

Réseaux connexionnistes

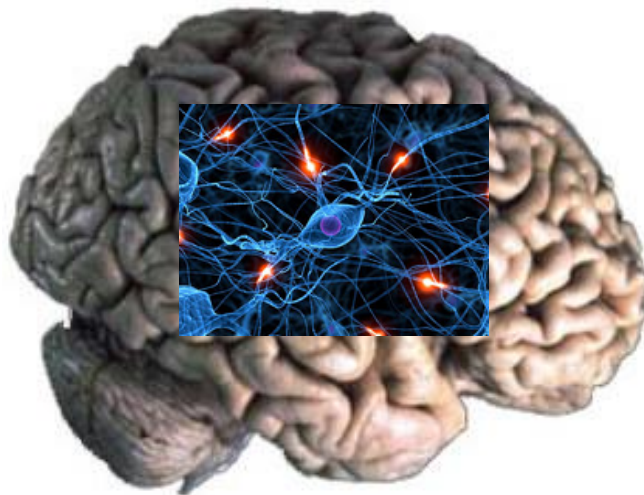
Cours électif « Intelligence Artificielle »

L'I.A. au début des années 60...

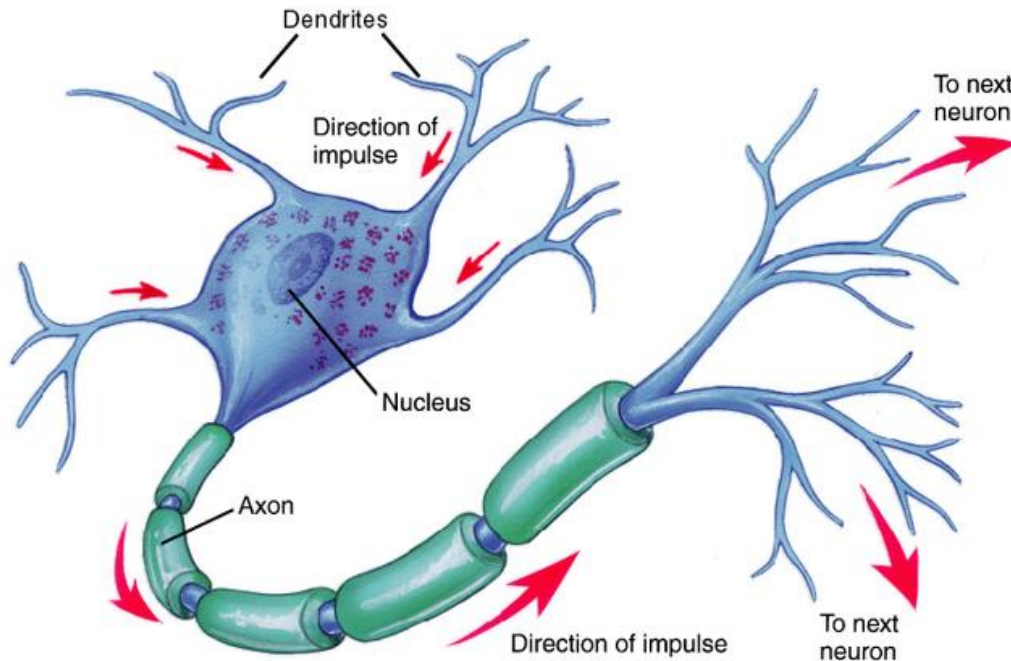
- Apparition des premières technologies d'Intelligence Artificielle
- Recherche de méthodes de raisonnement
 - générales à une classe de problèmes
 - capables d'apprentissage
 - fonctionnant en présence d'informations bruitées, vagues, ...
 - adaptatives
- Approche mimétique de l'intelligence humaine
 - Le fonctionnement physiologique du cerveau comme inspiration pour des réseaux connexionnistes

Le cerveau humain

- Récepteur des perceptions issues des 5 sens
- Capacité d'analyse et d'apprentissage
- Centre de contrôle des organes moteurs du corps humain
- Architecture composée de 10^{11} neurones reliés chacun à 100 autres neurones en moyenne



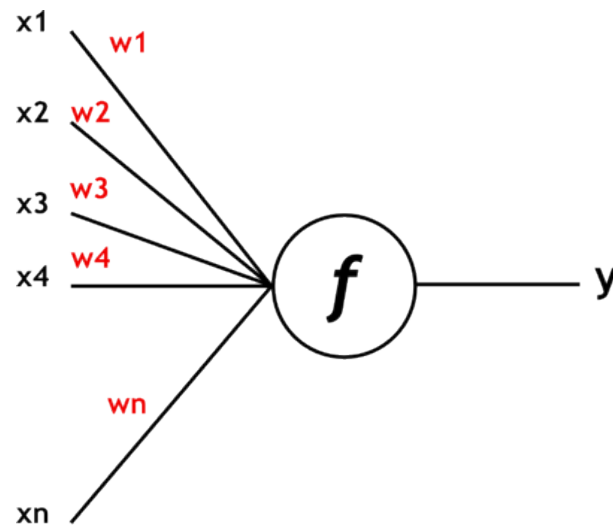
Le neurone biologique



McCulloch et Pitts (1943) ont modélisé le fonctionnement d'un neurone biologique et imaginé une reproduction artificielle mimétique du raisonnement humain

Le neurone formel

- Reproduction du fonctionnement biologique



- Le perceptron (Rosenblatt, 1958)

Le perceptron

Un perceptron a

- n entrées x_1, \dots, x_n avec n coefficients synaptiques, w_1, \dots, w_n
- 1 unique sortie avec un seuil θ

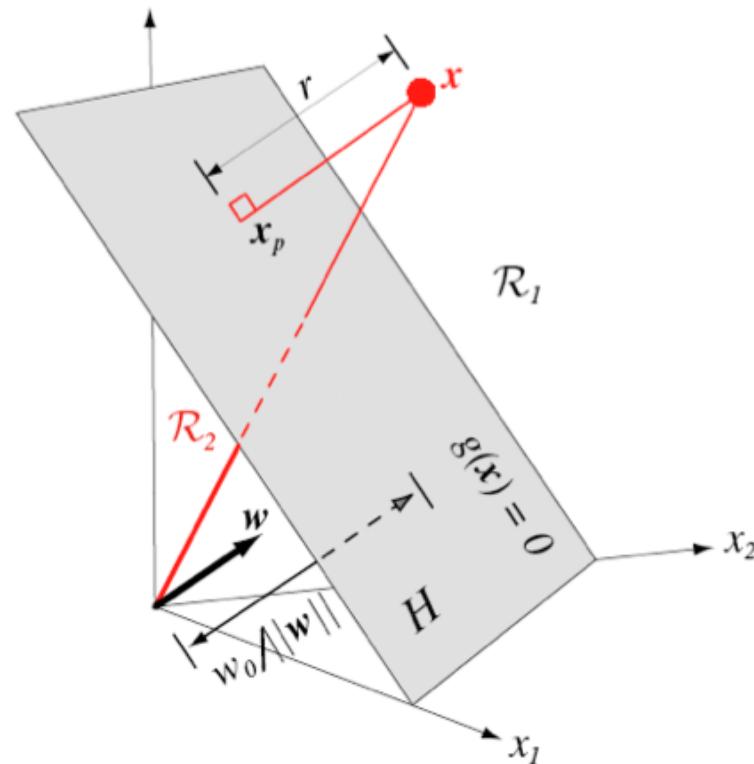
Potentiel post-synaptique : $\sum_i w_i * x_i$

La fonction d'activation donne la sortie en fonction du potentiel post-synaptique

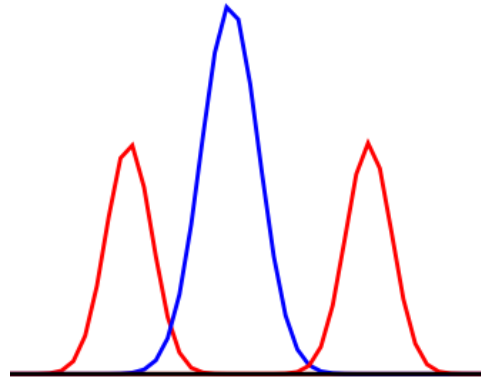
$$f(x) = \begin{cases} 1 & \text{si } x > \theta \\ 0 & \text{sinon} \end{cases}$$

Le perceptron (2)

Discrimination linéaire avec un hyperplan comme frontière de décision



Exemple : les eiders



Soit x un niveau de gris perçu

- Si $x < 1$ c'est un mâle
- Si $1 \leq x \leq 2$ c'est une femelle
- Si $x > 2$ c'est un mâle

Exemple tiré du cours de Laurent Miclet

Les eiders, résolution logique

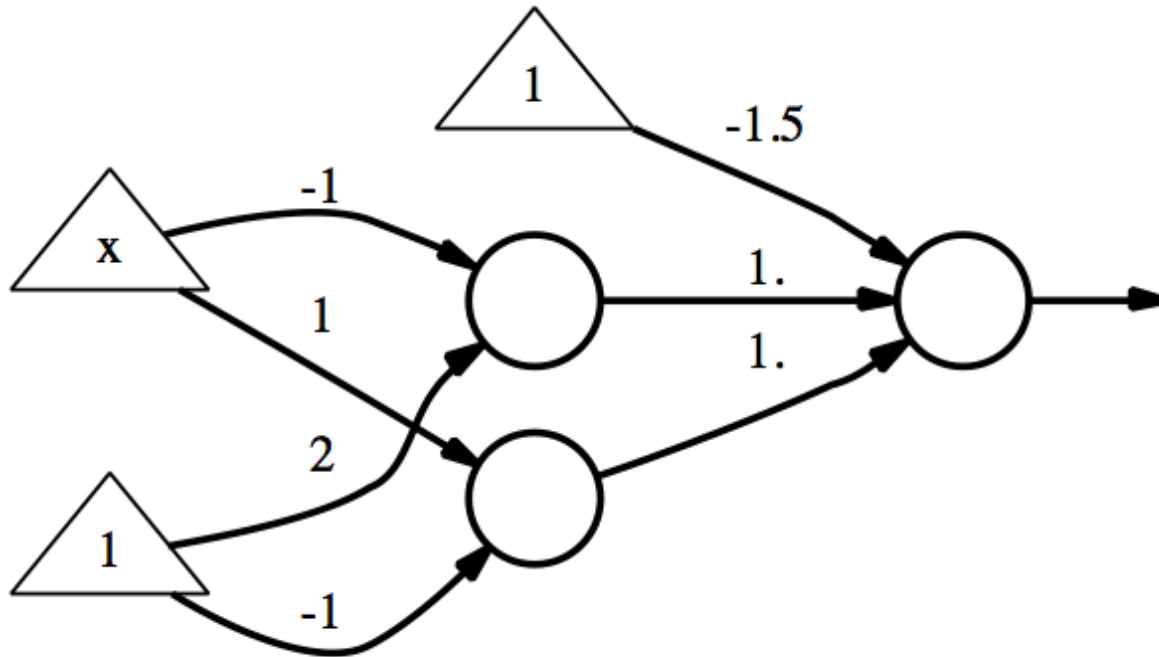
- Si $x \leq 2$ alors y_1 est vrai, faux sinon
- Si $x \geq 1$ alors y_2 est vrai, faux sinon

y1	y2	sexe
VRAI	VRAI	Femelle
VRAI	FAUX	Mâle
FAUX	VRAI	Mâle
FAUX	FAUX	impossible

Les eiders, résolution numérique

On attribue les valeurs 0 ou 1 à y_1 et y_2 et $z = y_1 + y_2 - 1,5$

Si z est positif c'est une femelle, sinon c'est un mâle



Limites du perceptron

En 1969, Marvin Minsky et Seymour Papert ont souligné les limites d'une approche connexionniste

- Incapacité à traiter des problèmes non linéaires
- Focalisation des fonds de recherche sur l'I.A. symbolique

Relance des travaux sur le connexionnisme au milieu des années 80

- Désillusions face aux attentes démesurées de l'I.A. symbolique
- Rétropropagation du gradient et perceptron multi-couches (Le Cun 1986, Rumelhart et al. 1986, Parker 1985)

Perceptron multi-couches : propagation

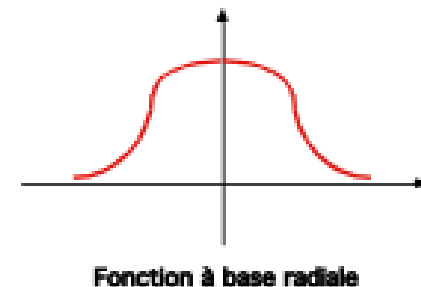
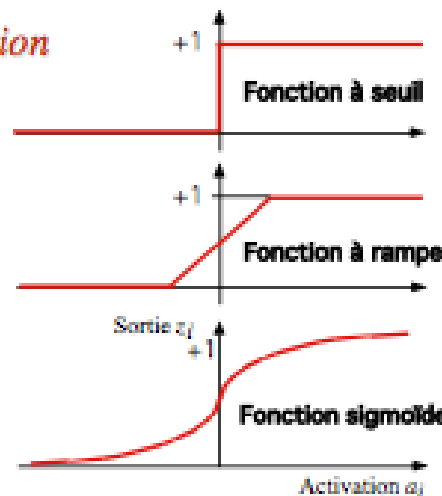
- Pour chaque neurone :

$$y_l = g\left(\sum_{j=0,d} w_{jk} \phi_j\right) = g(a_k)$$

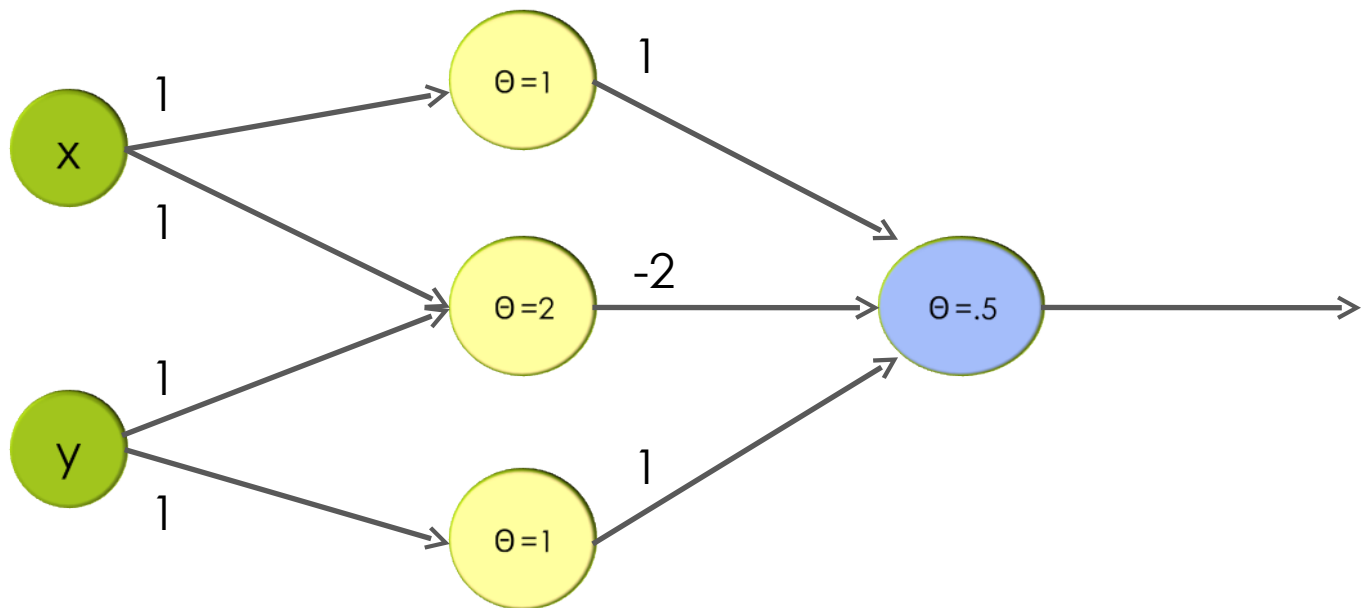
- w_{jk} : *poids* de la connexion de la cellule j à la cellule k
- a_k : *activation* de la cellule k
- g : *fonction d'activation*

$$g(a) = \frac{1}{1 + e^{-a}}$$

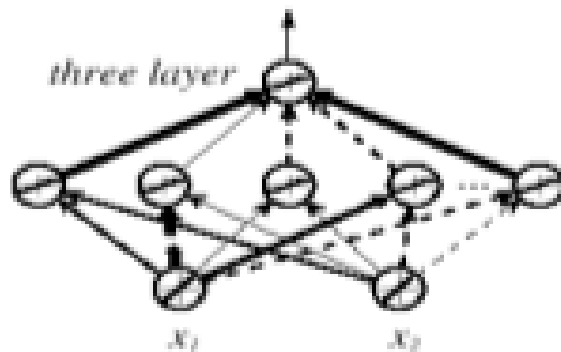
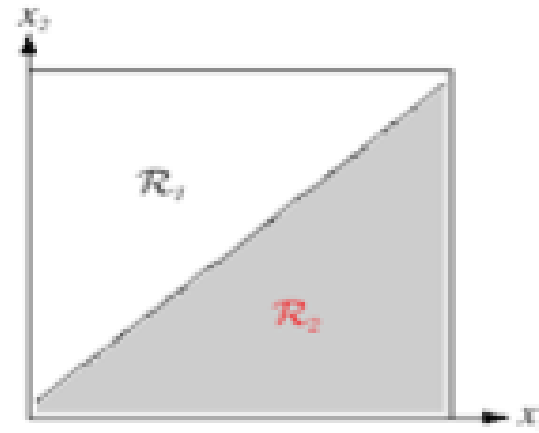
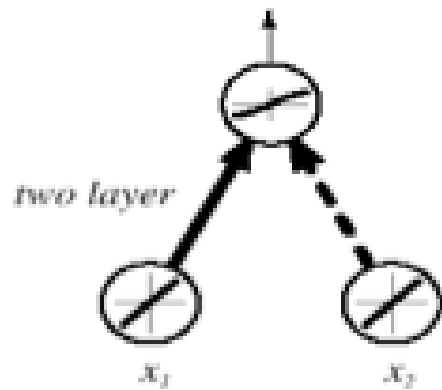
$$g'(a) = g(a)(1-g(a))$$



Exemple du XOR



PMC : puissance expressive



Mécanismes d'apprentissage

Un processus d'apprentissage a pour objectif de faire correspondre à un ensemble d'entrées, une sortie désirée

■ Apprentissage supervisé

- Le système dispose d'un échantillon d'entrées associées à une sortie désirée
- Entraînement du système pour ajuster ses paramètres à l'échantillon

■ Apprentissage non-supervisé

- Le système dispose d'un échantillon d'entrées sans sorties prévues
- Entraînement par une heuristique sur l'activité des entités

Apprentissage dans les PMC

- Objectif : trouver les poids permettant au réseau de réaliser une relation entrées-sortie spécifiée par des exemples
- Méthode :
 - Minimiser l'erreur produite en sortie des exemples issus de l'échantillon d'apprentissage
 - Algorithme de rétro-propagation du gradient
- Principe inductif : On fait l'hypothèse que l'échantillon d'apprentissage est représentatif des données réelles

Calcul de l'erreur (ADALINE)

- Si S est l'échantillon d'apprentissage, l'erreur d'un exemple s de S est

$$E(s) = (T_s - O_s)^2 / 2$$

Où T_s est la sortie attendue et O_s celle observée

- L'erreur globale du perceptron sur l'échantillon S est

$$E(S) = 1/|S| * \sum E(s)$$

- L'apprentissage consiste à **minimiser $E(S)$**

Rétro-propagation du gradient

- Les poids des entrées d'un neurone formel sont ajustés par une descente de gradient issu du calcul d'erreur de sa sortie
- $W_{i+1} = W_i + \eta \cdot E(s) \cdot X_i$
 - W_i est un poids à l'instant i
 - X_i est la valeur de l'entrée correspondante
 - $E(s)$ est l'erreur calculée en sortie
 - η est un paramètre de rapidité mais aussi de stabilité de l'apprentissage

Rétro-propagation du gradient : couches cachées

- Le problème est de déterminer la responsabilité de chaque connexion dans l'erreur
- Principe : estimer l'erreur d'une connexion en fonction de l'erreur sur la couche suivante
- Evaluation et utilisation des dérivées de l'erreur pour calculer les modifications des poids

PMC : applications

- Traitement d'images (reconnaissance de formes, d'écritures, ...)
- Traitement du signal (filtrage, compression, traitement de la parole, ...)
- Prédications (météorologie, bourse, consommations, ...)
- Diagnostic (industrie, médecine, ...)