

# Un modèle hebbien pour l'acquisition du lexique

Alexis Nasr & Arnaud Rey

26 octobre 2023

Ce stage s'inscrit à l'intersection de deux axes de recherche. Le premier est l'acquisition du lexique et plus précisément du modèle Parser [Perruchet98]. Le second est l'apprentissage associatif et plus précisément les réseaux de neurones hebbiens [Tovar18, Kuriscak15, Sejnowski99].

Parser est un modèle qui prend en entrée un flux de symboles (par exemple des phonèmes) et regroupe ces derniers en segments (par exemple des mots) en fonction de leur fréquence d'apparition. De manière plus précise, le modèle tient à jour un inventaire de segments auxquels il associe un poids. Lorsqu'un segment donné a tendance à se répéter, son poids augmente. Quand au contraire il ne se répète pas, son poids va décroître et lorsqu'il se situera en deçà d'un certain seuil, il n'apparaîtra plus dans l'inventaire des segments.

Les modèles de neurones hebbiens sont des réseaux de neurones artificiels dont l'apprentissage est fondé sur le principe de l'apprentissage associatif. Etant donné un vecteur d'entrée  $\mathbf{x}$  binaire de dimension  $d$ , le réseau se présente sous la forme d'un ensemble de neurones  $n_1 \dots n_d$  où le neurone  $n_i$  est connecté à l'entrée  $x_i$  et à l'ensemble des neurones  $n_j$  où  $i \neq j$ . Les connexions entre neurones sont pondérées par des poids  $w_{i,j}$ . Ce sont ces poids qui correspondent aux paramètres apprenables du réseau. Lorsqu'une entrée particulière est fournie au réseau, ce dernier met à jour la matrice des poids en fonction de la coactivation des neurones selon une règle de mise à jour de la forme suivante :  $w_{i,j}^{t+1} = w_{i,j}^t + f(x_i, x_j)$ .

La fonction  $f$  peut prendre des formes différentes, dans le plus simple des cas, elle est égale au produit  $x_i x_j$ , mais elle peut aussi prendre des formes plus complexes, permettant en particulier de diminuer la valeur d'un poids lorsque les neurones  $n_i$  et  $n_j$  ne sont pas activés simultanément.

Le stage consiste à reprendre l'implémentation (simple) de l'algorithme Parser et de remplacer le système de mise à jour des poids des segments par un réseau hebbien. Il s'agira de voir comment un tel réseau hebbien peut coder différentes formes de régularités (lexicales et sous-lexicales) à mesure que le réseau est exposé à un corpus de mots de grande taille.

## Bibliographie

- Perruchet98 Perruchet, P., & Vinter, A. PARSER : A model for word segmentation. *Journal of memory and language*, 39(2), 246-263, 1998.
- Tovar18 Tovar, Á. E., Westermann, G., & Torres, A. From altered synaptic plasticity to atypical learning : A computational model of Down syndrome. *Cognition*, 171, 15-24, 2018.
- Kuriscak15 Kuriscak, E., Marsalek, P., Stroffek, J., & Toth, P. G. Biological context of Hebb learning in artificial neural networks, a review. *Neurocomputing*, 152, 27-35, 2015.
- Sejnowski99 Sejnowski, T. J. The book of Hebb. *Neuron*, 24(4), 773-776, 1999.