



**HAL**  
open science

# Émergence de catégories par interaction entre systèmes d'apprentissage

Maxime Carrere, Frédéric Alexandre

► **To cite this version:**

Maxime Carrere, Frédéric Alexandre. Émergence de catégories par interaction entre systèmes d'apprentissage. Conférence Francophone sur l'Apprentissage Automatique (CAP), Jul 2013, Lille, France. hal-00826554

**HAL Id: hal-00826554**

**<https://inria.hal.science/hal-00826554v1>**

Submitted on 27 May 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Émergence de catégories par interaction entre systèmes d'apprentissage

Maxime Carrere<sup>1</sup> et Frédéric Alexandre<sup>1</sup>

<sup>1</sup>INRIA, Centre de recherche de Bordeaux

<sup>1</sup>Labri, UMR 5800, CNRS, Université de Bordeaux

<sup>1</sup>Institut des Maladies Neurodégénératives

27 mai 2013

## Résumé

Le modèle AGREL est un modèle connexionniste de catégorisation, reposant sur des observations comportementales et physiologiques qui indiquent que la création de catégories perceptives se fonde sur des phénomènes attentionnels et sur des critères d'erreur de prédiction de récompense. Ce modèle d'apprentissage par renforcement laisse ouverte la discussion sur l'origine du signal de récompense. Il impose également la supervision de la couche de sortie par la tâche de sélection de l'action. Le modèle que nous proposons étend ces travaux dans deux directions. D'une part, nous reposant sur des données biologiques, nous introduisons un réseau spécifique pour le calcul de l'erreur de prédiction, préfigurant l'action de l'amygdale. D'autre part, ce calcul dans un module séparé permet d'adapter le modèle AGREL spécifiquement pour la création des catégories nécessaires à la tâche, indépendamment de la sélection de l'action. Cette approche, dont la performance est évaluée par des tests classiques de discrimination, illustre la puissance de la structuration modulaire du cerveau, où la spécialisation de structures distinctes à des formes différentes d'apprentissage permet un plus grand pouvoir expressif et une grande flexibilité dans l'analyse et la représentation de l'information.

**Mots-clés** : Connexionnisme, Catégorisation hiérarchique, Modèle bio-inspiré.

## 1 Introduction

La catégorisation est une opération cognitive fondamentale d'organisation des connaissances. Elle résulte

d'un apprentissage et se traduit par un stockage d'information mobilisant la mémoire. Nous rassemblons nos connaissances en catégories pour optimiser nos ressources mnésiques limitées et pour retrouver ces connaissances à un moindre coût cognitif.

S'intéresser à la catégorisation s'inscrit donc à la fois dans le cadre du traitement automatique de l'information et de l'apprentissage automatique, mais aussi des neurosciences cognitives. Des travaux de modélisation neuronale bio-inspirée, tels que présentés ici, permettent de rassembler ces points de vue. En particulier, les neurosciences cognitives permettent de spécifier certains aspects de la catégorisation humaine.

Il est tout d'abord notable que la catégorisation s'observe dans toutes les cultures et qu'elle prend une forme hiérarchique, sous forme de taxinomie [RSW12] (par exemple, nous séparons les êtres vivants des objets inanimés, puis la faune de la flore et nous poursuivons en distinguant successivement les mammifères, les chiens, les différentes espèces, etc.). Cette hiérarchie se retrouve également, dans le cortex visuel par exemple, par une succession d'aires visuelles extrayant des indices de plus en plus haut niveau [TFT01], de neurones répondant à des orientations élémentaires dans l'aire visuelle primaire à des neurones faisant la discrimination animal/non animal dans les aires inférotemporales.

Par ailleurs, l'observation du déroulement et des critères d'apprentissage de ces catégories hiérarchiques nous donne des informations sur la stratégie utilisée pour optimiser nos ressources cognitives. Il apparaît tout d'abord que nous n'apprenons de nouvelles catégories que si nous y prêtons attention [AH93]. Ensuite, cet apprentissage mobilise différents types de mémoires, en commençant par la référence à des cas

particuliers avant de se transformer en l'extraction des indices pertinents à la différenciation de catégories [AS04].

L'accumulation de ces données issues du vivant a donné lieu à plusieurs travaux de modélisation, dont le modèle AGREL [RvO05], qui propose d'intégrer certaines de ces observations dans un réseau de neurones hiérarchique ayant une certaine plausibilité biologique. Nous décrivons ici les principales caractéristiques de ce réseau, avant d'en proposer des extensions qui améliorent cette plausibilité en relâchant des contraintes sur la nature de la supervision, tout en gardant d'excellentes capacités d'apprentissage et de traitement de l'information, comme le montrent les résultats expérimentaux rapportés ensuite, avant de conclure sur d'autres extensions envisagées.

## 2 Agrel

### 2.1 Présentation du modèle

AGREL (Attention-Gated REinforcement Learning) [RvO05] est un modèle d'apprentissage par réseau de neurones où la modification des poids est le résultat de la combinaison de deux signaux différents. Un signal de renforcement homogène dans l'ensemble du réseau indique la quantité de récompense obtenue après un essai, et un signal attentionnel limite l'apprentissage aux neurones impliqués dans la réponse au stimulus. Ce modèle requiert qu'il y ait dans la couche de sortie une unique réponse pour chaque stimulus et l'apprentissage opère en fonction de la validité de cette réponse. Les stimulus sont ainsi triés, rangés en catégories, suivant les réponses attendues. De plus la récompense indiquant si la réponse du réseau était la bonne est délivrée immédiatement après la réponse.

On présente au réseau  $P$  stimuli, qui peuvent être triés en  $C$  classes s'excluant mutuellement. Un stimulus est présenté à la couche d'entrée du réseau pour chaque essai. L'activité des neurones de la couche d'entrée est ensuite propagée aux éventuels niveaux intermédiaires cachés, puis aux  $C$  neurones de la couche de sortie. Pour chaque stimulus, un algorithme de type winner-take-all détermine le vainqueur de la carte de sortie, dont l'activité est mise à 1, les autres neurones de la carte de sortie voyant leur activité descendre à 0. Si la réponse du réseau est la bonne, il reçoit une récompense. Dans le cas contraire il n'en reçoit pas et n'est ainsi pas informé de la catégorie qu'il aurait dû choisir.

L'activité  $X_{1..N}$  des  $N$  neurones de la couche d'entrée est propagée au niveau caché composé de  $M$

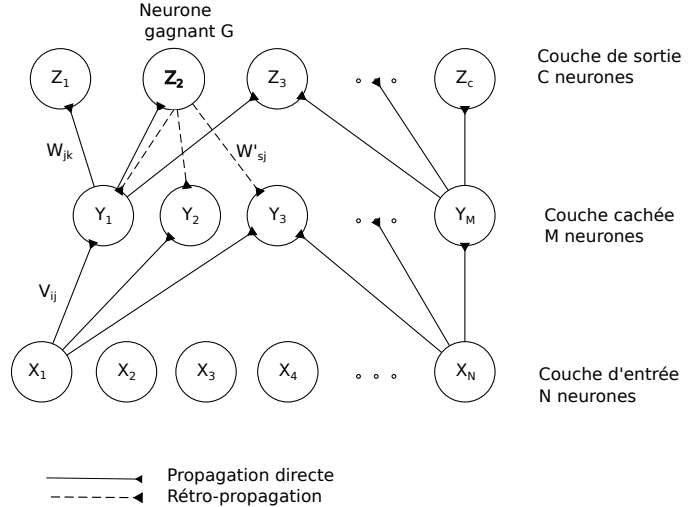


FIGURE 1 – Réseau à trois couches tel qu'utilisé dans Agrel pour la classification. La couche d'entrée propage l'activité aux neurones de la couche cachée via les connexions  $v_{ij}$ , qui eux-mêmes propagent leurs activités aux neurones de la couche de sortie par l'intermédiaire des connexions  $w_{jk}$ . Le neurone gagnant sur la couche de sortie,  $s$ , rétro-propage alors son activité aux neurones de la couche cachée via les poids  $w'_{sj}$ .

neurones d'activités  $Y_{1..M}$  par les connexions  $v_{ij}$  grâce à la fonction non linéaire suivante :

$$Y_j = \frac{1}{1 + \exp(-h_j)}, \text{ avec } h_j = \sum_{i=0}^N v_{ij} X_i \quad (1)$$

Chaque unité neuronale  $j$  possède ainsi un poids de biais  $v_{0j}$ . En cas de multiples couches cachées, cette opération est répétée; sans couche cachée, le réseau projette directement l'entrée sur la sortie. Les connexions  $w_{jk}$  propagent à leur tour l'activité du niveau caché au niveau de sortie d'activités  $Z_{1..C}$ . Le neurone gagnant, dont l'activité sera mise à 1, est alors tiré au sort selon la probabilité suivante :

$$Pr(Z_k = 1) = \frac{\exp(a_k)}{\sum_{k'=0}^C \exp(a_{k'})}, \text{ avec } a_k = \sum_{j=0}^M w_{jk} Y_j \quad (2)$$

Les neurones de sortie ont également un poids pour le biais  $w_{0k}$ . Si le réseau choisit la bonne sortie, la récompense est fixée à 1 dans AGREL. Les poids sont ensuite mis à jour suivant une loi Hebienne :

$$\Delta w_{jk} = \beta Y_j Z_k f(\delta), \text{ avec } \delta = \begin{cases} r - E(r) & \text{si } r = 1 \\ -1 & \text{sinon.} \end{cases} \quad (3)$$

Avec  $E(r)$  la quantité moyenne de récompense prédite, soit  $E(r) = Pr(Z_G = 1)$ , avec G le neurone gagnant de la couche de sortie. En raison de l’algorithme winner-take-all effectué sur la sortie, seuls les poids du neurone gagnant sont mis à jour, l’activité des autres neurones de sortie valant 0 après la compétition.

Les récompenses non prédites par le réseau sont particulièrement importantes pour l’apprentissage. Afin d’augmenter leur impact,  $\delta$  influence la plasticité au travers d’une fonction  $f(\delta)$ , qui va prendre des valeurs plus importantes quand  $\delta$  est proche de 1, c’est-à-dire quand une récompense survient sans avoir été prédite par le réseau.

$$f(\delta) = \begin{cases} \frac{\delta}{1-\delta}; & \text{si } \delta \geq 0 \\ \delta; & \text{si } \delta = -1 \end{cases}$$

Dans l’implémentation,  $f(\delta)$  est majorée par  $\frac{50}{\beta}$ . Les poids entre le niveau d’entrée et le niveau caché sont également modifiés selon une loi Hebbienne qui dépend de  $f(\delta)$ .

$$\Delta v_{jk} = \beta X_i Y_j f(\delta) f b_{Y_j} \quad (4)$$

avec  $f b_{Y_j}$  un terme rétro-propagé par le niveau de sortie, qui influence la plasticité :

$$f b_{Y_j} = (1 - Y_j) \sum_{k=1}^C Z_k w'_{kj} \quad (5)$$

où  $w'_{kj}$  correspond au poids du signal rétro-propagé par la couche de sortie. Ils seront ici pris égaux aux poids dans le sens direct  $w_{kj}$ , ce choix est par ailleurs argumenté dans AGREL [RvO05].

## 2.2 Analyse du modèle AGREL

Il est tout d’abord important de souligner que, dès son origine, le modèle AGREL a été proposé comme une version bio-inspirée du perceptron multi-couches et de sa règle d’apprentissage de rétropropagation du gradient de l’erreur [RM86], en l’adaptant de plus à l’apprentissage par renforcement. Le choix du cadre de l’apprentissage par renforcement est également motivé par l’inspiration de la biologie : l’apprentissage supervisé nécessite un signal d’erreur très précis dans sa mesure et dans son décours temporel et il se rencontre assez rarement au niveau biologique, alors que des signaux d’échec/succès ou plaisir/douleur sont plus fréquents dans un cadre comportemental et nettement plus simples à élaborer et diffuser largement dans le système nerveux, en utilisant en particulier le neurotransmetteur de la dopamine.

L’autre terme de la règle d’apprentissage correspond à un feed-back attentionnel, puisque les modifications de poids sont réalisées là où il y avait de l’activité dans le réseau. Sous cet angle, AGREL respecte également le résultat expérimental évoqué plus haut selon lequel l’apprentissage de catégories ne se produit que là où se porte l’attention [AH93].

Le choix de la structure multi-couches est dicté par la nature hiérarchique des catégories extraites mentionnée plus haut [TFT01]. Mais ce choix produit également une confrontation avec le problème du *credit-assignment* que l’on rencontre avec les structures multi-couches. L’algorithme de rétropropagation du gradient y a apporté une solution très efficace [RM86], mais au prix de mécanismes dont l’implausibilité biologique a été soulignée rapidement ; en particulier, la rétropropagation du terme d’erreur de la sortie vers l’entrée n’a pas trouvé de substrat biologique très convaincant. Une proposition essentielle du modèle AGREL est que, en passant d’un apprentissage supervisé à un apprentissage par renforcement, le terme d’erreur devient relatif à un signal de récompense et que ce dernier peut être proposé de manière diffuse à l’ensemble des couches du réseau, sans nécessiter sa rétropropagation de la couche de sortie vers l’entrée.

Cependant, Agrel propose que le codage de la récompense  $f(\delta)$  soit effectué dans le cerveau par un neuromodulateur tel que la dopamine. Or si des observations ont été faites [MB99] montrant que les cartes supérieures du cortex reçoivent effectivement de la dopamine (le cortex périrhinal, pouvant correspondre à la couche de sortie des modèles), les cartes inférieures n’en recevraient pas. Il était donc important de pouvoir attribuer à un autre mécanisme biologique la diffusion, à l’ensemble du réseau, d’un signal de récompense.

Nous expliquerons plus loin que nous proposons qu’il soit véhiculé par les connexions (bien avérées [LeD07]) de l’amygdale vers l’ensemble du cortex sensoriel. L’amygdale est une structure cérébrale ancienne, connue pour détecter des événements sensoriels biologiquement significatifs (des punitions ou des récompenses) et pour apprendre que certains autres traits sensoriels sont prédictifs de ces événements [LeD07] (apprentissage pavlovien). Ces détections conduisent ensuite à certains ajustements réflexes (réponse pavlovienne), mais aussi à la diffusion de messages plus élaborés à l’ensemble du système nerveux, soulignant en particulier l’importance de ces événements pour orienter certaines structures cérébrales (dont en particulier le cortex et l’hippocampe) vers un mode attentionnel approprié (porter attention vers un événement important mais encore

mal connu) ainsi que vers une mémorisation (apprendre les configurations de ces situations importantes). Parmi les messages diffusés par l'amygdale, les différences entre les récompenses qu'elle prédit et celles qu'elle reçoit ultérieurement sont particulièrement utiles à ces structures cérébrales.

Notre choix d'introduire l'amygdale dans le circuit modélisé est ainsi d'autant plus pertinent que l'amygdale est supposée calculer des erreurs de prédiction de récompense, qui sont des informations essentielles dans le cadre de l'apprentissage par renforcement. Ceci débouchera donc sur la conception modulaire du réseau que nous proposons, différenciant clairement la gestion de la récompense de celle de l'organisation hiérarchique des catégories, ce qui n'est pas fait dans AGREL. Un autre intérêt de faire gérer la récompense par un module séparé est que l'on va pouvoir considérer différents types de récompenses, pouvant être négatifs ou positifs (répulsifs ou attractifs) et de natures différentes (ex : boisson et nourriture), alors que AGREL considérerait uniquement un scalaire représentant la récompense comme une notion abstraite. Il est clair que, pour des apprentissages aussi bien répondant que opérant, plusieurs types de récompense doivent pouvoir être considérés (on doit pouvoir prédire en parallèle l'arrivée d'un choc électrique et de boisson).

Enfin, dans AGREL, la récompense est obtenue si la réponse du réseau correspond à celle attendue, i.e. si le neurone de sortie choisi par le winner-take-all est bien celui qui est attendu. AGREL motive ce choix en indiquant que la couche de sortie est en fait un ensemble d'actions, que l'on devrait sélectionner en fonction des stimuli perçus pour obtenir une récompense, ce qui conduirait à construire une hiérarchie de catégories pertinentes pour une sélection de l'action récompensée dans les couches intermédiaires, liant ainsi étroitement la catégorisation à l'apprentissage instrumental. Il semble cependant clair que la catégorisation est utilisée de manière plus large, également pour l'apprentissage répondant, lorsque l'on cherche à extraire les stimuli conditionnels prédisant des récompenses (appelées ici stimuli inconditionnels) [SD92]. Il peut donc être préjudiciable de faire trop étroitement dépendre la catégorisation de la sélection de l'action. Cet argumentaire milite là encore pour la création de modules différenciés, avec l'extraction de catégories assurée par le cortex postérieur sensoriel alors que l'organisation de l'action s'opère plutôt dans le lobe frontal du cortex. Ceci permet également de s'assurer une plus grande flexibilité du contrôle cognitif : lorsqu'on a extrait des catégories, on peut s'en servir comme critère de choix pour différentes tâches motrices et avoir lié la création

de ces catégories à un ensemble particulier d'actions en couche de sortie pourrait rendre plus difficile leur réutilisation pour d'autres tâches et d'autres actions. Dans le modèle que nous proposons, nous retirons la contrainte d'imposer une sélection de l'action en couche de sortie, en enlevant la contrainte d'imposer une supervision de l'unité de sortie gagnante comme critère d'apprentissage, comme nous le décrivons dans la section suivante.

### 3 Modèle avec interaction

Concernant son architecture, illustrée dans la Figure 3, le modèle que nous détaillons est donc essentiellement celui présenté dans AGREL (représentant le cortex), auquel vient se greffer un second réseau mimant le rôle de l'amygdale. Appuyé par des données anatomiques et physiologiques, le réseau correspondant à l'amygdale reçoit l'ensemble des informations sensorielles du cortex pour en extraire des stimuli prédictifs de récompense (CS : stimuli conditionnels) afin de prédire des récompenses qui peuvent être de natures différentes (US : stimuli inconditionnels), reçues directement par l'amygdale. En retour, l'amygdale peut envoyer ses erreurs de prédiction de récompense à l'ensemble des couches du cortex.

Les tâches effectuées par le réseau global sont des tâches de catégorisation : étant donnée une entrée, associée ou non à une récompense, le réseau doit apprendre à prédire quelle sera la récompense associée à cette entrée, et notamment apprendre à créer des catégories si celles-ci sont nécessaires à la prédiction. Le cortex débute l'apprentissage avec des unités de la couche d'entrée représentant des traits sensoriels (par exemple des couleurs et des formes) et un nombre limité de neurones dans les couches suivantes, ne permettant pas la représentation de toutes les combinaisons des traits sensoriels élémentaires. Les poids initiaux du réseau cortical étant aléatoires, l'ensemble des combinaisons initiales est relativement indifférencié et devra s'affiner au cours de l'apprentissage pour représenter des stimuli correspondant à des combinaisons de traits indispensables à la prédiction de récompenses. Un exemple classique de telles représentations conjonctives est celui du ou-exclusif (XOR), où la réponse en présence des traits présentés ensemble est différente de celle pour chacun des traits présenté séparément. Cet exemple est également connu comme non séparable linéairement, c'est à dire nécessitant la création d'une représentation intermédiaire.

Nous décrivons maintenant plus précisément le fonc-

tionnement du réseau représentant l'amygdale avant de présenter son interaction avec celui du cortex, permettant la création de ces catégories conjonctives dans les couches intermédiaires, afin de fournir en retour à l'amygdale une représentation de l'information qui lui permettra au mieux de prédire les récompenses.

### 3.1 Amygdale

Le rôle de l'amygdale est d'apprendre à prédire les associations entre stimulus (CS) et récompense (US). La structure et la règle d'apprentissage de notre modèle d'amygdale sont basés sur le modèle de Rescorla-Wagner [RW72]. Il s'agit d'un modèle à deux couches, proche d'un perceptron simple (figure 2). La couche d'entrée reçoit (de l'ensemble des couches du cortex) les différents traits sensoriels (simples et conjonctifs) et la couche de sortie représente les différents US. Dans un premier temps, cette couche sera évaluée à partir de l'activation de la couche d'entrée et représentera la prédiction de l'occurrence de ces US. Dans un deuxième temps, l'arrivée effective des US permettra de calculer l'erreur de prédiction pour chaque US.

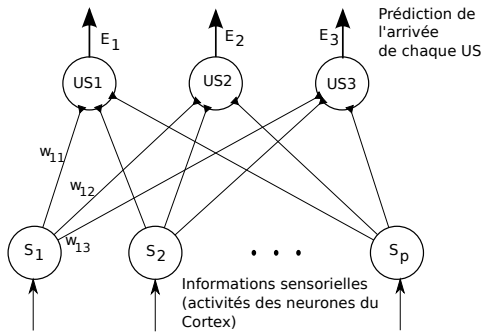


FIGURE 2 – Réseau modélisant l'amygdale. Pour chaque US, un neurone prédit son apparition en fonction des différents traits sensoriels en entrée. L'arrivée ultérieure éventuelle de l'US  $j$  permettra de calculer l'erreur de prédiction servant à mettre à jour  $w_{ij}$ .

Plus précisément, pour chaque US  $j$  à prédire, il existe un neurone dont l'activité  $E_j$  donne la prédiction d'apparition de l'US, c'est à dire la probabilité qu'un stimulus sensoriel soit suivi de cet US. Pour chacune des  $p$  entrées  $S_i$  de l'amygdale, et pour chaque US  $j$ , il existe un poids  $w_{ij}$  tel que :

$$E_j = \sum_i^p S_i w_{ij}$$

L'apprentissage des poids  $w_{ij}$  se fait proportionnellement à la différence entre la récompense reçue  $r_j$  et

celle effectivement prédite :

$$\Delta w_{ij} = \alpha S_i [r_j - E_j]$$

où  $\alpha$  est le taux d'apprentissage. Ce modèle d'apprentissage des associations CS-US, très simple, permet de renforcer les entrées fréquemment suivies d'une récompense et de diminuer les autres. En particulier, il ne peut pas s'adapter si certains entrées annoncent une récompense dans certains contextes mais pas dans d'autres ; il faudra dans ce cas là trouver des conjonctions de traits plus discriminantes.

### 3.2 Interactions avec AGREL

Les deux modèles, celui de l'amygdale (Rescorla-Wagner) et celui du cortex (AGREL), ont été mis en relation de la façon suivante (voir figure 3) :

- l'entrée de l'amygdale correspond à l'ensemble des couches sensorielles du cortex. Pour chacun de ces traits (simples ou conjonctifs) sensoriels, il existe donc une unité dans l'amygdale projetant par  $j$  connexions vers la couche de sortie de l'amygdale, avec  $j$  le nombre d'US représentés dans l'amygdale. Chaque connexion retient ainsi l'association entre l'activité de ce neurone et l'US correspondant.
- La sortie de l'amygdale, c'est-à-dire la prédiction des US en fonction de l'entrée, est utilisée dans AGREL pour calculer le signal global de renforcement  $\delta$ .

Un certain nombre de données bibliographiques (qu'il n'a pas lieu de détailler ici) peuvent venir apporter des arguments sur le rôle de l'amygdale dans l'apprentissage répondant pour la prédiction des US [CPHE02] et pour différents types de récompenses [SGZ03], sur l'entrée massive d'afférences du cortex sensoriel [PA10], sur sa taille importante permettant de supporter une telle combinatoire [SGZ03], sur le rôle modulateur de l'amygdale sur le processus de mémorisation d'autres structures dont le cortex [PC01].

Dans notre modèle, l'amygdale ne sort pas une unique prédiction  $E$  mais une prédiction par US  $E_j$ . Or AGREL reçoit un signal de récompense global sans faire de différence suivant quel US est responsable du signal. Dans notre modèle, nous obtenons  $\delta$  en prenant le maximum de l'erreur de prédiction faite pour l'amygdale pour chaque US :

$$\delta = r_m - E_m$$

$$\text{avec } m \text{ tel que } |r_m - E_m| = \max_j (|r_j - E_j|)$$

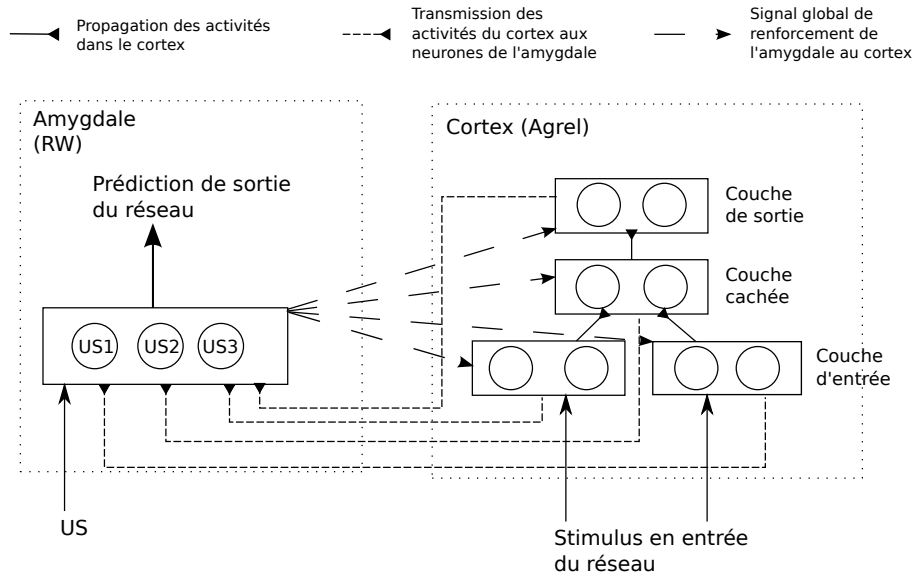


FIGURE 3 – Le modèle est composé de deux parties, l’amygdale et le cortex. Les stimuli en entrée sont présentés au cortex, qui propage les activités dans ses différentes couches. Une flèche d’un ensemble de neurones vers un autre indique sur ce schéma que tous les neurones du premier ensemble envoient leurs activités à tous les neurones du deuxième ensemble. Les activités du cortex sont envoyées à l’amygdale, qui s’en sert pour effectuer la règle de Rescorla-Wagner (RW), et ainsi prédire les probabilités d’apparition pour chacun des US. L’erreur de prédiction est ensuite envoyée de l’amygdale au cortex, sous forme de signal global, pour guider son apprentissage.

Intuitivement, cela revient à dire que nous améliorons nos prédictions les unes après les autres, en se focalisant à chaque fois sur la plus mauvaise. L’utilisation de la valeur absolue permet d’éviter qu’une légère erreur de prédiction positive sur un CS ne masque une importante erreur négative (US prévu alors qu’il n’y en avait pas).

Un autre changement effectué est que, dans AGREL, le réseau apprend uniquement quand la réponse donnée est la bonne, i.e. si le bon neurone est sélectionné en sortie. Le cas échéant,  $f(\delta)$  est fixé à -1. Dans notre modèle, seule l’erreur la plus importante est prise en compte, que  $\delta$  soit ou non positif. Notamment, quand  $\delta$  est négatif, sa valeur est quand même une indication sur l’importance de l’erreur produite par l’amygdale. Afin de conserver cette information, la fonction  $f(\delta)$  prend les valeurs suivantes :

$$f(\delta) = \begin{cases} \frac{\delta}{1-\delta}; & \text{si } \delta \geq 0 \\ \frac{\delta}{1+\delta}; & \text{sinon} \end{cases}$$

Selon la même logique que lorsque  $\delta$  est positif, si  $\delta \leq 0$ ,  $f(\delta)$  va augmenter l’impact des erreurs de prédiction négatives proches de -1, c’est-à-dire quand une récompense prédite ne survient pas. De manière similaire à AGREL,  $f(\delta)$  est minorée par  $\frac{-50}{\beta}$  quand  $\delta$  tend vers -1.

Le winner-take-all probabiliste à la sortie d’AGREL peut également être remplacé par un maximum, comme la sortie d’AGREL ne correspond plus dans notre modèle à la probabilité que le stimulus soit récompensé. Ainsi le gagnant est le neurone ayant l’activité la plus forte, au lieu d’en tirer un, suivant une probabilité proportionnelle à leur activité. La méthode non probabiliste permet une convergence plus rapide, mais enlève la possibilité d’exploration au réseau. Si une catégorie est apprise et convient *convenablement*, le réseau ne va pas explorer de temps à autre une autre décision, pour voir si celle-ci est meilleure. Dans les illustrations présentées par la suite, nous avons gardé la méthode probabiliste. Ce modèle permet ainsi de ne plus faire d’apprentissage supervisé pour la création de catégories comme dans AGREL.

## 4 Illustrations

Nous avons réalisé différents tests classiques de catégorisation, le but étant à la fois de vérifier la capacité de l’amygdale à correctement prédire les récompenses, mais aussi s’assurer de la cohérence des représentations créées dans le cortex. Notre but dans un premier temps est de prouver le bon fonctionne-

ment du modèle, et en particulier de montrer qu’il respecte bien certaines propriétés comme la création de catégories minimales. Nous cherchons à reproduire des expériences similaires faites avec des humains et des animaux en prise de décision, et qui permettront, par des études lésionnelles, de préciser le rôle de l’amygdale car les sujets lésés font des erreurs caractéristiques.

#### 4.1 Problème linéairement séparable

Ce premier problème est un problème linéairement séparable, autrement dit qui ne nécessite en théorie pas l’usage du cortex. Notre réseau est composé d’une carte de 2 neurones en entrée, d’une carte intermédiaire de 2 neurones et d’une carte de sortie, également composée de 2 neurones. Trois stimuli étaient présentés à l’entrée du réseau. Le stimulus 11 était récompensé, mais pas les 01 et 10. Ces stimuli étaient présentés de manière alternée au réseau.

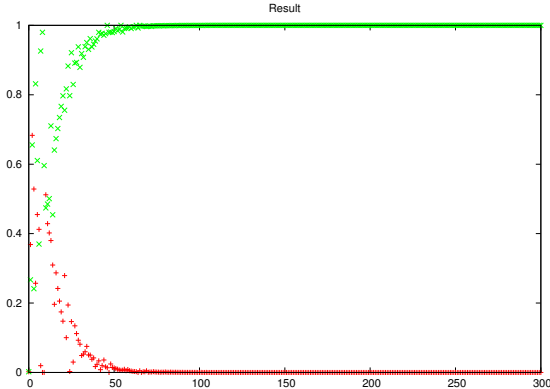


FIGURE 4 – Évolution de la sortie de l’amygdale au cours de l’apprentissage. Cette figure affiche les probabilités que la récompense survienne après présentation des stimuli 10 (en rouge ; pas de récompense) et 11 (en vert ; récompense). On peut remarquer un départ de convergence à la 10<sup>ème</sup> itération, qui n’aboutit pas. La convergence survient cependant très rapidement.

La première observation que l’on peut faire est sur le temps de convergence de la figure 4. Celui-ci est très bref, et correspond au temps de convergence d’un perceptron classique. De plus, le changement de pente dans l’évolution des poids du cortex (figure 5) montre également que celui-ci n’avait pas encore convergé lorsque la convergence de l’amygdale a provoqué l’arrêt de l’apprentissage dans le cortex. Les probabilités de sortie du cortex sont de plus restées indépendantes du stimulus d’entrée (i.e. quel que soit le stimulus en entrée, le neurone 1 de la couche de sortie répond avec

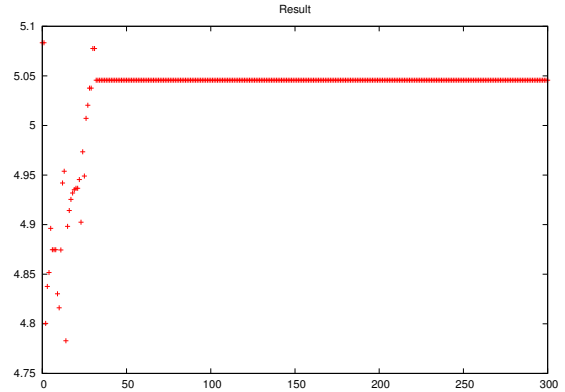


FIGURE 5 – Évolution des poids de la sortie du cortex au cours de l’apprentissage. Affiche l’évolution du poids d’un neurone de la couche de sortie au cours de l’apprentissage. Le changement de pente brutal, au moment de la convergence de l’amygdale, montre que la convergence du cortex était encore en cours lorsque l’amygdale a convergé.

une probabilité de 0.60 et le deuxième 0.4). Ainsi, notre système n’a pas créé de catégorie lorsque cette création n’était pas requise.

#### 4.2 XOR

Un problème classique de catégorisation non linéaire, nécessitant donc dans notre modèle un apport du cortex, est le problème du “ou exclusif”, ou XOR. Dans cette partie, nous obligeons de plus le réseau, par son architecture, à se construire les représentations nécessaires à la résolution du XOR sur sa couche de sortie. Notre réseau est composé de deux cartes d’entrée. Chacune de ces cartes est reliée à une carte correspondante dans la couche cachée. Puis une couche de sortie rassemble les informations en provenance des deux cartes de la couche cachée. Toutes les cartes sont composées de 2 neurones. Une spécificité de ce réseau est que seule la carte de sortie peut recevoir et faire la synthèse de toutes les informations passées en entrée. Dans le cas du XOR, seule cette carte peut donc apprendre les catégories nécessaires à sa résolution. Les stimuli 10-10 (première carte-deuxième carte) et 01-01 sont récompensés. Les stimuli 10-01 et 01-10 ne sont pas suivis de récompense. En prenant par exemple 10 pour coder 1 et 01 pour coder 0, on retrouve bien le problème classique du XOR pour différencier deux catégories.

Comme le montrent les figures 6 et 7, la convergence prend de nombreuses itérations, au contraire du problème linéairement séparable présenté ci-dessus.



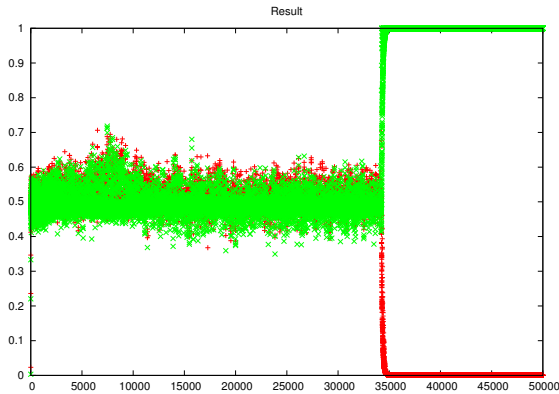


FIGURE 6 – Évolution de la sortie de l’amygdale au cours de l’apprentissage. La figure affiche les probabilités que la récompense survienne après présentation des stimulus 10-01 (en rouge; pas de récompense) et 10-10 (en vert; récompense). La convergence met de nombreuses itérations avant d’arriver. Cependant, une fois amorcée, le modèle converge très vite.

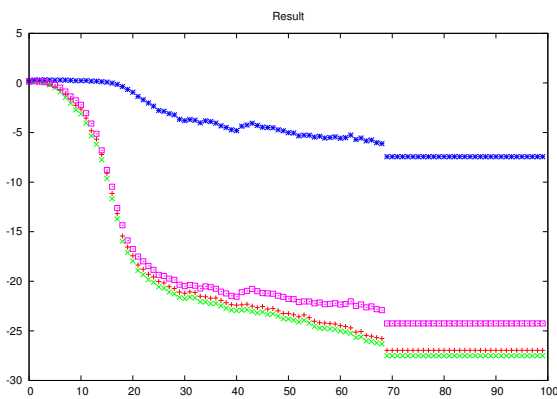


FIGURE 7 – Évolution des poids de la sortie du cortex au cours de l’apprentissage. Affiche l’évolution des poids d’un neurone de la couche de sortie au cours de l’apprentissage. Dans un souci de lisibilité, un point correspond à la valeur d’un poids toutes les 5000 itérations. Au contraire de l’amygdale, la convergence se fait de manière continue, et s’arrête lorsque l’amygdale a convergé.

La solution utilisée par le système pour résoudre le problème est la suivante : différencier un des 4 cas dans le cortex en utilisant la couche supérieure. En tenant compte de cette catégorie nouvellement créée, il ne reste donc que 3 entrées à séparer dans le problème du XOR, ce qui devient donc linéairement séparable, comme vu à l’exemple précédent, et l’amygdale parvient donc bien à donner les probabilités correctes, comme le montre la figure 6.

### 4.3 XOR (variante)

Dans cette partie, nous testons une variante du XOR plus complexe, afin de tester les performances du réseau lorsque tous les stimuli sont récompensés, la différence se faisant sur la nature de la récompense reçue. Le réseau est composé de deux cartes de 2 neurones en entrée, d’une carte intermédiaire de 2 neurones et d’une carte de sortie, également composée de 2 neurones. Deux récompenses différentes, associées aux US1 et US2, sont présentées au réseau. L’US1 survient après la présentation des stimuli 1010 et 0101. L’US2 survient quant à lui suite à la présentation des stimuli 1001 et 0110. Comme pour l’illustration précédente, il s’agit bien d’un problème de type XOR. Cependant, la structure du réseau n’impose pas de condition sur la couche devant créer les catégories.

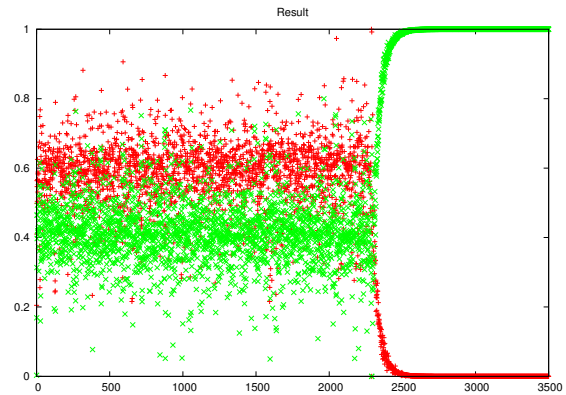


FIGURE 8 – Évolution de la sortie de l’amygdale au cours de l’apprentissage. Affiche les probabilités que l’US1 survienne après présentation des stimuli 1001 et 0110. Tant que le cortex n’a pas formé de catégorie utile au XOR, l’amygdale ne peut prévoir correctement les réponses à donner. Une fois que le cortex a convergé et créé la catégorie nécessaire, l’amygdale converge rapidement.

Comme le montre la figure 8, l’amygdale n’arrive pas à converger tant que le cortex n’a pas créé les catégories nécessaires. Lorsque celui-ci les a créées, l’amygdale converge rapidement et prédit correctement l’arrivée des US. La reconnaissance du XOR se fait grâce à un unique neurone de la couche cachée, qui a appris à répondre spécifiquement au “et” (0101). A partir de ce neurone, l’amygdale a appris à effectuer un “ou” normal sur les entrées, auquel vient se soustraire le résultat du neurone “et”, lui permettant ainsi de répondre correctement au XOR. Cependant, en raison de la complexité du problème, le temps de convergence du réseau total reste important. De plus, ce même temps de convergence, sur le même problème,

peut varier énormément, entre 100 et 50 000 itérations, pour une moyenne de 5500 itérations. Une piste pour corriger l'importance de cet aléa pourrait être de changer l'algorithme de winner-take-all, de façon à ce qu'il tienne moins compte du facteur chance dans ses choix, mais cela pourrait exposer le réseau au risque de ne pas trouver de réponse, car ne pratiquant pas suffisamment d'exploration.

## 5 Discussion

Dans cet article, nous avons proposé plusieurs extensions au modèle de catégorisation AGREL [RvO05], visant à augmenter sa plausibilité biologique pour le calcul du signal d'erreur et pour la représentation des catégories. Le modèle qui en résulte extrait des catégories par interaction entre deux systèmes complémentaires, préfigurant deux structures cérébrales, le cortex et l'amygdale, respectivement impliquées dans la représentation des catégories et le calcul des erreurs de prédiction. Nos résultats expérimentaux montrent que ce système permet bien une catégorisation des entrées, même dans des tâches complexes où uniquement l'un des deux systèmes ne saurait suffire. De plus, en cohérence avec les données observées dans la biologie, les catégories qu'il apprend à former sont à la fois celles importantes pour l'organisme, i.e. suivies d'une récompense, et dont la construction s'avère nécessaire, i.e. quand la différenciation est nécessaire à la prédiction correcte. Ainsi, il évite les problèmes de dimensionnalité liés à la catégorisation, en ne stockant pas toutes les combinaisons d'entrées possibles ou rencontrées, mais en ne se créant uniquement que celles qui sont pertinentes à son apprentissage.

Par rapport au modèle AGREL seul, il possède l'avantage de ne pas nécessiter de supervision, et donc n'impose pas dès le début une connaissance des catégories nécessaires à la résolution du problème. La création de ces catégories devient donc non déterministe, en fonction des tirages des gagnants effectués dans les tours précédents. L'aspect non déterministe devient donc plus important dans notre modèle, ce qui amène, dans des problèmes difficiles, de grandes différences de temps de convergence entre deux expériences, comme rencontré dans notre dernière illustration. En affectant le signal global de renforcement  $f(\delta)$  à l'amygdale, notre modèle propose également une explication biologiquement plausible de la provenance d'un tel signal.

Notre modèle, s'il est systémique, n'est pas

symétrique dans ses interactions. Les interactions du cortex vers l'amygdale ont pour but de "permettre" la prévision de récompense en fournissant à l'amygdale les catégories sur lesquelles baser cette prédiction. L'action de l'amygdale sur le cortex, c'est à dire le signal global de renforcement émis par l'amygdale, a elle pour but de "guider" l'apprentissage et la création de catégories, en disant au cortex quand il est pertinent d'apprendre. Étant donnée cette information, le cortex peut alors créer des catégories nécessaires, grâce au signal attentionnel présent dans Agrel.

Un certain nombre d'extensions sont envisagées pour ce modèle. Des extensions à des jeux de données plus importants sont étudiées. Les résultats préliminaires indiquent un bon comportement du modèle sur des corpus d'exemples de plus grande dimensionnalité et plus réalistes (données réelles et bruitées). D'autres extensions visent à améliorer sa proximité avec les caractéristiques du vivant, ce qui se traduira par la complexification du réseau, en y ajoutant d'autres modules. Tout d'abord, nous avons insisté ici sur les apprentissages répondants et opérants, en soulignant que AGREL est trop exclusivement orienté vers l'opérant alors que les signaux de récompense sont également traités par le répondant. La structure et les exemples que nous avons présentés ici sont cependant principalement dédiés au répondant et il serait donc important de pouvoir montrer que cette même base de gestion du renforcement et de construction des catégories peut également s'appliquer à l'opérant. Ceci nécessitera d'ajouter de nouveaux modules à notre système (mais des modèles existants sont disponibles [GPR01]), correspondant à la représentation des actions (cortex préfrontal) et à leur sélection (ganglions de la base).

On peut également noter que le délai entre les CS et l'US est pour le moment stéréotypé, dans la mesure où l'activation du réseau suppose que l'US arrive juste après l'évaluation de la sortie de l'amygdale, permettant ainsi de calculer l'erreur de prédiction. On sait cependant, dans la réalité, que nous sommes capables de repérer des corrélations entre CS et US sur des intervalles de temps plus long. Mais on sait également que l'amygdale n'a pas directement cette capacité d'apprendre des durées variables, mais que c'est plutôt l'inhibition d'autres structures, en particulier le cortex préfrontal [CPHE02], qui s'adapte pour déclencher la réponse de l'amygdale au moment désiré, ce que l'on pourra également considérer si on introduit des modules relatifs au cortex préfrontal.

Enfin, nous avons rapporté que l'apprentissage de nouveaux CS était parfois long, alors qu'ils peuvent aussi être appris rapidement dans la réalité. Nous

avons déjà entamé une autre étude impliquant une autre structure cérébrale à apprentissage rapide, l'hippocampe. Nous comptons à court terme rapprocher ces deux modèles pour étudier les interactions réciproques entre ces deux modes de création de catégories, par la création d'une hiérarchie selon une étude statistique lente et par l'apprentissage par coeur de cas particuliers importants [MMO95].

## Références

- [AH93] M. Ahissar and S. Hochstein. Attentional control of early perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, 90(12) :5718–5722, June 1993.
- [AS04] F. Gregory Ashby and Brian J. Spiering. The neurobiology of category learning. *Behavioral and cognitive neuroscience reviews*, 3(2) :101–113, June 2004.
- [CPHE02] Rudolf N. Cardinal, John A. Parkinson, Jeremy Hall, and Barry J. Everitt. Emotion and motivation : the role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience & Biobehavioral Reviews*, 26(3) :321–352, May 2002.
- [GPR01] K. Gurney, T. J. Prescott, and P. Redgrave. A computational model of action selection in the basal ganglia. i. a new functional anatomy. *Biological Cybernetics*, 84(6) :401–410, May 2001.
- [LeD07] Joseph LeDoux. The amygdala. *Current Biology*, 17(20) :R868–R874, October 2007.
- [MB99] Elisabeth A. Murray and Timothy J. Bussey. Perceptual-mnemonic functions of the perirhinal cortex. *Trends in Cognitive Sciences*, 3(4) :142–151, April 1999.
- [MMO95] J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly. Why there are complementary learning systems in the hippocampus and neocortex : insights from the successes and failures of connectionist models of learning and memory. *Psychological review*, 102(3), July 1995.
- [PA10] Luiz Pessoa and Ralph Adolphs. Emotion processing and the amygdala : from a 'low road' to 'many roads' of evaluating biological significance. *Nature reviews. Neuroscience*, 11(11) :773–783, November 2010.
- [PC01] M. G. Packard and L. Cahill. Affective modulation of multiple memory systems. *Curr Opin Neurobiol*, 11(6) :752–6, 2001.
- [RM86] D.E. Rumelhart and J. L. McClelland. *Parallel Distributed Processing : Exploration in the Microstructure of Cognition*. Cambridge, MA : MIT Press, 1986.
- [RSW12] Diogo Rendeiro, João Sacramento, and Andreas Wichert. Taxonomical Associative Memory. *Cognitive Computation*, pages 1–21, 2012.
- [RvO05] Pieter R. Roelfsema and Arjen van Ooyen. Attention-gated reinforcement learning of internal representations for classification. *Neural computation*, 17(10) :2176–2214, 2005.
- [RW72] R.A. Rescorla and A.R. Wagner. A theory of pavlovian conditioning : Variations in the effectiveness of reinforcement and nonreinforcement. In *Classical Conditioning II : Current Research and Theory*, pages 64–99. Appleton Century Crofts, 1972.
- [SD92] N. Schmajuk and J. DiCarlo. Stimulus configuration, classical conditioning and the hippocampus. *Psychological Review*, 99 :268–305, 1992.
- [SGZ03] D. Sander, J. Grafman, and T. Zalla. The human amygdala : an evolved system for relevance detection. *Reviews in the Neurosciences*, 14(4) :303–316, 2003.
- [TFT01] Simon J. Thorpe and Michèle Fabre-Thorpe. Seeking Categories in the Brain. *Science*, 291(5502) :260–263, January 2001.