliser

Extraire l'essence des observations  $\Rightarrow$  généraliser

- Principe du Rasoir d'Occam



- Résultats de Kolmogorov/Solomonoff 1964
- Résultats de Vapnik 1972

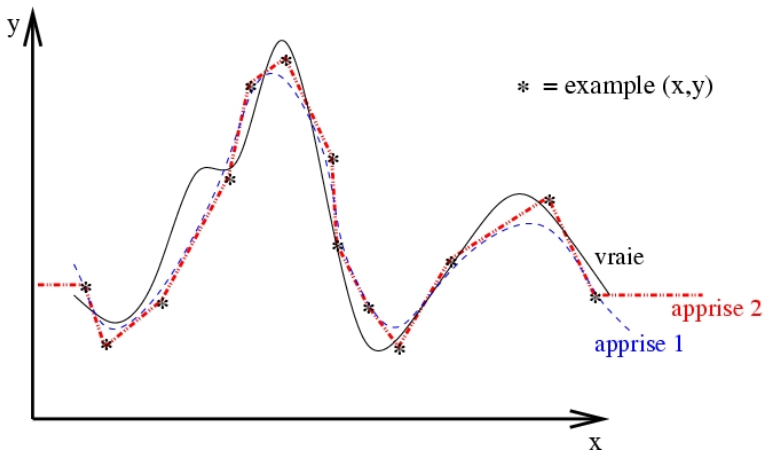


$$\min_f \frac{1}{n} \sum_{i=1}^n \text{erreur}(x_i, f(x_i)) \neq \min_f E[\text{erreur}(x_i, f(x_i))]$$



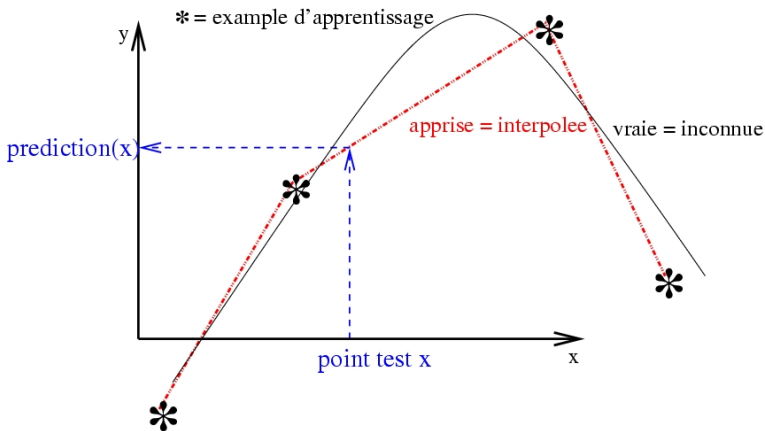
# Généraliser en 1D

C'est facile si la fonction n'a pas beaucoup de variations.



# Généraliser localement

Ça fonctionne bien avec la bonne représentation : où la notion de voisinage fonctionne.




# Situations où généraliser localement échoue

Apprentissage :



Test :



- Distance euclidienne pixel à pixel inappropriée
- Klingons utilisent un espace de représentation plus abstrait