

Chapitre 3

Modélisation et Simulation des réseaux neuronaux

Avant-propos Les neurosciences en général constituent un domaine privilégié d'application des méthodes de modélisation et de simulation depuis les niveaux moléculaires jusqu'aux niveaux les plus intégrés (niveaux cognitif et comportemental). Cet exposé se limitera à la modélisation et à la simulation des réseaux nerveux entreprises dans un but d'élucidation à la fois des comportements d'ensemble et des mécanismes élémentaires qui les sous-tendent. Après un bref historique concernant la modélisation des réseaux nerveux, nous décrirons les principes de la modélisation et de la simulation de ces réseaux.

3.1 Bref historique

→ Transparent Fig. Mc CULLOCH et PITTS

La notion de neurone formel apparaît au début des années 40 avec les travaux théoriques de MC CULLOCH et PITTS (voir Annexe). Ces auteurs postulent l'existence d'une logique neuronique, des fonctions logiques élémentaires (négation, somme ou produit logique par exemple) sont réalisées par des réseaux paucineuroniques ; les fonctions plus complexes par un agencement de tels circuits élémentaires. A la même époque sont jetées les bases de l'informatique et VON NEUMANN définit la structure des premiers ordinateurs. Les deux disciplines s'influencent mutuellement et cette époque marque le début d'un compagnonnage qui persiste jusqu'à nos jours et dont nous reparlerons plus loin. Neurobiologistes et psychologues, mathématiciens et physiciens s'intéressent aux machines intelligentes : celles qui sont le fruit de l'évolution des organismes biologiques pour les premiers, celles que l'homme invente et construit pour les seconds.

→ Transparent Perceptrons

Au début des années 60, FRANK ROSENBLATT propose une classe de modèles neuromimétiques, les Perceptrons, capables de réaliser des catégorisations sur la base d'un apprentissage (Cf. annexe pour la description sommaire d'un per-

ceptron). Les tentatives d'exploitation de cet algorithme neuromimétique dans le domaine de l'intelligence artificielle montre des lacunes dans les compétences des perceptrons (incapacité à réaliser les fonctions logiques comme le "ou exclusif"). La démonstration de ces limitations par MINSKY et PAPERT entraîne alors l'abandon par les chercheurs en intelligence artificielle de la solution réseau neuromimétique, ou *CONNEXIONNISME* au bénéfice d'une algorithmique logique classique. Des chercheurs, tels les physiciens J. HOPFIELD et T. KÖHÖNEN, relanceront dans les années 70 l'intérêt pour le connexionnisme en proposant des modèles de réseau intéressants. Ce renouveau est lié, d'une part, à un relatif essoufflement de la démarche classique, et d'autre part à la nécessité de renouveler les architectures d'ordinateurs pour aller vers plus de parallélisme (à l'instar des réseaux de neurones naturels). Les enjeux industriels et économiques considérables liés aux applications potentielles du connexionnisme explique la quantité croissante des travaux effectués dans ce domaine actuellement.

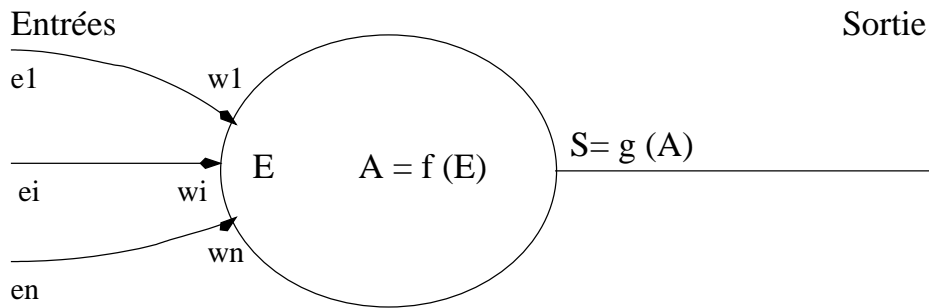
3.2 Modélisation des réseaux neuronaux

Un réseau neuromimétique est composé d'un ensemble fini de cellules ou neurones formels interconnectés et en relation avec le monde extérieur. L'architecture du réseau est définie

- par le nombre et la nature de ces éléments (à noter, comme nous l'avons déjà souligné, que ce nombre lui-même peut être une variable, avec la naissance et la mort de cellules {-se rappeler le jeu de la vie de Conway-} pour tenir compte des lois naturelles ou pour une recherche d'un nombre optimal de cellules apte à remplir une fonction dans le cas des ingénieurs)
- par la topologie et la nature des connexions qui lient ces éléments entre eux et au monde extérieur.

3.2.1 Neuromimes et équation neuroniques

Le neurone formel ou neuromime est caractérisé globalement par la relation qui unit ses entrées (soma neuronique et arborisations dendritiques) et sa sortie (axone). Cette relation peut être conçue comme une succession de fonctions : fonction(s) d'entrées, fonction d'activation et fonction de sortie. La fonction d'entrée E régit la manière de prendre en compte les différents signaux e_i parvenant au neuromime, la plus simple consistant à considérer cette fonction comme la somme pondérée des signaux afférents à tout instant ; les facteurs de pondération w correspondent alors aux efficacités synaptiques (forces ou poids synaptiques)

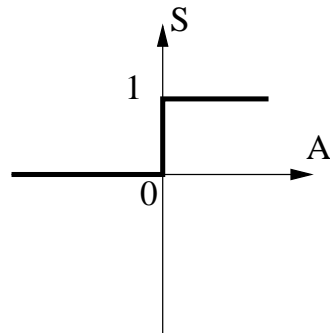


$$E_j(t) = \sum_i e_{ij}(t) \times w_{ij}(t)$$

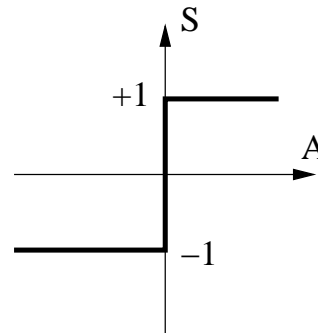
La fonction d'activation A traduit l'activité du neuromime au temps t , qui peut dépendre de E et de l'état propre de l'élément à cet instant (par exemple réfractorité, adaptation etc...). Cette fonction peut être non linéaire pour mimer, par exemple, des effets de saturation. Mais de manière simple on prend souvent :

$$A_j(t) = E_j(t) \text{ ou encore } A_j(t) = E_j(t) - a$$

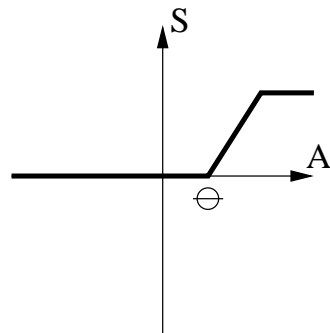
La fonction de sortie S détermine enfin la valeur, à tout instant, du signal de sortie de la cellule considérée, signal qui pourra constituer à son tour une entrée pour d'autres cellules. De nombreuses fonctions peuvent être envisagées, la figure ci-après montre les graphes de quelques unes de ces fonctions.



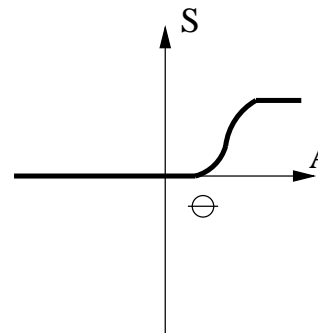
Fonction de HEAVISIDE



Fonction Signe



Fonction Affine à seuil



Fonction Sigmoide

Dans ces exemples de fonctions S , Heaviside et Signe sont des fonctions binaires de A , alors que les fonctions Affine et Sigmoide permettent des sorties graduées. Ces dernières offrent par ailleurs des sorties avec des limitations : niveau de saturation que la sortie ne peut pas dépasser, et niveau seuil Θ qui est la valeur minimale que l'activation A doit dépasser pour que la sortie ne soit pas nulle.

D'autres formalisations du neurone que celles qui sont illustrées ici, sont bien entendu possibles : voir par exemple le modèle de PEARSON et EDELMAN qui fait intervenir des fonctions d'intégration différentes (au niveau dendritique) pour les entrées excitatrices ou inhibitrices, des fonctions 'shunt', 'retard', fonction π etc. (transparent)

Remarque : Comparées aux activités de sortie des neurones biologiques (activités que l'on enregistre au niveau axonal notamment), ces deux dernières fonctions nous proposent des activités continues graduées alors que les activités naturelles sont de type 'tout ou rien', impulsions (potentiels d'action ou spikes). Il faut donc les considérer comme reflétant les variations de fréquence de la décharge neuronale (fréquence des impulsions), qui dépend elle-même des potentiels somatiques qui, eux, sont gradués. On verra par la suite qu'il est possible également de simuler le comportement tout-ou-rien des neurones naturels

(neurones formels dits 'à spike').

3.2.2 Jonctions synaptiques et équations mnémoniques.

L'architecture du réseau est spécifiée en précisant le détail de sa connectivité (emploi du terme de connectique par les ingénieurs pour désigner le câblage). Celle-ci établit d'une part, l'existence de jonctions entre les éléments, ou entre les éléments et la périphérie du réseau (afférences et/ou efférences), et d'autre part la nature de ces jonctions (excitatrices ou inhibitrices) ainsi que leur force.

La "force" d'une jonction représente sa capacité à modifier l'activité de la membrane post-synaptique, on parle également de *poids*, de *coefficient* ou encore d'*efficacité synaptique*. Les liaisons peuvent être envisagées comme des entrées à poids fixe ou variable. Le caractère évolutif des poids des jonctions confère au réseau sa plasticité d'ensemble et permet notamment l'apprentissage, c'est-à-dire l'acquisition par le réseau d'une structure fonctionnelle qui se superpose à sa structure "anatomique". L'existence d'une "plasticité synaptique" se modélise en établissant la règle, ou équation mnémonique, ou encore loi de plasticité, qui permet de calculer la modification des coefficients synaptiques (notés généralement "*w*", pour weight). Cette loi peut prendre en compte dans ses formulations les plus sophistiquées l'existence des composantes élémentaires de la synapse et notamment les différents canaux ioniques ou les réserves en précurseur. Mais elle se résume le plus souvent, dans les modèles actuels, à une relation entre poids synaptique et activité pré- (*e* : signal entrant), et post-synaptique (*E*, *A* ou *S*). Une des règles les plus communément utilisées, applique le principe de HEBB, stipulant que la synapse se modifie en fonction de la conjonction des activités pré et postsynaptiques. D.O. HEBB écrit * :

"Quand un axone de la cellule A est assez proche pour exciter une cellule B et quand, de façon répétée et persistante, il participe à son activation, un certain processus de croissance ou un changement métabolique s'installe dans une cellule ou dans les deux tel que l'efficacité de A, en tant qu'elle est une des cellules qui active B, est augmentée".

D.O. Hebb, "The organization of Behaviour" (New-York, John Wiley & Sons Inc.)(1949)

Cette proposition, (qui constitue sous cette forme une sorte de modèle dialectique), peut bien évidemment adopter des formes mathématiques diverses telle que par exemple :

$$\Delta w_{ij}(t) = \alpha \cdot f[w_{ij}(t) \cdot e_i(t) \cdot s_j(t)]$$

Cette équation définit la variation du coefficient synaptique *w* de la liaison unissant les neurones *i* et *j*, où *e_i* et *s_j* correspondent respectivement aux activités pré et post-synaptique, *α* correspondant à un facteur de sensibilité synaptique (permet de régler le taux de variation du coefficient synaptique). Notez que la fonction *f* n'est pas définie ici, nous en verrons des formes précises dans les exemples étudiés dans les chapitres suivants. Des lois de plasticité du type de celle proposée par HEBB permettent aux réseaux neuromimétiques qui en sont munis de manifester des propriétés d'auto-organisation, c'est-à-dire qu'elles confèrent à ces assemblées neuroniques la possibilité d'acquérir une structure fonctionnelle au cours d'un apprentissage autonome (c'est-à-dire sans "maître", sans intervention extérieure).

Un autre type de règle de plasticité fait intervenir la sortie "désirée" d valeur souhaitée par un opérateur extérieur au réseau, le "maître" en quelque sorte; celui-ci peut être un opérateur humain, un autre réseau... Il s'agit dans ce cas des règles dites de *Réduction d'écart* (Delta rule, ou Règle de WIDROW-HOFF du nom de ses inventeurs) entre l'activité de sortie de la cellule et une activité souhaitée. Cette règle intervient dans des réseaux dont l'apprentissage est dit supervisé, impliquant une intervention extérieure au réseau lui-même (par exemple Perceptrons de ROSENBLATT) :

$$\Delta w_{ij}(t) = \alpha \cdot [d_i(t) - s_j(t)] \cdot e_i(t)$$

où l'on voit que lorsque $d_i = s_j$ alors $d_i - s_j = 0$, la variation de w devient donc nulle : *la valeur désirée est atteinte*.

Il est important de noter que le choix de ces lois de plasticité peut déterminer la nature du réseau dans son ensemble : muni d'une loi de Hebb le réseau est capable d'évoluer sans intervention extérieure, il peut s'auto-organiser (cf. par exemple le réseau topologique de Köhonen).

3.3 Simulation numérique

3.3.1 Algorithmique

Comme il a été dit précédemment, la simulation numérique a pour objectif de tester le modèle au moyen d'un ordinateur digital. Pratiquement la procédure consiste à écrire un programme permettant de calculer les états successifs des différents éléments du réseau pour un jeu de paramètres choisis. Les variables d'états, dont l'évolution suit les lois préalablement définies lors de la modélisation, seront consignées dans des tables. Selon la manière de gérer l'évolution du temps (nous parlons ici du temps fictif de la simulation) on distingue des simulations synchrones ou asynchrones :

3.3.1.1 Simulation synchrone

Dans le cas le plus simple, celle d'une simulation de type synchrone (ou périodique), les variables d'états, comme les niveaux d'activité de chaque élément ou les poids synaptiques de chacune des jonctions, sont actualisées à chaque itération (c'est-à-dire chaque pas temporel). Le programme présente alors la structure décrite sommairement ci-après :

1. Initialiser variables et paramètres
2. avancer l'horloge à $t + \tau$
3. définir la valeurs des afférences
4. calculer l'activité des neurones
5. actualiser les poids synaptiques
6. afficher le(s) graphe(s) de l'état
7. si la durée prévue n'est pas atteinte, reprendre en (3)
8. fin de la simulation (sauvegarde des données)

3.3.1.2 Simulation asynchrone

Dans le cas d'une simulation de type asynchrone, seules sont actualisées les variables d'état des éléments où il se "passe quelque chose", où se produit un événement comme la survenue d'un potentiel d'action au niveau d'une jonction à un instant donné. Les instants où doivent se produire ces événements sont consignés dans un échéancier, où ils sont classés séquentiellement selon un ordre temporel croissant (le plus précoce en tête). La simulation progresse alors de manière apériodique d'événement en événement et seules sont actualisées les variables d'état des éléments qui sont le siège de l'événement. Cela veut dire que l'horloge qui définit le temps de la vie du réseau progresse par bonds d'ampleur variable. Cette actualisation peut être elle-même à l'origine d'événements futurs qui sont alors consignés à leur tour dans l'échéancier ; les événements qui viennent d'être traités sont éliminés, effacés de l'échéancier. (Ainsi par exemple le potentiel de membrane d'un élément de type neurone ne sera actualisé que lors de la survenue d'un P.A. sur une de ses entrées ; si cet événement permet de faire décharger le neurone, alors on peut prévoir pour les cellules auxquelles ce neurone est connecté en sortie, des événements futurs qui devront donc être inscrit dans l'échéancier).

L'algorithme de ce type de simulation est le suivant :

1. Initialiser, lire et classer dans l'échéancier les événements extérieurs ;
2. avancer l'horloge au temps prévu du 1er événement de l'échéancier ;
3. traiter cet événement ;
4. l'effacer de l'échéancier ;
5. si cet événement a engendré de nouveaux événements, les inscrire dans l'échéancier ;
6. s'il y a encore des événements dans l'échéancier et si la durée prévue de la simulation n'est pas dépassée, recommencer en (2), sinon :
7. fin de la simulation (sauvegarde des données).

Chapitre 4

Exemples de Réseaux neuromimétiques

4.1 Réseau discriminateur(B. ANS, J. HÉRAULT, C. JUTTEN)

- Ce modèle de réseau est intéressant à étudier pour au moins deux raisons :
- il montre qu'un réseau simple, ne mettant en jeu que peu de neurones formels (a minima 2 cellules) peut accomplir une tâche complexe que les algorithmes classiques ne savent pas réaliser ;
 - il est un bon exemple de l'intérêt des échanges interdisciplinaires

4.1.1 Le questionnement

En neurosciences, l'élaboration de modèles de réseaux nerveux est un moyen de répondre à un questionnement, il est intéressant de développer celui qui a conduit à l'élaboration du réseau discriminateur. Le signal produit par un certain nombre de capteurs sensoriels biologiques est souvent porteur d'une information ambiguë. C'est le cas notamment des fuseaux neuromusculaires qui codent, sous forme impulsionnelle, à la fois la longueur et la vitesse de l'élongation musculaire (cf. § 0.6). Ce codage est effectué de manière sensiblement différente par des terminaisons primaires de gros diamètre (fibre Ia) et par des terminaisons secondaires de diamètre moyen (fibre II). Par ailleurs, le codage par les fuseaux de l'état musculaire en terme de longueur et de vitesse d'élongation se complique en raison de l'existence d'un contrôle efférent de sensibilité statique et/ou dynamique par les motoneurones α ou β . Ce contrôle modifie la sensibilité à la composante statique ou dynamique de l'étirement musculaire.

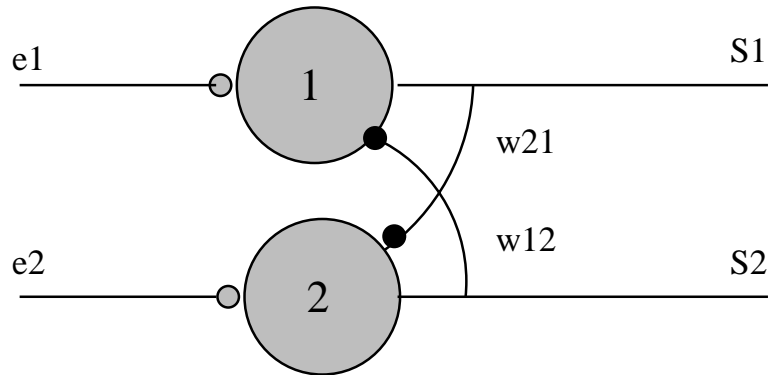
Cette relative complexité du signal afférent issu des fuseaux neuromusculaires conduit à se poser la question suivante : si l'élaboration des comportements sensori-moteurs nécessite la connaissance par le système de l'une ou l'autre de ces deux composantes (vitesse d'élongation ou longueur), y-a-t'il une structure

susceptible de séparer ces deux informations à partir du signal fusorial? {cette question se pose aussi bien dans le cadre de la compréhension des mécanismes de contrôle et de régulation motrice, que dans le domaine de la perception kinesthésique.}

Un première proposition formulée par des physiologistes, consiste à supposer l'existence d'un réseau simple effectuant la soustraction du signal des fibres II à celui des fibres Ia. Cette proposition est basée sur le fait que les fibres secondaires produisent un signal dans lequel la présence de l'information 'position' est plus importante qu'elle ne l'est dans le signal des fibres Ia, où domine davantage l'information 'vitesse'. Il reste à déterminer, pour que cette hypothèse soit recevable, comment le réseau qui opère la séparation "connaît" les coefficients optimaux permettant de la réaliser au mieux. (Ces coefficients doivent rendre compte des proportions relatives des deux données dans le signal composé). Une réponse consiste à supposer l'existence d'une plasticité synaptique qui par un processus auto-adaptatif modulerait la force synaptique de la jonction inhibitrice (celle qui réalise la "soustraction" du signal II), pour tendre vers cette valeur optimale. Ainsi par auto organisation le système tend vers un état stable où s'effectue correctement la séparation des deux composantes. Ce modèle que nous avons proposé (ANS et GILHODES, C.R.A.S.) est composé d'une seule cellule afférentée par les fibres primaire et secondaire, cette dernière formant une jonction inhibitrice adaptative. Ce modèle ne répond toutefois que partiellement à la question puisqu'une seule des deux composantes est produite en sortie du réseau. Le modèle que nous allons étudier en détail est une sorte de généralisation de cette première proposition.

4.1.2 Description du modèle

Le réseau discriminateur est formé de deux neurones en inhibition mutuelle. Chacun de ces deux neurones reçoit en entrée une afférence porteuse d'un signal sensoriel, soit e_1 pour le neurone 1 et e_2 pour le neurone 2.



Chacun de ces signaux e_1 et e_2 est la résultante d'un mixage, somme pondérée d'informations primaires x_1 et x_2 (équations 1 et 2). Les coefficients a_{ij} sont des

constantes telles que, par exemple, $a_{11} > a_{12}$ et $a_{21} < a_{22}$, c'est-à-dire que l'information x_1 est prédominante dans e_1 alors que x_2 est dominante dans e_2 (la situation inverse est bien évidemment possible).

$$e_1 = a_{11}.x_1 + a_{12}.x_2 \quad (1)$$

$$e_2 = a_{21}.x_1 + a_{22}.x_2 \quad (2)$$

Les équations neuroniques ci-après (3 et 4), définissent les sorties de chaque cellule en fonction des entrées et des inhibitions latérales, compte tenu des coefficients synaptiques w_{12} et w_{21} qui caractérisent les jonctions de 1 vers 2 et de 2 vers 1 respectivement.

$$s_1 = e_1 - w_{12}.s_2 \quad (3)$$

$$s_2 = e_2 - w_{21}.s_1 \quad (4)$$

Une étude mathématique élémentaire nous permet de déterminer les valeurs de ces coefficients qui donnent au réseau la capacité de séparer x_1 et x_2 à partir des entrées e_1 et e_2 .

$$(1) \text{ et } (3) \quad s_1 = e_1 - w_{12}(e_2 - w_{21}.s_1)$$

$$(2) \text{ et } (4) \quad s_2 = e_2 - w_{21}(e_1 - w_{12}.s_2)$$

$$s_1 = e_1 - w_{12}.e_2 + w_{12}.w_{21}.s_1$$

$$s_2 = e_2 - w_{21}.e_1 + w_{21}.w_{12}.s_2$$

$$s_1 = \frac{e_1 - w_{12}.e_2}{1 - w_{12}.w_{21}}$$

$$s_2 = \frac{e_2 - w_{21}.e_1}{1 - w_{12}.w_{21}}$$

posons $d = 1 - w_{12}.w_{21}$

$$s_1 = \frac{a_{11}.x_1 + a_{12}.x_2 - w_{12}(a_{21}.x_1 + a_{22}.x_2)}{d}$$

$$s_2 = \frac{a_{21}.x_1 + a_{22}.x_2 - w_{21}(a_{11}.x_1 + a_{12}.x_2)}{d}$$

$$s_1 = \frac{x_1(a_{11} - w_{12}.a_{21}) + x_2(a_{12} - w_{12}.a_{22})}{d}$$

$$s_2 = \frac{x_1(a_{21} - w_{21}.a_{11}) + x_2(a_{22} - w_{21}.a_{12})}{d}$$

pour $s_1 = f(x_1)$ et $s_2 = g(x_2)$:

$$x_2(a_{12} - w_{12}.a_{22}) = 0 \rightarrow a_{12} - w_{12}.a_{22} = 0$$

$$x_1(a_{21} - w_{21}.a_{11}) = 0 \rightarrow a_{21} - w_{21}.a_{11} = 0$$

$$\rightarrow w_{12} = \frac{a_{12}}{a_{22}}$$

$$\rightarrow w_{21} = \frac{a_{21}}{a_{11}}$$

Ces valeurs idéales pour une séparation correcte des deux signaux primaires sont inconnues du réseau et résultent d'une analyse mathématique, et ne peuvent être calculées qu'à condition de connaître les valeurs des coefficients a . Le réseau doit s'approcher des ces valeurs selon un processus autoadaptatif, c'est-à-dire sans intervention extérieure. Ce processus est fondé sur une loi de plasticité synaptique selon l'équation suivante :

$$\frac{\partial w_{ij}}{\partial t} = \alpha.s_i^3 \cdot \arctan[\beta(s_j - \widehat{S}_j)]$$

α et β sont des paramètres, et \widehat{S}_j constitue un facteur de traînage tel que, à $t + 1$:

$$\widehat{S}_j = \widehat{S}_j + \tau \cdot (s_j - \widehat{S}_j)$$

4.1.3 Simulations

Elles montrent que le réseau "apprend", en modifiant progressivement ses coefficients synaptiques, à séparer les composantes x_1 et x_2 et les restitue sur les sorties s_1 et s_2 . Cette performance s'obtient après un certain nombre d'itérations au cours desquelles on présente en entrée du réseau les signaux mixés. La simulation montre également que la modification de la matrice des coefficients de mixage en cours d'apprentissage réoriente l'évolution des coefficients synaptiques vers les nouvelles valeurs optimales.

4.1.4 Généralisation

Le réseau tel qu'il vient d'être décrit réalise la séparation de 2 sources indépendantes mélangées sur 2 capteurs. C. JUTTEN et J. HÉRAULT ont montré qu'il était possible de généraliser ce principe au cas de n sources mixées sur n capteurs; le réseau comporte alors n neurones en inhibitions mutuelles. Un exemple intéressant est celui des capteurs de position angulaire de 2 caméras braquées à chaque instant sur un même point de l'espace. Les données primitives sont les coordonnées x , y et z du point dans l'espace, et chaque capteur angulaire de la position des caméras est une fonction trigonométrique de ces coordonnées. Le réseau après apprentissage produit en sortie des signaux proportionnels à ces coordonnées cartésiennes pour chacun des points visés. Il n'est pas interdit de faire l'analogie de système avec le système visuel où les yeux remplacent les caméras et où la proprioception des muscles extra-oculaires rend compte de la position des globes oculaires. Un réseau simple du type de celui que nous venons de décrire peut rendre compte du repérage d'un point dans l'espace extérieur sur la base des afférences extra rétiniennes. Bien entendu cette proposition reste à démontrer.

4.1.5 Applications

Ce type de réseau trouve de nombreuses applications dans le domaine du traitement du signal. Parmi ces applications citons l'élimination de la diaphonie dans la lecture des pistes d'enregistrements stéréophoniques. Ce réseau existe sous forme de circuit intégré où la vitesse de convergence est extrêmement rapide (par rapport notamment à l'exploitation digitale). Des applications ont également été proposées pour le traitement des signaux sonar.

4.2 Modèle de mémoire associative (T. KÖHÖ- NEN)

Les processus de mémorisation, d'apprentissage, constituent un des problèmes majeurs de la neurobiologie et des sciences de la cognition. Le modèle du

Finlandais T. KÖHÖNEN, est une proposition pour un mécanisme de stockage, d'engrammation mnésique, non localisé. Dans une mémoire de type localisé chaque information est en un lieu donné de l'espace mémoire, lieu qui lui est propre ; dans ce cas pour retrouver l'information il suffit de connaître le lieu, c'est à dire l'adresse de l'information. Ce modèle présente un mode alternatif de stockage, et montre comment plusieurs informations peuvent être enregistrées dans un même espace neuronique, dans une même population de synapses. Chaque information est restituée lors de la présentation en entrée d'une partie de cette information, la partie est associée au tout. Dans ce cas on parle de mémoire autoassociative {les mémoires de type hétéroassociatif sont celles qui associent des informations différentes : par exemple les informations a, b,... sont associées lors d'un apprentissage à m, n... respectivement ; la présentation de a en entrée du réseau provoquera la production de m en sortie. cf. par exemple le réseau de ANDERSON que nous n'étudierons pas ici}

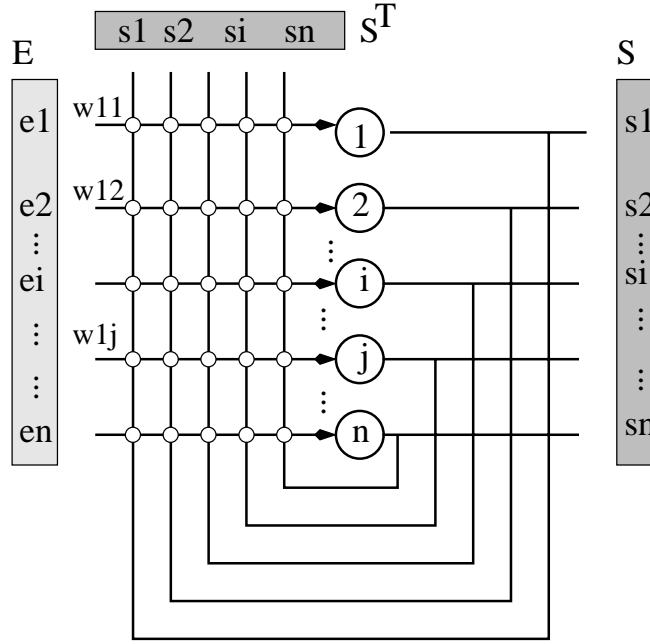
Nous décrirons successivement son architecture (connectique), les lois neuroniques et mnémoniques (plasticité), puis les résultats des simulations et enfin les commentaires que ce modèle suscite.

4.2.1 Architecture

La figure de la page suivante décrit schématiquement la structure du réseau. Il est composé d'une matrice (couche) de neurones formels, c'est donc un couche à deux dimensions dont seule une rangée a été représentée sur la figure pour la clarté du schéma. Chaque neurone i reçoit en entrée :

une afférence e_i qui lui est spécifique, (la topologie des afférences correspond à la topologie de la couche, on peut se représenter cet ensemble d'afférences comme des fibres provenant d'une couche de photorécepteurs rétiniens ; la rétinotopie est conservée au niveau de la couche) ;

des collatérales récurrentes, chaque neurone est relié à tous les autres dans la couche, ainsi qu'à lui-même. Ces jonctions récurrentes sont munies d'une loi de plasticité régissant les poids w_{ij} de ces synapses.



4.2.2 Équations

Chaque neurone produit à chaque instant un signal qui est la somme pondérée de ses entrées, le coefficient de l'entrée e_i étant égal à 1. La sortie du neurone i au temps $t + \tau$ est égale à :

$$s_i(t + \tau) = e_i(t) + \sum_j w_{ij}(t) \cdot s_j(t)$$

où w_{ij} est le coefficient synaptique de la jonction de j vers i . τ est le pas d'incrémentement de l'horloge rythmant la simulation.

La loi de plasticité s'apparente à une loi de Hebb, la variation du coefficient synaptique dépend à la fois de l'activité de la fibre récurrente, c'est à dire l'activité de sortie s_j de la cellule formant la jonction, et de l'activité du neurone post synaptique s_i :

$$\Delta w_{ij}(t) = \alpha \cdot \Delta t \cdot s_i(t) \cdot [s_j(t) - \beta]$$

α est un coefficient d'adaptation, sorte de facteur de sensibilité synaptique et β un niveau de base positif.

1

¹En notation matricielle ces équations s'écrivent : $S = E + WS$ et $\Delta W = \alpha \cdot \Delta t \cdot S(S - \beta)^T$ où E et S sont des vecteurs (respectivement représentatifs des entrées et des sorties) et W une matrice comportant les coefficients synaptiques (L'exposant T indique l'opération de transposition).

4.2.3 Simulations

Elles consistent à présenter en entrée E du réseau des images, par exemple des photos d'identité de différentes personnes. A chaque présentation on calcule l'activité stabilisée du réseau, et on applique ensuite la règle de plasticité; tous les coefficients synaptiques sont donc modifiés à chaque présentation. C'est la phase d'apprentissage, d'engrammation du réseau. Dans une seconde phase, on présentera des images et on calculera l'état d'activité de l'ensemble des neurones, sans faire intervenir la règle de plasticité. On observe à chaque fois ce que le réseau produit en sortie, et si l'on fait une représentation dans le plan des niveaux de sortie de la matrice des neurones on obtient une image qui est la réponse du réseau à l'image proposée en entrée. C'est la phase de test, ou phase de restitution, qui va permettre de constater les effets de l'apprentissage.

On observe que si l'on présente en entrée une photo comportant un partie d'une photo apprise (par exemple la bande comportant les yeux), alors le réseau restitue la photo initiale dans son intégralité. En d'autres termes avec une partie de l'information le réseau reconstruit le tout, c'est en ce sens que ce réseau est qualifié d'autoassociatif. On observe également qu'une photo sévèrement bruitée (sujet quasiment non reconnaissable) est restituée améliorée en sortie (proche de l'original appris). D'autres simulations effectuées par KÖHÖNEN montrent que la qualité des restitutions dépend du nombre de visages mémorisés (160 ou 500).

4.2.4 Commentaires

Certaines contraintes existent pour qu'un tel apprentissage soit possible, il est nécessaire notamment que les patterns soient orthogonaux (produit scalaire nul ou proche de zéro). Cette contrainte peut trouver sa solution dans les réseaux nerveux naturels grâce au prétraitement effectué par des réseaux à inhibition latérale récurrente comme on en observe dans les voies sensorielles (par ex. au niveau des relais rétiniens), réseau dont une des propriétés est d'accroître les contrastes. Cette contrainte d'orthogonalité n'invalide donc pas nécessairement le modèle.

Par contre, le modèle est défaillant si l'on présente en entrée des fragments d'images décalés par rapport à l'original (translation, rotation... etc). Dans ce cas la restitution est impossible. On peut toutefois imaginer que l'image engrammée puisse être construite à partir d'invariants qui ne dépendent pas de la position absolue de l'objet visualisé. En conséquence, à condition de le compléter de réseaux de prétraitement (accroissement des contrastes, extraction des invariants par ex.), ce modèle reste une représentation possible des mécanismes de mémorisation distribuée (non localisée).

4.3 Réseau Topologique de T. Köhonen (1982)

4.3.1 Problématique

De nombreuses structures du système nerveux central sont le siège de repré-

T. KÖHÖNEN, Self-organized formation of topologically correct feature maps. *Biological Cybernetics* 43 (1982)