



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Une modélisation fonctionnelle du cortex: la colonne corticale

Aspects visuels et moteurs

THÈSE

présentée et soutenue publiquement le 9 Avril 1990

pour l'obtention du

Doctorat de l'Université de Nancy I
(Spécialité Informatique)

par

Frédéric Alexandre



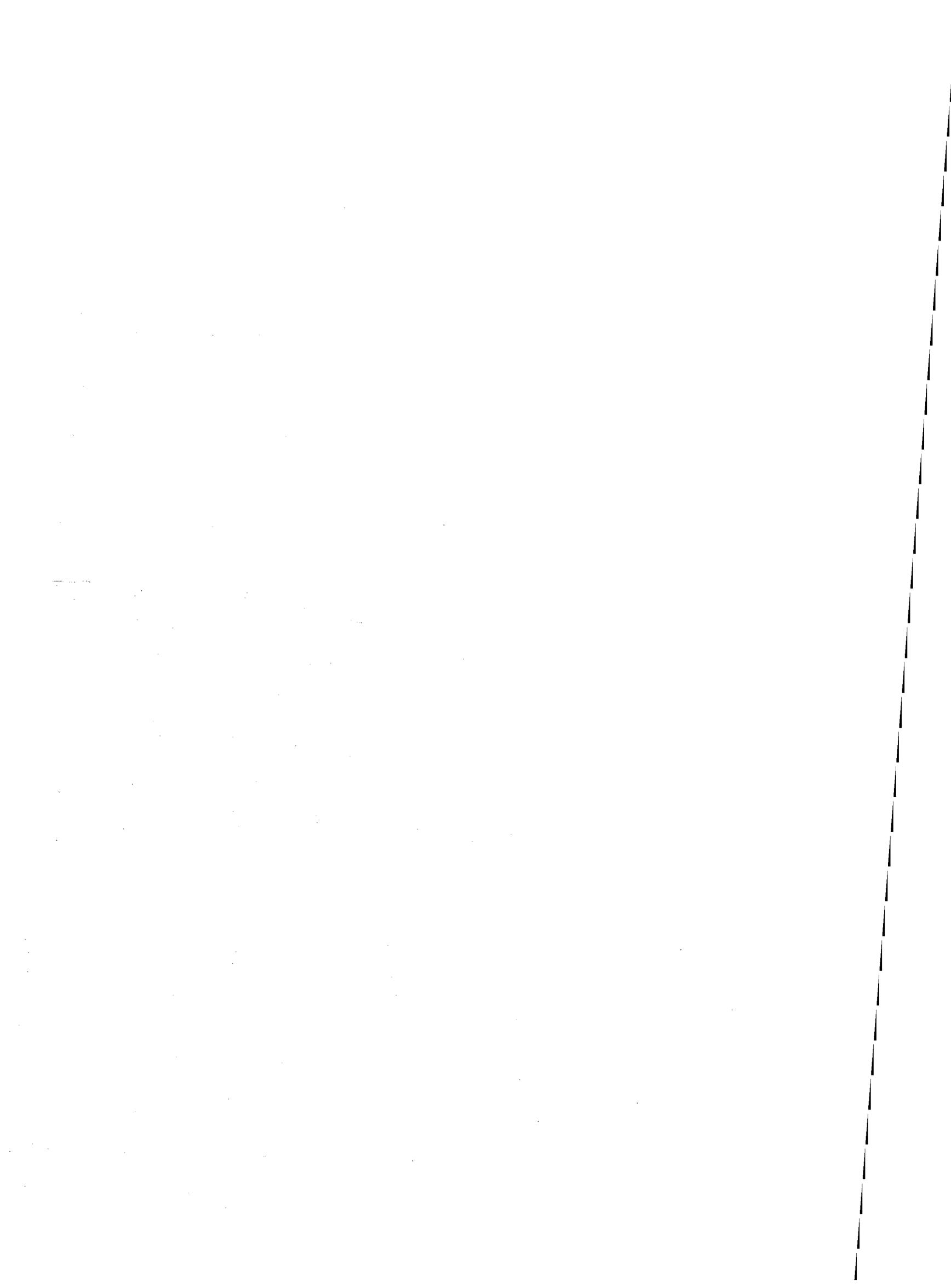
Composition du jury :

Président : J.M. PIERREL

Rapporteurs : J. HÉRAULT
M. RUSINOWITCH

Examineurs : J.P. AUBIN
J.P. HATON

Invité : Y. BURNOD



Cette thèse a été, pour moi, bien plus qu'un travail; elle m'a montré que, en plus des sciences exactes, les sciences humaines existent, sont passionnantes et jouent un rôle primordial.

Elle est dédiée à tous ceux qui, d'une manière ou d'une autre, ont permis qu'elle se déroule ainsi.

Elle est donc dédiée à mes parents et à ma famille; ils m'ont permis de cultiver ma très grande curiosité qui m'a fait découvrir et me passionner pour ce sujet.

Elle est également dédiée à mes amis, dont ma compagne; à leur tour, ils supportent, voire même encouragent, toutes les conséquences de cette passion dans la vie que nous menons ensemble.

Elle est enfin dédiée à tous les autres; à ceux que le sujet intéresse et à ceux qui n'en ont rien à f...

pour leur façon d'être.

Remerciements

Je tiens à remercier:

- M. Jean-Marie Pierrel, professeur à l'Université de Nancy I, responsable du pôle parole de l'équipe Reconnaissance des Formes et Intelligence Artificielle, qui me fait l'honneur de présider ce jury. Je lui suis reconnaissant pour l'intérêt qu'il a toujours porté, lors des réunions de l'équipe, à mes travaux et, en particulier, à mes propositions dans le domaine de la reconnaissance de la parole.
- M. Jeanny Hérault, professeur à l'Institut National Polytechnique de Grenoble, pour le grand honneur qu'il me fait d'avoir accepté de juger mon travail. En effet, sa qualité de chercheur, depuis 1970, dans le domaine du connexionnisme et sa reconnaissance nationale et internationale, rendent ses avis extrêmement précieux pour le chercheur débutant que je suis. Qu'il trouve ici l'expression de mon respect et de ma gratitude pour l'intérêt dont il fait preuve à l'égard de mon travail.
- M. Michael Rusinowitch, maître de conférence à l'Université de Nancy I, qui a accepté d'être rapporteur de ma thèse, malgré ses occupations de recherche et d'enseignement. Je le remercie pour l'intérêt qu'il a porté à ce sujet et pour les discussions que cet intérêt a engendré.
- M. Jean-Pierre Aubin, professeur à l'Université de Paris Dauphine, qui a accepté de siéger dans ce jury et qui m'a fait ainsi le grand honneur de s'intéresser à ma thèse.
- M. Jean-Paul Haton, professeur à l'Université de Nancy I, qui a encadré et dirigé mon travail de recherche. Je ne le remercierai jamais assez pour l'intérêt sans faille qu'il a, depuis le début, toujours porté à mes recherches ni pour le temps qu'il a su me consacrer, malgré toutes les responsabilités qu'il assume. Il est pour beaucoup dans mon apprentissage du métier de chercheur et ses qualités scientifiques et humaines sont pour moi un exemple.
- M. Yves Burnod, chargé de recherche à l'Institut National de la Santé Et de la Recherche Médicale, qui est à l'origine du modèle de la

colonne corticale et à qui je dois ma formation dans le domaine de la Neurobiologie. Qu'il trouve ici l'expression de tous les sentiments qu'ont suscité, chez moi, son intelligence, sa culture, sa grande patience et son amitié. J'espère de tout coeur que notre collaboration pourra continuer dans les années à venir.

Je tiens également à remercier tous mes collègues du CRIN pour l'ambiance dans laquelle ils me permettent de travailler. Je pense en particulier aux membres de la jeune équipe connexionniste dont, bien évidemment, Frédéric Guyot avec qui je viens encore de passer trois ans dans des conditions qui ne sont plus vraiment du travail.

Je destine enfin un remerciement tout spécial à Martine, Nadine, Isabelle, Agnès et tou(te)s les autres, qui m'ont apporté leur aide administrative et technique et leur disponibilité chaleureuse.

Sommaire

CHAPITRE I: INTRODUCTION

1. Intelligence Artificielle	1
-Vision	2
-Audition	3
-Raisonnement	3
-Bilan	4
2. Sciences Cognitives	5
-Neurobiologie	5
-Psychologie	7
3. Coopération	8
4. Démarche	10
-Choix d'un modèle	10
-Plan général	11
-Avertissement	12

CHAPITRE II: TROIS NIVEAUX D'ORGANISATION DU CORTEX

1. Introduction au système nerveux	13
-Le système nerveux périphérique	14
-Le système nerveux central	15
-Les voies d'information	17
-Trois niveaux d'organisation du cortex	20
2. Le neurone	22
-Le neurone vivant	22
-Le neurone formel	26
-Les réseaux à couches	30
-Les réseaux à satisfaction de contrainte	34
-Les autres modèles	37
-Discussion	39
3. Les aires	42
-Les types d'aires	42
-Le réseau	45
4. La colonne corticale	52
-Réalité biologique de la colonne	52
-Architecture et connectivité	56
-Etats d'activation	64
-Règles de fonctionnement	66
-Apprentissage	71
-Implantation du modèle	75
-Conclusion sur la modélisation	82

CHAPITRE III: APPLICATIONS

1. Vision	85
-Présentation du problème	85
-Le codage	87
-Le retour sur l'information	93
-Réalisation	97
-Discussion	105
2. Reconnaissance de la parole	108
-Présentation du problème	108
-La division	109
3. La motricité	114
-Présentation du problème	114
-Le codage	115
-L'auto-organisation fonctionnelle	122
-Discussion	129
4. Interprétation d'images	131
-Présentation du problème	131
-Les aires	131
-Les colonnes	133

CHAPITRE IV: CONCLUSION

1. Bilan	135
2. Perspectives	139
3. Conclusions personnelles	143
ANNEXE	145
BIBLIOGRAPHIE	148

Avertissement

La partie théorique présentée dans ce mémoire est le fruit d'un travail commun entre Frédéric Guyot et moi-même. C'est pourquoi les deux premiers chapitres de cette thèse se trouvent également dans (Guyot, 1990). En revanche, les applications issues de cette théorie ont été partagés entre aspects auditifs et visuels proposés dans (Guyot, 1990) et aspects visuels et moteurs développés au chapitre III de ce mémoire.

CHAPITRE I:

INTRODUCTION

”Modélisation fonctionnelle du cortex”; ces trois mots, qui symbolisent trois années de recherche, révèlent implicitement quelle a été notre démarche. Mais avant de lever le voile sur nos travaux, il convient d’introduire quelques notions élémentaires (mais fondamentales) sur divers aspects des sciences cognitives...

1 Intelligence Artificielle

De tous les outils inventés par l’homme pour faire face aux problèmes du monde extérieur, l’informatique est sans doute l’un des plus fascinants, puisqu’elle ne le seconde plus sur le plan physique, mais sur le plan intellectuel. Plus particulièrement, l’Intelligence Artificielle s’est donné comme objectif la simulation de fonctions intellectuelles humaines.

De façon assez générale, un banal dictionnaire nous livre les définitions suivantes:

Intelligence: Faculté de donner un sens, de s’adapter aux circonstances, de résoudre des problèmes.

Analysons un peu plus:

Sens: Ensemble des **représentations** que suggère un énoncé,

Circonstances: Ensemble des **faits** qui accompagnent un évènement,

Résoudre: Trouver la solution, **décider**.

Intuitivement, cette séparation en différentes expressions de l’intelligence semble bien artificielle: comment ”décider” sans appréhender ”l’énoncé”

ou sans tenir compte des "évènements" extérieurs? Plus pragmatiquement, l'Intelligence Artificielle traditionnelle essaie d'élaborer des systèmes sensibles au contexte et propose une séparation fonctionnelle de l'intelligence. Entre autres, on distingue classiquement: la simulation de fonctions visuelles, de fonctions auditives, le raisonnement, la génération de plans.

Là non plus, cette distinction n'est pas satisfaisante, mais elle a le mérite de séparer le problème de l'intelligence suivant une certaine logique: la *fonctionnalité*.

Depuis de nombreuses années, les informaticiens ont développé une pléiade de méthodes pour s'attaquer à des problèmes spécifiques, bien compartimentés dans leur domaine.

1.1 Vision

Les efforts les plus anciens ont porté sur la reconnaissance d'images; les applications potentielles sont innombrables, de la lecture automatique de documents à la conduite automatique de véhicules. La quantité d'informations codées visuellement est telle que l'informatique peut être d'un grand secours. Le problème ici est de donner un sens aux éléments visuels, et plus particulièrement de mettre un nom sur un flux d'informations. Nous parlerons par la suite de "capacités associatives" lorsqu'on associera un mot à des données. L'aphorisme "diviser pour régner" va encore aider l'informaticien dans sa tâche: il va lui permettre de décomposer le travail visuel de l'ordinateur en plusieurs phases, séquentielles ou, mieux, parallèles:

- Une phase de prétraitement de l'information qui consiste essentiellement en une mise en forme des données (lissage, extraction de contours, segmentation en régions...).
- Une phase d'extraction d'indices qui permet de sélectionner les éléments importants, voire structurels, parmi la quantité astronomique d'informations (1 Méga octets pour une image) et qui propose un codage plus compact (extraction de primitives pertinentes comme les segments, les arcs ou les angles...).

- Une phase de reconnaissance proprement dite qui consiste, la plupart du temps, à comparer le codage final de la forme examinée avec des *modèles* généraux et prédéfinis des formes à reconnaître.

On pourrait, pour chacune de ces phases, citer des dizaines d'algorithmes, plus ou moins complexes et plus ou moins performants, dont l'accumulation au fil des années montre bien la complexité du problème; les progrès sont durs à acquérir, malgré les efforts de nombreux chercheurs à travers le monde. On peut distinguer, entre autres, deux problèmes majeurs: (i) l'acquisition des données qui est loin d'être parfaite et génère des bruits généralement importants et difficilement modélisables, (ii) la grande diversité des instanciations d'un même concept, qui en rend la modélisation difficile.

1.2 Audition

Les applications de la reconnaissance auditive sont également très larges: de la machine à dicter à la commande vocale d'outils divers, l'intérêt de la communication orale réside, entre autres, dans sa rapidité. Le problème ressemble étrangement à la reconnaissance visuelle (de façon générale, on parle de reconnaissance de formes), puisqu'il s'agit là aussi d'associer un sens à un flux d'information. On retrouve donc le même schéma: séparation en trois phases de traitement et même problématique. La variabilité inter et intra locuteur, le bruit de fond et l'acquisition sont autant d'obstacles à des performances comparables à celles de l'homme.

1.3 Raisonnement

L'Intelligence Artificielle fait appel à des notions plus abstraites, plus symboliques (mais le problème n'est pas forcément plus complexe à manipuler que la mémorisation associative), et utilise une formalisation des processus différente. On distingue typiquement des événements, des faits, qui représentent une vision du monde extérieur dans lequel le système doit évoluer et des mécanismes de déduction qui, à partir des faits, permettent de converger vers des solutions. Le diagnostic médical, l'interprétation de données ou la prise de décisions sont de bons exemples de systèmes de raisonnement couramment abordés en Intelligence Artificielle. Malheureusement, les problèmes ont tendance à s'accumuler: les faits en provenance de l'extérieur

sont issus de systèmes dont l'analyse n'est pas forcément absolue; les règles de déductions, souvent modélisées par des experts du domaine d'analyse considéré, ne sont pas toujours compatibles; qui plus est, la connaissance intuitive de l'expert est rarement perçue, et donc modélisable, consciemment. Quoique plus récemment étudié que les fonctions perceptives, le raisonnement a déjà inspiré de nombreuses approches, de plus en plus fines et complexes, sans pour autant aboutir à des systèmes réellement opérationnels.

1.4 Bilan

En résumé, la méthodologie employée a compartimenté l'intelligence en un certain nombre de domaines isolés, tandis que des techniques bien spécifiques évoluaient en parallèle pour résoudre ponctuellement certains problèmes. Cependant les performances plafonnent.

On se rend d'abord compte que les différents domaines ont besoin les uns des autres; comment reconnaître la parole sans utiliser des connaissances lexicales ou sémantiques? L'homme lui-même n'y parvient que partiellement; il faut donc faire coopérer des systèmes qui sont, malheureusement, souvent incompatibles.

Ensuite, il est clair que, quel que soit le domaine considéré, on retrouve toujours la même difficulté: il s'agit de bâtir des modèles de la connaissance qui puissent servir de trames de réflexion ou d'éléments de comparaison. Les écueils sont nombreux, et s'appellent généralisation, structuration, concepts flous, temps....

Les systèmes à base de connaissances doivent être à même d'intégrer ces notions dans leur représentation interne. La tâche n'est guère facilitée par le manque d'outils mathématiques susceptibles de modéliser et de rendre compte de telles aptitudes; la rigueur et la logique cartésienne ont fait évoluer les mathématiques dans une voie incompatible avec la souplesse et la plasticité de l'intelligence humaine.

Un premier pas a été fait dans la représentation de connaissances floues en développant de nouvelles techniques basées sur des méthodes logiques (logique floue; Zadeh, 1988), ensemblistes (Zadeh, 1979) ou probabilistes (Nilsson, 1986; Shafer & Logan, 1987). L'approche probabiliste nous semble être la plus à même de modéliser cette notion car la coïncidence de probabilités et statistiques dans les systèmes à base de connaissances fait émerger

un nouveau concept primordial: l'apprentissage. Quoi de plus normal, en effet, que de laisser le système générer ses propres modèles de connaissances à partir d'exemples variés et représentatifs, plutôt que de lui imposer une structuration rigide et inadaptée de l'information ?

Le deuxième concept qui va guider notre recherche, est l'unification des multiples domaines de réflexion en Intelligence Artificielle qui, non seulement butent sur des difficultés communes, mais sont aussi extrêmement interdépendants dans leurs traitements. Notre objectif n'est certes pas de révolutionner l'Intelligence Artificielle, mais plutôt de prendre du recul par rapport à la masse considérable de connaissances accumulées sur le sujet ces dernières années, et d'apporter notre modeste contribution à une approche globale du problème.

2 Sciences Cognitives

Puisque nous parlons d'unification et de globalité, remontons encore la hiérarchie structurelle des sciences et quittons pour un moment Intelligence Artificielle et Informatique. Il existe épistémologiquement d'autres domaines scientifiques qui s'intéressent de très près à l'intelligence (humaine); il nous semble constructif de rechercher dans les approches neurobiologiques et psychologiques des concepts transposables en Intelligence Artificielle. Cette approche tri-disciplinaire des mécanismes cognitifs, riche d'enseignements, constitue la clé de voûte de notre réflexion.

La compréhension des sciences passe par une introduction à leur vocabulaire et à leurs concepts à travers, par exemple, un historique du domaine étudié.

2.1 Neurobiologie

Après le foie et le coeur, les scientifiques ont commencé à soupçonner le cerveau d'être le siège de l'intelligence (Changeux, 1983); plus tard, l'évolution instrumentale va permettre des progrès rapides comme la découverte de la cellule, et plus spécifiquement de la cellule nerveuse: le neurone. C'est alors que vont débiter les études neurobiologiques du système nerveux. Les moyens actuels d'investigation sont nombreux, et contrairement à l'opinion habituellement répandue, ils ont déjà permis de dégager une somme de

connaissance très appréciable sur le fonctionnement intime du cerveau. Les examens au microscope de coupes histologiques colorées permettent une étude statique et structurelle des cellules; les micro-électrodes permettent d'enregistrer dynamiquement le fonctionnement d'une seule cellule vivante, *in situ*, en réponse à des stimuli bien précis; les caméras à positrons sont capables d'enregistrer le fonctionnement global du cerveau avec une résolution inférieure à *5mm*. Tous ces instruments ont stimulé l'imagination des chercheurs et des milliers d'observations expérimentales sont disponibles comme support indispensable à leurs travaux. On peut alors se demander pourquoi le fonctionnement de ce superbe système n'a pas été élucidé depuis longtemps; quelles sont les limitations qui, une fois de plus, viennent entraver la marche de la découverte?

Nous en distinguerons au moins trois:

1. Même si le neurone est déjà bien schématisé, il reste une multitude de mécanismes annexes, plus ou moins bien connus, qui se combinent les uns aux autres et dont l'importance pour le traitement de l'information est très difficilement évaluable. La dualité électrique-chimique du fonctionnement est pour beaucoup dans les problèmes d'interprétation de ces multiples micro-phénomènes. Cette complexité est telle que les neurobiologistes hésitent à proposer des modèles de neurones, conscients de la grande incertitude qui entoure son fonctionnement.
2. Si, en informatique, il se passe rarement longtemps entre l'énoncé d'une théorie et sa validation expérimentale par simulation, les neurobiologistes se débattent avec des constantes de temps beaucoup plus grandes. L'expérimentation et l'observation doivent passer par des étapes terriblement minutieuses et fastidieuses. Le temps de conditionnement d'un animal à effectuer une tâche cognitive est fonction de sa difficulté; la préparation des dispositifs et la mise en place d'observations, les multiples manipulations pour mettre en évidence des résultats significatifs et la nécessité de la reproductibilité ralentissent les recherches; une expérimentation simple s'entreprind généralement sur des années !
3. De manière moins concrète mais plus décisive, il est extrêmement difficile (quel que soit le domaine) de faire des prévisions de compor-

tement à un niveau macroscopique, en partant d'observations locales microscopiques. Ce phénomène de césure (Hamburger, 85) entre les deux extrêmes peut être illustré par un exemple célèbre en thermodynamique: comment déduire la loi d'équilibre des gaz parfaits ($PV=nRT$) à partir du comportement d'une molécule d'un gaz ? La résolution fait intervenir plusieurs étapes très intéressantes:

- il faut d'abord posséder un système mathématique qui puisse décrire les phénomènes observables, en l'occurrence les statistiques,
- modéliser le comportement individuel des éléments microscopiques,
- modéliser leurs interactions parallèles et réciproques;
- des outils mathématiques permettent enfin de déduire un comportement global.

Le problème se précise: il faut *modéliser* les micro-phénomènes et leurs interactions pour comprendre le cerveau dans sa globalité comportementale, avec tous les problèmes qu'implique une schématisation: hypothèses, réductionnisme et niveaux de description. La neurobiologie peut descendre jusqu'aux explications moléculaires du fonctionnement d'une cellule, mais la césure est trop grande pour qu'on puisse actuellement en tirer des interprétations macroscopiques. Des modèles de plus haut niveau de description (à l'échelle du neurone ou, plus récemment, de la colonne corticale) commencent à être proposés et demandent à être soumis à une validation expérimentale.

2.2 Psychologie

L'approche psychologique de l'intelligence humaine emprunte exactement le chemin inverse: c'est à partir d'observations globales du comportement d'un sujet placé dans des conditions spécifiques, que l'on va proposer des modèles de la pensée. Il faut noter ici l'extrême richesse de cette approche qui a déjà apporté à l'Intelligence Artificielle les bases nécessaires à l'élaboration des systèmes experts qui constituent depuis quelques années les techniques les plus efficaces dans le domaine du raisonnement. Mais l'apport de la psychologie est aussi considérable dans les domaines perceptifs.

Ce n'est que depuis le début de ce siècle que la psychologie s'est imposée comme une science exacte, utilisant des méthodes de formalisation et de modélisation cartésiennes basées sur l'expérimentation. Il est intéressant de comparer la dispersion des branches de la Psychologie et de l'Intelligence Artificielle. Plus particulièrement, on parlera de psychologie génétique pour l'étude du développement de l'enfant, branche dédiée à l'apprentissage, de psychologie cognitive pour modéliser le raisonnement et de psychologie perceptive pour l'étude des divers comportements en relation avec le système nerveux efférent (Danset, 1983; Piaget, 1975;...). Une approche pluridisciplinaire ne peut pas négliger de telles informations, mais le travail est rude:

1. Le premier problème réside dans l'interprétation des données expérimentales, phase de réflexion qui doit éviter toute subjectivité; l'introspection est ainsi bien souvent contestable.
2. Les modèles proposés sont rarement informatisables, car les concepts utilisés sont très flous et susceptibles de nombreuses interprétations.
3. La psychologie essaie de modéliser et donc de schématiser des comportements qui sont fonction d'une multitude de paramètres et résultat de nombreux processus complexes. Les liens avec la neurobiologie sont souvent inexistantes sinon fantaisistes; c'est la problématique inverse: passer du macroscopique au microscopique.

Il est indéniable que la psychologie fournit des éléments de réflexions essentiels par son approche macroscopique. De plus, il existe des quantités considérables de résultats expérimentaux qui, quelles que soient leurs interprétations théoriques, peuvent être utilisés dans le cadre d'une théorie plus générale.

3 Coopération

Il est clair que l'Intelligence Artificielle (I.A.), la Neurobiologie et la Psychologie abordent le problème de l'intelligence sous trois angles différents, et qu'une collaboration interdisciplinaire ne peut être que très fructueuse. De

par la complémentarité de leurs approches, les sciences cognitives s'enrichissent et se répondent mutuellement. Les hypothèses, les idées et les modélisations diverses proposées dans un domaine précis peuvent être confrontées et comparées selon différentes perspectives; la validation d'une théorie particulière tend de plus en plus à se faire en accord avec les autres disciplines. Chacune des branches cognitives détient des morceaux d'un puzzle qui se recoupent ou s'imbriquent, et qui conduisent obligatoirement à la description fonctionnelle du cerveau humain.

Le travail peut s'effectuer sur plusieurs plans: il faut avant tout établir une communication interdisciplinaire basée sur un vocabulaire commun accepté par tous, réaliser des comparaisons entre les diverses théories existantes, voir comment elles se situent les unes par rapport aux autres, déterminer les problèmes communs et décider de méthodes pluridisciplinaires et complémentaires pour les résoudre.

On pourrait citer d'autres branches des sciences potentiellement intéressantes pour comprendre et simuler l'intelligence:

- La génétique peut expliquer la genèse du système nerveux et proposer des schémas de construction simples et fonctionnels; n'oublions pas que toute l'information nécessaire pour construire un cerveau est contenue dans nos gènes.
- L'automatique se propose d'identifier, de modéliser des "boîtes noires" en corrélant des signaux de sorties avec des entrées bien spécifiques.
- L'électronique fournit toute la puissance technologique et théorique qui permet déjà aujourd'hui de concevoir des ordinateurs.
- L'optique permettra peut-être un jour de concevoir des systèmes de taille très importante, dépassant de très loin les systèmes électriques.

Nous ne doutons pas qu'il sera un jour nécessaire, pour progresser dans quelque domaine que ce soit, d'emprunter des voies pluridisciplinaires qui, en s'affranchissant des catégorisations artificielles et désuètes des sciences, pourront unifier et simplifier des approches souvent complémentaires.

4 Démarche

Notre objectif est d'élaborer un système à base de connaissances applicable aussi bien à la perception qu'au raisonnement et doté de grandes capacités d'apprentissage. Nous disposons pour cela de bases informatiques, neurobiologiques et psychologiques qui proposent, chacune à leur façon, des éléments de réponse à divers problèmes cognitifs.

4.1 Choix d'un modèle.

Deux possibilités se dégagent: combiner des modèles entre eux ou choisir un modèle particulier et l'enrichir grâce aux autres disciplines. Notre choix s'est porté sur un modèle biologique bien particulier : la colonne corticale, dont la modélisation la plus poussée (et de loin) a été proposée par Y. Burnod en 1981 (Burnod, 1988). Plusieurs raisons ont motivé notre décision:

- ce modèle particulier présente l'avantage de fournir une interprétation de phénomènes comportementaux évolués, très liée aux théories de psychologie expérimentale,
- il se situe à un niveau d'intégration supérieur aux modèles du neurone,
- il permet, de par sa fonctionnalité évoluée, de se rattacher aux notions fondamentales de l'Intelligence Artificielle.

Il semble donc tout à fait adapté à une approche tri-disciplinaire; la neurobiologie sert de base pour le modèle, l'I.A. est utilisée pour la mise au point fonctionnelle d'un système cohérent, et les données expérimentales de la psychologie servent à élaguer différentes options et à déterminer des valeurs de paramètres souvent cruciales (Guyot, Alexandre, Haton, 1989).

C'est en collaboration avec Y. Burnod que nous avons travaillé sur cette approche nouvelle, en ayant pour objectifs des implantations informatiques indispensables pour mettre au point la théorie, la valider et analyser ses conséquences.

4.2 Plan général.

Notre démarche sera la suivante:

Nous présenterons avant tout une modélisation du cerveau humain pour trois niveaux d'organisation possibles:

- le neurone à la modélisation duquel les informaticiens ont déjà beaucoup travaillé,
- la colonne corticale qui, nous l'avons déjà annoncé, sera notre unité de base pour des simulations de comportements globaux,
- et l'aire corticale, où s'élaborent des fonctionnalités de très haut niveau.

Cette modélisation partira toujours de données neurobiologiques pour aboutir à une description formelle et programmable.

Ces bases étant posées, nous nous pencherons sur les applications informatiques réalisées selon une double démarche : aborder les fonctions humaines les unes après les autres et ajouter à chaque fois un nouveau concept à notre modélisation. On rencontrera ainsi les quatre chapitres suivants:

- **La reconnaissance visuelle.**
On étudiera d'abord le codage de l'information visuelle par le système nerveux. Il sera ensuite question de mémoire associative, et de feedback.
- **La reconnaissance auditive.**
Bien que les données neurobiologiques sur le cortex auditif ne soient pas aussi riches que pour le cas visuel, nous proposerons un mécanisme possible. Puis, nous introduirons dans le modèle un deuxième mécanisme d'apprentissage et surtout la notion de temps.
- **La motricité.**
Elle est plus complexe que la perception et met en jeu d'autres structures neuronales aussi primordiales que le cortex. Il y sera également question de connectivité locale et d'indépendance en taille.

- **Coopération.**

Enfin, nous présenterons un système qui associe deux axes perceptifs, l'auditif et le visuel, avec deux mécanismes complémentaires de reconnaissance, et ce pour une interprétation d'images.

Une discussion, un bilan, et des perspectives conclueront cette étude.

4.3 Avertissement.

Nous nous adressons à un public varié et nous passons fréquemment d'une discipline à l'autre. Les exposés théoriques et expérimentaux resteront souvent superficiels ou dépassés aux yeux du spécialiste; inversement, le néophyte risquera d'être submergé par le volume des connaissances nouvelles. De plus la quantité d'information, les données parfois contradictoires et les théories difficilement transposables ne nous permettent pas de prétendre à une étude exhaustive. Nous nous en excusons à l'avance: telle est la contrepartie des études pluridisciplinaires sur un sujet aussi complexe que le Cerveau.

CHAPITRE II:

TROIS NIVEAUX D'ORGANISATION DU CORTEX

1 Introduction au système nerveux

Une étude sur le cortex en particulier ne doit pas masquer le système nerveux en général. En effet, le cortex, si performant soit-il, n'a de sens que couplé aux autres structures nerveuses afin de former un ensemble capable de canaliser les informations sensorielles et de les faire converger sur les neurones qui agissent sur les muscles ou les glandes.

Cette introduction au système nerveux sera d'abord anatomique: nous aborderons succinctement le système nerveux périphérique, trait d'union entre le cerveau et le reste du corps. Puis, partant de la moëlle épinière vers le cortex, nous allons énumérer les différentes parties du cerveau dans un ordre proche de celui dont s'est servi l'évolution pour aboutir à l'homme. Nous y parlerons des fonctions remplies par ces différentes structures nerveuses, certaines structures étant très spécialisées (ex: la rétine), d'autres moins (ex: le corps strié), avant de dégager les principales voies d'information menant au cortex.

En effet, un des buts d'une telle description est de justifier le choix de notre étude en montrant le rôle central et évolué que le cortex joue dans le cerveau. Un autre est plus introductif: il s'agit d'appréhender une certaine logique fonctionnelle dans un réseau de structures comprenant le

cortex, logique que l'on retrouvera dans des réseaux plus fins compris dans le cortex. Enfin, le but principal est pragmatique et relève d'une approche plus informatique que biologique: le passage en revue des structures, de leur interdépendance et de leurs fonctions va nous fournir un premier typage des entrées et des sorties d'informations de notre objet "cortex". Cet intérêt premier pour les contraintes sur les flux d'informations, ajouté à nos compétences et au cadre de ce travail, excusera la vision tronquée et certainement trop fonctionnelle que l'on va donner ici du cerveau. Que le biologiste excuse nos raccourcis saisissants et que l'amateur éclairé se reporte aux références biologiques proposées dans le texte.

1.1 Le système nerveux périphérique

Le système nerveux périphérique, illustré dans la Figure 1, se compose de deux parties (Vincent, 1986).

Végétatif	orthosympathique
	parasympathique
Somatique	motoneurones
	neurones sensitifs

Figure 1. Système nerveux périphérique.

Citons, pour mémoire, le système végétatif ortho et para-sympathique chargé du fonctionnement des organes. Par ailleurs, on trouve le système sensorimoteur somatique, ensemble de nerfs correspondant aux prolongements des motoneurones et des neurones sensitifs. Ce système, beaucoup plus intéressant pour notre étude, sera repris plus longuement, partie par partie, en tête des applications à venir. On peut cependant le décrire très schématiquement. Les motoneurones, depuis la moëlle épinière, innervent tous les muscles squelettiques. Les neurones sensitifs fournissent une représentation du corps par proprioception et, par extéroception, du monde extérieur, analysé au moyen de cinq capteurs spécialisés: la peau, le nez, la langue, l'oreille et l'oeil. Cette information est envoyée au cerveau et plus particulièrement au thalamus et au cortex par des voies correspondant aux nerfs visuels, auditifs...

1.2 Le système nerveux central

Les quelque dizaines de milliards de neurones qui composent le cerveau humain sont décrits ici en cinq parties correspondant à leur répartition topographique, de la moëlle épinière vers le cortex, ainsi qu'à la chronologie de leur apparition au cours de l'évolution (Nauta & Feirtag, 1984). Cette hiérarchisation est utilisée sur l'axe vertical de la Figure 2. L'autre axe représente l'emboîtement des ensembles neuronaux. Nous passerons volontairement très vite sur les structures dont la fonction ne nous a pas semblé primordiale dans le cadre de notre travail (centres responsables de la faim, de la soif, des rythmes cardiaques, respiratoires...).

Cerveau antérieur	Télocéphale	néocortex
		corps strié
		système limbic
	Diencéphale	thalamus
		hypothalamus
Tronc cérébral	Mésencéphale	
	Rhombencéphale	cervelet
		pont de varole
	Myélocéphale	
Moëlle épinière		

Figure 2. Système nerveux central.

1. **Le myélocéphale** ("cerveau proche de la moëlle") ou bulbe rachidien est l'élément nécessaire à la vie, puisqu'il assure la rythmicité cardiaque et respiratoire, et règle la pression artérielle. La majorité de ces fonctions siègent dans la formation réticulée du bulbe.
2. **Le rhombencéphale** ("cerveau en forme de losange"). La protubérance annulaire (ou pont de Varole) est le lieu de passage des fibres verticales qui unissent les différents niveaux du système nerveux central. D'autre part, le cervelet, branché en dérivation le long des voies motrices et sensibles, assure la coordination motrice.

3. **Le mésencéphale** ("cerveau du milieu") contient des centres relatifs au sommeil et au rêve (locus coeruleus), au tonus musculaire (noyau rouge) et à la coordination du mouvement des yeux (colliculus).

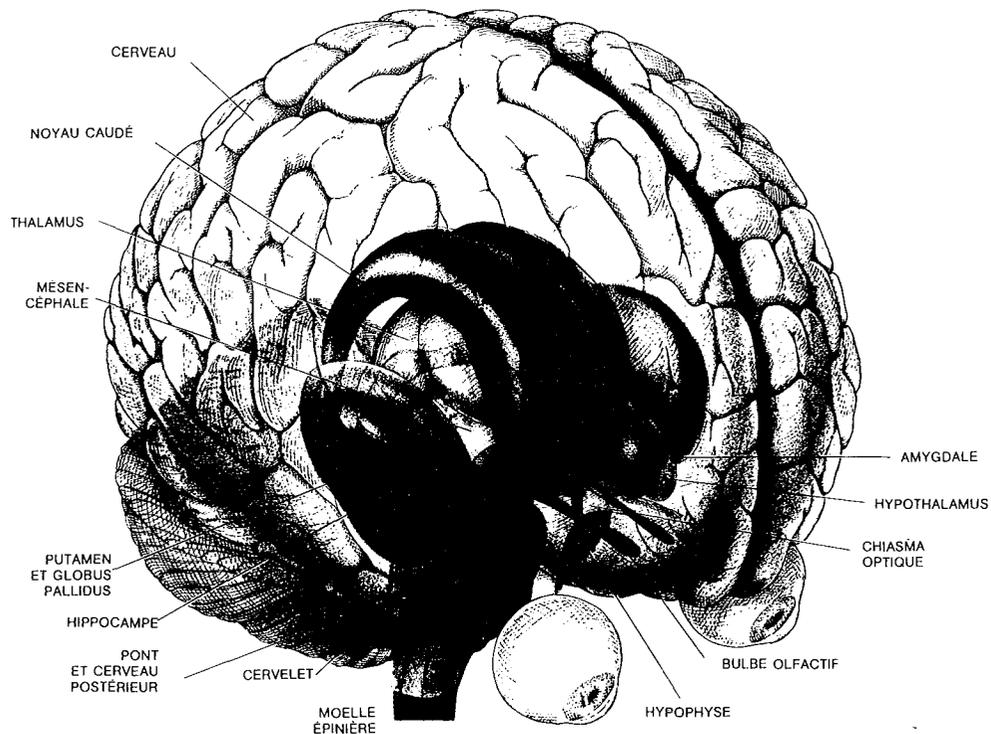


Figure 3. Le cerveau (d'après Nauta & Feirtag, 1984).

4. **Le diencéphale** ("cerveau intermédiaire") est la région centrale du cerveau antérieur. Il constitue l'aboutissement du cerveau des poissons et devient, chez les espèces plus évoluées, un relais privilégié de l'information. Il comprend:

- Le thalamus ("lit" du diencéphale) contient un certain nombre de noyaux dédiés et peut être comparé à un standard téléphonique miniature vers lequel convergent toutes les informations sensibles et sensorielles en provenance du système nerveux périphérique (sauf pour l'odorat). Le thalamus, chez les vertébrés

ayant un cortex, redistribue sélectivement sur les aires du cortex des informations spécialisées sensoriellement et en retourne d'autres vers les aires corticales motrices.

- L'hypothalamus (situé "sous" le thalamus) remplit, de manière schématique, quatre fonctions régulatrices, relatives à la température, la soif, la faim et la sexualité.

5. Le télencéphale ("cerveau avancé")

- Le système limbique (le "bord"), intermédiaire au diencephale et aux hémisphères cérébraux et construit autour de l'odorat, regroupe des structures essentielles à la mémorisation. Schématiquement, on peut dire que le septum joue un rôle dans l'agressivité, l'amygdale relie sensations et souvenirs, les tubercules mamillaires renforcent la mémoire des faits récents et l'hippocampe stocke et consolide les traces mnésiques du cortex. Ces quatre structures coopèrent pour établir un support physique à la mémoire et à l'apprentissage d'une part, et à l'humeur et à l'affectivité d'autre part.
- Le corps strié, comprenant le striatum et le pallidum, joue un rôle d'intégration spatiale de l'information sur les mouvements corporels.
- Le néocortex, que nous appellerons désormais cortex ("écorce"), comprend, chez l'homme, une grande partie des neurones du système nerveux et commande l'activité intellectuelle et cognitive (par opposition aux activités viscérales et émotives vues précédemment). L'installation d'un néocortex dominant sur un antique système nerveux et hormonal correspond, suivant l'expression du prix Nobel François Jacob, "à l'installation d'un moteur à réaction sur une vieille charrette à cheval" (Jacob, 1981).

1.3 Les voies d'information

Paradoxalement, notre description du cortex a été jusqu'à présent concise relativement à celle d'autres structures dont on ne voit pas, a priori, le rôle qu'elles pourraient jouer en Intelligence Artificielle. D'une part, cette

énumération avait pour but de replacer, au moins une fois, le cortex dans le contexte du système nerveux. D'autre part, nous allons maintenant entreprendre une description fonctionnelle de cette structure de manière plus systématique.

1.3.1 Les entrées du cortex

Le cortex, avec l'aide des autres structures coopérantes, participe à de nombreuses fonctions qui peuvent être, et c'est heureux pour l'anatomiste, localisées de manière topographique. Intéressons-nous donc aux différentes parties de ce manteau cortical qui recouvre les autres parties du cerveau.

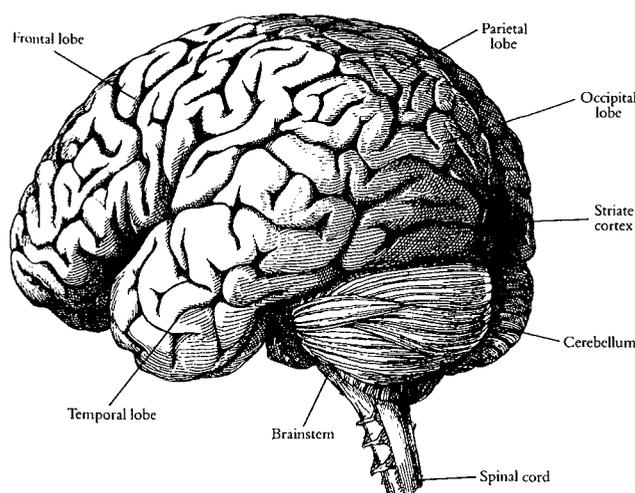


Figure 4. Régions corticales. (d'après Hubel, 1988)

Au début du siècle, Brodmann établit une division du cortex en régions (Brodmann, 1909) qui reste encore valable de nos jours, uniquement au moyen d'études anatomiques de cytoarchitecture, c'est-à-dire de répartition et de relation des différents types de neurones à travers le cortex. Il est remarquable de constater que cette même division peut être obtenue en examinant les liaisons du cortex avec les diverses structures coopérantes. Ainsi, le corps genouillé latéral (CGL), noyau du thalamus recevant les afférences du nerf optique, se projette dans la région occipitale du cortex et réalise les aires visuelles. Le corps genouillé médian (CGM) reçoit les afférences auditives, se projette dans la région temporale et crée les aires

auditives. Il en est de même pour le noyau VL et les aires corporelles et motrices dans la région centrale et pariétale.

C'est aussi avec cette même région centrale pariétale que le cervelet établit ses relations avec le cortex. Quant à la formation réticulée et au corps strié, ils se projettent plutôt sur la région frontale du cortex. Ces différentes coopérations permettent déjà une première compréhension intuitive de la contribution corticale. On voit que les messages visuels, auditifs, kinestésiques et moteurs seront traités en des endroits différents. On peut imaginer que la région liée au cervelet se spécialisera dans la motricité. De même, des structures comme la formation réticulée ou le corps strié, génératrices de rythmes, vont apporter aux aires frontales une sensibilité au temps et à ses différentes échelles, indispensable à la génération de programmes sensorimoteurs.

Ces premières données jetées en vrac donnent le ton: *le cortex se spécialise en différentes modalités fonctionnelles d'une manière topographique, directement liée à la localisation de ses liens avec les autres structures coopérantes.*

1.3.2 Les sorties du cortex

Les liaisons entre structures nerveuses coopérantes et cortex ne sont pas à sens unique. Si, pour l'instant, l'apport d'informations au cortex par des centres qui effectuent des prétraitements ou qui fournissent une fonction très spécialisée nous semble primordiale, il ne faut pas négliger la réciproque dont on verra l'utilité par la suite.

En effet, le cortex, lui aussi, envoie des informations. Il vient tout de suite à l'esprit les commandes envoyées par le cortex, via les motoneurones, aux divers effecteurs qui nous permettent de nous déplacer ou d'agir sur le monde extérieur. Mais le cortex fournit aussi d'autres messages dont les conséquences, si elles sont moins visibles, n'en sont pas moins vitales.

Faisons en tout de suite un rapide survol. Par ses connexions réciproques, le cortex peut sélectivement demander aux structures coopérantes d'affiner ou au contraire d'occulter certaines informations. Nous retrouverons cette propriété quand nous parlerons, dans les chapitres suivants, de feedback sur les noyaux du thalamus. Le cortex peut aussi participer à l'apprentissage de ces structures, c'est-à-dire fournir, de manière consciente, des réponses mémorisables par la structure qui, une fois mûre, lui

servira de "mémoire de masse" de manière inconsciente. Cette propriété est illustrée, dans les deux premières années de la vie d'un petit homme, par l'apprentissage de son cervelet qui lui permet ensuite, de manière inconsciente, de marcher, de jouer au tennis ou de devenir champion de ski...

Enfin, juste retour des choses, le cortex informe ses informateurs trop spécialisés. Par exemple, le corps strié et le cervelet ont besoin de connaître le contexte environnemental pour réaliser leurs fonctions. C'est le cortex, bien informé de ce contexte par le thalamus, qui le leur fournit.

1.3.3 La communication interne dans le cortex

Cette dernière mention aux relations réciproques cortex-structures coopérantes nous amène tout naturellement à préciser un autre point dans le tableau cortical. Nous venons de voir que le cortex, lié au corps strié par sa partie frontale, peut l'informer de données recueillies en parties occipitales ou temporales. C'est donc qu'il n'y a pas de cloisons imperméables entre les différentes fonctionnalités du cortex.

C'est vrai, une grande partie de la communication du cortex est interne. Qui plus est, toutes les parties du cortex y participent et certaines de manière exclusive. Nous nous en rendons ici compte: tout un pan de l'ensemble des propriétés du cortex a, pour l'instant, été laissé dans l'ombre. Nous avons tout d'abord voulu considérer le cortex comme un tout, afin de mieux le replacer dans le contexte du système nerveux et d'en dégager une première esquisse architecturale et fonctionnelle. Cette approche globale et parfois globalisante ayant, nous l'espérons, porté ses fruits, il est temps qu'elle cède la place à une analyse plus fine et détaillée des propriétés anatomiques et fonctionnelles du cortex. Pour cela, nous allons faire un voyage en trois étapes dans les profondeurs du cortex.

1.4 Trois niveaux d'organisation du cortex

1.4.1 Le neurone

Le neurone est l'élément anatomique de base du système nerveux, selon un niveau de description classique. Avant de parler de la modélisation de cette cellule, nous en ferons une description biologique, afin d'avoir toujours à l'esprit le contexte de notre étude. Nous pourrons alors passer à

l'informatique et décrire les propriétés des principales approches connexionnistes. Puis, après une synthèse sur l'intérêt de ces modèles, nous pourrions présenter le nôtre.

1.4.2 Les aires

A un niveau macroscopique, nous allons traiter du réseau entre les aires corticales, à l'intérieur du cortex. Il ne va pas s'agir d'une simple énumération de régions corticales ayant telle ou telle propriété, mais d'un modèle logique et fonctionnel de la répartition des aires, de leur contributions réciproques et, nous le verrons, des propriétés isomorphiques qui s'en dégagent et qui offrent un support aux fonctions d'adaptation que nous allons traiter.

1.4.3 La colonne corticale

A un niveau de description intermédiaire entre le neurone et les aires, la colonne corticale est l'élément fonctionnel de base du cortex. Nous allons proposer une modélisation de cette unité fondée sur des observations biologiques structurelles (architecture et connectivité). L'aspect dynamique de l'unité (fonctionnement et mémorisation) sera abordé au travers d'une interprétation cognitive des données neurobiologiques.

2 Le neurone

De nombreuses modélisations des fonctions du cerveau humain ont été tentées à partir de la définition des neurones formels, afin de reproduire les structures de base du cerveau (Shaw & Palm, 1988).

Un premier échec de cette technique dans les années 60 (Minsky & Papert, 1969; Rosenblatt, 1962) a fait préférer l'approche symbolique de l'intelligence artificielle, qui consiste à modéliser la pensée à l'aide d'un ensemble de symboles et de règles logiques. Les systèmes experts en sont notamment le résultat. Cependant, la lenteur de ces techniques séquentielles, l'atomisation des connaissances et des règles de production et la difficulté de définir un apprentissage efficace ont décidé certains chercheurs à réserver l'approche symbolique pour les tâches de haut niveau et à revenir, pour les tâches plus perceptuelles, à l'approche connexionniste qui ne connaît pas ces limitations. Ce retour au connexionnisme a été encouragé par l'amélioration des techniques anciennes et la résolution de problèmes conceptuels.

Cependant, avant de parler des neurones formels et des différents types de réseaux qui les utilisent, il est bon d'étudier les caractéristiques et les propriétés du neurone réel, en tant que cellule vivante spécialisée.

2.1 Le neurone vivant

Le neurone vivant possède de nombreuses propriétés que l'on retrouve chez les autres cellules. Cependant, en tant que cellule spécialisée du système nerveux, le neurone a acquis cinq fonctions qui lui sont propres (Robert, 1982):

1. recevoir, en entrée, des signaux provenant d'autres neurones,
2. intégrer ces entrées,
3. engendrer un influx nerveux,
4. le conduire,
5. le transmettre, en sortie, à d'autres neurones.

Cette spécialisation des fonctions du neurone n'a pu se réaliser qu'avec la création de structures cellulaires particulières: le corps cellulaire, ses prolongements (les dendrites et l'axone), une membrane capable d'engendrer et de propager l'influx nerveux, et des synapses pouvant transmettre l'information de neurone à neurone (Stevens, 1984).

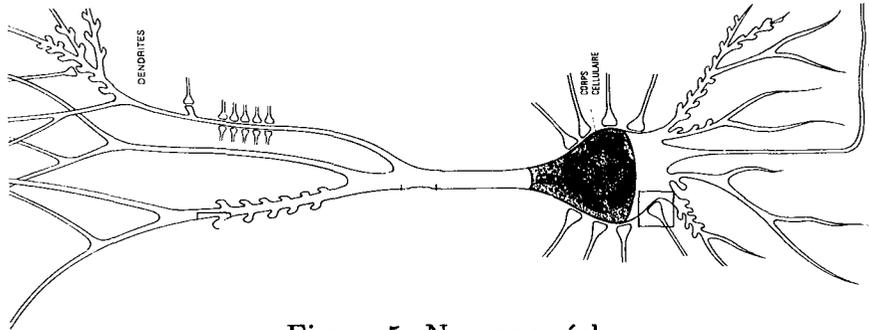


Figure 5. Neurone réel.

2.1.1 Fonctionnement

Sur ces structures, les cinq fonctions spécialisées du neurone sont réalisées de la manière suivante:

Réception

La réception des signaux en provenance d'autres neurones est effectuée par les dendrites. Il s'agit d'un ensemble d'extensions du corps cellulaire extrêmement ramifiées en épines dendritiques, réalisant des points de contact avec les neurones voisins (un neurone peut recevoir des informations de centaines d'autres neurones). Il n'y a pas continuité d'un neurone à l'autre: les points de contact (ou synapses) sont des extrémités spécialisées comprenant un bouton présynaptique, un espace synaptique et une membrane postsynaptique appartenant à un autre neurone.

Lorsque l'influx nerveux arrive dans un bouton présynaptique, un certain nombre de vésicules, renfermant une substance chimique appelée neuromédiateur, est libéré dans l'espace synaptique; ces molécules sont alors captées par la membrane postsynaptique d'une dendrite où va naître un influx nerveux qui se propagera vers le corps cellulaire, si le neuromédiateur est excitateur de ce neurone; sinon, il y aura inhibition. Les dendrites collectent donc l'information, c'est-à-dire recueillent excitation ou inhibi-

tion selon le schéma: influx nerveux électrique-médiation chimique-influx nerveux électrique.

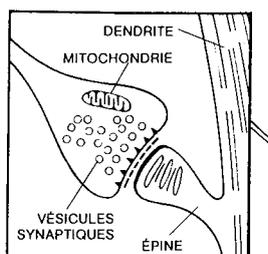


Figure 6. Synapse.

Le message véhiculé par l'influx nerveux se présente sous la forme de potentiels d'action, de fréquence variable mais d'amplitude et d'intensité fixes. Seul le nombre d'influx par seconde varie et code l'information selon le principe de la modulation de fréquence. On le sait, ce codage est insensible au brouillage électromagnétique et assure une haute fidélité du message, ce qui est heureux dans un environnement où les champs électriques peuvent atteindre 100 kV/cm (sans parler de l'environnement humain de plus en plus envahi de phénomènes électromagnétiques de toutes sortes).

Intégrations spatiale et temporelle

L'intégration des entrées en vue d'engendrer un influx nerveux comme réponse à ces entrées est réalisée par le corps cellulaire (ou soma) selon deux modalités.

Il y a d'une part sommation temporelle qui est fonction de la fréquence de l'information reçue. Plus la fréquence présynaptique est grande (et donc plus le nombre de molécules de transmetteur libérées est grand), plus la fréquence du potentiel postsynaptique est élevée.

Il y a d'autre part sommation spatiale. Il s'agit d'un mécanisme par lequel tous les influx provenant d'autres neurones, qu'ils soient inhibiteurs ou excitateurs, s'annulent ou se renforcent afin de donner à la membrane du neurone la réponse représentant la somme de ces événements, par un potentiel "excitateur" (dépolarisé) ou "inhibiteur" (hyperpolarisé).

Conduction, transmission

C'est en effet l'axone, attribut unique pour un neurone, qui achemine la réponse vers les sites concernés, selon la règle du tout ou rien: en dessous

d'un certain seuil de potentiel, l'axone ne produit pas de potentiel d'action; au dessus, il conduit l'information.

Les propriétés de la membrane du neurone lui permettent de conduire l'influx nerveux, c'est-à-dire de propager un déséquilibre ionique, au moyen d'éléments tels que pompes à ions ou canaux ioniques qui règlent la différence de concentration ionique entre les milieux intérieur et extérieur. Ainsi, l'axone peut conduire un influx nerveux vers les zones cibles du neurone sur des distances parfois très longues.

A son extrémité, l'axone se ramifie et chaque ramification "fait synapse" avec d'autres neurones. Un neurone peut ainsi transmettre en sortie son information selon le principe chimico-électrique: de l'autre côté des synapses, les dendrites des neurones cibles attendent les molécules de neurotransmetteurs qui vont créer la réponse postsynaptique.

Variabilités

Voici donc schématisées les principales fonctions caractéristiques du neurone et ses structures qui les réalisent. C'est sur ces bases neurobiologiques que s'effectue la description du neurone formel qui va déboucher sur les réseaux et le connexionnisme. Toutefois, avant de définir "le" neurone formel, revenons une dernière fois sur "les" neurones réels du cortex.

Evidemment, tous les neurones corticaux fonctionnent selon les principes que nous venons d'énoncer. Pourtant, aucun biologiste ne vous parlera "du" neurone cortical tant les différences sont nombreuses entre les variétés de neurones dans le cortex.

Ces différences sont dûes à la taille des neurones, à leur forme, à leurs positions respectives dans les différentes couches qu'ils forment dans le cortex, au genre inhibiteur ou excitateur de leurs connexions...

Les caractéristiques fondamentalement différentes des principales classes de neurones du cortex illustrent cette variété (Szentagothai, 1973): les neurones pyramidaux sont de grands neurones excitateurs avec un grand axone et sont connectés avec d'autres neurones pyramidaux (dans les couches supérieures) ou avec les structures coopérantes (dans les couches inférieures). Les interneurones sont des neurones se projetant à courte distance. Ils peuvent être excitateurs (étoilés, bipolaires), ou inhibiteurs (en panier, en chandelier, à double bouquet). Ils peuvent agir latéralement (en panier) ou transversalement (en chandelier, étoilé, bipolaire, à double bouquet).

Ainsi, une fois de plus, la biologie nous propose des modèles mais modère nos tendances généralisatrices en nous offrant une diversité de réalisations

que l'on doit garder à l'esprit si l'on ne veut pas avoir une vision trop simpliste des choses.

Il n'en reste pas moins que cette collection de neurones possède toutes les propriétés que l'on a vues plus haut et que, pour cette raison, ils ne renieraient pas leur petit frère artificiel: le neurone formel.

2.2 Le neurone formel

Précisons tout de suite que, comme son nom l'indique, le neurone formel n'est pas une simulation, mais une modélisation ou une formalisation du neurone réel (Davalo & Naim, 1989; Shaw & Palm, 1988)

Les différents types de neurones formels que l'on va définir par la suite vont dépendre, pour un système donné, de plusieurs paramètres.

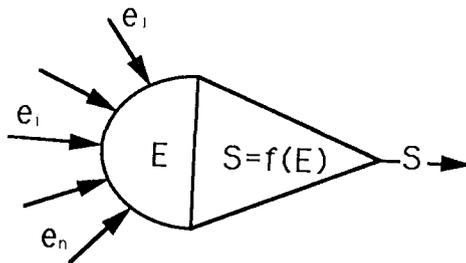


Figure 7. Neurone formel.

2.2.1 Evaluation des entrées

Ce calcul est effectué sur les entrées $(e_i)_{i=1,n}$, qui peuvent être binaires ou réelles. Il correspond aux sommations spatiales et temporelles et fournit l'entrée totale E .

E peut être une fonction booléenne des entrées (e_i) . Sinon, on utilise des paramètres de pondération W_i permettant à l'entrée totale d'être une fonction linéaire $E=\sum W_i e_i$ ou affine $E=\sum W_i e_i - a$.

2.2.2 Evaluation de la sortie

La fonction d'activation f définit l'état du neurone et donc sa sortie, en fonction de son entrée totale E . Cette fonction correspond au seuil de conduc-

tion de l'axone. On la représente généralement par une fonction impaire et croissante de l'axone pouvant être:

- une fonction binaire à seuil

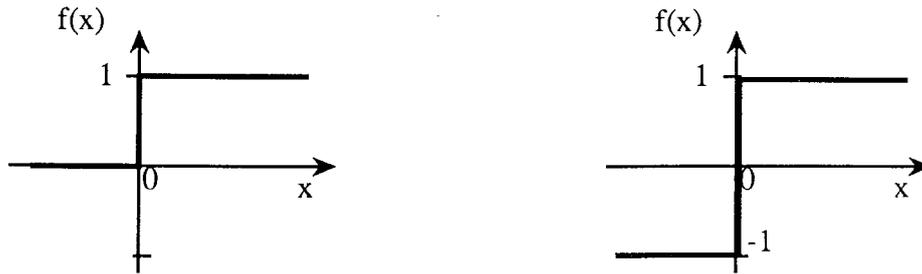


Figure 8. Fonctions à seuil (Heaviside, signe).

Remarque: la fonction de Heaviside n'est pas impaire.

- une fonction linéaire à seuil(s) comprenant des domaines linéaires et des domaines seuillés

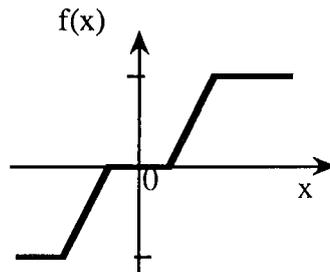


Figure 9. Fonction linéaire à seuils.

- une fonction stochastique:

$f(x)=1$ avec la probabilité $1/(1 + e^{-x/T})$

$f(x)=0$ sinon.

- une fonction sigmoïde:

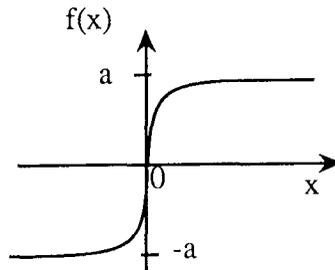


Figure 10. Fonction sigmoïde.

$$f(x) = a(e^{kx} - 1)/(e^{kx} + 1)$$

Ce modèle général du neurone formel nous permet maintenant de définir simplement le premier modèle de neurone mis au point il y a plus de 40 ans par MacCulloch et Pitts (MacCulloch & Pitts, 1943): il s'agit d'un automate booléen, avec des entrées et des sorties binaires. La fonction sur les entrées est linéaire et la fonction d'activation à seuil est représentée par:

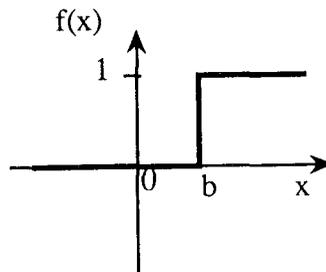


Figure 11. Fonction de MacCulloch & Pitts.

Ce premier modèle de neurone est encore couramment utilisé aujourd'hui dans plusieurs types de réseaux. Toutefois, avant de décrire différents modèles de réseaux, nous devons tout d'abord spécifier d'autres paramètres qui, en plus de la description interne du neurone formel et de ses caractéristiques propres, permettent de définir un réseau (Rumelhart & MacClelland, 1986).

2.2.3 Les règles de fonctionnement

Ces règles dépendent du type de neurone formel employé et définissent l'état initial des neurones ainsi que l'ordre de leur évaluation: l'état initial peut être aléatoire, forcé (défini par l'extérieur) ou mixte. L'ordre d'évaluation des neurones peut être parallèle: à chaque cycle, tous les neurones sont évalués d'après les paramètres du cycle précédent. Le mode d'évaluation peut être séquentiel: à chaque cycle, un seul neurone est activé. Le neurone peut être choisi aléatoirement ou à partir d'un ordre prédéfini (cette technique permet de définir un parallélisme asynchrone). Il est enfin possible de choisir une combinaison de ces deux modes de fonctionnement.

2.2.4 Les règles de propagation

Ces règles sont déterminées par la structure et la nature des connexions: le réseau peut être structuré en couches avec une liaison totale de couche à couche et avec ou sans liaison à l'intérieur d'une même couche (sans liaison le plus souvent). Le nombre de couches est variable. Le réseau peut être entièrement connecté (chaque neurone est lié à tous les neurones du réseau, y compris lui-même). Enfin, ces connexions peuvent être monodirectionnelles ou bidirectionnelles.

2.2.5 Les règles d'apprentissage

Dans tous les réseaux, l'apprentissage s'effectue en modifiant les poids W_i des connexions entre les neurones et la mémorisation dépend, à un instant donné, de la force de ces liens. Cette modification est réalisée selon deux principes d'inspiration différente.

1. Inspiration biologique:

Il s'agit de la règle de Hebb (Hebb, 1949) qui a proposé, il y a une quarantaine d'années, le mécanisme de mémorisation suivant: un neurone ne mémorise une association que si les signaux émis par une source présynaptique se produisent en même temps qu'une décharge dans l'élément postsynaptique. Dans un réseau de neurones formels, cette règle est appliquée en renforçant la connexion de deux neurones activés au même moment.

2. Inspiration mathématique:

Il s'agit alors de considérer que le réseau effectue une fonction de transfert sur les entrées pour calculer les sorties. Les paramètres de cette fonction, c'est-à-dire les poids des connexions, sont modifiés en terme de minimisation de fonctions de coûts, définies entre la sortie réelle et la sortie désirée du réseau. Il s'agit d'un apprentissage supervisé.

2.2.6 Environnement

Il est généralement important de connaître la destination du réseau, à savoir la ou les fonctions qu'il est censé remplir. Classiquement, les réseaux sont utilisés comme mémoires associatives. Leur rôle peut être de reconstruire une information bruitée (auto-association) ou d'associer deux types d'informations différents (hétéro-association). Ceci permet de déboucher sur des applications de reconnaissance des formes, de traitement du signal, de classification, de traitement d'images... Un autre domaine d'action des réseaux est la satisfaction de contraintes avec des applications à l'optimisation, l'aide à la décision, la prévision...

Ces caractéristiques générales et leurs différentes expressions étant posées, elles vont nous permettre de manière beaucoup plus simple de décrire les principales classes de réseaux de neurones couramment utilisées.

Nous ne pouvons pas parler ici de manière exhaustive de toutes les architectures connexionnistes, proposées dans une littérature très abondante et souvent utilisées pour des applications très ponctuelles. Il existe cependant des modèles "de base" (Bienestock, 1986; Kohonen, 1988b) que nous avons choisi de regrouper selon trois thèmes.

2.3 Les réseaux à couches

Une structuration en couches tend à rendre compte des différents traitements que l'on peut effectuer en cascade sur un ensemble d'informations. Ces informations sont proposées sur une couche extrême, appelée couche d'entrée; elles sont ensuite traitées par un nombre variable de couches intermédiaires ou couches cachées. Le résultat apparaît sur l'autre couche extrême, la couche de sortie.

Généralement, ces réseaux sont utilisés pour des tâches d'association (hétéro-association dans la majorité des cas) et sont, pour cette raison, souvent appelés réseaux associatifs. Etudions en quelques uns.

2.3.1 Le perceptron

Historiquement, le perceptron (Rosenblatt, 1962) est le premier réseau "sérieux" à avoir été proposé.

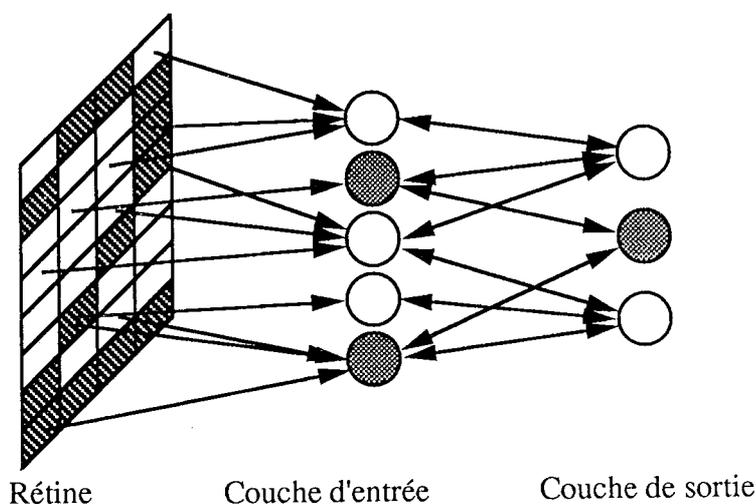


Figure 12. Perceptron.

Il comprend une couche d'entrée et une couche de sortie. Les connexions entre ces deux couches sont modifiables et bidirectionnelles. Les unités ont des sorties binaires. Les unités de la couche de sortie réalisent une fonction à seuil. C'est la couche de sortie qui remplit la tâche de classification. Les unités sont évaluées parallèlement. Les connexions sont modifiées par un apprentissage supervisé selon le principe de correction d'erreurs: si une unité de la couche de sortie n'est pas dans l'état désiré, toutes les liaisons que cette unité réalise avec les unités de la couche d'entrée sont augmentées ou diminuées selon le type d'action que l'unité de la couche d'entrée réalisait sur l'unité de sortie considérée, dans un sens permettant de favoriser une

réponse exacte. Analytiquement, cette modification potentielle de poids synaptique peut être décrite par:

$$W_i = W_i + k(d - s)e_i$$

avec d: sortie désirée, s: sortie réelle, et k: réel positif représentant le pas de modification des poids.

De par son rôle d'unité linéaire à seuil, chaque unité de la couche de sortie réalise une partition de l'espace de ses entrées en deux classes séparées par l'hyperplan d'équation

$$\Sigma_i W_i e_i - \text{seuil} = 0$$

Ainsi, n unités de sortie permettent de séparer linéairement 2^n classes.

2.3.2 Les autres réseaux monocouches

On appelle en fait réseaux monocouches les réseaux dont seuls les liens entre les deux dernières couches sont modifiables. Le perceptron en est un exemple. Il a été employé pour apprendre à réaliser une classification à partir d'exemples qui lui sont proposés et qui lui permettent de converger, si le problème est linéairement séparable. La règle de Widrow-Hoff (Widrow & Hoff, 1960) est une amélioration qui permet de réaliser un apprentissage plus robuste. En effet, on l'a vu, la modification des poids dans le perceptron a lieu tant que la sortie réelle n'est pas identique à la sortie désirée. Autrement dit, dès que la somme pondérée seuillée est égale à la sortie désirée, il n'y a plus modification des poids. Widrow et Hoff proposent de modifier les poids tant que la somme pondérée **non seuillée** est différente du résultat attendu. Ceci correspond à choisir un hyperplan qui ne soit pas trop proche d'une des deux classes.

Cette amélioration ne permet pas aux réseaux monocouches de surmonter leur limitation principale qui est de ne résoudre que les problèmes linéairement séparables (on cite classiquement l'exemple du "ou exclusif" qui ne peut pas être réalisé par un réseau monocouche).

Cette limitation pourrait être surmontée par l'utilisation de réseaux multicouches comportant alors plusieurs couches modifiables. Mais dans ce cas, l'apprentissage tel qu'on l'a vu pour le perceptron n'est plus efficace: en effet, quand on connaît la sortie réelle et la sortie désirée, dans

quel sens modifier un poids d'une couche lointaine? Autrement dit, comment connaître l'impact de la modification d'un poids d'une couche donnée dans une couche qui n'est pas voisine? Ainsi, tant qu'il n'a pas été possible de répercuter un signal d'erreur mesuré sur la couche de sortie sur plusieurs couches successives, les réseaux multicouches étaient inutilisables. Heureusement, les années 80 ont vu la naissance de l'algorithme de rétropropagation du gradient...

2.3.3 Les réseaux multicouches

L'unité choisie est du même genre que pour le perceptron. Seule différence, la fonction à seuil cède la place à une fonction sigmoïde de la forme:

$$f(x) = k(e^{kx} - 1)/(e^{kx} + 1)$$

Or, les fonctions sigmoïdes sont dérivables...

L'apprentissage va consister, pour minimiser l'erreur commise, à réaliser une descente du gradient, c'est-à-dire à calculer la dérivée de l'erreur totale commise par rapport à un poids particulier. Un neurone possédant une fonction d'activation dérivable (comme les sigmoïdes) permet de réaliser simplement cette opération (LeCun, 1987; Fogelman Soulié, 1987).

La mise au point de cet algorithme a entraîné une utilisation intensive des réseaux multicouches dans des domaines très variés (Kienker, Sejnowski, Hinton, Schumacher, 1986; Bourlard & Wellekens, 1987; Rumelhart & MacClelland, 1986; Sejnowski & Rosenberg, 1986;...). Il n'en reste pas moins que cette technique possède des inconvénients et des limitations pas encore surmontés. Il n'a pas encore été possible de déterminer automatiquement un certain nombre de paramètres fondamentaux pour un réseau multicouche.

Tout d'abord, sa taille: le nombre et la taille des couches cachées n'ont pas été reliés au type de problème à traiter (nombre de classes, nombre d'exemples...) et doivent être "devinés" à chaque fois par l'utilisateur.

Ensuite, le pas du gradient, c'est-à-dire la valeur de la modification apporté à chaque opération, est aussi à définir "à la main", avec le risque de voir le réseau osciller si le pas est trop grand ou de ne pas avoir la patience d'attendre la convergence s'il est trop petit.

Ceci nous amène à l'inconvénient majeur de l'algorithme de rétropropagation du gradient: le temps de calcul est souvent prohibitif et ne peut pas

être prévu à l'avance. De nombreuses études ont été faites pour accélérer cet algorithme qui demeure néanmoins un des plus répandus et utilisés dans le domaine du connexionnisme.

2.4 Les réseaux à satisfaction de contrainte

Les deux modèles que nous présentons ici diffèrent des réseaux à couches, par leur domaine d'inspiration aussi bien que par leur architecture et leur fonctionnement. En effet, si les réseaux à couches ont été d'inspiration biologique de par leur structure qui schématise les différentes couches de traitement du cerveau, puis mathématique par des algorithmes tels que celui de la rétropropagation du gradient, les réseaux que nous décrivons ici trouvent leurs origines dans les sciences physiques.

2.4.1 Le modèle de Hopfield

Ce modèle a été proposé par un physicien (Hopfield, 1982), par analogie avec la théorie des verres de spin. D'une part, ceci a permis d'intéresser beaucoup de physiciens au connexionnisme, d'autre part, les liens de ce modèle avec la physique statistique lui ont apporté toute la puissance de la théorie et des outils de calcul de ce domaine, ce qui a permis d'étudier en profondeur ses capacités.

Tout d'abord, l'unité utilisée est le neurone de Mac Culloch et Pitts. A la différence des réseaux à couches, le réseau de Hopfield n'a pas de structure. Ses neurones sont totalement interconnectés (y compris à eux-mêmes) par des liens bidirectionnels et symétriques. Il n'y a donc pas à proprement parler d'entrée et de sortie: l'état du système à un moment donné correspond à l'état de tous les neurones. Un réseau de Hopfield peut, en fait, être considéré comme une mémoire adressable par son contenu. En effet, ce réseau fonctionne par auto-association: on propose aux neurones différents états qui sont mémorisés selon la règle de Hebb (le fonctionnement est généralement séquentiel avec tirage aléatoire). Les états mémorisés au cours de la phase d'apprentissage deviennent ainsi des états attracteurs. La phase de reconnaissance consiste à proposer un état bruité et à évaluer le réseau jusqu'à ce qu'il se stabilise, c'est-à-dire jusqu'à ce que l'état proposé, attiré par les bassins d'attractions créés par apprentissage, s'identifie à un des états appris. Ainsi, à la différence de l'apprentissage qui demande une

seule modification des liens par élément, la reconnaissance est dynamique puisque le réseau effectue des itérations jusqu'à stabilisation.

C'est cette évolution du réseau vers un état stable, selon une fonction d'énergie qui décroît vers un minimum local, qui a permis à Hopfield l'analogie avec les verres de spin. Ceux-ci possèdent en effet un ensemble d'états stables correspondant à des états d'énergie minimale. Le retournement d'une partie des spins fait soudain passer le système à un autre état de basse énergie (Stein, 89).

A l'aide de ces outils théoriques, des études intensives sur ce modèle ont précisé ses capacités aussi bien que ses limites.

Tout d'abord, les capacités de mémorisation d'un tel réseau ont été évaluées: un réseau de N neurones peut mémoriser $0,144N$ états stables, sous réserve d'une certaine "distance" entre ces états, le cas optimal correspondant à des états orthogonaux. Au delà de cette limite, survient ce qu'on appelle une catastrophe et qui en est réellement une, puisque le réseau peut oublier, à ce moment là, tout ce qu'il avait appris auparavant...

De plus, la règle de Hebb est une règle purement locale et peut introduire sans le savoir, des états stables parasites (correspondant à des minima locaux formés par la création de plusieurs bassins d'attraction). Il existe de nombreuses études visant à élargir les bassins d'attraction des états stables et à éviter une trop grande attraction de ces états parasites (Krauth & Mézard, 1987).

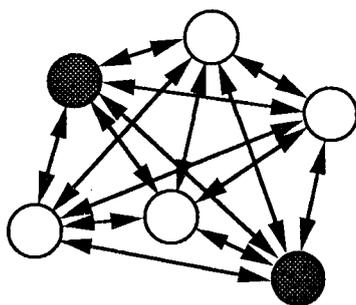


Figure 13. Réseau de Hopfield.

Tous ces éléments font du réseau d'Hopfield un modèle très simple et intéressant dans ses limites d'utilisation. Il est à noter que son fonctionnement, théoriquement limité à l'auto-association, peut être élargi avec un peu d'astuce: l'hétéroassociation est possible si l'apprentissage consiste à lui

soumettre des couples d'éléments et la reconnaissance à lui faire compléter des couples où il manque un élément. De plus, ce modèle est aussi largement utilisé dans des problèmes d'optimisation comme celui du voyageur de commerce, en lui présentant de manière astucieuse une représentation du parcours et en fixant les connexions de manière à lui faire minimiser différentes fonctions d'énergie (Hopfield & Tank, 1985).

2.4.2 La machine de Boltzmann

Il n'en reste pas moins qu'un des inconvénients majeurs des réseaux de Hopfield, ainsi que des réseaux multicouches, est la stabilisation anticipée du réseau dans un minimum local parasite.

Ce problème peut être résolu par l'algorithme du recuit simulé. Dans un réseau de Hopfield, un neurone évalue son nouvel état en faisant la somme pondérée de ses entrées. Si cette somme dépasse un seuil donné, l'état sera 1 et 0 sinon.

Selon l'algorithme du recuit simulé, les neurones fonctionnent selon une règle stochastique et effectuent leur changement d'états par tirage au sort selon les probabilités suivantes:

Soit d la différence entre la somme pondérée des entrées et le seuil. Alors la probabilité que le neurone soit dans l'état 1 vaut:

$$1/(1 + e^{-d/T})$$

et dans l'état 0:

$$e^{-d/T}/(1 + e^{-d/T})$$

On remarque que, quand T tend vers 0, ce neurone stochastique se comporte comme un neurone de Mac Culloch et Pitts.

En fait, T correspond à une température et l'algorithme du recuit simulé utilise, comme son nom l'indique, une analogie thermodynamique. En cristallographie, le recuit consiste à réaliser un cristal très pur en le créant à haute température et en le faisant ensuite progressivement refroidir. Les molécules ont une possibilité de mouvement qui décroît rapidement avec la température. Ainsi obtient-on une meilleure organisation du cristal qu'en le figeant trop rapidement. L'analogie avec les réseaux de neurones est immédiate: plutôt que de figer immédiatement l'état du réseau dans ce qui peut

être un minimum local parasite, on commence par évaluer le réseau avec le paramètre T grand. Ceci permet aux neurones d'effectuer d'importants sauts d'énergie qui peuvent sortir le réseau d'états parasites. Puis le réseau est réévalué en diminuant T progressivement. $T=0$ permet de retrouver la configuration réelle du paysage d'énergie et le réseau dans un état stable différent d'un minimum local.

Cet algorithme peut être utilisé dans un réseau de Hopfield. Dans ce cas, la trop grande localité de la règle de Hebb est compensée par le paramètre T , qui rend compte de l'agitation globale du réseau, et va permettre, par sa diminution progressive, de stabiliser le réseau au minimum global, avec une bonne probabilité d'éviter les minima locaux.

La machine de Boltzmann consiste en une version étendue du réseau de Hopfield avec recuit simulé et avec unités cachées. Il s'agit d'un réseau de Hopfield comprenant N neurones dont certains sont cachés et ont pour rôle de faire des corrélations d'ordre supérieur à 2 sur les neurones visibles. Ce réseau fonctionne selon deux modes. Un mode forcé consiste à forcer un exemple sur les unités visibles et à laisser les unités cachées évoluer selon l'algorithme du recuit simulé. Un mode libre permet à tout le réseau d'évoluer selon cet algorithme et de mesurer les distributions de probabilités entre les unités afin de modifier les poids et de faire tendre cette distribution vers une représentation exacte des exemples appris: $1/n$ pour chacun des n exemples appris et 0 pour les autres configurations.

On le voit, cet algorithme est relativement complexe et ne peut être décrit en quelques lignes. Le lecteur intéressé pourra se reporter à (Hinton, Sejnowski, Ackley, 1984; Azencott, 1989). Précisons toutefois que, par sa lourdeur, la puissance de calcul et le temps qu'il demande, cet algorithme n'est guère employé aujourd'hui, en attendant des circuits intégrés spécifiques.

2.5 Les autres modèles

Il ne s'agit pas ici d'une partie où l'on cite, pour mémoire, des travaux marginaux, mais d'un ensemble de modèles très intéressants qui correspondent, à notre avis, à une voie prometteuse du connexionnisme, débarassé de considérations purement mathématiques ou physiques. Le point commun de ces approches est leur plus grande plausibilité biologique, ce qui ne les empêche pas de "tenir la route" mathématiquement.

2.5.1 Le modèle de Kohonen

Kohonen est l'exemple typique du chercheur qui sait extraire, pour un problème donné, ses fondements biologiques et ses contraintes techniques tout en proposant une formulation mathématique solide.

Ainsi, son neurone formel, en plus des caractéristiques classiques déjà vues, comprend un terme de pertes non linéaires correspondant au phénomène de saturation du neurone, ce qui implique des bornes à ses états et à sa rapidité de variation (Kohonen, 1984a). De même, pour l'apprentissage, la règle de Hebb est modifiée afin de tenir compte des phénomènes d'oubli (il suffit de diminuer le poids des connexions qui ne servent pas). Par ailleurs, ces réseaux sont structurés selon le principe des cartes topologiques que l'on trouve dans le cortex (Kohonen, 1984b). Ainsi, des neurones voisins codent des phénomènes voisins. Par rapport au monde extérieur, ces phénomènes correspondent à des stimuli spécifiques dont on a pu observer le codage sélectif dans le cortex: des traits orientés pour une application visuelle, des fréquences particulières pour le cas auditif. En ce qui concerne les liens entre neurones, ceux-ci ne sont pas uniformes: des neurones voisins dans une même carte auront des liens respectant le principe d'inhibition latérale, des neurones de différentes couches ont des liens correspondant au traitement qu'ils effectuent (Kohonen, 1988a). On le voit, il est difficile de parler d'un modèle de Kohonen, puisque architecture et fonctionnement vont dépendre de l'application réalisée (il en est de même pour le système nerveux). Il suffit donc de se souvenir que les travaux de Kohonen correspondent à une étude sérieuse du problème considéré et à la création de fonctions connexionnistes ad hoc permettant de spécifier un réseau pouvant traiter ce problème.

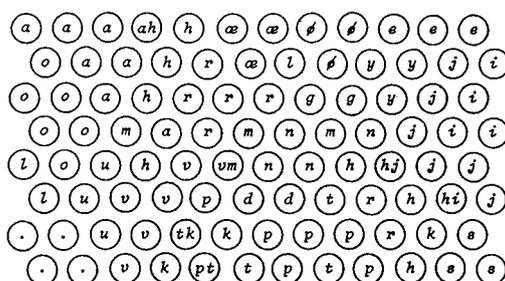


Figure 14. Carte topologique (d'après Kohonen, 1984b).

2.5.2 D'autres modèles

Enfin, parmi les dizaines de réseaux proposés, il en existe quelques uns qui ne se prétendent pas aussi généraux que les références "rétropropagation" ou "Hopfield". mais qui ont néanmoins marqué la communauté par les réalisations qu'ils présentaient. Parmi ceux-ci, citons le néocognitron de Fukushima (Fukushima, 1980), système hétéroassociatif dédié à la reconnaissance de formes et très largement inspiré des mécanismes de traitements visuels décrit par les neurobiologistes.

Citons aussi le très intéressant modèle de mémoire associative dynamique et adaptative développé au LIMSI par D. Béroule (Béroule, 1985). Schématiquement, ce modèle consiste à définir des unités qui créent des chemins représentant différents signaux. Ces unités sont rangées hiérarchiquement en couches, ce qui autorise l'intégration consécutive de séquences. Ce modèle est initialement proposé pour le traitement de la parole, mais peut traiter d'autres données ayant une dimension temporelle.

2.6 Discussion

Ce tour d'horizon des principales classes de modèles connexionnistes actuellement développés, nous a permis de prendre la mesure de leurs forces et de leurs faiblesses.

Une première constatation s'impose: les réseaux possèdent une indéniable capacité d'apprentissage et de mémorisation, ce qui leur donne un atout certain. De plus, ces réseaux peuvent être employés pour de nombreuses applications telles que la reconnaissance de formes, le traitement du signal, la satisfaction de contraintes et l'aide à la décision, qui utilisent les différentes propriétés des réseaux (réseaux associatifs, mémoires adressables, réseaux symétriques...).

Cette étude nous a également permis de dégager les trois principaux domaines d'inspiration du connexionnisme, leurs avantages et leurs inconvénients respectifs.

Chronologiquement, le premier est naturellement la biologie. Il est en effet naturel de s'inspirer des vrais neurones pour les modéliser. La biologie fournit une masse considérable de données et de mécanismes qui sont de mieux en mieux connus. En retour, le biologiste est prêt à discuter des modélisations et à étudier les implications de telle ou telle hypothèse. Ce-

pendant, de chaque côté, le langage et les attentes sont souvent différents; la formalisation choque souvent le biologiste qui travaille sur l'exemple, tout comme une série de données contradictoires plonge le modélisateur dans l'embarras. En conséquence, les modèles biologiques ont souvent un domaine d'application très restreint ou peuvent vite ressembler à du "bricolage" consistant à mettre bout à bout une série de mécanismes biologiques.

Ces raisons ont poussé les chercheurs à élaborer une théorie unifiée avec l'aide des mathématiques. Le grand succès de cette approche a très certainement été l'algorithme de rétropropagation du gradient qui a résolu un problème fondamental. De plus, cette technique permet aujourd'hui d'aborder et de réaliser de nombreuses applications, sans longues études préalables, grâce à des programmes "clés en main". Cependant cette approche mathématique ne sait pas maîtriser les propriétés émergentes du système: la taille et le nombre des couches cachées ne sont pas directement reliés à la complexité du problème. Le pas du gradient nécessite encore un choix empirique. La réduction du temps d'apprentissage fournit également un sujet de recherche supplémentaire.

Parallèlement à ces deux techniques, l'approche physique a suscité de nombreux espoirs. En effet, le formidable arsenal d'outils de la physique statistique permet une bonne maîtrise du réseau et de ses propriétés émergentes. Cependant, la formulation théorique d'un problème n'implique pas sa résolution. On se trouve alors partagé entre le piège du minimum local et la longueur de la méthode du recuit simulé.

Enfin, beaucoup d'observateurs de tous horizons font grief au connexionnisme en général, dont ils attendaient qualitativement plus: dans les trois approches, les capacités d'abstraction se sont révélées limitées et la représentation du temps dans la dynamique des réseaux en est à ses premiers balbutiements (Gori, Bengio, DeMori, 1989; Waibel, Hanazawa, Hinton, Shikano, Lang, 1989).

Où en est la recherche aujourd'hui? Chaque approche a ses détracteurs et ses inconditionnels qui font de nombreuses recherches parallèles mais pas forcément convergentes. D'autres attendent les progrès de la technologie qui devraient bientôt nous fournir une machine de complexité comparable au cerveau humain. Il reste cependant à trouver un modèle de fonctionnement pour ces neurones.

Notre approche est la suivante: un retour aux sources de l'inspiration biologique nous a fait rencontrer Y. Burnod. Nous avons été séduits par son

modèle biologique théorique qui essaie de dépasser les propriétés locales des neurones, pour déboucher sur des fonctions corticales plus globales. Cependant, les premières applications nous ont vite fait renoncer au "tout biologique". L'étude conjointe des autres domaines se révèle très fructueuse: les sciences exactes peuvent se montrer très utiles pour comprendre et généraliser un modèle. Nous n'avons pas parlé de psychologie dans cette partie car son approche est plus macroscopique que le connexionnisme. Cependant, de nombreux modèles psychologiques existent et peuvent permettre une étude des propriétés émergentes d'un modèle connexionniste. De même, l'approche symbolique, par sa connaissance des modèles de raisonnement, peut se révéler une source d'inspiration indispensable.

Ainsi, nous ne voyons aucune raison de ne pas travailler de concert avec les autres chercheurs du domaine cognitif et des neurosciences qui ont, après tout, une quête semblable à la nôtre: comprendre un système connexionniste capable de réaliser des fonctions humaines.

3 Les aires

Nous nous intéressons maintenant à un premier réseau pouvant modéliser le cortex. Nous le verrons, il en existe d'autres, à des niveaux de description différents, qui amèneront leur lot de contraintes et de propriétés à la mesure de leur échelle. Ce premier réseau, pour le qualifier, reste très global par le type des fonctions qu'il va émuler. Ces fonctions, par leur nature et leurs arrangements relatifs, ne manqueront pas de nous faire penser à une approche de psychologie cognitive. Pourtant, c'est encore de neurobiologie et d'anatomie qu'il va être question pour commencer.

3.1 Les types d'aires

Nous l'avons vu, dès 1909, Brodmann, par des études histologiques et anatomiques, découpe un hémisphère cortical en une cinquantaine de régions qu'il numérote. Cette numérotation sert encore aujourd'hui à la dénomination de ces régions, mais nous ne l'utiliserons pas ici. En effet, Y. Burnod a repris cette étude dans son travail de modélisation du cortex et a transformé cette cinquantaine de régions disparates en un ensemble de 24 aires fonctionnellement homogènes. Cette homogénéité permet une nomenclature plus fonctionnelle et logique qu'un simple numéro. Nous allons donc l'utiliser.

A l'aide d'un schéma d'un hémisphère cérébral permettant la localisation de ces aires (Figure 15), nous allons procéder à une énumération de ces aires par catégories. Elles sont définies par le type d'information que reçoivent les aires puisque, il est important de le répéter, la fonction des régions corticales n'est pas définie d'une manière quelconque, mais imposée par la nature de l'information qui l'alimente. (De là à dire que c'est l'organe qui crée la fonction...).

Ce typage par les données nous permet de diviser l'ensemble des aires en trois groupes:

3.1.1 Les aires sensorimotrices

Ces aires reçoivent les signaux sensoriels du monde extérieur ou envoient les ordres moteurs aux effecteurs. Elles sont donc situées aux deux extrémités de la chaîne stimuli-traitement-réponse (S-h-R). Ces aires reçoivent leurs

informations de façon somatotopique relativement à différentes parties du corps (rétine, cochlée, muscles, peau...).

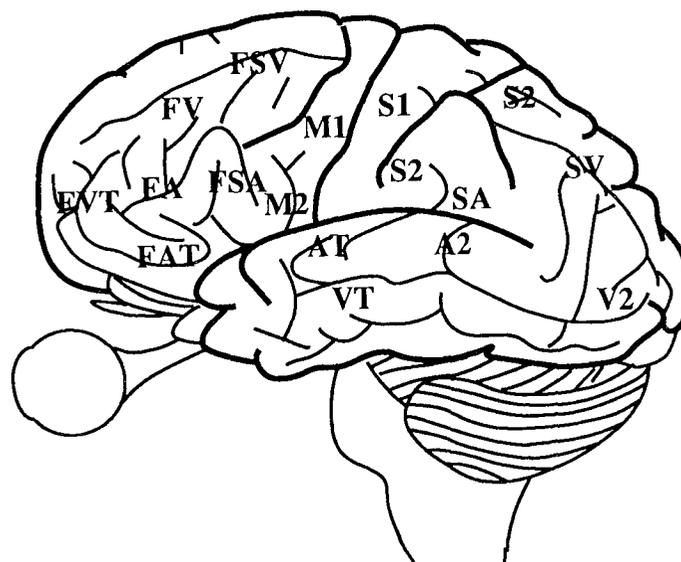


Figure 15. Les aires corticales.

Ainsi, parmi les aires sensorielles, V2 reçoit l'information Visuelle, via le thalamus, de façon rétinotopique (c'est à dire que la rétine se projette sur le cortex visuel en conservant la propriété de voisinage). De même, à partir de la cochlée, A2 recueille les informations Auditives selon une disposition tonotopique. Selon le même principe, S2 offre une représentation Somatotopique de la peau et de la position des muscles tandis que, parallèlement, l'aire motrice M2 constitue une base Motrice des différents mouvements possibles. Le principe de somatotopie est bien illustré par la Figure 16, représentant la répartition des afférences de S2. Pour rendre la nomenclature plus claire, précisons que ces indices "2" renvoient à des aires de traitement dites secondaires, par opposition aux aires primaires qui, comme nous le verrons dans les chapitres suivants, effectuent un prétraitement de mise en forme de l'information. Pour rester fidèle au modèle de Burnod, relevons aussi l'existence des aires nommées SH, TH et O. Ces aires se font le reflet du monde intérieur, c'est à dire plus hormonal que sensoriel (au sens où on l'entend généralement) et correspondent respectivement à la réception des informations venant de l'hypothalamus, de l'hippocampe et des capteurs

olfactifs. Le lecteur comprendra aisément que l'on passe sous silence, pour l'instant, de telles fonctions dans un travail d'Intelligence Artificielle.

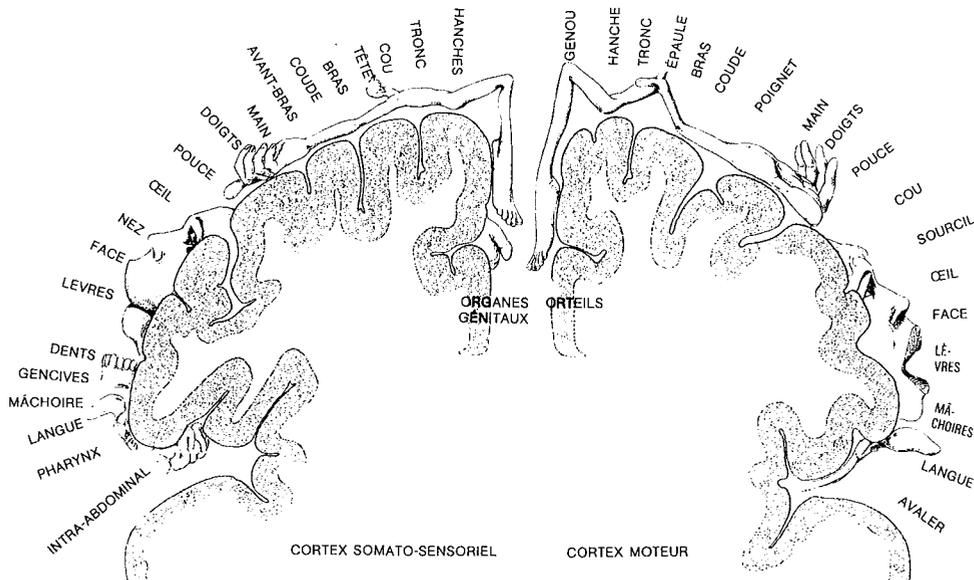


Figure 16. L'homonculus (d'après Geshwind, 1984).

3.1.2 Les aires associatives

On l'a vu précédemment, les aires, typées par la dimension du monde extérieur qu'elles perçoivent, ne sont pas cloisonnées. De plus, certaines ont pour fonction, de par leurs connexions spécifiques, de relier les aires sensori-motrices et, par là même, d'établir des correspondances entre les différentes représentations du monde extérieur. Ainsi, les différentes dimensions d'un objet, visuelles, auditives, tactiles ou émotionnelles, vont générer, par leurs combinaisons, une représentation distribuée de cet objet. Les aires associatives proprement dites portent bien leur nom. SA, qui associe S2 et A2, et SV (pour S2 et V2) sont les aires dites pariétales et lient les activités musculaires respectivement à l'audition et à la vision. Ces associations, nous le verrons dans les deux derniers chapitres, permettront de guider la main par l'oeil ou l'appareil phonatoire par l'oreille.

De manière similaire, les aires dites temporales VT et AT associent l'aire O et ses connotations limbiques, pour ne pas dire émotionnelles, avec respectivement V2 et A2. Pour schématiser, ces aires vont servir à mettre

un nom sur les flux visuels et auditifs. Autrement dit, les aires pariétales s'orientent vers l'aspect spatial des informations et les aires temporales vers leur aspect symbolique.

3.1.3 Les aires frontales

Enfin, dernières aires et non les moindres puisqu'elles occupent tout le lobe frontal, les aires frontales sont aussi qualitativement importantes car elles représentent la dernière innovation de l'évolution, poussée à son extrême chez l'homme. La nomenclature est simple; en effet, chacune des aires vues précédemment a son pendant Frontal: FV pour V2, FVT pour VT, FSA pour SA et ainsi de suite. Là encore, il est possible de se faire une idée de la fonction de ces aires d'après le type des communications qu'elles supportent. D'une part, on l'a vu plus haut, les aires frontales sont sensibilisées aux différentes échelles de temps par leurs liaisons avec le corps strié et la formation réticulée. D'autre part, chaque aire frontale reçoit des informations d'une aire sensorielle ou associative. Enfin, les aires frontales suivent le modèle de ces aires pour leurs interconnexions: si V2 et VT communiquent, il en sera de même pour FV et FVT. Tout ceci permet cette première interprétation des aires frontales: elles reprennent le contexte sensoriel et associatif vu précédemment et ajoutent à ces atomes de connaissance leurs interdépendances temporelles. Ainsi vont apparaître plans d'action, hiérarchie, raisonnement et contraintes temporelles qui sont bien l'apanage de l'homme.

3.2 Le réseau

3.2.1 Deux représentations graphiques

Ayant énuméré les aires cérébrales et leurs connexions, il reste à donner une représentation de ce réseau. Ici encore, nous l'empruntons à Y. Burnod (Burnod, 1988) qui l'a d'abord représenté sous une forme "fermée", selon une disposition relativement proche de celle du véritable cortex. Ce schéma, avec sa forme de deux losanges concentriques, met l'accent sur deux coordinations sensorimotrices essentielles: la liaison entre l'oeil et la main pour le losange extérieur permet à l'homme la maîtrise de son environnement tandis que la relation entre l'appareil phonatoire et l'oreille, pour le

losange intérieur, lui offre le langage. L'autre attribut essentiel qui apparaît sur ce schéma a pour l'instant été passé sous silence; des connexions entre les deux losanges (V2 lié à A2, SV lié à SA, ...) relient les complexes auditifs et visuels et permettent à l'homme, lors de l'apprentissage du langage, d'associer une forme visuelle et une forme auditive (son nom, par exemple).

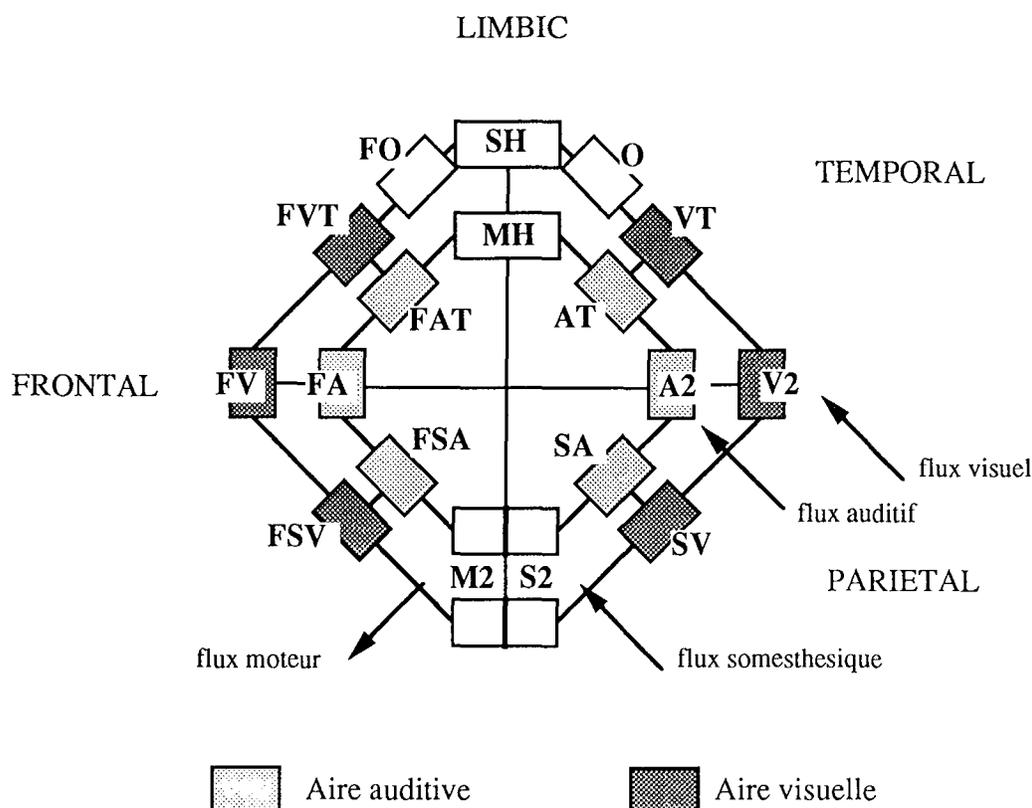


Figure 17. Forme fermée.

Cependant, cette schématisation pose certains problèmes. Par exemple, les liens entre V2 et FV ou O et FO n'apparaissent pas clairement. Cet inconvénient, certes mineur, ajouté aux avantages apportés par la seconde schématisation que nous allons présenter, nous fera choisir par la suite cette représentation sous la forme "ouverte". Précisons tout de suite que ces deux schémas sont équivalents; il est possible de passer de l'un à l'autre en coupant la partie supérieure des deux losanges et en les dépliant. Il s'agit en fait de deux projections bidimensionnelles du même réseau tridimensionnel,

modèle du cortex, qui, nous le verrons bientôt, possède d'intéressantes propriétés hypercubiques. Le premier avantage de cette représentation est en fait réductionniste et spécifiquement lié à l'utilisation particulière que nous voulons faire de ce modèle. En effet, toutes les aires qui n'ont pas d'intérêt dans l'optique immédiate de nos simulations (cortex limbique: FH, TH, SH, MH, FO et O) sont situées exclusivement dans la partie supérieure de la forme ouverte, qui sera négligée par la suite.

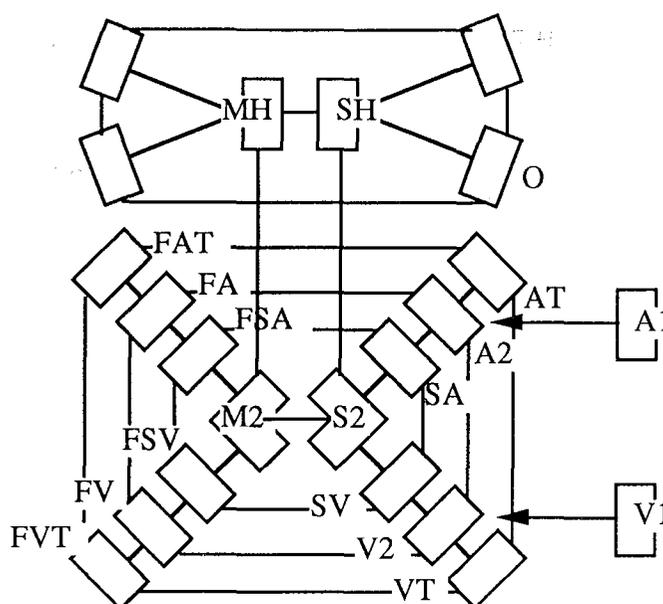


Figure 18. Forme ouverte.

La géométrie de la partie inférieure du schéma met en évidence les circuits privilégiés de l'information dans le réseau.

3.2.2 Les circuits de l'information

Les quatre demi-diagonales de la forme ouverte (Figure 19) correspondent à des circuits d'information représentant des fonctions directement liées aux aires qui composent ces diagonales. Ainsi sont localisées l'audition, la vision, la phonation et la manipulation.

Les deux premières fonctions représentent l'ensemble des actions associatives et les deux dernières tous les plans d'actions impliquant ces actions élémentaires.

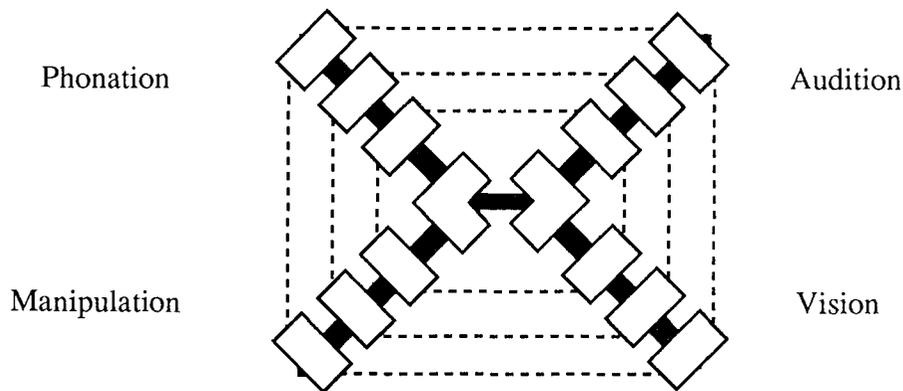


Figure 19. Axes privilégiés.

De manière complémentaire, les quatre circuits rectangulaires du schéma (Figure 20) relient les mêmes aires que celles qui participaient aux fonctions précédentes, mais l'ordre et la disposition de ces liens font apparaître quatre nouveaux chemins pour l'information. Ces chemins relient, chacun à son niveau, l'information sensorielle et la réponse motrice, selon les différentes composantes temporelles apportées par le cortex frontal.

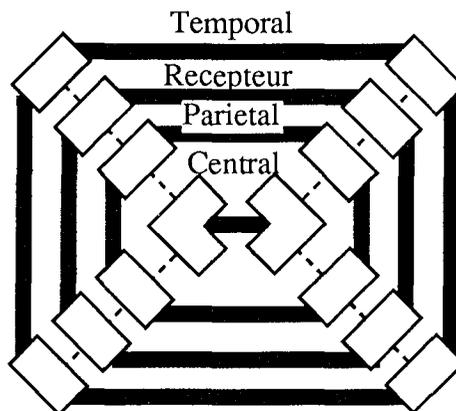


Figure 20. Circuits privilégiés.

1. Le circuit central, relation directe entre M2 et S2, correspond à des actions dites réflexes.

2. Le circuit pariétal correspond aux séquences de positionnement du corps par rapport aux informations visuelles et auditives.
3. Le circuit récepteur permet, d'après l'information visuelle ou auditive, de contrôler les séquences d'exploration visuelle.
4. Le circuit temporal propose un contrôle moteur plus global, fondé sur des traitements visuels ou auditifs à plus long terme et plus symboliques.

3.2.3 Symétries hypercubiques

Enfin, il existe une dernière propriété géométrique, peut-être la plus importante, qui, plusieurs fois, est apparue en filigrane.

On a vu que chaque aire associative ou sensorielle avait sa symétrique frontale, et réciproquement. Chaque aire appartenant au domaine visuel est connectée à sa symétrique auditive, et réciproquement. Enfin, ne l'oublions pas, ce modèle porte, pour l'instant, sur un hémisphère cérébral qui possède, lui aussi, son symétrique. Toutes ces propriétés de symétries géométriques et fonctionnelles (Figure 21) nous amènent naturellement à la notion d'hypercube.

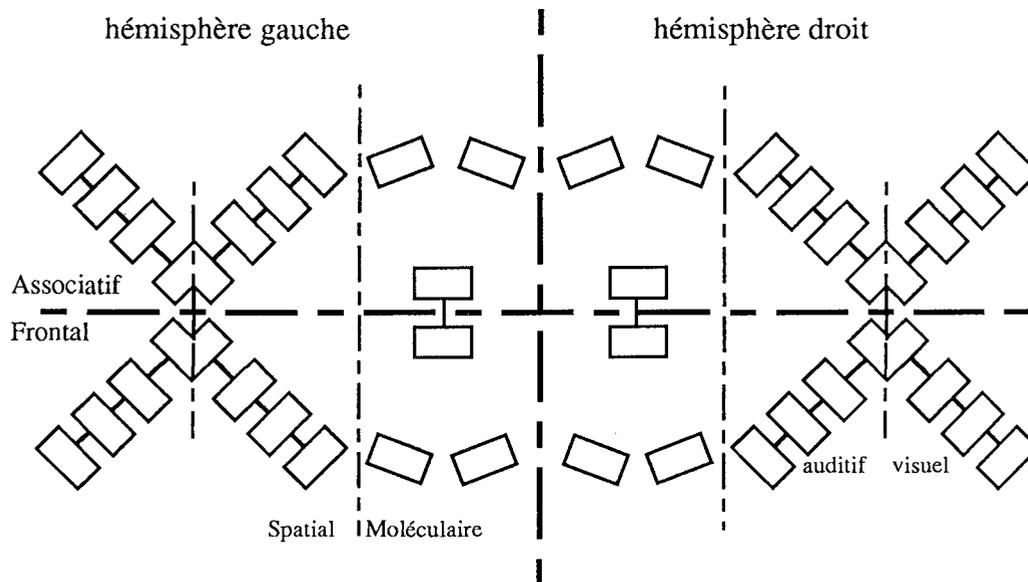


Figure 21. Symétries.

En effet, une aire quelconque du réseau est connectée à un nombre constant d'aires:

- selon la symétrie droite-gauche
- selon la symétrie associatif-frontal
- selon la symétrie visuel-auditif
- à deux aires voisines, sur les diagonales du schéma, une dans la direction du pôle somatique, l'autre dans la direction inverse.

Notons, pour être complet, qu'en plus de ces cinq liaisons pour chaque aire, on peut en définir une sixième selon une symétrie "environnemental-moléculaire" si on prend le réseau complet avec ses aires limbiques (SH, MH, ...).

Cet ensemble de cinq connexions symétriques par aire, défini indépendamment de l'aire considérée, permet, avec un nombre réduit de connexions totales, de trouver un chemin minimum pour relier deux aires. Cette propriété hypercubique, ajoutée à l'existence, dans un tel réseau, de plusieurs voies parallèles potentielles entre les entrées sensorielles et les sorties motrices, est primordiale pour un apprentissage structuré efficace. En effet, d'une part, les symétries créent des liens directs entre les différentes représentations du monde extérieur, d'autre part, les connexions ordonnées permettent de combiner les différentes représentations suivant plusieurs modalités. Cet isomorphisme fonctionnel entre les circuits d'information favorise le transfert d'apprentissage de relations en relations similaires dans les aires symétriques.

Cette description fonctionnelle du cortex nous permet maintenant, pour illustrer son fonctionnement en contexte, de citer l'exemple de la pomme proposé par le Professeur Vincent (Vincent, 1986):

Une pomme est posée sur la table. Mon cortex visuel la voit; mon cortex temporal associatif dit: "elle a l'air bonne". Mon cortex pariétal conclut: "je vais la manger". Mon cortex frontal dit alors: "je vais la porter à la bouche et la croquer", ce que fait mon cortex moteur, sous le contrôle vigilant de mon cortex somesthésique. Et tout mon cerveau se régale!

3.2.4 Réseau: mode d'emploi

D'une manière plus pragmatique, cette localisation des fonctions sur des aires spécifiques va nous permettre d'isoler un ensemble d'aires ainsi que leurs relations afin d'en étudier la fonction globale.

La simulation d'une fonction donnée passera d'abord par l'extraction du réseau général des caractéristiques suivantes: tout d'abord, il faudra trouver l'axe géométrique qui supporte cette fonction et énumérer les aires et les relations qui le composent. Puis les structures coopérantes imposeront une distribution de l'information à la surface de certaines aires, qu'il sera bon d'étudier en détail.

Cette construction particulière étant faite, il conviendra ensuite d'insuffler à ce sous-réseau sa logique générale de fonctionnement, de connectivité, de transmission et d'apprentissage, pour enfin réaliser la fonction adaptative.

De telles études seront faites en détail, selon ce principe à chaque fois répété, dans les chapitres suivants pour des fonctions visuelles, auditives, motrices et mixtes.

Il nous reste auparavant à préciser plusieurs points. La structure du réseau ayant été décrite à un niveau d'organisation élevé, il nous faudra parler de la logique générale, interne aux aires. Pour cela, il va nous falloir descendre encore plus bas dans le niveau de description du cortex et de son fonctionnement. En effet, nous avons jusqu'à présent parlé des fonctions corticales sous un angle que l'on pourrait plus qualifier de psychologique que d'informatique. On a vu comment le cortex, par ses spécialisations imposées par la localisation de l'entrée et de la sortie des flux d'information, génère de manière interne un réseau permettant à ces flux d'interagir et de créer les principales fonctions sensorimotrices d'adaptation humaines. Nous allons maintenant devoir considérer les éléments qui permettent à une aire de recevoir l'information, de la traiter et de la distribuer par le jeu des symétries. Ces éléments, constitutifs des aires corticales et constitués de neurones, sont connus sous le nom de colonnes corticales.

4 La colonne corticale

4.1 Réalité biologique de la colonne

Historiquement, c'est Lorente de No (Lorente de No, 1938) qui, le premier, a proposé un modèle d'organisation fonctionnelle des neurones du cortex en colonnes verticales, à partir de ses propres études de la connectivité intracorticale. De nombreux neurobiologistes se sont ensuite intéressés, sous divers aspects, à cette structure "cristalline" particulière. Citons, parmi tant d'autres, Szentagothai (Szentagothai, 1973),

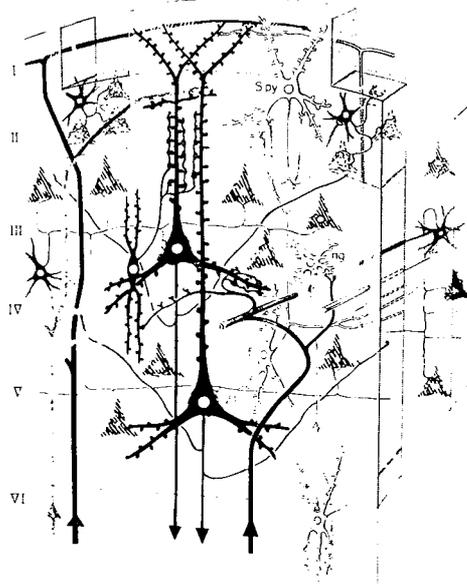


Figure 22. Cellules de la colonne (d'après Szentagothai).

qui propose une description cytologique très précise des neurones du cortex et de leur connectivité, Hubel et Wiesel, qui ont mis les colonnes corticales en évidence dans le cortex visuel du chat et du singe (Hubel & Wiesel, 1962; Hubel & Wiesel, 1977) en soulignant l'aspect traitement de l'information et enfin Mountcastle, qui, après des observations sur le cortex somato- sensoriel du chat puis sur le cortex pariétal (Mountcastle, 1957; Mountcastle, Lynch, Georgopoulos, Sakata, Acuna, 1975), généralisa le concept.

4.1.1 Szentagothai

Dans son article intitulé "Synaptology of the visual cortex" (Szentagothai, 1973), Szentagothai aborde l'aspect anatomique du cortex visuel et donne une interprétation possible de toutes les données biologiques en terme d'organisation colonnaire.

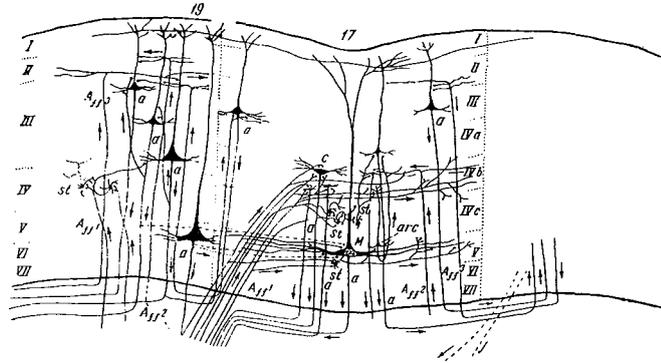


Figure 23. Connectivité (d'après Szentagothai).

Son étude très précise et remarquablement documentée décrit les différents types de cellules rencontrées ainsi que leurs afférences. Ceci est bien illustré par les deux schémas précédents, extraits de l'article.

Nous reviendrons plus loin sur ces deux schémas qui contiennent des informations indispensables pour l'élaboration du modèle. Szentagothai conclut: "L'organisation des afférences sensorielles semble être en agrément total avec le concept de l'organisation colonnaire des aires sensorielles corticales."

4.1.2 Hubel et Wiesel

C'est en explorant les aires visuelles du chat, qu'Hubel et Wiesel (Hubel & Wiesel, 1962) ont pu montrer que les neurones répondent de façon sélective à l'orientation, et que cette fonctionnalité se retrouve localement sur des groupements d'une centaine de neurones répartis dans l'épaisseur du cortex. Des pénétrations perpendiculaires à la surface montrent une invariance dans la fonctionnalité visuelle de la zone et des pénétrations obliques montrent au contraire une évolution discrète de la sélectivité (figure 24).

L'organisation topologique des aires sensorimotrices du cortex était déjà connue et plus particulièrement la rétinotopie des aires visuelles. Ces chercheurs se sont alors penchés sur le problème de localisation des champs récepteurs de la colonne corticale, et ont mis en évidence des groupements de colonnes corticales sensibles à une même position du stimulus, mais pour des orientations diverses .

"Quand nous avons entrepris ce type d'expérience pour la première fois en 1961, le résultat fut si surprenant que nous avions du mal à le croire. Au lieu d'une suite d'orientations aléatoires, nous avons trouvé une succession présentant un ordre stupéfiant. Chaque fois que l'électrode avançait, ne serait-ce que de 25 à 50 micromètres (millièmes de millimètre), l'orientation optimale changeait d'un petit pas angulaire de dix degrés en moyenne; les pas se succédaient dans le même sens, soit dans le sens des aiguilles d'une montre, soit dans le sens inverse" (Hubel & Wiesel, 1978).

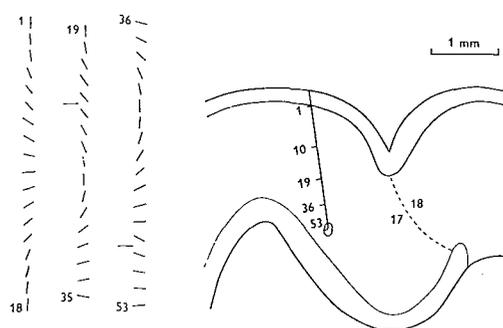


Figure 24. Pénétration oblique.

La rétinotopie, localement différenciée par sélectivité de l'orientation, permet de représenter plus de deux dimensions sur une surface bidimensionnelle. Le cortex manipule au moins trois types de variables, deux pour les coordonnées (x,y) du champ visuel, et un pour l'orientation. Mais les conclusions de ces deux chercheurs vont plus loin: "Dans l'état actuel de nos connaissances sur le cortex visuel primaire, il est évident que l'on peut considérer un îlot d'environ un millimètre carré de surface et de deux millimètres d'épaisseur comme un élément de base du cortex. Connaître l'organisation de ce petit morceau de tissu, c'est connaître l'organisation de toute l'aire 17. Le cortex tout entier n'est vraisemblablement qu'un pavage répétant

un module élémentaire.(...) Cependant, il apparaît clairement qu'une organisation en fins réseaux périodiques est une caractéristique générale de l'architecture du cortex cérébral" (Hubel & Wiesel, 1978).

4.1.3 Mountcastle

Mountcastle a d'abord mis en évidence des paramètres de fonctionnement des colonnes corticales dans le cortex sensori-moteur, directement liés à la nature des entrées. Dans le cortex pariétal, plus lié aux mouvements, il observe que les sorties des colonnes sont, elles aussi, organisées fonctionnellement, pour des tâches de préhension, de manipulation ou de scrutation.

Nous ne pouvons pas résister au plaisir de citer ce passage savoureux, tiré de ses conclusions (Mountcastle, 1978): "L'idée générale est la suivante: les grands ensembles neuronaux du cerveau, à savoir les aires corticales, mais aussi le lobe limbique, le ganglion de la base, ou le thalamus, sont constitués par la juxtaposition de circuits de neurones locaux de taille et de fonctionnement similaires, du moins dans une aire donnée. Chaque module effectue des traitements sur ses entrées, les répercute sur ses sorties, et ces traitements sont uniquement déterminés par la provenance des afférences".

4.1.4 Présentation

Cette présentation historique de la colonne corticale va céder la place à sa modélisation, qui est au coeur de nos travaux de simulation. La description de cette unité a été énoncée par Y. Burnod qui, à partir de nombreuses données neurobiologiques, a proposé une formalisation de ses diverses caractéristiques. Nous n'avons malheureusement pas trouvé dans la littérature de description concurrente qui permette une mise en oeuvre informatique; aussi restons-nous tributaires de ce modèle. Cependant, et le lecteur pourra lui-même juger la cohérence des interprétations, nous présenterons pour chaque paragraphe les données neurobiologiques qui ont motivé la modélisation choisie, avant d'en énoncer la formalisation.

Typiquement, un réseau d'unités parallèles peut être entièrement décrit par la donnée d'un petit nombre de caractéristiques (Rumelhart & Mc Clelland, 1986; Shaw, Silverman, Pearson, 1988). Nous commencerons par une description statique (architecture et connectivité) puis dynamique (états d'activation, règles de fonctionnement et de mémorisation).

4.2 Architecture et connectivité

4.2.1 Données Neurobiologiques

Délimitation spatiale de la colonne corticale

Le cortex humain peut être représenté comme un feuillet bidimensionnel dont l'épaisseur ne dépasse pas 2mm pour une surface de $0,25m^2$; les différentes aires corticales s'en partagent la surface. Une chaussée pavée est constituée par la juxtaposition bidimensionnelle d'éléments cubiques; il en est de même pour le cortex et la colonne; cette dernière se situera donc dans une interprétation strictement bidimensionnelle du cortex.

La colonne corticale est haute de 2mm et pour un diamètre d'environ 50 à 70 microns; ceci correspond à un maximum d'interconnexions. Elle détermine ainsi un motif répétitif d'une centaine de neurones indépendamment de l'aire corticale ou même de l'espèce animale observée. Plus généralement, suivant le niveau de définition requis par la fonction corticale envisagée, les colonnes se regroupent en maxi-colonnes correspondant à une fonctionnalité sensorimotrice bien précise.

Une organisation en six couches

Il est maintenant bien admis que, perpendiculairement à la surface, le cortex est organisé en six couches (Bullier, 1983) d'épaisseur variable suivant le type d'aire (variations cytoarchitecturales).

Soulignons ici la variabilité de l'épaisseur de la couche IV du cortex suivant l'aire corticale considérée; très importante pour les aires sensorielles, elle diminue pour les aires associatives et plus encore pour les aires motrices, au bénéfice des couches III et V.

Isolons une colonne du reste du cortex; dans sa hauteur, nous retrouvons les six couches, avec une répartition assez précise de cellules nerveuses de différents types et des schémas de connectivité réguliers.

Types de cellules

Les plus grandes cellules du cortex, les neurones pyramidaux, ont pour principale caractéristique l'extension verticale de leur arbre dendritique qui atteint la surface du cortex indépendamment de la position de leur soma. On les trouve essentiellement dans les couches II et III ainsi que V et VI. Ce sont les principaux responsables des connexions à longue distance dans le cortex, grâce à la projection de leurs axones vers des régions lointaines.

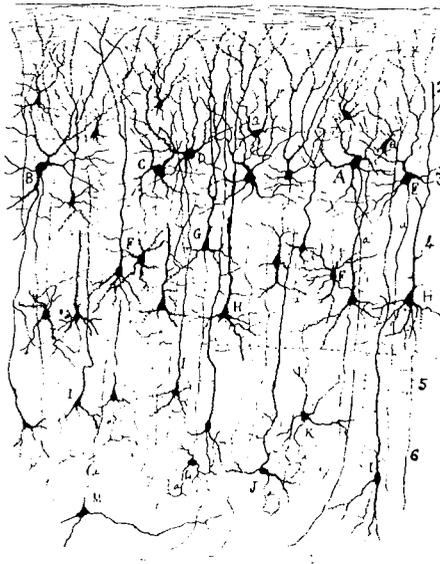


Figure 25. Les 6 couches du cortex

Les autres neurones, appelés interneurones, sont inhibiteurs ou excitateurs. Leur connectivité et leur influence sont très locales. On les regroupe grossièrement en cinq types: étoilé, bipolaire, en panier, en chandelier et à double bouquet. La figure 26 qui modélise la répartition de ces types, est à rapprocher de la figure 22, croquis anatomique de la colonne.

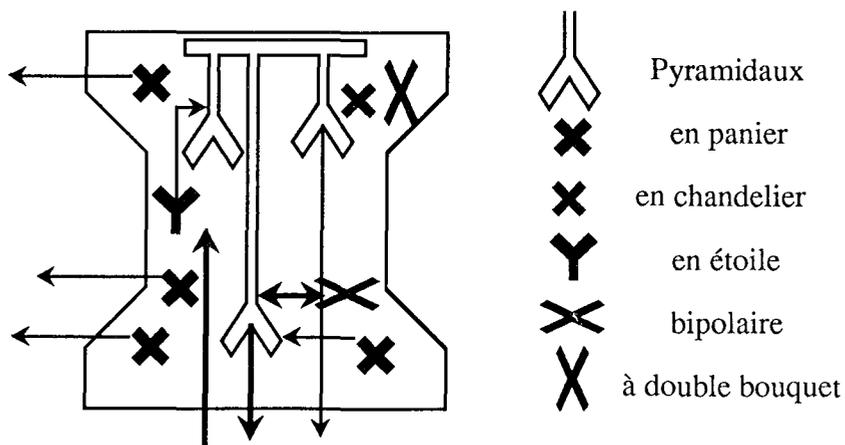


Figure 26. Modélisation des neurones de la colonne.

On peut considérer les neurones pyramidaux comme intégrant des informations de provenances diverses et les interneurons comme mémorisant les différentes configurations fonctionnelles.

Connexions

La figure 27 esquisse un schéma de connectivité entre les colonnes:

- Les couches I, II et III (les couches "supragranulaires" hautes) vont recevoir les axones des neurones pyramidaux supérieurs d'autres modules corticaux. Si on analyse plus finement la provenance de ces axones afférents, on remarque une structuration hiérarchisée par la nature de la symétrie, illustrée par la figure 27. Réciproquement, les neurones pyramidaux des couches hautes vont se projeter vers d'autres modules corticaux.

ENTREE:

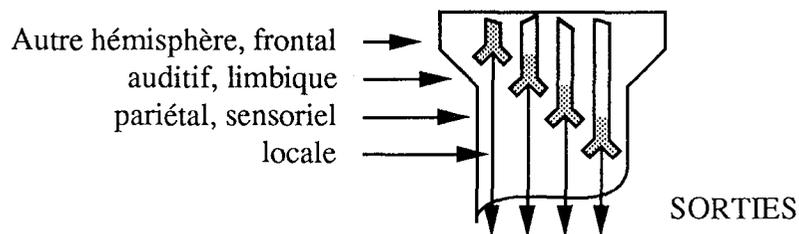


Figure 27. Hiérarchisation des entrées de la couche haute.

- La couche IV reçoit généralement des afférences en provenance de l'extérieur du cortex de diverses structures coopérantes (ex: thalamus) et en provenance des aires sensorielles primaires pour les aires sensorielles secondaires.
- Les couches V et VI envoient généralement des connexions hors du cortex, sur les mêmes structures coopérantes. Soulignons encore le rôle particulier joué par le thalamus, qui constitue la voie d'échange avec le cortex la plus importante: la plupart des informations sensorielles et motrices y transitent.
- Mentionnons encore les interneurons des couches II, IV et V qui établissent des connexions réciproques avec les colonnes voisines.

4.2.2 Modélisation

Que peut-on déduire de ces données afin de construire un automate cellulaire ?

Principes d'architecture:

Les entrées et les sorties des neurones pyramidaux sont nettement localisables sur l'axe vertical de la colonne (épaisseur du cortex). De haut en bas, on distingue les couches suivantes (Szentagothai, 1975; Mountcastle, 1978; Jones, 1981):

- Des couches regroupent les entrées et les sorties de colonnes d'autres aires, connectées suivant le principe des symétries du cortex, ainsi que de colonnes de la même aire (Szentagothai, 1975; Jones, 1981).
- Une couche intermédiaire reçoit les entrées en provenance du monde extérieur (Zeki & Shipp, 1988; Van Essen & Maunsell, 1983)
- Une couche basse effectue les sorties vers le monde extérieur (Zeki & Shipp, 1988; Van Essen & Maunsell, 1983)

D'autre part, deux colonnes voisines partagent un certain nombre d'interneurones qui vont générer des interactions locales.

C'est pourquoi, nous proposons, dans la partie suivante, de distinguer, pour notre unité de base, les quatre types de connexions réciproques suivant:

- **des liens avec des unités dans d'autres aires (liens lointains)**
- **des liens avec des unités de la même aire (liens locaux)**
- **des liens avec le monde extérieur (liens externes)**
- **des liens avec des unités voisines (liens voisins)**

La Figure 28 montre la représentation graphique que nous ferons dorénavant de la colonne et ses liens.

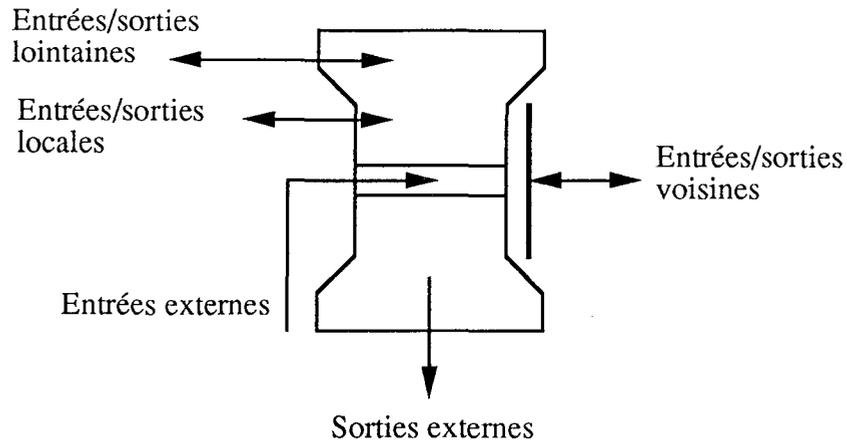


Figure 28. Entrées/sorties de la colonne corticale.

Principes de connectivité:

Plus précisément, pour chaque type de message échangé, il convient de spécifier le nombre et la répartition spatiale des connexions.

1. Une unité sera, de façon minimale, connectée avec ses quatre voisins par des liens réciproques répartis selon la Figure 29.

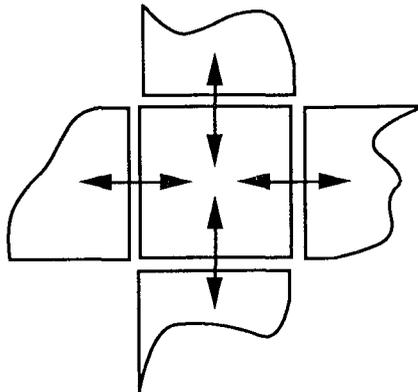


Figure 29. Liaisons voisines.

2. En ce qui concerne les connexions entre unités d'aires différentes

(connexions lointaines), la géométrie du réseau simulé définira le nombre d'aires et donc le nombre de ces liens.

Le modèle de connectivité d'une aire C1 à une aire C2 respecte les principes suivants:

- Les aires sont connectées selon une symétrie hypercubique (chapitre III),
- Une unité de C1 est liée à une unité de C2, son image,
- Pour deux unités voisines de C1, leurs images dans C2 sont voisines.

D'inspiration biologique, réduites à leur plus simple expression, ces trois règles s'exploitent aisément; la Figure 30 représente, pour un réseau à deux aires, une implantation possible.

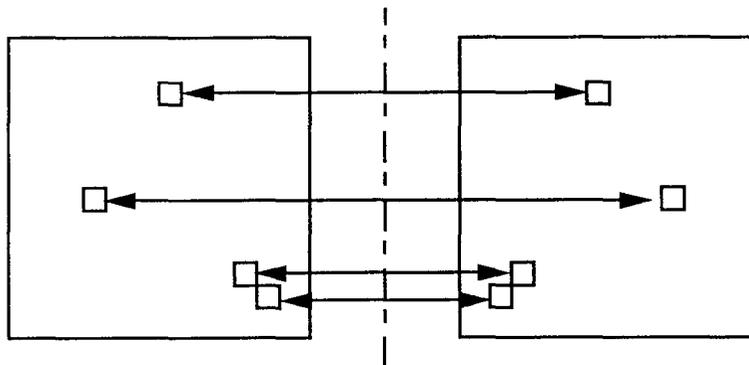


Figure 30. Connexions dans un réseau à deux aires.

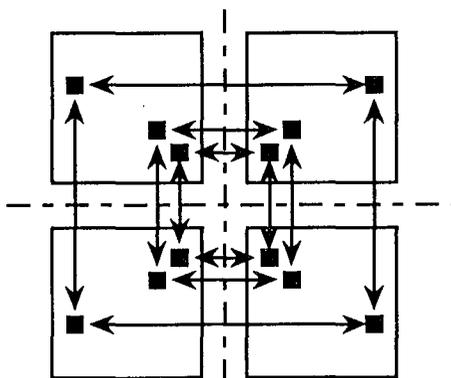


Figure 31. Connexions dans un réseau à quatre aires.

En prenant comme image l'unité symétrique par rapport à un axe, on respecte bien le principe des voisinages; la figure 31 illustre la connectivité pour quatre aires reliées deux à deux par des symétries de type hypercubique.

De façon générale, dans un réseau à 2^n aires, une colonne est reliée à n autres colonnes dans n autres aires par des connexions lointaines.

3. Les entrées/sorties avec des colonnes d'une même aire s'effectuent suivant un schéma simple, comparable à celui des entrées voisines.

4. Il est nécessaire que le réseau ait accès à une représentation du monde extérieur sur laquelle ses unités puissent effectuer des entrées/sorties locales. Pour une unité donnée, ces entrées/sorties s'effectuent sur un voisinage connexe de la surface bidimensionnelle qui propose un codage initial de l'information. Nous appellerons cette surface "carte de codage". La connectivité respecte les deux principes suivants:

1. Une colonne d'une aire A1 est connectée à un seul voisinage de la carte de codage correspondante.
2. Deux colonnes voisines sont connectées à deux voisinages voisins.

Explorons un peu plus la nature des connexions externes d'une colonne et d'un voisinage. On peut émettre l'hypothèse d'un câblage a priori définissant le voisinage comme composé d'une partie centrale excitatrice et d'une partie périphérique inhibitrice. Ce schéma de câblage est compatible avec les explorations neurobiologiques du cortex visuel primaire (Gilbert & Wiesel, 1981).

La figure 32 représente toutes les entrées/sorties de la colonne C de l'aire A1 dans un réseau à deux symétries (quatre aires):

On trouve exactement:

- 2 connexions lointaines
 - 4 connexions locales
 - 4 connexions voisines
 - 4 connexions externes sur un voisinage 2x2,
- soit au total, 14 connexions par unité.

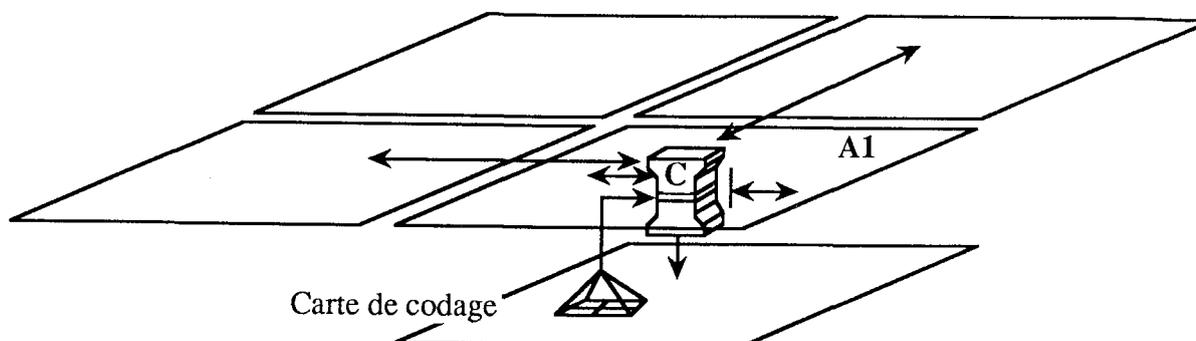


Figure 32. Résumé des connexions.

4.2.3 Remarques

- Il faut souligner, avant tout, le nombre limité de connexions (Mountcastle, 1978; Szentagothai, 1975), qui ne dépend pas de la taille des aires. Il est proportionnel à $\log_2 n$, où n est le nombre des aires. On ne retrouvera donc pas l'explosion combinatoire des réseaux classiques totalement interconnectés. Il est cependant nécessaire que deux modules quelconques puissent échanger des informations, même s'ils ne sont pas physiquement reliés. C'est par le biais des connexions locales, où l'information va pouvoir circuler, que nous résoudrons le problème.
- En choisissant pour image d'une unité sa symétrique par rapport à un axe de symétrie de l'hypercube, on privilégie certaines recombinaisons d'information. Ainsi, dans l'aire pariétale SV qui combine les représentations visuelles (V2) et somesthésiques (S2), des connexions relient la main et la vision centrale; d'autres connexions relient les zones de plus en plus périphériques de ces deux domaines sensoriels. Nous verrons l'importance de ce codage pour le guidage de la main par l'oeil dans une tâche de préhension (chapitre III.3). D'autres illustrations de ce principe seront données par la suite (chapitre III.2 et III.4).
- La nécessité de définir, a priori, des schémas de connectivité pour les entrées/sorties externes peut sembler contraignante; elle est en fait

très riche. On va pouvoir agir sur la spécification des aires avant apprentissage, et faire converger chaque aire vers une dimension de la représentation des données. Ainsi, au chapitre 6, pour une représentation visuelle, la différence de taille des voisinages fait converger une partie du réseau vers une représentation précise mais locale (identification d'une forme) et l'autre partie vers une représentation grossière mais globale (localisation d'une forme).

- Cette séparation en quatre types d'entrées/sorties, lointaines, locales, voisines et externes, est une idée neuve dans le domaine du connexionnisme. Classiquement, un neurone formel effectue un calcul indifférencié sur ses entrées. Notre unité, par contre, est en mesure de distinguer la provenance d'un message. Cette capacité va générer plusieurs mécanismes d'apprentissage et de fonctionnement, dépendant du type du message traité.

4.3 Etats d'activation

Nous venons de le voir, la connectivité et la structure du réseau sont très contraignantes: connectivité réduite à l'extrême et entrées/sorties externes prédéterminées sur des aires de codage. Cette étude descriptive statique de notre unité de base a déjà bien esquissé ce que sera l'architecture de nos réseaux. Abordons l'étude dynamique par le problème primordial des états d'activation.

4.3.1 Données neurobiologiques

Hubel et Mountcastle ont effectué des mesures de potentiel d'action des neurones pyramidaux dans les aires sensorielles (Hubel & Wiesel, 1962; Mountcastle, Lynch, Georgopoulos, Sakata, Acuna, 1975). Ils observent des fréquences de 50 à 100 Hz lorsque l'animal est correctement stimulé (Hubel & Wiesel, 1977; Mountcastle, Andersen & Motter, 1981; Evarts & Tanji, 1974). Des fréquences du même ordre de grandeur ont été mises en évidence dans d'autres régions du cortex, lors de mouvements volontaires par exemple.

On observe des fréquences plus faibles (de 5 à 10 Hz) lors de tâches qui ne génèrent pas de comportement visible, comme dans les phases d'attention

sélective (Mountcastle, Andersen & Motter, 1981; Fuster, 1977).

Quand les neurones pyramidaux des couches hautes s'activent à haute fréquence, il se produit le phénomène "d'avalanche" (Lorente de No, 1938) qui propage cette activité aux couches basses, responsables des sorties vers l'extérieur.

Trois états généraux, distincts et stables se dégagent:

1. un état où toute la colonne est très activée,
2. un état où seules les couches hautes sont moyennement actives,
3. et bien sûr un état sans activité.

4.3.2 Modélisation

En première approximation, nous définirons trois états, notés **E0**, **E1** et **E2** (inhibé, peu activé et très activé). Plus finement, ces trois états discontinus sont modulés par des variations continues.

L'introduction de ces petites variations ne remplace pas les trois états, mais les affine pour émuler des mécanismes plus complexes. Nous simulerons donc l'état d'une unité par deux valeurs: un état E_i et une petite variation locale C .

4.3.3 Remarques

- Il est clair que les trois états de la colonne vont définir une logique trivaluée et peuvent être interprétés comme oui ($E2$), non ($E0$) et peut-être ($E1$).
- D'un point de vue comportementaliste, $E1$ peut s'interpréter en terme de recherche ou de sensibilisation et $E2$ en terme d'action ou de reconnaissance. Cette interprétation psychologique va nous être très utile pour comprendre et mettre au point les règles d'activation et de mémorisation. On verra, par la suite, toute la nouveauté apportée par l'état intermédiaire $E1$.
- Les trois valeurs $E0$, $E1$, et $E2$ définissent qualitativement l'état de la colonne; la valeur C le définit quantitativement. Par la suite, nous définirons l'usage de ces différentes valeurs pour les logiques de fonctionnement et d'apprentissage.

4.4 Règles de fonctionnement

Notre objectif est de définir le comportement d'un automate lors de processus fonctionnels. Plus précisément, il s'agit de définir des règles qui permettent de calculer l'état des sorties d'un automate en fonctions des messages reçus, des apprentissages associés ou de paramètres divers (type de l'entrée, type de l'aire..).

4.4.1 Données neurobiologiques

Quand une colonne reçoit deux stimuli, en couches haute et intermédiaire, ses neurones pyramidaux, sous l'effet d'une amplification mutuelle, sont très activés.

L'effet des entrées thalamiques (couche intermédiaire) sur l'activité de la colonne dépend du type de l'aire considérée. La réponse à une entrée thalamique est fonction de l'épaisseur de la couche IV et de sa densité d'interneurones, plus grande pour les aires sensorielles qu'associatives et motrices.

Des stimulations corticales sont insuffisantes pour déclencher un haut niveau d'activation de la colonne; cependant, elles peuvent augmenter la réponse de la colonne à une entrée thalamique. Cet effet d'amplification peut s'interpréter en terme d'attente comportementale ou d'anticipation.

Des colonnes fortement actives ont un effet d'inhibition latérale sur leurs voisines; deux colonnes voisines peuvent cependant être coactivées fortement (Mountcastle, 1978; Hubel & Wiesel, 1977).

4.4.2 Modélisation

Exprimons les paramètres qui vont influencer l'état d'un automate. Il semble premièrement nécessaire de le faire dépendre des messages qu'il reçoit d'autres automates ou du monde extérieur, c'est à dire des entrées voisines, locales, lointaines et externes. De plus, l'apprentissage, quel qu'il soit, modifie certainement la réponse de l'unité. Nous avons aussi fait état de modifications cytoarchitecturales qui déterminent l'influence des entrées externes selon le type de aire (sensorielle, associative ou motrice). Enfin, puisque la modélisation des états prend en compte l'aspect temporel des données, il semble naturel de tenir compte de l'état présent de l'automate avant de le modifier.

Plus formellement, l'état d'une unité (modélisé par E_i et C) dépend donc de paramètres:

-variations cytoarchitecturales (V_c),

-apprentissage (A_p)

et de variables qui évoluent temporellement:

-entrées externes (E_e, C_e), lointaines (E_c, C_c), locales (E_l, C_l) et voisines (E_v, C_v),

-état précédent de l'automate (P_s).

Le modèle auquel nous sommes arrivés actuellement est basé sur une hypothèse réductionniste: la séparation et l'indépendance entre d'une part un mécanisme de transmission de l'information de proche en proche, nécessaire pour satisfaire le principe de connectivité limitée, et d'autre part un mécanisme logique, support du raisonnement. Ces deux mécanismes sont distribués et constituent la logique interne de la colonne. La transmission de proche en proche s'effectuera à l'intérieur d'une même aire (aussi appelée carte) et sera supportée par les liens locaux. La logique de raisonnement dépendra des entrées lointaines et externes. **Cette dernière est illustrée, figure 33, par une table de vérité qui donne les sorties externes et lointaines en fonction des entrées.** Dans cette table, seules les configurations impliquant une modification d'état sont mentionnées.

Entrées		Etat Précédent	Sorties		Commentaires	Dépend de:
Lointaine	Externe		Lointaine	Externe		
E0	E0		E0	E0		
E0	E1		E2	E2	Sensorielle	Type de l'aire
			E1	E0	Associative	
			E0	E0	Motrice	
E1	E0		E0	E0	Inhibition	Apprentissage
			E1	E0	Decl. cond.	
			E2	E2	Decl. incond.	
E1	E1		E2	E2	Amplification	
E2		E1	E1	E0	Retour	Apprentissage
E2		E2	E2	E2	Reactivation	

Figure 33. Table d'entrée/sortie (entrées lointaines et externes).

Opérations sur les entrées externes

La force des connexions d'une colonne avec un voisinage externe est définie a priori, dans le but d'extraire des invariants spécifiques. La colonne réalise ainsi une fonction de transfert sur le voisinage externe proposé. Par exemple, dans le cortex visuel, la fonction de transfert des colonnes sensibles à l'orientation peut se représenter par un masque composé d'une partie centrale excitatrice et d'une partie périphérique inhibitrice (Figure 34).

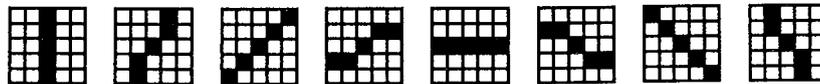


Figure 34. Masques thalamiques.

Pour un signal donné, la fonction de transfert mesure la distance entre le signal et le masque. Cette mesure détermine la valeur de l'entrée externe: E2 pour une distance nulle, E0 pour une distance trop grande et E1 dans le cas intermédiaire (colonne 2 de la table). La ligne 2 illustre une propriété importante de l'organisation corticale, la variabilité cytoarchitecturale. Suivant le type de l'aire considérée, la fonction de l'automate s'interprète différemment: pour les aires sensorielles, c'est la détection d'un indice dans un flux d'information; pour les aires motrices, c'est le déclenchement d'une action et pour les aires associatives, c'est la mise en correspondance de deux espaces de représentation.

Opérations sur les entrées lointaines

Faisons d'abord abstraction de la connectivité locale pour proposer une logique de fonctionnement apte au raisonnement et à la mémorisation associative. On peut modéliser la résolution d'un problème par une résolution parallèle ou séquentielle de sous-problèmes. Cette définition récursive s'achève par la résolution de problèmes atomiques de mouvements (sortie de la réponse) ou de présence de stimuli (entrée des données). Supposons que le réseau soit capable de discerner les sous-problèmes susceptibles de conduire à la résolution d'un problème plus général (voir section apprentissage). Dès lors, il est possible de spécifier une logique de fonctionnement à partir de l'interprétation des états de l'automate, considéré comme un problème potentiel local:

dans l'état E0, problème non résolu, faux,

dans l'état E2, problème résolu, vrai,
dans l'état E1, problème à résoudre, hypothèse.

Seul l'état E1 est un état de déséquilibre, pour lequel le réseau doit apporter une solution. La mémorisation permet au réseau de propager cet état de déséquilibre vers les automates susceptibles de conduire à la résolution, et de converger vers des problèmes atomiques résolubles. C'est pourquoi, à un niveau atomique et sous l'influence d'une requête E1, le système doit être capable de déclencher une action (E2) ou de vérifier un stimulus (E2) (ligne 3).

Opérations sur les entrées locales

Nous avons supposé pour élaborer cette table que deux modules quelconques pouvaient communiquer entre eux, malgré la connectivité réduite et locale. Nous proposons maintenant une logique de transmission de l'information d'un automate à un autre, de proche en proche.

Les connexions locales vont jouer un rôle particulier dans cette propagation à travers le réseau.

Le signal propagé s'affaiblit au fur et à mesure de son éloignement de l'automate émetteur; la variable C modélise une petite variation autour de la valeur d'un état et est utilisée comme fonction d'éloignement. L'information transite d'automate en automate par les entrées/sorties locales, jusqu'à son objectif. La transmission consiste à faire décroître pas à pas la variable C (selon une suite décroissante u). On peut ainsi modéliser la transmission par une table d'entrée/sortie, où l'on distingue trois cas: automate émetteur, transmetteur et récepteur:

Etat précédent	Entrées Locales	Sortie locale	Entrée lointaine	Commentaires
	E1,C	E1,C-1		Transmission
E1	E2	E2		Mémorisation
Ei		Ei, Cmax		Emission
	E1,C		E1	Réception

Figure 35. Table d'entrée/sortie (entrées locales).

Par apprentissage, un automate mémorise la valeur de l'éloignement (C) d'un message reçu. Il est alors capable de localiser l'émetteur précisément et d'échanger avec lui des informations sans connexions directes.

Opérations sur les entrées voisines

Des colonnes voisines partagent le même voisinage externe mais y effectuent des fonctions de transfert différentes. Les opérations sur les entrées voisines sont décrites dans la table 36.

Etat précédent	Entrées voisines	Sortie lointaine	Sortie externe	Commentaires
E1,E0	E2	E0	E0	Inhibition lat.
E2	E2	E2	E2	Réactivation

Figure 36. Table d'entrée/sortie (entrées voisines).

La première ligne illustre le principe d'inhibition latérale, mis en place entre deux colonnes voisines, de fonctions de transfert exclusives.

Dans le cas contraire, ligne 2, il y a réactivation réciproque permettant une mémorisation locale.

4.4.3 Remarques

- Il est d'abord important de constater que peu de configurations dépendent de l'apprentissage. La plupart correspond à un traitement stéréotypé de l'information. Ceci nous semble plus plausible dans la perspective de simulation de fonctions humaines.
- La conjugaison de plusieurs entrées de même type n'est pas prévue par le modèle neurobiologique; nous avons défini la règle suivante: pour un type d'entrée donné, l'automate comptabilise les messages reçus en fonction de leurs effets. L'état le plus demandé est retenu; en cas d'égalité, on reporte la décision en mettant l'unité à E1 (hypothèse à vérifier).
- Il est évident que l'hypothèse de séparation entre transmission et logique nous prive de propriétés intéressantes; certaines applications, basées sur une signification topologique de l'information montreront l'intérêt que peut avoir une dépendance des deux mécanismes (phénomènes de calibrage, indépendance de la taille).

- Il serait idéal de réunir les différentes tables d'entrées/sorties en une seule, plus complète et plus riche. Ainsi, on pourrait par exemple intégrer par apprentissage une dépendance de C' et C'' , des inhibitions de la transmission par les entrées lointaines...

4.5 Apprentissage

Pour chaque type d'entrées, nous allons trouver un mécanisme d'apprentissage différent. Pour les entrées externes, il permet une extraction d'invariants, pour les entrées lointaines, une décomposition d'un problème en sous-problèmes, pour les entrées locales, une création de chemin de propagation et pour les entrées voisines, une généralisation ou une différenciation de concept. Les quatre mécanismes vont suivre la même logique: le réseau n'approfondit son apprentissage qu'en cas de besoin.

4.5.1 Données neurobiologiques

L'apprentissage découle des modifications synaptiques et des variations d'excitabilité des cellules (Delacour, 1987). Les différents récepteurs des neuro-transmetteurs permettent de réaliser deux sortes de couplages entre cellules. Les récepteurs au glutamate (Barrionuevo & Brown, 1983), neuro-transmetteur des neurones pyramidaux, favoriseraient un couplage temporel, alors que les récepteurs des neuro-transmetteurs des inter-neurones (principalement le GABA, inhibiteur) ne tiendraient compte que de la simultanéité des activités.

A plus haut niveau, différentes études ont montré la plasticité des cartes sensorielles dans le cortex. Des déprivations sensorielles modifient l'organisation des cartes sensorielles (Merzenich, Jenkins, Middlebrooks, 1984).

A la naissance, les aires associatives sont moins spécifiées que les aires sensorimotrices; l'apprentissage y fait apparaître des zones fonctionnelles de taille variable, organisées en bande (Woolsey, 1958; Hubel & Wiesel, 1962).

4.5.2 Apprentissage sur les entrées externes

Bien que câblées a priori, les connexions externes peuvent jouir d'une certaine plasticité, et évoluer par apprentissage. Leur fonction principale est

de proposer un codage de l'information qui rende compte des invariants significatifs du monde extérieur. L'apprentissage est assez limité et a pour objectif une meilleure adaptation du masque d'entrée au stimulus réel. Trois conditions sont nécessaires pour déclencher l'apprentissage: il faut une petite différence entre le masque et le stimulus ($E_e = E_1$), un message d'erreur qui indique que ce dernier est significatif ($E_c = E_1$) et que l'automate ne soit pas inhibé par un voisin ($E_v \neq E_2$).

Dans ce cas, l'automate doit adapter les coefficients de son masque d'entrée au stimulus réel; on peut imaginer de nombreuses façons d'effectuer les changements. **La plus plausible à nos yeux est d'implanter un mécanisme Hebbien qui augmente la valeur des connexions avec les unités actives et diminue les autres.** Il suffit que les changements ne soient pas trop rapides pour éviter des oscillations du réseau d'un exemple à l'autre.

Remarque

Il s'agit d'un mécanisme d'apprentissage très local qui, pour les applications réalisées, n'a guère amélioré les performances. La figure 37 montre les masques après apprentissage sur des données visuelles, très proche de la modélisation initiale.



Figure 37. Masques déterminés par apprentissage.

4.5.3 Apprentissage sur les entrées lointaines

Il convient de mettre au point un mécanisme susceptible de rendre compte de la décomposition d'un problème en sous-problèmes validants. Cet apprentissage n'est effectif que pour une situation de déséquilibre (E_1 , problème à résoudre).

Comment établir un lien de "sous-problématique" ou, en d'autres termes, comment discerner, dans les schémas d'activité entre deux unités, l'unité indispensable à la résolution de l'autre? L'idée qui vient naturellement à l'esprit est d'utiliser le temps comme lien de causalité entre deux événements. Un problème B toujours résolu avant un problème A est très

certainement un sous-problème de A. Il sera en tout cas considéré comme tel.

Nous allons donc évaluer et mémoriser les schémas d'activité temporels entre deux automates A et B et définir des coefficients statistiques P0 et P2 pour différents schémas:

P2 mesure la probabilité que l'automate B soit dans l'état E2 avant que A ne le soit,

reciproquement, P0 mesure la probabilité d'avoir B à E2 avant A à E0;

P0 et P2 varient de 0 (jamais) à 1 (toujours). A partir de ces deux valeurs, il est maintenant possible de définir plus précisément l'effet de l'apprentissage sur une unité B excitée par A à E1. La Figure 38 indique, pour différentes valeurs de P0 et P2, les conséquences sur le fonctionnement de la colonne, limitées au troisième cas de la table 33.

Coefficients		Sorties		Commentaires
P2	P0	Lointaine	Externe	
?	?	E0,E1	E0	Aléatoire
0	1	E0	E0	Inhibition
>0	<1	E1	E0	Decl. Cond.
1	0	E2	E2	Decl. Incond.

Figure 38. Table de mémorisation.

Le protocole d'apprentissage se décompose en trois phases:

1. On infère un problème sur une aire de codage qui provoque des déséquilibres (E1) dans le réseau,
2. on force les réponses sur une autre aire de codage en sortie (E2)
3. on met à jour les coefficients P0 et P2.

Lors d'une résolution, le même problème va déclencher récursivement les sous-problèmes appris qui vont finalement proposer des solutions.

Remarques

- Pour des réseaux à plus d'une symétrie, l'ordre dans lequel les aires apprennent est d'une importance capitale; avant de chercher à résoudre un problème, il faudra déjà savoir résoudre ses sous-problèmes. Il semble d'ailleurs que le système nerveux humain suive des règles de maturation séquencées (Piaget, 1975).

C'est pourquoi un réseau doit suivre un protocole d'apprentissage strict qui dépend de la nature même des questions qu'il est censé traiter. Ce processus doit décomposer les différentes étapes de résolutions en apprentissages de plus en plus complexes.

- Il est important, pour la suite, de constater que le mécanisme d'apprentissage sur les entrées lointaines ne fournit aucune généralisation, mais seulement une logique de résolution. Le système risque de converger vers une représentation des connaissances très redondante où des instanciations légèrement différentes d'un même exemple vont générer des représentations multiples du même concept.

4.5.4 Apprentissage sur les entrées locales

L'apprentissage sur le troisième type d'entrées va permettre de mémoriser la localisation d'un message en provenance de la même aire. L'automate sera ensuite capable de reconnaître la provenance d'un message.

En première approximation, il suffit de mémoriser la valeur même de la variable C qui indique l'éloignement de l'émetteur lorsqu'un message significatif pour l'automate est reçu. L'automate possède alors une liste d'émetteurs potentiels, chacun associés à des coefficients probabilistes P2 et P0.

Lorsqu'un automate reçoit un message d'une entrée locale, il compare la valeur de l'éloignement avec celles mémorisées; si le message lui est destiné, il en tient compte à un niveau plus interne (E_c), sinon, il le propage par sa sortie locale.

4.5.5 Apprentissage sur les entrées voisines

Cependant, cette première approximation ne permet aucune généralisation topologique sur les localisations des automates émetteurs. C'est pourquoi, **il vaut mieux mémoriser un intervalle d'éloignement.** La logique

de l'apprentissage va être modifiée, trop de généralisation menant systématiquement à une mauvaise différenciation. Nous allons donc définir deux mécanismes d'apprentissage complémentaires sur les entrées voisines: généralisation et différenciation topologique donneront toute sa puissance à l'apprentissage entre automates d'une même aire.

Plusieurs règles gèrent ce que l'on peut aussi appeler la fusion/division des intervalles d'éloignements:

- en cas de confusion entre plusieurs solutions, chaque automate actif et impliqué dans une mauvaise proposition voit ses intervalles divisés,
- un automate qui ne reçoit jamais de message d'une zone mémorisée élimine l'intervalle correspondant de sa liste,
- pour un automate donné, les intervalles topologiquement voisins et de fonctionnement synchrone sont fusionnés en un seul.

Remarques

- On peut facilement étendre le processus à deux dimensions, on parlera alors plutôt de zone que d'intervalle bidimensionnel d'éloignement.
- Ce mécanisme tend à rendre compte informatiquement de processus réels dans les aires corticales; il s'agit de simuler les regroupements fonctionnels de colonnes en maxi-colonnes (fusion) ou réciproquement les inhibitions latérales qui peuvent différencier des groupements de colonnes (division; Burnod, 1988).
- C'est par une succession de divisions et de fusions que le système va différencier et généraliser topologiquement les indices pertinents d'un concept.

4.6 Implantation du modèle

4.6.1 Structure de données

L'algorithmique nous enseigne que, pour un programme donné, il est bon de commencer par spécifier la structure de données, puis de définir les fonctions qui les manipulent. Il est clair que la complexité des fonctions dépend de la manière dont est stockée l'information.

En ce qui concerne la colonne, les différents points de la modélisation que nous venons d'exposer nous ont amenés à faire les choix suivants:

D'une part, pour une colonne donnée, il existe un certain nombre d'informations qui dépend des relations avec d'autres unités (tels que les apprentissages), d'autre part, chaque relation doit pouvoir accéder à un ensemble d'informations communes (tel que l'état d'activation).

C'est la raison pour laquelle, la structure de données qui représente une colonne contient ces données communes et ses relations. Nous stockons dans chacune de ces relations les informations qui lui sont propres et auxquelles les autres relations n'ont donc pas accès.

En langage C, cette structure de données peut se représenter comme suit:

```

struct colonne
{
char type;
struct relation *liens_voisins;
struct relation *liens_lointains;
struct relation *liens_locaux;
struct relation *liens_externes;
int tempactive;
char entree_lointaine,
entree_voisine,
entree_locale,
entree_externes;
char sortie_lointaine,
sortie_voisine,
sortie_locale,
sortie_externes;
char precedente_sortie_lointaine;
char precedente_sortie_locale;
int nb_E2;
};

struct relation
{
int nbcoups2;
int nbcoups0;
int experience;
int depot;
int borne_inf_x, borne_sup_x;

```

```

int borne_inf_y, borne_sup_y;
struct colonne *soma;
struct relation *recepteur;
struct relation *suivant;
};

```

Pour la structure colonne, le *type* représente le type de l'aire de la colonne, motrice, sensorielle ou associative. Les *struct relation *liens_* renvoient à des listes chaînées de connexions avec les relations correspondantes des colonnes cibles. Ces relations sont structurées suivant la provenance de l'information (voisine, locale, externe, lointaine). De même, les *entrées_* et les *sorties_* respectent cette différenciation, et contiennent la valeur de l'état correspondant. Le calcul de l'activité sur les entrées lointaines nécessite la valeur de la *précédente sortie lointaine*; il en est de même pour les liens locaux, eux aussi temporels. Enfin, les coefficients probabilistes P0 et P2 sont calculés à partir du nombre de fois où la sortie lointaine de l'unité s'est trouvée dans l'état E2 (*nb_E2*).

La structure des relations doit être suffisamment générale pour représenter les différents types d'apprentissages. C'est pourquoi, nous avons défini le champ *nbcoups* pour évaluer, en rapport avec le *nb_E2* de la structure colonne, les coefficients probabilistes P0 et P2. La mesure de l'*expérience* s'est révélée nécessaire pour n'accorder crédit qu'aux relations suffisamment "vieilles" et dont la valeur des coefficients d'apprentissage reflétait une mesure et non une initialisation. L'introduction de ce champ constitue une "astuce informatique" afin de réduire les calculs inopportuns mais n'est pas nécessaire à la convergence du réseau. C'est la raison pour laquelle nous n'avons pas fait mention de cet aspect lors de la modélisation. Le champ *depot* permet à une relation de disposer d'une "boite aux lettres" où la relation de la colonne cible qui lui correspond peut déposer son appel sans que l'on ait à évaluer tout de suite les conséquences de cette transmission. Les champs *borne_inf* et *borne_sup* permettent, pour les relations locales, de définir les intervalles d'éloignement selon deux directions que nous avons appelé x et y. Enfin, il est évident qu'une relation doit pouvoir accéder à sa colonne pour y chercher ou y déposer des informations. Le pointeur **soma* indique l'adresse de cette colonne. De même, **recepteur* pointe sur la relation adéquate de la colonne cible. Finalement, la création et l'utilisation d'une liste chaînée de relation (et ceci pour les quatre types de relations) dépendent du pointeur **suivant* qui permet d'avoir accès, si besoin est, à

une nouvelle relation du même type.

4.6.2 Algorithmes

Parmi les nombreuses fonctions définies dans les programmes, nous n'en retiendrons que trois, qui décrivent les mécanismes de fonctionnement essentiels. Les fonctions annexes de visualisation, d'initialisation ou de prétraitement ne seront donc pas détaillées ici.

Création du réseau

La géométrie du réseau simulé est déterminante pour la création des différents liens qui relient les unités. A partir de la donnée des différentes aires impliquées dans le processus à simuler, il faut construire les listes chaînées de relations correspondantes. Suivant la connectivité décrite précédemment, et sachant que les aires sont représentées sous forme de tableaux bidimensionnels de structures "colonne", l'algorithme de la fonction création peut se décrire comme suit:

Fonction Creation:

/* Pour chaque aire */

Pour a variant de 1 à nb_aire **faire:**

/* Pour chaque colonne */

Pour i variant de 1 à taille_aire **faire:**

Pour j variant de 1 à taille_aire **faire:**

/* liens voisins */

Pour k variant de i-1 à i+1

et l variant de j-1 à j+1 **faire:**

connecte(colonne[a][i][j].liens_voisins,colonne[a][k][l])

Fin pour

/* liens locaux */

Pour k variant de i-1 à i+1

et l variant de j-1 à j+1 **faire:**

connecte(colonne[a][i][j].liens_locaux,colonne[a][k][l])

Fin pour

/* lien externes */

t = taille_champ_recepteur/2

Pour k variant de i-t à i+t

```

    et l variant de j-t à j+t faire:
        connecte(colonne[a][i][j].liens_externes, carte_de_codage[a][k][l])
    Fin pour
    /* lien lointains */
    Pour chaque aire b connectée avec a faire:
        connecte(colonne[a][i][j].liens_lointains, colonne[b][i][j])
    Fin pour
Fin pour
Fin pour
Fin pour

```

Avec:

Fonction Connecte(liste_de_liens, colonne):
créer une relation
initialiser cette relation avec l'adresse de colonne,
ajouter la relation dans la liste_de_liens

Activation

L'activation du réseau prend en compte le coté dynamique du modèle. Il s'agit ici principalement de déterminer le nouvel état des différentes sorties des unités en fonction de leurs entrées. Sur le principe, ces renseignements nous sont donnés par les différentes tables d'entrées/sorties définies dans la modélisation. Il reste à préciser la manière dont cette évaluation va être entreprise: ici, l'ordre d'évaluation des colonnes est parallèle; à chaque cycle, toutes les unités sont évaluées d'après les paramètres du cycle précédent.

Fonction Activation:

Pour chaque aire **faire:**

Pour chaque colonne c **faire:**

Pour chaque type de lien l **faire:**

Pour chaque relation r **faire:**

affecter l'état de la sortie.l de c dans dépôt de r

Fin pour

Fin pour

Fin pour

Pour chaque colonne c **faire:**

Pour chaque type de lien l **faire:**
 évaluer l'état le plus demandé
 à l'aide de la table correspondante
 affecter cet état à la sortie de c correspondante
Fin pour
Fin pour
Fin pour

Apprentissage

Selon le modèle que nous proposons ici, l'apprentissage fait partie intégrante de l'évaluation dynamique du réseau et se déroule donc en même temps que cette dernière. Pour des raisons de simplicité, et sans aucune influence sur le résultat, nous définissons à part les procédures d'apprentissage que nous appelons après chaque activation du réseau. Il s'agit alors, pour chaque type de liens, de regarder les échanges d'état E2 et d'en tirer les conséquences sur la modification des coefficients d'apprentissage. Comme nous l'avons précisé lors de la modélisation, l'apprentissage sur les entrées externes n'améliore guère les performances, aussi ne l'avons nous pas intégré dans cette version de l'implantation.

Fonction Apprentissage:

Pour chaque aire **faire:**

Pour chaque colonne **faire:**

 /* liens lointains */

 /* P2 et P0 sont évalués dans la fonction d'activation, */

 /* avec $P2 = \text{nbcoups2}/\text{nb_E2}$ et $P0 = \text{nbcoups0}/\text{nb_E2}$ */

Si sortie_lointaine = E2 **alors:**

 nb_E2 = nb_E2 + 1

Pour chaque lien_lointain **faire:**

Si recepneur→soma→precedente_sortie_lointaine = E2 **alors:**

 recepneur→nbcoups2 = recepneur→nbcoups2 + 1

Si recepneur→soma→precedente_sortie_lointaine = E0 **alors:**

 recepneur→nbcoups0 = recepneur→nbcoups0 + 1

Fin pour

Fin Si

 /* liens locaux */

```

Si sortie_lointaine = E2 alors:
  nb_E2 = nb_E2 + 1
  Pour chaque lien_local faire:
    Si recep_teur->soma->precedente_sortie_locale = E2 alors:
      recep_teur->nbcoups2 = recep_teur->nbcoups2 + 1
    Si recep_teur->soma->precedente_sortie_locale = E0 alors:
      recep_teur->nbcoups0 = recep_teur->nbcoups0 + 1
  Fin pour
Fin Si
Si sortie_lointaine = E1 alors:
  Pour chaque lien_local faire:
    Si recep_teur->soma->precedente_sortie_locale = E2 alors:
      modifier borne_inf_x, borne_sup_x, borne_inf_y, borne_sup_y
      pour inclure Cx, Cy
    Fin pour
  Fin si
  /* liens voisins */
Si sortie_voisine = E2 alors:
  Pour chaque lien_voisin faire:
    Si recep_teur->soma->sortie_voisine = E2 alors:
      recep_teur->nbcoups2 = recep_teur->nbcoups2 + 1
    Si recep_teur->soma->sortie_voisine = E0 alors:
      recep_teur->nbcoups0 = recep_teur->nbcoups0 + 1
  Fin pour
Fin Si
Fin pour
Fin pour

```

4.6.3 Remarques sur l'implantation

Il est évident que les structures de données et les algorithmes proposés ici ne sont qu'un choix de notre part, ne peuvent se substituer en aucune manière à la modélisation qui précède et en sont uniquement une implantation possible. La forme de cette implantation est largement influencée par le langage que nous avons utilisé et par la formation d'algorithmique que nous avons reçu. C'est peut-être cette formation qui nous a incité, par mo-

ment, à effectuer des "raccourcis saisissants". Ceux-ci correspondent à des petites différences entre le modèle et l'implantation que nous nous sommes accordées lorsque cela simplifiait beaucoup notre travail de programmation sans avoir trop de conséquences sur le comportement général du réseau. De plus, nous tenons à rappeler que cette implantation correspond à la dernière version du modèle (au moment où nous rédigeons ce rapport) et que les premières applications ont été réalisées à l'aide d'une structure de données beaucoup plus pauvre.

4.7 Conclusion sur la modélisation

Il est clair que ce que nous proposons n'est qu'un modèle, c'est à dire une conception plus ou moins formelle qui tend à rendre compte de phénomènes réels. Dès lors, nous ne saurions prétendre à une parfaite adéquation avec la nature; par principe, un modèle ne peut pas représenter exactement le processus qu'il simule. Cette première approche ne prétend pas aboutir à un système complet et achevé mais propose plutôt des mécanismes possibles pour des caractéristiques particulières des réseaux neuro-mimétiques. C'est essentiellement sur les règles d'activation et de mémorisation qu'il reste à peaufiner la logique temporelle, parallèle et locale pour s'orienter vers une approche plus globale. En effet, si la neurobiologie décrit assez précisément la structure statique du système, les sources de connaissances sur les phénomènes dynamiques sont assez limitées. Ce n'est que grâce à une interprétation des éléments en terme de logique et/ou de psychologie cognitive que nous pouvons émettre des hypothèses sur les caractéristiques non décrites par la biologie.

De nombreux perfectionnements peuvent être apportés au modèle dynamique; il semble en effet qu'il y ait une logique sous-jacente dont la mise en évidence permettra d'unifier les tables d'entrées/sorties et les mécanismes d'apprentissage pour les quatre types d'entrées. L'interprétation proposée n'est cependant pas assez stricte pour qu'un passage de l'un à l'autre soit déjà possible.

Cette première modélisation fonctionnelle du cortex, malgré ses faiblesses, a surtout mis l'accent sur un certain nombre de principes qui nous semblent indispensables à la mise au point d'une unité connexionniste orientée vers la simulation des comportements humains et compatible avec les données neurobiologiques. Ces quatre principes devront nous guider par la

suite et être sous-jacents à toute nouvelle modélisation:

1. principe de connectivité locale:

Il nous semble actuellement vain de développer des modèles qui nécessitent une interconnexion totale des unités d'une couche ou d'un réseau; ce n'est ni biologiquement plausible, ni simulable à une échelle pouvant faire émerger des comportements non triviaux.

2. Principe d'intégration temporelle

Un grand nombre de processus humains (compréhension de la parole, explorations visuelles, mouvements) et de comportements (anticipation, imitation) sont fortement liés au temps. Le traitement de l'information dans l'automate doit être implicitement temporel et capable de générer des processus parallèles et séquentiels.

3. Principe de décomposition fonctionnelle

Le réseau doit être capable de travailler sur plusieurs flux d'informations (visuel, auditif,...). Ces flux organisent une décomposition en aires fonctionnelles spécialisées par leurs entrées externes. Le fonctionnement parallèle de ces aires induit des comportements qui dépendent de leur connectivité. De plus, des bouclages externes de l'information par l'intermédiaire des fonctions sensori-motrices (j'entends ce que je dis, je vois ce que j'écris) doivent compléter et renforcer l'apprentissage.

4. Principe de haut niveau d'intégration

L'automate doit localement avoir un fonctionnement complexe pour émuler des processus fonctionnels plus évolués que ceux proposés par les neurones formels.

Traditionnellement les systèmes hybrides font intervenir un réseau connexionniste pour les traitements de bas niveau, sur lequel vient se greffer, un peu artificiellement, un système symbolique quelconque. Notre approche est opposée: l'hybridation symbolique/traitement du signal s'effectue de façon unifiée à l'intérieur même de l'unité "symbolico-connexionniste".

Suivre ces principes nous conduit à une modélisation plus ou moins précise et rigoureuse. Comment juger des capacités cognitives du modèle?

La validation peut s'effectuer grâce à des simulations pour des problèmes particuliers. Dans les chapitres qui suivent, notre démarche sera multiple:

- présentation d'une application cognitive particulière,
- introduction d'une caractéristique du modèle et des mécanismes induits,
- donnée des résultats expérimentaux validants.

L'automate s'enrichira des propriétés abordées, de simulation en simulation.

Les simulations vont soulever des problèmes spécifiques liés, non pas à la modélisation du cortex, mais à celle des structures coopérantes qui codent les entrées/sorties. Leur importance est fondamentale, puisque les aires corticales se spécialisent d'après leurs entrées et que le comportement global en dépend. Nous aurons donc à revenir sur des données expérimentales biologiques ou psychologiques qui nous donneront quelques pistes pour l'élaboration des codages initiaux.

CHAPITRE III:

APPLICATIONS

1 Vision

1.1 Présentation du problème

La première application que nous avons décidé de réaliser concernait la reconnaissance de formes. A cela, plusieurs raisons peuvent être données:

-Notre appartenance à l'équipe RFIA (Reconnaissance de Formes et Intelligence Artificielle), ainsi que notre stage de DEA sur un sujet de reconnaissance de formes nous avaient donné une expérience et un certain goût pour les problèmes d'analyse et de traitement d'images.

-Comme nous l'avons mentionné dans le chapitre précédent, le traitement d'images est un sujet de prédilection du connexionnisme et nous y vîmes là un bon moyen de confrontation de nos résultats et de ceux, largement diffusés, d'applications semblables utilisant des techniques différentes.

-L'application ne requiert pas d'interprétation mais demande uniquement des propriétés de mise en correspondance. Or, dans le modèle théorique que nous venons de décrire, cette propriété n'a pas été explicitement démontrée. Il s'agissait donc de mesurer les capacités de mémoire associative de notre modèle.

La reconnaissance automatique de caractères, qu'ils soient imprimés ou manuscrits, correspond à un réel besoin dans un monde qui utilise et qui produit de plus en plus de documents écrits. Analysons d'abord le domaine d'application et ses difficultés qui sont bien connues de ceux qui ont déjà utilisé des algorithmes de reconnaissance de formes classiques.

Pour une réalisation connexionniste de cette application, le problème consiste à faire de l'hétéro-association puisqu'il s'agit de mettre en corres-

pondance une image contenant une information (par exemple une lettre de l'alphabet) avec sa signification (par exemple: "ceci est un A"). Des algorithmes classiques de traitement d'images et de reconnaissance de formes ont déjà abordé le problème (voir par exemple Baird & Thompson, 1987; Baird, 1988; Bozinovic & Srihari, 1989; Berthod & Maroy, 1974). Ils demandent une description quantitative et qualitative de chacun des signes à reconnaître, ainsi que des traitements sur l'image, capables d'extraire les éléments de cette description. Ils se sont tous heurtés à la difficulté principale de ce genre d'application: la variabilité des données. En effet, le but d'une telle réalisation est la lecture de textes imprimés ou même manuscrits. Nous n'aborderons pas ici le problème de la localisation des lignes et des différents caractères qui la composent. Quand le caractère a été localisé, sa reconnaissance est rendue difficile par les différentes polices qui ont pu être employées, sa taille qui peut changer, des effets de translation ou de rotation ainsi que des pertes d'informations dues à un défaut d'impression ou de saisie de l'information, tout ceci dans le cas de textes imprimés. Dans le cas de textes manuscrits, à tous ces problèmes viennent s'ajouter la variabilité due à un seul scripteur, selon les conditions physiques et psychiques dans lesquelles il écrit, ainsi que la variabilité d'écriture d'un scripteur à l'autre. Ainsi, toutes ces manières différentes de représenter un même caractère rendent très difficiles une description morphologique fiable a priori. De plus, un homme à qui l'on demande de réaliser cette tâche ne pourra s'empêcher d'y introduire une partie d'interprétation alors qu'il s'agit ici uniquement de reconnaissance. Il est possible, de plus, qu'il reste totalement insensible à des propriétés morphologiques importantes qu'il réalise, en tant que lecteur, de manière inconsciente.

Ainsi, un des intérêts de notre méthode et du connexionnisme en général, est qu'aucune description n'est fournie a priori. C'est le système qui apprend seul à partir d'un ensemble d'exemples qui lui est fourni et qui doit lui permettre de généraliser cet apprentissage pour reconnaître d'autres exemples qu'il n'a pas appris.

Cet apprentissage qui doit permettre au système d'extraire lui-même les éléments pertinents, utiles pour différencier les différents signes à reconnaître, n'est pas suffisant pour analyser l'image du premier coup. Il doit s'accompagner, selon nous, d'une possibilité de retour sur l'information permettant plusieurs hypothèses possibles pour reconstruire l'image et surmonter ainsi la variabilité de l'information proposée. Ce retour sur l'information

est une propriété des systèmes biologiques connue sous le nom de "feedback". Ce n'est d'ailleurs pas le seul emprunt à la biologie que nous avons fait. En effet, l'étude de la colonne corticale et du réseau des aires est insuffisante pour réaliser une application visuelle. Nous avons du également faire une étude bibliographique et interroger différents spécialistes du système visuel des êtres vivants afin de mieux connaître les propriétés des systèmes périphériques (rétine, nerf optique, thalamus) qui vont permettre ensuite l'action des colonnes des différentes aires visuelles.

Nous allons donc tout d'abord résumer brièvement la masse considérable d'information que nous avons récolté sur ce sujet. Nous pourrons alors analyser les conséquences de l'application de tels principes en informatique. Nous tenterons ensuite de voir plus en détails comment la propriété de retour sur l'information s'inscrit dans ce cadre. Nous parlerons enfin de l'implantation de l'application, de la procédure suivie, des tests et des résultats obtenus, pour enfin tirer les leçons de cette première application.

1.2 Le codage

Dans le cadre de notre application, une étude des systèmes visuels des êtres vivants n'a pas pour ambition de modéliser les différents traitements que le cortex effectue sur l'information visuelle et encore moins de simuler des organes aux capacités remarquables tels que l'oeil, mais plutôt de s'inspirer des propriétés et des techniques d'organisation de tels systèmes afin de réaliser un codage initial de l'information qui respecte une certaine logique. Autrement dit, si on considère une colonne corticale comme un automate réalisant une fonction donnée, les questions auxquelles cette étude doit répondre sont les suivantes. Comment et sous quelle forme est proposée l'information visuelle aux aires corticales? Quelles sont les aires qui réalisent la reconnaissance et comment sont-elles interconnectées? Quelle est la taille de ces aires?...

1.2.1 Données biologiques

La rétine est le capteur spécifique de l'information visuelle (Imbert, 1983) et réagit à des caractéristiques telles que la luminosité, la forme, la couleur ou le déplacement des objets. Elle effectue de plus une série de prétraitements adaptatifs. Deux sortes de cellules spécialisées, les bâtonnets et

les cônes, convertissent les messages lumineux en messages électriques et permettent déjà, par leur localisation et leur densité relatives, de séparer la fonction visuelle en vision périphérique, plus sensible au mouvement, et vision centrale, qui reconnaît et analyse les objets. Pour cette application, nous ne parlerons donc que de vision centrale et attendrons la partie III.4 pour traiter cette dualité essentielle entre les deux fonctions visuelles.

Les deux nerfs optiques issus des deux rétines se rencontrent et se dirigent vers le thalamus de manière à y conduire les messages correspondant au champ visuel droit dans le corps genouillé latéral gauche et gauche dans le CGL droit (Hubel & Wiesel, 1978). Cette opération est faite de manière à conserver une certaine identité à l'information. En effet, il n'y a pas mélange des signaux venant des deux yeux et, qui plus est, les fibres viennent d'une même rétine de manière topologique. Le même processus est décrit pour la projection du CGL vers le cortex visuel primaire (aire 17 selon Brodmann). En effet, de très nombreuses études ont montré la nature et la disposition de l'information reçue dans la couche IV de l'aire 17 (Hubel & Wiesel, 1962; Imbert, 1983).

Une première analyse peut porter sur la provenance de cette information. On observe alors deux phénomènes.

-Selon le principe de rétinotopie, les fibres géniculées se projettent sur l'aire visuelle primaire en conservant la topologie de la rétine.

-Les informations provenant de l'oeil droit et de l'oeil gauche ne sont pas mélangées et se projettent dans des zones distinctes.

Il nous est ainsi possible de diviser le cortex visuel primaire en éléments de 1mm^2 recevant la même information de la même partie du champ visuel (Hubel & Wiesel, 1978). Une partie de champ visuel ainsi délimité est appelé champ récepteur. Une analyse plus fine d'un de ces éléments nous révèle alors un troisième phénomène:

-par le biais d'interneurones excitateurs et inhibiteurs, chaque colonne réalise une fonction de transfert sur le flux d'information provenant du champ récepteur considéré. Cette fonction de transfert peut être interprétée comme un masque apposé sur le champ récepteur correspondant. Ce masque comprend une partie centrale excitatrice et une partie périphérique inhibitrice. La forme de cette partie centrale permet à la colonne de détecter un segment orienté précis. D'une colonne à l'autre, cette orientation préférentielle varie régulièrement dans un sens déterminé (Hubel & Wiesel, 1977).

On peut alors donner une première interprétation de l'analyse que le cortex visuel primaire effectue sur l'image qui lui est soumise. Il s'agit, pour lui, de traduire l'image en un ensemble de segments orientés, localisés précisément. Les mêmes types d'architecture et de connectivité sont décrits pour le cortex visuel secondaire; mais, cette fois, l'information arrivant dans la couche IV provient de l'aire visuelle primaire (Van Essen & Maunsell, 1983). Il y a donc encore conservation de la rétinitopie. De plus, les masques réalisés par les colonnes de l'aire visuelle secondaire leur permettent de rendre compte de la coactivation d'un ensemble déterminé de colonne de l'aire visuelle primaire. Ces masques, rapportés aux indices qu'ils extraient de l'information recueillie par la rétine, peuvent rendre les colonnes de l'aire visuelle secondaire révélatrices de la présence de segments orientés de longueur plus importante.

L'image que nous venons de donner du cortex visuel est extrêmement partielle. En effet, celui-ci est le siège de nombreuses autres analyses aussi pertinentes et réalisées de manière aussi astucieuse que celles que nous venons de décrire (couleur, mouvement...; Poggio & Koch, 1987). Il nous a semblé utile ici de nous limiter aux éléments dont nous nous sommes servis pour réaliser nos programmes. Les implications de l'utilisation de tels codages dans un programme informatique sont discutées ici.

1.2.2 Analyse informatique

L'analyse par un informaticien de l'ensemble de données que nous venons d'énumérer conduit à une série de constatations qu'il est utile de percevoir avant de concevoir un modèle implantable:

Certains neurones (dans le système oculomoteur, par exemple) ont une sortie linéaire: leur taux de décharge est proportionnel à un paramètre scalaire (la rotation des yeux, par exemple). Cependant, dans le meilleur des cas, la sortie est discriminable avec un taux de 1/1000 (Ballard, 1986). Ce taux est insuffisant pour le cortex visuel et l'ensemble des paramètres qu'il faut y représenter. C'est la raison pour laquelle les neurones n'utilisent plus un codage en fréquence, mais un codage que l'on pourrait qualifier de spatial: les dimensions à représenter sont partitionnées en valeurs discrètes avec une tolérance exprimée par un intervalle et l'on trouve ainsi des colonnes qui répondent uniquement si leur entrée appartient à cet intervalle (Barlow, 1972). A un niveau plus fin de représentation, le codage fréquen-

tiel est conservé pour indiquer le degré de correspondance de l'entrée avec la valeur discrète représentative de l'intervalle. On a pu remarquer comment cette représentation permet de créer facilement des fonctions plus complètes à l'aide de connexions conjonctives (Feldman & Ballard, 1982). Mentionnons ici que cette technique de codage va se retrouver dans la plupart des cellules du cortex.

La description que nous venons de faire des différents niveaux de traitement de l'information visuelle illustre une solution intéressante retenue par l'évolution. Une manière triviale de réaliser les schémas de câblage consiste à interconnecter toutes les combinaisons de stimuli possibles (N^2 connexions entre N capteurs élémentaires). Cette solution est possible chez la libellule qui possède très peu de capteurs (Olberg, 1981). En revanche, le nombre total de connexions disponibles chez l'homme permettrait à peine d'interconnecter les deux nerfs optiques. Il a donc fallu trouver une autre solution pour lui permettre de bouger, parler, entendre.... Il s'agit de l'introduction de niveaux de traitements intermédiaires et indépendants appelés cartes. Ainsi, au cours de la phylogénèse, le nombre de cartes augmente alors que le nombre relatif de connexions diminue.

Si cette dernière remarque peut être illustrée par l'exemple de cartes analysant spécifiquement la couleur, il n'en reste pas moins que de telles cartes peuvent avoir à représenter un nombre assez grand de paramètres à leur surface. Ainsi, nous avons vu que l'aire visuelle primaire est rétinotopique (position x,y), détecte les orientations et traite sélectivement l'information venant de chaque oeil. Le problème est alors de représenter ces quatre dimensions sur l'architecture bidimensionnelle de l'aire corticale. La solution retenue est la suivante: deux paramètres sont privilégiés et sont représentés sur les axes principaux de la carte. On les appellera indices primaires (Ballard, 1986). Il s'agit, dans le cas visuel, des coordonnées rétiniennes. Les autres paramètres sont représentés en indices secondaires pour chaque valeur possible des indices primaires. La figure 39 schématise cette distribution. Remarquons que cette solution est rendue possible par le codage spatial, par intervalles.

Le volume de la boîte crânienne étant très limité, la Nature s'est efforcée de limiter au maximum le nombre d'unités à employer au profit des connexions qui prennent relativement moins de place. Selon la distribution que nous venons de décrire, le nombre d'unités augmente exponentiellement avec le nombre de paramètres représentés dans une aire: des dimen-

sions pouvant prendre chacune N valeurs différentes nécessitent N^k unités. L'utilisation de champs récepteurs diminue ce nombre. Si ces champs récepteurs ont un diamètre D défini par rapport à la résolution maximale, le nombre d'unités devient N^k/D^2 .

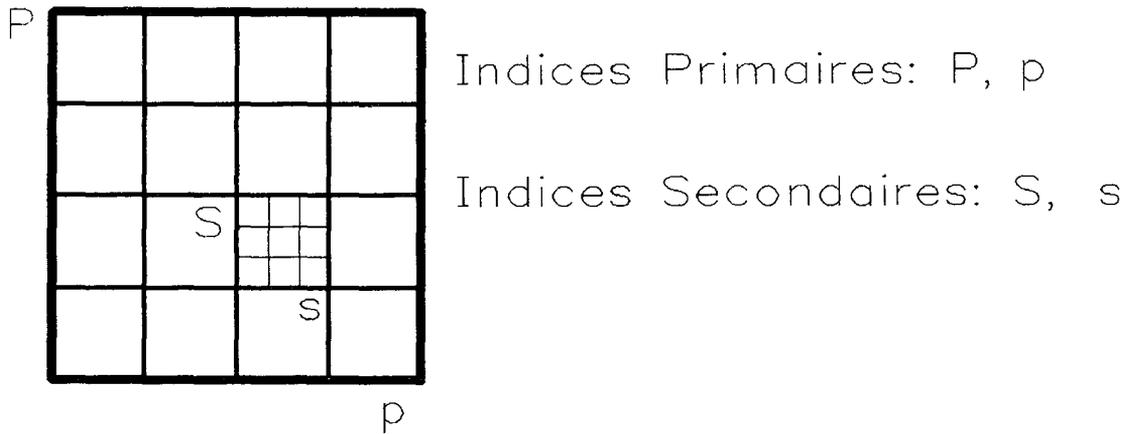


Figure 39. Indices primaires et secondaires.

Cependant, le codage devient alors beaucoup trop grossier. Il existe une solution intermédiaire très avantageuse: il s'agit d'utiliser des champs récepteurs avec recouvrement partiel.

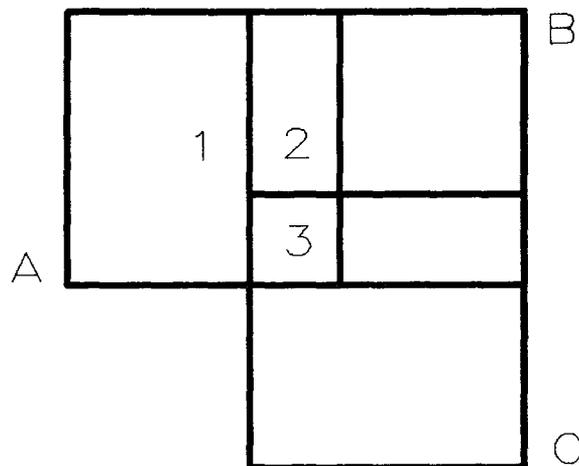


Figure 40. Recouvrement.

Si r est le taux de recouvrement de ces champs, le nombre d'unités nécessaire est $N^k/((1-r)D)^2$. Comme le montre la figure 40 pour $D=3$ et $r=1/3$, la précision est alors remarquable. Si le stimulus est en 1, A est active seule. En 2, A et B sont actives; en 3, A, B et C sont actives. Cette technique de codage a été retenue bien qu'elle présente à priori une légère perte d'information, comme l'illustre la figure 41.

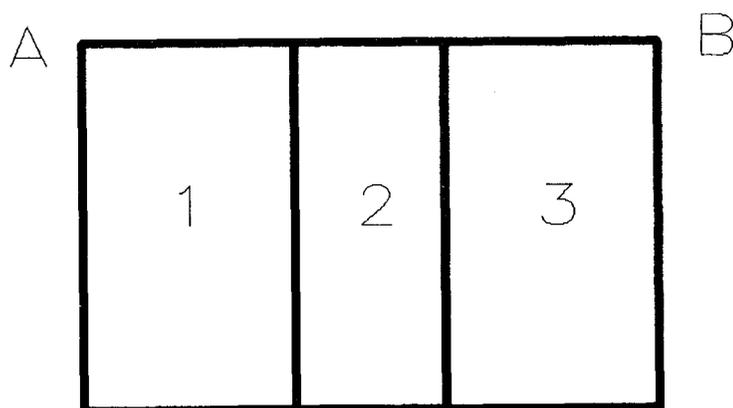


Figure 41. Inconvénient du recouvrement.

La présence de stimuli en 1 et 3 rend A et B actives. Ceci pourrait aussi être interprété comme un stimulus en 2. Cependant, cette propriété dans un système distribué n'est pas forcément gênante. Elle correspond en fait à un codage flou mais pas inexact puisque le stimulus 2 est, par sa position et par son seuil, intermédiaire de 1 et 3. De plus, un empilement de champs récepteurs réalisé par une suite de traitements hiérarchiques permettra aux unités de plus haut niveau, par ce phénomène de divergence, d'avoir accès à une plus grande partie du champ visuel avec, en contrepartie, un certain flou qui est loin de nuire à la généralisation que doivent effectuer les unités.

On vient de le voir, la Nature cherche à faire des économies mais privilégie cependant une certaine richesse de l'information représentée. Dans le même ordre d'idées, les premiers biologistes ont été surpris par le rapport élevé qui existe entre le nombre de neurones du cortex visuel et le nombre de neurones qui partent de la rétine alors que, a priori, le cortex ne crée pas d'information. C'est vrai, mais son rôle est de rendre explicites des informations implicites. Ce travail nécessite un accroissement du nombre

des unités et nous voulons insister une fois de plus ici sur les conséquences de ce codage sur les unités. Dans le cadre visuel, une colonne représente de manière explicite une certaine caractéristique de l'image perçue. L'activité d'une colonne devra donc être traduite par: "il y a un segment orienté à 15 degrés perçu par l'oeil gauche à telle position". C'est la raison pour laquelle une colonne peut (et doit) être considérée comme un **but**.

1.3 Le retour sur l'information

Il semble important de mettre l'accent ici sur cette caractéristique essentielle de la colonne corticale. Cette propriété est très intéressante pour le traitement des données; elle a été très souvent décrite dans le système nerveux en général et dans le cortex en particulier. Cependant, elle n'est à peu près jamais utilisée en connexionnisme, car cette dimension supplémentaire s'ajoute mal au formalisme mathématique existant (voir cependant Fukushima, 1980; Grossberg, 1988). Nous voulons donc montrer ici toutes les capacités qu'apporte ce mécanisme dans le cadre d'une application de mise en correspondance. Dans notre modèle, le retour sur l'information dépend intimement de la manière dont l'information est extraite puisqu'il utilise les mêmes voies, mais à rebours. Il est donc intéressant d'étudier ces deux mécanismes complémentaires à travers leurs supports physiques et leurs logiques avant d'en donner une interprétation argumentée.

1.3.1 Le support

Dans un réseau à couches effectuant une mise en correspondance, le retour sur l'information est réalisé par la rétroaction (ou feedback) de la couche de sortie vers la couche d'entrée. Réciproquement, l'information est extraite par proaction (ou feedforward) de la couche d'entrée vers la couche de sortie. Van Essen (Van Essen & Maunsell, 1983) a longuement étudié ces deux voies d'information dans le cortex. Cette étude, ramenée à notre modèle impliqué dans un travail de mise en correspondance, permet d'identifier ces voies: il s'agit des liaisons externes et lointaines. Ainsi, si nous excluons les liaisons locales et voisines, le squelette de réseau ainsi obtenu, représenté par la figure 42, permet d'étudier le support de la pro et de la rétro-action.

Dans le sens de la pro-action, l'information est transmise par une cascade de champs récepteurs. Il y a convergence de l'information en entrée

vers les unités de sortie. Cette convergence permet à chaque unité de sortie de recevoir des informations de la carte d'entrée dans son ensemble avec une connectivité réduite. Dans le sens de la rétro-action, le même principe de structure en champs récepteurs permet potentiellement à chaque unité de sortie d'effectuer un retour sur l'information n'importe où sur la carte d'entrée, par divergence. Des études anatomiques de ces voies dans le cortex ont montré que, globalement, le rapport convergence/divergence égale 1 (Van Essen & Maunsell, 1983).

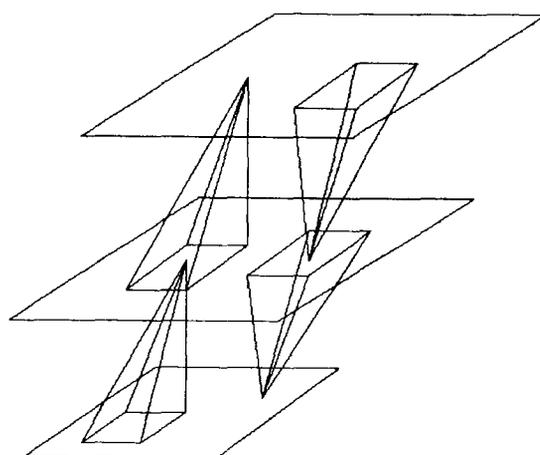


Figure 42. Pro-action et rétro-action.

Autrement dit, si une colonne d'une carte est liée à n colonnes d'une autre carte, il en sera réciproquement de même. Cela veut-il dire que la nature accorde autant d'importance à la rétro-action qu'à la pro-action? Nous nous sommes en tous cas efforcés de respecter cette propriété par le choix adéquat de la taille des champs récepteurs.

1.3.2 La logique

La logique de transmission est bien résumée par les tables 33, 35 et 36. Il nous semble donc plus important de mettre l'accent ici sur les conséquences de l'apprentissage sur une première interprétation du rôle de ces deux mécanismes.

Soient A et B deux colonnes interconnectées appartenant à deux aires différentes telles que B puisse envoyer une rétro-action sur A (par exemple B est une unité de sortie et A appartient à l'aire associative). On a vu que les liens réciproques et non symétriques entre deux unités sont évalués de la manière suivante: quand une unité reçoit E2, elle réévalue la probabilité P2 si elle est à E2 et P0 si elle est à E0.

Pour l'unité B, P2 est la probabilité que B soit à E2 quand A est à E2. P2=1 signifie que, quand A est à E2, B est toujours à E2. Dans ce cas, A implique B; A est suffisant pour déclencher la réponse B. Si P0=1, B est toujours à E0 quand A est à E2; A implique non B; A est suffisant pour inhiber B.

L'apprentissage contenu dans A est tout aussi intéressant: si P2=1, A est toujours à E2 quand B est à E2. A est nécessaire à B, donc l'absence de A suffira à inhiber B (non A implique non B). P0=1 signifie que A est toujours à E0 quand B est à E2. A est donc suffisante pour inhiber B (A implique non B).

Apprentissage	Interpretation
$P2(A/B)=1$	$\text{non}A \implies B$
$P0(A/B)=1$	$A \implies \text{non}B$
$P2(B/A)=1$	$A \implies B$
$P0(B/A)=1$	$A \implies \text{non}B$

Figure 43. Interprétation de la logique d'apprentissage.

L'angle logique sous lequel nous venons de revoir l'apprentissage nous permet maintenant la description suivante: dans le sens de la pro-action, A peut inhiber B par sa présence (E2) si son P2=1, ou par son absence (E0) si P0=1.

Lors d'une phase de reconnaissance, on le verra, toutes les unités de sortie sont mises dans l'état E1 correspondant à une hypothèse, ou une demande. Dans le sens de la rétro-action B peut être déclenché par A à E2 si son P2=1 ou inhibé par A à E2 si P0=1. Si une telle relation existe et que l'état de A ne permet pas cette action, B va envoyer, grâce à son état E1, une rétro-action sur A lui demandant de réévaluer son état.

Toujours dans la même direction, A répercute cette demande vers la carte d'entrée. Cette demande permettra, si le stimulus en entrée est incertain (état E1), de confirmer le stimulus et de répercuter les conséquences de cette confirmation vers la sortie.

1.3.3 Interprétation

L'analyse que nous venons de faire permet de dégager un mot-clé commun à ces deux mécanismes: logique. Ils utilisent en effet tous les deux la logique structurelle et relationnelle de l'information.

Dans le sens de la pro-action, la logique structurelle est utilisée dès les premières unités qui essaient d'extraire des segments orientés de l'image. A l'opposé de certains autres modèles (Fogelman Soulie, Gallinari, Le Cun, Thiria, 1989), notre modèle qui travaille sur la logique structurelle de l'information, est incapable de classifier des vecteurs aléatoires. A notre sens, ceci n'est pas inquiétant pour un modèle dirigé vers la simulation de fonctions humaines...

La logique relationnelle dans le sens de la pro-action permet de dégager les conditions nécessaires au déclenchement ou à l'inhibition de réponses.

Réciproquement, la rétro-action permet de vérifier les conditions suffisantes pouvant mener à des décisions semblables. De plus, la rétro-action peut se transmettre vers les couches d'entrées afin de demander une vérification sur la structure de l'information.

On voit donc, dès cette première application, que la colonne corticale a des capacités ne se limitant pas à la mise en correspondance. Nous avons déjà mentionné que nous décrivons ici une application de reconnaissance de formes et ne cherchons pas à faire d'interprétation de l'image analysée. Ainsi, si une forme est ambiguë, il ne s'agira pas de dire qu'il s'agit d'un "u" et non d'un "v", car elle est précédée de "q". Nous venons néanmoins de voir que si, dans le cadre de cette application, nous ne pouvons pas utiliser le contexte, nous pourrions néanmoins utiliser la logique de l'information. Ce qui distingue un caractère alphabétique d'un vecteur aléatoire, c'est que l'information est regroupée selon des directions précises en traits continus servant à tracer ce caractère. Il s'agit de la logique structurelle des signes alphabétiques. Leur logique relationnelle passe par la connaissance, pour chaque signe, de l'ensemble des traits qui le composent. Il nous paraît donc évident qu'une unité connexionniste qui, par apprentissage, évalue et utilise

ces deux logiques, aura un comportement et des capacités plus intéressantes que des unités qui, par le calcul, évaluent des hyperplans pouvant séparer des classes, indépendamment de la nature de l'information traitée (Davallo & Naim, 1989; Oja, 1989). C'est ce que nous allons nous efforcer de montrer dans la partie suivante qui détaille la réalisation du réseau et les résultats obtenus.

1.4 Réalisation

La nature de l'application nous désigne tout naturellement la fraction du réseau des aires que nous allons simuler. Il s'agit de l'axe visuel temporal. De cet axe, nous extrayons une aire de traitement visuel (V2) et une aire associative (VT). Nous avons également ajouté une aire de traitement visuel primaire (V1, correspondant à l'aire 17 du cortex).

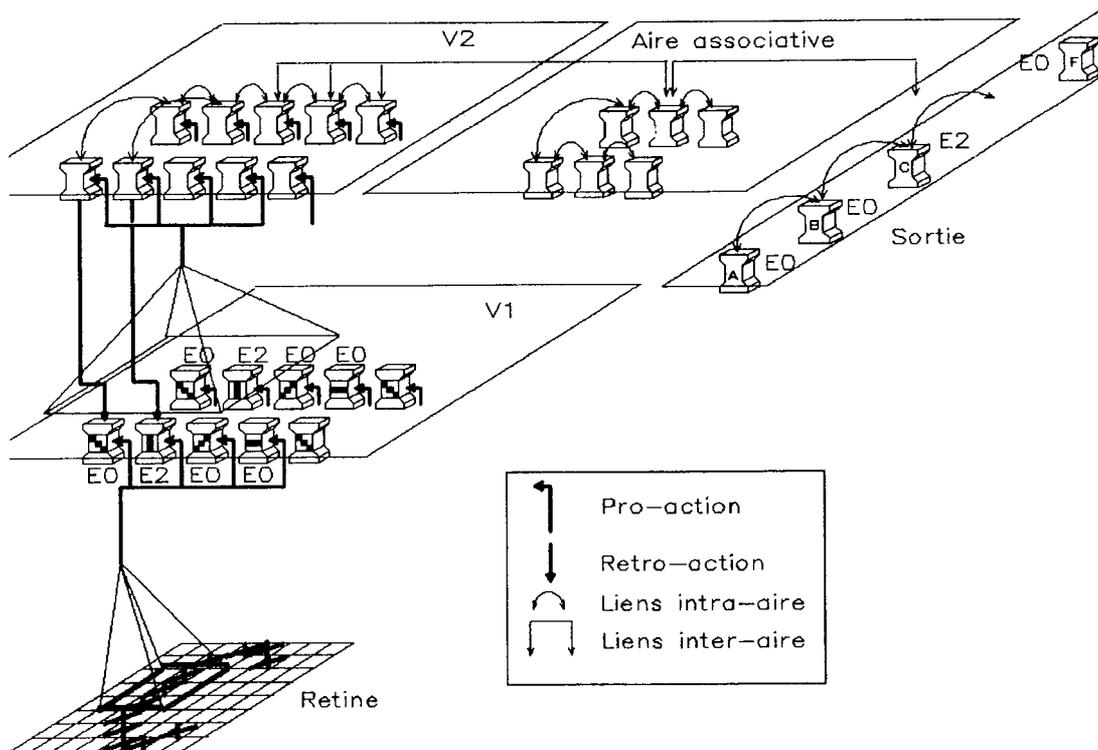


Figure 44. Réseau pour la reconnaissance des caractères.

Cette application avait deux buts: évaluer les capacités de mémoire associative du modèle et comparer ce modèle avec les autres techniques dans des conditions expérimentales semblables (Fogelman Soulie, Gallinari, Le Cun, Thiria, 1989). Ces raisons, ajoutées à la complexité du problème, nous ont fait renoncer à une sortie motrice distribuée au profit d'une "aire de la signification" regroupant les sorties possibles. De même, la couche d'entrée a été choisie de taille identique à celle décrite dans les autres réalisations connexionnistes. Ces différents éléments du réseau, représentés dans la figure 44, sont maintenant décrits plus précisément.

1.4.1 La structure du réseau

La couche d'entrée est une représentation très grossière de la rétine. Son rôle est d'être le support de l'information et de l'envoyer vers les aires corticales. Cette information est une image binaire représentant un caractère alphabétique digitalisé sur une matrice 8x8. La taille de cette matrice a été choisie de manière à correspondre aux données de simulations portant sur le même problème. Les caractères seront représentés dans cette matrice, de manière à occuper toute la place, avec une orientation donnée (figure 45). Autrement dit, nous ne traitons pas ici des problèmes d'invariance en taille, rotation et translation.

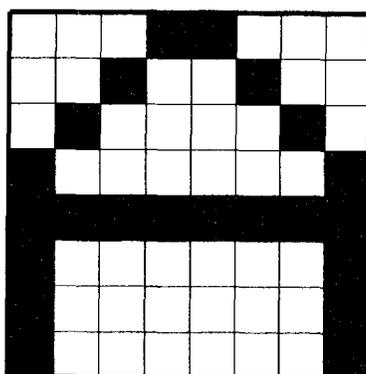


Figure 45. Exemple de caractère.

L'aire visuelle primaire (V1) reçoit l'information à sa surface de manière rétinotopique (indices primaires). Les indices secondaires consistent

en des variations d'orientation pour une position donnée. Chaque colonne de V1 possède des connexions externes avec un voisinage de la carte de codage (il s'agit ici de la couche d'entrée) correspondant à un champ récepteur 3x3. La fonction de transfert des champs récepteurs réalise une détection de segments orientés au moyen des masques suivants:



Figure 46. Masques pour la détection d'orientation.

La relative simplicité du signal visuel analysé ici nous a conduit à détecter uniquement ces quatre directions principales. Ainsi, pour une position rétinotopique donnée, les indices secondaires demandent la définition de quatre colonnes, chacune cherchant à détecter une direction précise dans leur champ récepteur commun.

C'est entre ces quatre colonnes que vont se développer les connexions voisines. Les connexions locales sont définies entre les colonnes ayant un champ récepteur voisin et une fonction de transfert détectant la même orientation. On verra comment, dans l'aire visuelle secondaire, ces liens permettront de détecter des segments orientés plus longs. Notons ici que des champs récepteurs voisins se recouvrent partiellement au tiers. Les connexions lointaines sont réalisées avec des colonnes de V2 que l'on va étudier maintenant.

L'aire visuelle secondaire (V2) effectue sur l'aire visuelle primaire le même type de traitement que cette dernière réalise sur la couche d'entrée. Il y a donc, là aussi, rétinotopie en indices primaires et orientation en indices secondaires. Les remarques du paragraphe précédent sont donc valables pour les colonnes de V2 qui ont leurs entrées externes sur V1, détectent ainsi des segments orientés et ont des relations voisines et locales avec d'autres colonnes de V2 selon le même principe. L'accent doit être mis ici sur la cascade de détections locales qui aboutissent, dans V2, à un apprentissage rendant compte, indépendamment d'un signe alphabétique particulier, de la logique structurelle de l'information. Enfin, des liaisons lointaines sont réalisées avec un ensemble de colonnes de l'aire associative.

La couche de sortie va être décrite d'abord pour rendre l'explication de l'aire associative plus simple. Cet ensemble de colonnes qui fournit la réponse du réseau, est composé de 26 unités représentant chacune une lettre de l'alphabet. Les liaisons voisines développées entre ces unités seront inhibitrices (un signe ne peut être à la fois un A et un Z). Ces liaisons apprises permettront, lors du succès d'une hypothèse, d'inhiber les hypothèses concurrentes. Enfin, chaque colonne de cette couche a des connexions lointaines avec l'aire associative de manière à avoir accès à l'ensemble de la couche d'entrée.

L'aire associative se trouve entre l'aire visuelle secondaire et la couche de sortie avec lesquelles elle développe des connexions lointaines. La relation lointaine avec l'aire secondaire permet aux colonnes de l'aire associative de détecter, en indices secondaires, des angles, combinaison locale sur V2 de deux segments orientés. Des liaisons voisines sont établies dans l'aire associative entre les colonnes représentant ces indices secondaires. Les liaisons voisines sont réalisées entre colonnes ayant, sur V2, des champs récepteurs voisins.

Ayant vu la structure du réseau, nous pouvons maintenant décrire son comportement dynamique et en donner une interprétation.

1.4.2 Le fonctionnement du réseau

La structure du réseau ayant été décrite, les différentes règles du modèle théorique devraient suffire à spécifier le comportement du réseau. Nous allons néanmoins le décrire ici pour deux raisons. Tout d'abord, les simulations d'un modèle parallèle sur une machine séquentielle demandent des éclaircissements. Ensuite, il est intéressant, d'un point de vue cognitif, de décrire un fonctionnement non plus en termes de règles produites par des tables de vérité mais plutôt en termes comportementalistes.

Démarche suivie

Nous avons retenu les caractéristiques suivantes pour permettre la simulation. Au début de l'activation du réseau, l'état initial des unités est forcé en entrée (l'image) et en sortie (voir plus loin). Les autres unités ont un état initial nul. L'ordre d'évaluation des unités est parallèle: à chaque cycle, chaque unité est évaluée d'après les paramètres du cycle précédent, y compris, comme le spécifient les tables de vérité, son propre état précédent. Nous décrivons maintenant les cycles que nous faisons suivre au

réseau dans les deux cas de fonctionnement: l'apprentissage et la reconnaissance. Lors d'une session d'apprentissage, une image est proposée sur la couche d'entrée. Toutes les unités de la couche de sortie sont forcées à E0, sauf l'unité correspondant à la solution apprise, qui est mise à E2 (but réalisé). Pour ne pas privilégier la pro-action par rapport à la rétro-action, nous évaluons ensuite, et dans cet ordre: V2, V1, la sortie et l'aire associative. Nous effectuons ce cycle quatre fois pour que l'information puisse se propager dans tout le réseau.

Les mêmes cycles sont effectués lors de la reconnaissance. Seule différence, l'image à reconnaître est proposée en entrée et toutes les unités de sortie sont mises dans l'état de recherche active E1.

Interprétation

Dans le cas de la reconnaissance, l'aire associative, informée par les aires sensorielles, est capable d'inhiber les colonnes de la sortie dont les indices nécessaires à leur déclenchement sont absents. En retour, cette même aire associative fait évoluer, à partir de l'état de demande E1 des colonnes de la sortie, la distribution de l'activité dans les aires sensorielles vers une répartition déjà apprise. La configuration la plus probable va ainsi se dégager et faire basculer, via l'aire associative, la colonne correspondante de l'aire de sortie dans l'état stable E2 qui inhibera les colonnes concurrentes.

Lors de l'apprentissage, les mêmes cycles sont effectués afin que le système mémorise les relations entre les colonnes de toutes les aires. C'est ainsi que l'on voit apparaître selon les aires et les connexions concernées, la logique structurelle correspondant aux invariants du monde extérieur et la logique relationnelle qui permet plus particulièrement l'apprentissage des caractères alphabétiques.

Résultats

Le premier corpus d'apprentissage est constitué d'une police de 26 caractères majuscules représentés par une matrice 8x8 en valeurs binaires. Lors de l'apprentissage, chaque lettre est présentée deux fois au réseau.

Le deuxième corpus d'apprentissage reprend le premier en changeant, aléatoirement, pour chaque lettre, 6, 9 et 12 pixels (Fogelman Soulie, Gallinari, Le Cun, Thiria, 1989)

Le corpus de test est construit à partir des lettres originales en changeant aléatoirement n pixels (n variant de 0 à 25). Lors des essais, le réseau reconnaît parfaitement les lettres apprises (quand $n=0$). Quand le bruit augmente, les performances décroissent et l'on note (figure 47) que l'écart

entre les deux couches est relativement faible. On peut interpréter ce fait comme une bonne capacité de généralisation du réseau.

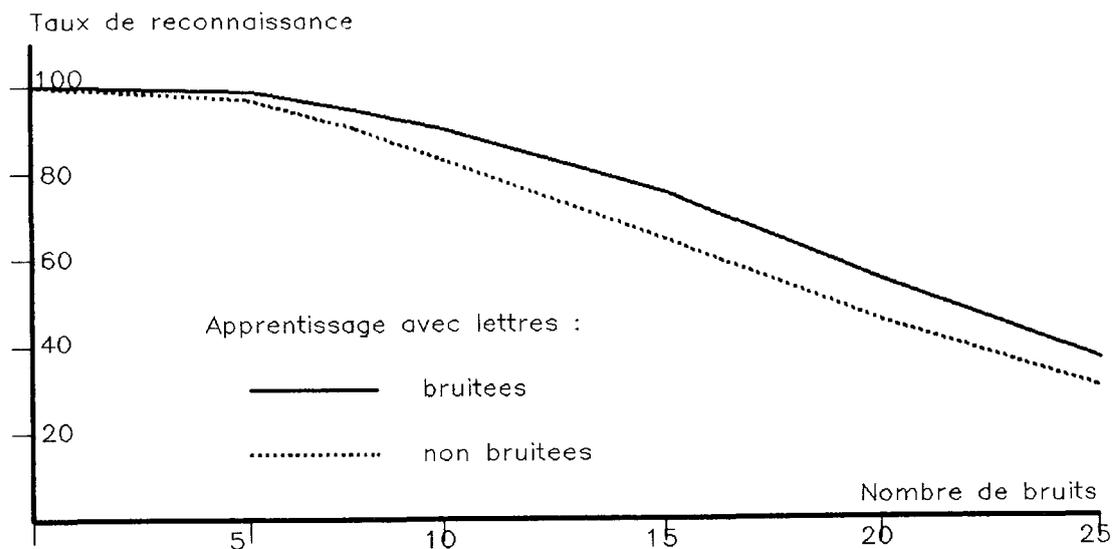


Figure 47. Résultats en reconnaissance de caractères.

Notons ici une propriété très intéressante du modèle. L'apprentissage des 26 lettres dure approximativement 10 secondes sur un SUN 3/50. A l'époque où nous avons réalisé cette simulation, la méthode de rétropropagation du gradient nécessitait un apprentissage de plusieurs heures pour arriver à des résultats similaires (Fogelman Soulie, Gallinari, Le Cun, Thiria, 1989). Aujourd'hui, les recherches d'amélioration de l'algorithme permettent un apprentissage en quelques dizaines de minutes. Notons également que, pour les références citées, le corpus ne portait que sur les 18 premières lettres de l'alphabet (pas de confusion O, Q; U, V).

Nous avons suivi fidèlement la procédure de cette expérience classique qui est, nous le pensons, un bon test des capacités de mémoire associative d'un réseau. Cependant, l'interprétation de cette application comme simulation de fonctions humaines n'est pas évidente. La plupart des erreurs commises lors de la reconnaissance des lettres bruitées était également faite par l'expérimentateur humain; les autres étaient morphologiquement logiques (confusion du E et du F, du O et du Q...). En effet, les données d'origine sont contenues dans une matrice 8x8. Comme on peut inverser

jusqu'à 25 pixels dans cette image, les modifications possibles sont bien supérieures à la distance qui sépare certaines lettres, comme l'illustre la figure 48.

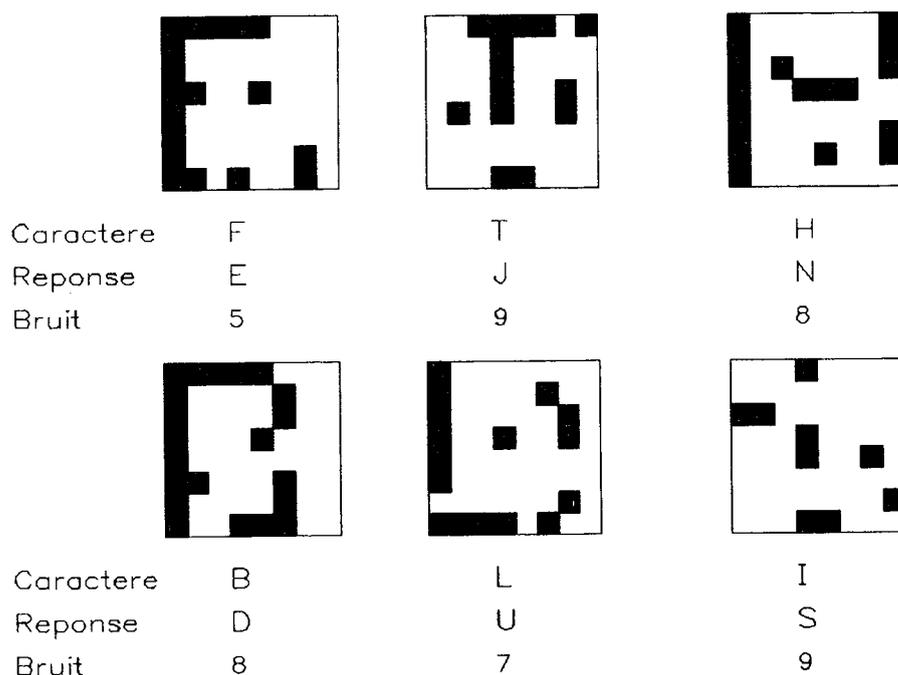


Figure 48. Exemples de confusion.

C'est la raison pour laquelle nous avons "détourné" cette application vers une utilisation qui nous paraissait plus représentative des fonctions humaines: la couche d'entrée n'est plus une image digitalisée sur une matrice 8x8 mais la saisie, à la souris, d'un caractère majuscule manuscrit. Cette application tourne, est testée, et apprend depuis deux ans. Elle fournit, selon l'expérimentateur, des résultats de reconnaissance variant autour de 80%. Elle illustre également une capacité intéressante du modèle: après deux ans d'apprentissage, une lettre peut être désapprise en seulement 4 ou 5 présentations. La figure 49 propose quelques saisies à la souris que le réseau est capable de reconnaître.

Nous avons également fait un autre emploi de ce réseau pour un premier essai de reconnaissance de la parole. Entre le cas visuel et le cas auditif,

seul le codage initial (la couche d'entrée) change et va typer le réseau en l'orientant vers l'apprentissage de différentes logiques structurelles et relationnelles. L'application porte sur la reconnaissance des spectrogrammes des 10 chiffres.

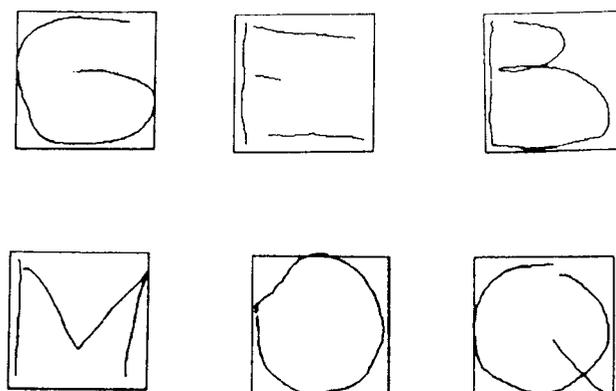


Figure 49. Exemples de succès.

Ce sont donc des spectrogrammes avec 32 canaux de fréquence de 0 à 8000 Hz qui sont présentés sur la couche d'entrée. Dans les aires, les indices primaires représentent la fréquence et le temps alors que les indices secondaires détectent des variations d'énergie pour un intervalle de fréquence et de temps donné (Knudsen, du Lac, Esterly, 1987).

Le corpus d'apprentissage et de test a été réalisé auprès de trente personnes qui ont prononcé chacune 4 fois les 10 chiffres.

Les essais ont été réalisés selon des procédures classiques:

Pour un essai monolocuteur, le réseau apprend une fois seulement les quatre versions des 10 chiffres; il reconnaît les mots appris et commet une erreur de 3% sur les mots tests (Elman & Zipser, 1987).

Lors des essais multilocuteurs, le réseau apprend les 40 spectrogrammes de 2 à 20 personnes. La figure 50 représente les taux de reconnaissance du réseau sur les personnes ayant participé à l'apprentissage (courbe supérieure) et sur des personnes différentes (courbe inférieure). Pour un grand nombre de locuteurs appris, ces deux courbes convergent vers un taux de reconnaissance commun (Peeling, Moore, Tomlimson, 1986).

Nous notons ici aussi la même rapidité d'apprentissage du réseau: sur un SUN 3/50, 40 spectrogrammes sont appris en 15 secondes.

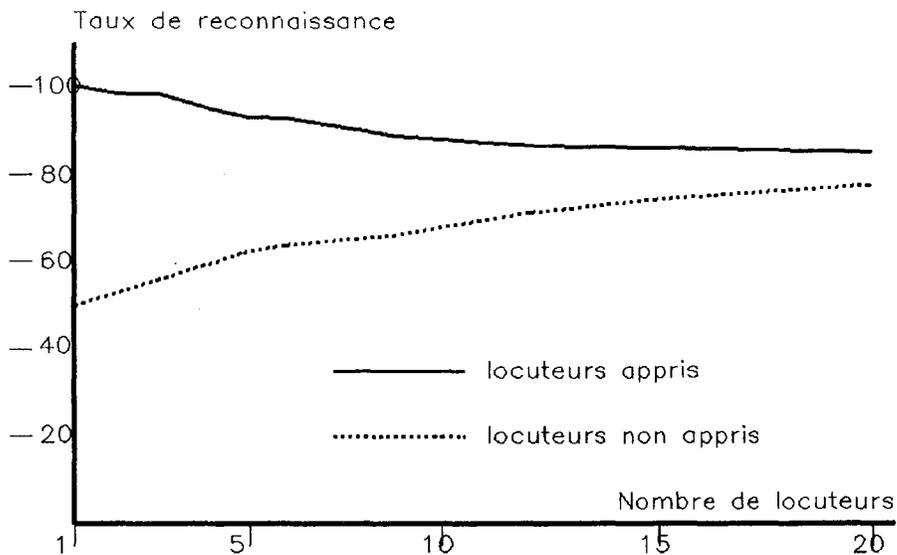


Figure 50. Résultats pour la reconnaissance de mots isolés.

1.5 Discussion

Cette étude et l'expérience acquise au cours de cette première application nous permettent d'en tirer maintenant les enseignements.

La leçon, probablement la plus importante, que nous avons apprise, n'est pas spécifiquement relative au sujet que nous traitons mais concerne la conception que nous avons de notre métier. Nous avons jusqu'alors fait des études bibliographiques sur ce sujet et avons développé le modèle de Y. Burnod dont nous étions très satisfaits. Pour nous, écrire un programme revenait uniquement à valider, ou à prouver, nos idées théoriques. Nous nous sommes en fait rendu compte que, pour un modèle à cette échelle, un programme est un révélateur très cruel de toutes les failles du modèle. A cette époque, il était moins complet que celui que nous venons de présenter mais nous paraissait intellectuellement au point. Le fait d'avoir à spécifier dans un programme tous les paramètres du réseau nous a révélé tous les oublis et les généralisations hâtives que nous avons commis.

Cette application nous a donc obligés à clarifier nos idées et à formuler des questions beaucoup plus précises aux biologistes. Cette remarque de

fond peut d'ailleurs s'appliquer de la même manière aux autres applications, aux publications ou aux thèses.

D'un point de vue plus pragmatique, cette remise en cause ayant permis la réalisation du programme, son rôle validant a pu jouer. Nous avons pu tout d'abord proposer une réalisation qui affirmait les capacités de mémoires associatives du modèle de colonne corticale. Quant aux comparaisons avec les autres modèles existants, les taux de reconnaissance sont honorables, sans plus. Nous avons déjà expliqué que, selon nous, cette procédure de brouillage des lettres entraîne de toutes façons un plafonnement des réussites qui diminuent quand le bruit augmente. De plus, il ne serait pas surprenant qu'un modèle centré uniquement sur l'apprentissage de la répartition de pixels puisse obtenir de meilleurs résultats qu'un modèle apprenant la logique de l'information qui lui est proposée et la représentant afin de l'utiliser éventuellement pour d'autres tâches. La vision que nous avons du connexionnisme nous pousse plus, comme nous le verrons dans les parties suivantes, à développer ces autres capacités, qu'à faire une course au pourcentage sur un genre d'applications qui, de toutes façons, est réalisable par des classifieurs classiques.

Plus intéressante par contre est l'analyse des propriétés de l'apprentissage. Certes, il est frappant de constater un temps d'apprentissage très réduit en comparaison des autres modèles. Il est plus instructif, à notre avis, d'en parler qualitativement: en fait, ces performances s'expliquent par le fait qu'il suffit de présenter une ou deux fois un modèle au réseau pour qu'il l'apprenne, tout comme il est possible de désapprendre, avec 4 ou 5 contradictions, un réseau instruit depuis deux ans. Sans que cet argument puisse servir de preuve, il nous semble cependant que ces caractéristiques sont plus conformes aux normes humaines que les quelques dizaines de milliers d'exemples nécessaires à la convergence d'un perceptron multicouches...

On le voit, notre appréciation du problème nous rend plus sensibles à la simulation des fonctions humaines qu'à la réalisation d'un bon classifieur. C'est pourquoi nous allons maintenant aborder sous cet angle les faiblesses et les améliorations à apporter qui concernent un tel réseau.

A notre avis, sa faiblesse réside dans sa faible tolérance à la translation, à la rotation et à l'agrandissement de l'information à reconnaître. De telles études ont été faites avec le néocognitron de Fukushima (Fukushima, 1980) qui possède une structure en couches proche de notre réseau. Selon nous, la colonne a potentiellement les capacités de résoudre un tel problème mais,

nous le verrons lors de la quatrième application, cette résolution passe par un changement dans le réseau des aires employées.

Le dernier point de la discussion porte sur l'application en reconnaissance de parole que nous avons réalisé avec ce réseau. Cette application nous a confronté avec le phénomène de la parole. Tout d'abord, ce phénomène a une connotation temporelle très forte qui requiert, pour une bonne reconnaissance, une grande tolérance vis à vis de cette dimension. Cette remarque remet en cause l'axe temporel en indice primaire. De plus, il existe des éléments structurels tels que les formants ou les frictions qui doivent être reconnus comme tels avant la reconnaissance du mot qui les contient. C'est ainsi que nous en sommes venus à considérer cette application comme une reconnaissance visuelle de spectrogramme et que nous avons décidé de réaliser une étude plus sérieuse sur le phénomène de la parole (Guyot, Alexandre & Haton, 1989; Dingeon, Alexandre, Guyot, Haton, 1989). C'est cette étude et les modifications qu'elle a impliquées que nous présentons rapidement dans la partie suivante.

2 Reconnaissance de la parole

2.1 Présentation du problème

Dans l'application précédente, l'extraction de l'information est assurée par un empilement de champs récepteurs. Dans le cas de l'application visuelle, le genre de l'information pertinente à extraire permet le choix de champs récepteurs de taille constante. On verra, dans la quatrième partie de ce chapitre, qu'un autre besoin en terme de finesse de l'information peut nous amener à prendre une autre taille pour les champs récepteurs; cependant, là aussi, cette taille peut être décidée une fois pour toute.

Les difficultés que nous avons rencontré, lors du premier essai de reconnaissance de la parole que nous venons de décrire, sont de deux types.

Le premier se rapporte directement à la taille des champs récepteurs. En effet, le phénomène de la parole peut être décrit au moyen d'indices tels que les formants, les bursts, les frictions... Or, ces indices ont des tailles et des positions en fréquence très changeantes d'un phonème à l'autre. Une reconnaissance de phonème doit pouvoir passer par une analyse très fine pour certains indices, tout comme par une détection dans un domaine très large pour d'autres indices. Cette relation qu'un réseau de reconnaissance de la parole doit pouvoir effectuer entre chaque phonème de la couche de sortie et différents domaines de fréquence révélateurs de ce phonème dans la couche d'entrée n'est pas adaptée à des champs récepteurs de taille fixe.

Pour surmonter cette difficulté, on pourrait penser à prédéfinir "à la main" un ensemble de champs récepteurs adapté au moyen d'une étude phonétique. Outre son caractère fastidieux, cette tâche ne serait pas optimale. Il est en effet extrêmement difficile de décider des limites fréquentielles exactes des domaines d'influence des phonèmes. De plus, cette connectivité ne serait pas transportable à une autre application et devrait être remise en cause à la moindre modification dans l'ensemble des phonèmes à reconnaître.

Ces arguments, ajoutés aux indices que nous avons sur la manière dont le cortex réalise ce genre de tâche, nous ont donc amenés à définir, à partir de données neurobiologiques, un mécanisme de différenciation automatique, par apprentissage. Ce mécanisme général que nous avons d'ailleurs réemployé pour d'autres applications va être décrit dans un moment.

Le deuxième type de difficultés induit par notre première approche est

relatif aux dimensions que nous avons reportées sur les axes de la carte d'entrée. En effet, le choix de la dimension temporelle n'est pas judicieux pour un phénomène aussi sujet à variations dans le temps que la parole. Nous avons donc étudié puis proposé une nouvelle carte d'entrée dont les indices primaires sont tous les deux fréquentiels; le temps est pris en compte par la dynamique du réseau. Cette étude, ainsi que la description du réseau, de son utilisation et de ses performances sont proposées en détails dans (Guyot, 1990; Digeon, Alexandre, Guyot, Haton, 1989). Nous mettons uniquement l'accent dans cette partie sur le très important phénomène de différenciation automatique, ou division.

2.2 La division

2.2.1 Données neurobiologiques

La différenciation topologique ou détermination par apprentissage de liens pertinents est permise par la très grande plasticité des aires associatives (Hubel & Wiesel, 1977). Un remodelage de ces cartes autour de stimuli significatifs lors d'un travail de conditionnement permet de relier ces stimuli avec la réponse souhaitée (Fregnac & Imbert, 1984; Merzenich, Jenkins, Middlebrooks, 1984).

2.2.2 Réalisation informatique

Au début de l'apprentissage, l'aire associative n'est pas différenciée. L'auto-organisation fonctionnelle de cette carte est réalisée par une série de divisions effectuées automatiquement quand elles sont nécessaires. A n'importe quel niveau de division, l'opération s'effectue en trois étapes:

- Une série de phonèmes est présentée et permet la mémorisation dans chaque zone des coefficients P0 et P2.

- Pour un signal donné, le réseau fait une tentative de reconnaissance.

- En cas de mauvaise réponse, certaines zones vont être divisées.

Ce sont les zones en déclenchement conditionnel avec ce phonème (P0 et P2 différents de 0 ou de 1) qui vont être divisées, chacune en deux sous-zones dont on peut espérer qu'elles seront plus révélatrices d'un indice du phonème (inhibition ou déclenchement inconditionnel). Ce phénomène est illustré par la figure 51.

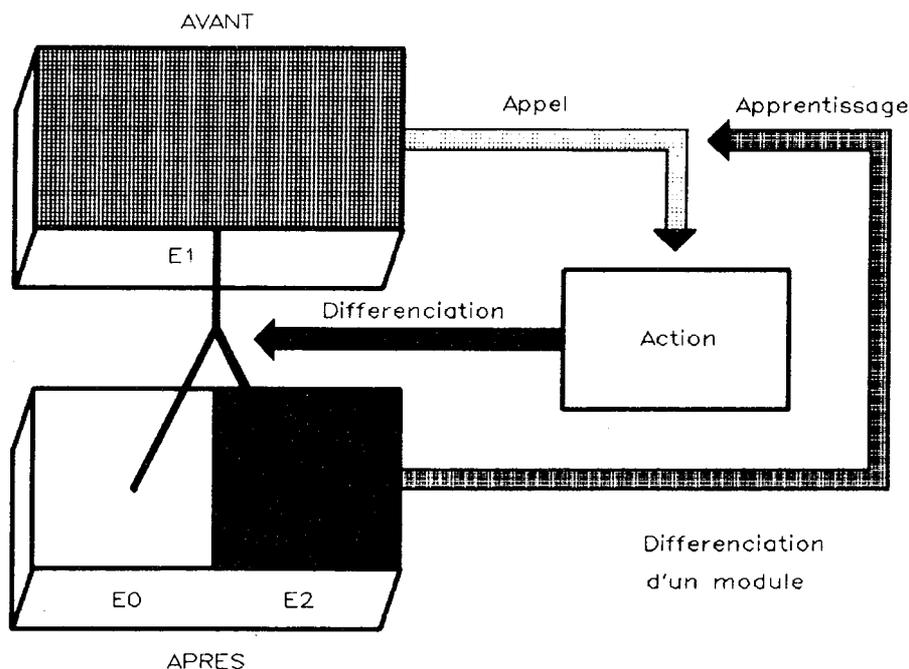


Figure 51. Principe de la division.

L'apprentissage se poursuit alors sur une aire dont la taille n'a pas changé (une zone se divise en deux sous-zones de taille moitié) mais enrichie de modules supplémentaires (un module divisé peut toujours, par exemple, déclencher inconditionnellement une réponse pour un autre phonème). Ce phénomène fait tendre cette aire vers un état de contraste maximum en définissant automatiquement, pour chaque région, la précision suffisante pour discriminer les indices qui y sont représentés. L'opération porte uniquement sur les indices permettant de différencier un phonème des autres. Le niveau de précision n'est affiné que lorsque l'état de division ne suffit plus à rendre compte des indices pertinents pour la différenciation. Cette remarque est bien illustrée par les figures 52 et 53 qui représentent les divisions de l'aire associative pour la différenciation de 4 voyelles simples (figure 52) et de 7 voyelles qu'il est plus difficile de différencier (figure 53).

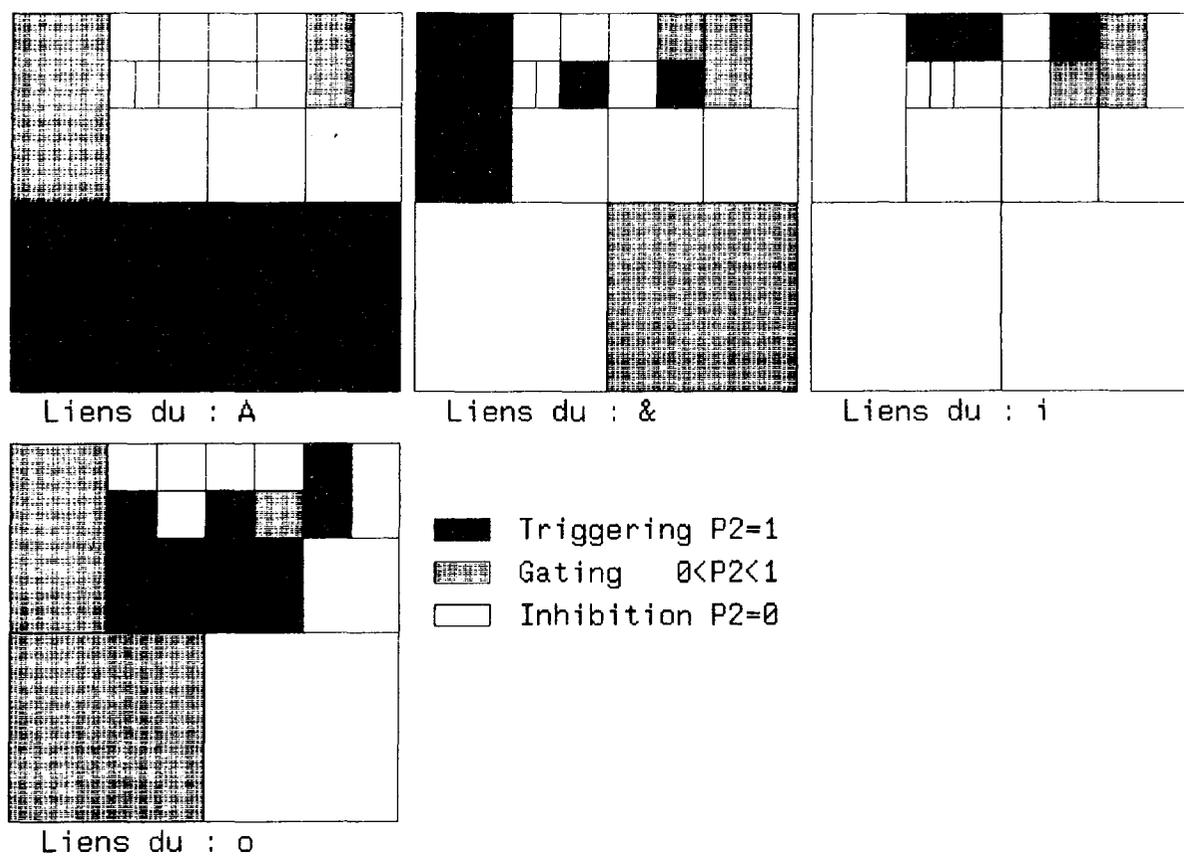


Figure 52. Reconnaissance de quatre voyelles.

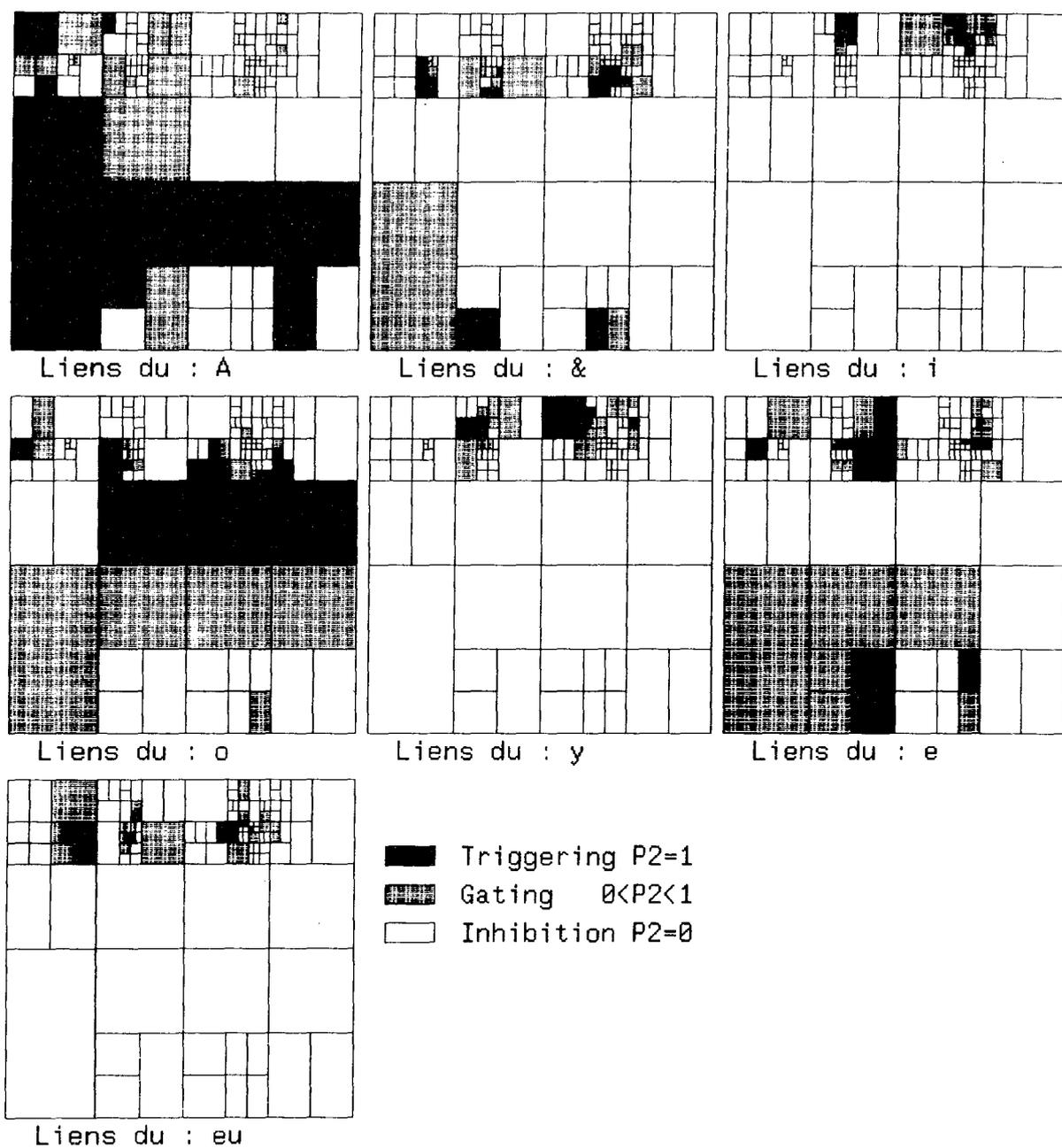


Figure 53. Reconnaissance de sept voyelles.

Le mécanisme de division que nous venons de présenter n'est pas spécifiquement dédié à la reconnaissance de la parole. Il s'agit en fait d'un mécanisme général permettant de réaliser une auto-organisation fonctionnelle à l'intérieur d'une carte. Il y a "auto"-organisation car les choix de division sont automatiques et ne dépendent pas de la position des zones mais de leur activité. Ce mécanisme est fonctionnel car il permet de dégager les ensembles de colonnes dont les fonctions conduisent à la définition des indices permettant de séparer les différentes classes proposées au réseau.

3 La motricité

3.1 Présentation du problème

Nous nous sommes efforcé tout au long de cet exposé de présenter la colonne corticale comme une base possible de toutes les fonctions corticales. Nous avons illustré l'éventail de ses possibilités par des applications visuelles ou auditives. Il est donc logique que nous poursuivions ce tour d'horizon des facultés humaines par le mouvement. Il est à noter que le pôle somesthésique constitue, après les pôles visuels et auditifs, la troisième voie d'entrée du réseau cortical présenté au chapitre II.

La diversité des fonctions réalisées par la colonne réside dans la diversité des domaines d'application de ces fonctions. Elle réside également dans leurs capacités et leurs propriétés. Nous avons jusqu'à présent parlé de capacité associative et d'auto-organisation. Nous voulons maintenant aborder les techniques de résolution de problèmes avec satisfaction de contraintes. Ce domaine, illustré par le problème du voyageur de commerce, peut être abordé par des techniques algorithmiques classiques ou par certains modèles connexionnistes (Kirkpatrick, Gellat, Vecchi, 1983).

Nous n'avons pas voulu étudier ce problème très classique mais nous avons plutôt cherché à simuler une fonction d'adaptation humaine de plus bas niveau, mais tout aussi intéressante. La motricité nous offre une application possible, le guidage de la main par l'oeil (Hinton, 1984; Saerens & Soquet, 1989), d'autant plus intéressante qu'elle demande la mise en relation de deux domaines sensori-moteurs différents. En effet, jusqu'à présent, les applications consistaient à effectuer la reconnaissance d'informations visuelles ou auditives; il va s'agir ici de faire se rencontrer l'information visuelle et la représentation somesthésique afin que, par apprentissage, se dégage une fonction d'adaptation permettant de placer la main sur une cible repérée visuellement.

L'intérêt applicatif de cette réalisation peut sembler moins important que la reconnaissance de la parole ou la reconnaissance de forme. On peut toutefois penser au domaine de la robotique ou, *mutatis mutandis*, à la satisfaction de contraintes. Nous avons de plus pensé qu'il était utile de proposer ici cette réalisation afin d'introduire des caractéristiques du modèle qui permettent de dépasser le stade de la mise en correspondance et introduisent les premiers supports du raisonnement à l'intérieur d'un ré-

seau.

L'analyse et la clarification de certaines caractéristiques du modèle va cependant s'accompagner de l'émergence d'autres problèmes qui avaient jusqu'à présent, été peu exprimés. Notre modèle se veut un modèle de cortex et non un modèle de cerveau. Il va alors se poser le problème de la prise en compte et de la simulation de structures coopérantes. Nous verrons ainsi que le cervelet joue un rôle important dans la motricité (Marr, 1969; Eccles, Ito, Szentagothai, 1967) et qu'il est impossible de le représenter par des colonnes corticales.

Ces remarques introductives étant faites, nous allons aborder l'analyse du problème par la présentation du réseau des aires impliquées et par le codage qu'il va nécessiter afin de représenter l'information. Nous étudierons également les moyens de représenter, parallèlement à ce réseau, l'action du cervelet sur cette information. Le codage étant posé, nous y introduirons alors la logique qui va permettre la résolution de problèmes avec satisfaction de contraintes. Cette logique sera supportée par les techniques de recombinaison et d'arbres d'appels. Il sera alors possible de présenter l'application particulière de guidage de la main par l'oeil avant de conclure sur les possibilités que nous offre l'introduction de tels mécanismes dans un réseau connexionniste.

3.2 Le codage

La fonction dont nous voulons décrire ici une réalisation possible a une localisation précise dans le réseau des aires présenté au chapitre II: il s'agit de l'axe visuel pariétal composé des aires V2, SV et S2. Cet axe a accès à l'aire de la motricité M2 directement par S2 ou par l'intermédiaire de l'aire frontale FSV. En plus de ces aires corticales, la fonction de guidage de la main par l'oeil nécessite la prise en compte de structures coopérantes telles que le cervelet ou le colliculus (responsable du mouvement des yeux) et des flux extra-corticaux des informations visuelles et somesthésiques. Cet ensemble forme un réseau représenté par la figure 54. Il va être d'abord précisé au moyen d'une série de données biologiques. Ensuite, une analyse plus formelle étudiera les conséquences de telles représentations dans un programme informatique.

3.2.1 Données biologiques

Données extra-corticales

Le colliculus est une structure neuronale liée aux aires visuelles et aux muscles oculaires. Cette structure transforme la position d'un stimulus en intensité de contraction des muscles des yeux (Robinson, Goldberg, Stanton, 1975). Cette fonction pourrait être réalisée par une seule couche câblée proposant, pour chaque position sur la rétine, la réponse musculaire permettant de fixer ce point. Chez l'homme, cette structure est composée de deux couches (la deuxième prenant en compte ou demandant un mouvement de la tête); nous n'en retiendrons qu'une pour nos simulations.

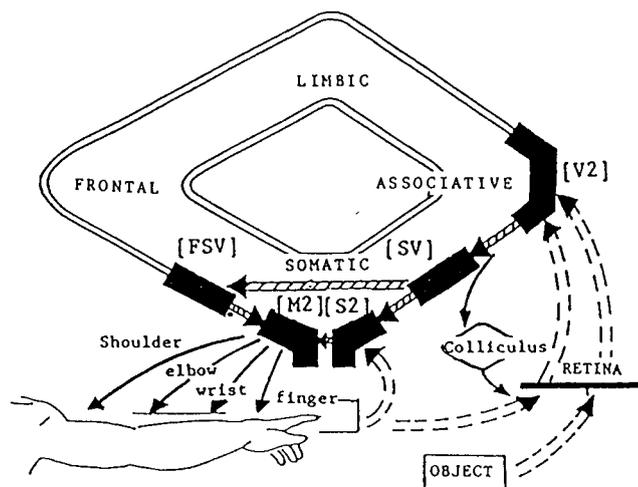


Figure 54. Réseau pour la motricité (d'après Burnod).

Le cervelet est une très grosse structure neuronale et peut être considérée comme une "mémoire de masse" de la fonction motrice du cortex (Eccles, Ito, Szentagothai, 1967). De manière très schématisée, la partie du cervelet qui nous intéresse pour cette application est illustrée par la figure 55.

Cette structure matricielle est composée de trois types de neurones ayant chacun un rôle précis. Les cellules de Purkinje sont la sortie du cervelet vers les aires motrices du cortex (M2). En retour, le cervelet reçoit deux types d'informations: des renseignements sur la situation sensori-motrice générale (venant par exemple de SV) sont recueillies par les fibre parallèles (ou fibres moussues); l'activation proprioceptive spécifique (venant par exemple

de S2) permet une mesure du déséquilibre et est amenée par les fibres montantes.

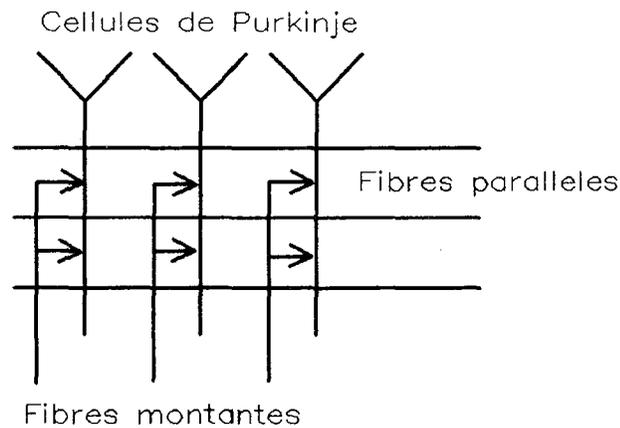


Figure 55. Modélisation du cervelet.

La structure matricielle du cervelet va lui permettre d'apprendre à reconnaître les situations sensori-motrices générales liées à une activation proprioceptive particulière (Ito, Sakurai, Tongroach, 1982). Après apprentissage, l'arrivée de telles situations commandera, via les cellules de Purkinje, une demande au cortex d'une anticipation posturale prévenant le déséquilibre.

Données corticales

Le point que nous voulons mettre en avant ici est la structuration en bandes des aires corticales (Woolsey, 1958; Hubel & Wiesel, 1962). Nous avons parlé pour le moment de la structure bidimensionnelle des aires dont les deux dimensions supportent les indices primaires. Cependant, une analyse fonctionnelle de cette représentation fait généralement apparaître la prépondérance d'un des indices primaires.

Ainsi, l'aire S2 regroupe un ensemble de bandes dont chacune représente une partie du corps. Ces bandes sont disposées somatotopiquement et permettent ainsi la représentation d'un homoncule (voir figure 16 ou Kaas, Nelson, Sur, Merzenich, 1979). Parallèlement à cette représentation sensorielle du corps, l'aire motrice M2 développe une représentation en bandes similaire. Dans le cas de cette aire, c'est une distribution somatotopique des muscles qui est ainsi effectuée. Dans l'aire visuelle V2, une distribution

rétinotopique n'empêche pas une déformation de la représentation. Cette déformation permet, ici aussi, l'apparition de bandes dans V2. Ces bandes correspondent aux différentes distances du centre de la rétine. La rétinotopie impose une disposition en distances croissantes, de la fovea jusqu'à la vision la plus périphérique (Hubel & Wiesel, 1962; Imbert, 1983; Jeannerod, 1974).

Un autre point important à souligner ici est la relation directe entre la surface de représentation corticale et l'importance fonctionnelle de l'indice représenté (Woolsey, 1958). C'est ainsi que l'on trouve une surface de représentation nettement supérieure à sa taille relative pour la main dans S2 et M2 et pour la fovea dans V2. Ces données structurelles ayant été énoncées, nous pouvons maintenant fournir des données plus fonctionnelles sur les relations entre les aires étudiées.

Les liens réciproques entre M2 et S2 permettent l'apprentissage de l'image corporelle. Le but d'un mouvement déclenché dans M2 sera une stimulation corporelle dans S2 qui servira ainsi de contrôle. La réciproque permettra de déclencher un mouvement réalisant une attente dans S2. Par exemple, lors d'une fonction de préhension, c'est une demande active (E1) des colonnes représentant le contact des doigts avec un objet qui guide le mouvement. De manière plus générale, on peut dire que le but recherché organise principalement les mouvements volontaires (Bernstein, 1967). Les liens de S2 et V2 avec SV vont permettre un apprentissage par imitation de l'image visuelle. Les modules de SV qui effectuent une mise en correspondance des stimuli visuels et somesthésiques vont devenir des buts partiels pour un mouvement de positionnement. L'arrangement des liens de SV avec S2 et V2 est révélateur de la logique corticale. La corrélation des zones de S2 et V2 n'est pas quelconque, mais inclut des relations implicites entre la rétinotopie et la somatotopie. Partant du principe que, en général, un objet repéré visuellement dans la fovea pourra être directement manipulé et qu'une position visuelle de plus en plus périphérique demandera généralement l'aide de muscles de plus en plus distaux (épaule, hanches, jambes,...), la recombinaison de S2 et V2 dans SV est disposée ainsi: SV permet de relier directement la main et la fovea et les régions de plus en plus éloignées des représentations visuelles et somesthésiques entre elles.

La fonction de guidage de la main par l'oeil est réalisée en deux étapes. Il y a d'abord fixation des récepteurs visuels sur l'objet (axe oculaire et convergence). Le but cortical est ensuite l'image de la main sur la fovea

à la distance de convergence. Le placement de la main est réalisé en deux temps. Il y a tout d'abord une composante balistique proposée par le cervelet qui amène la main à proximité de l'objet. Le cortex effectue ensuite les régulations plus fines permettant d'atteindre l'objet et de le saisir.

3.2.2 Analyse informatique

Nous allons maintenant reprendre ces données biologiques afin d'étudier les contraintes et les avantages qu'elles impliquent dans un programme informatique.

Si, a priori, la simulation du colliculus ne pose pas trop de problème (une colonne de V2 effectue une sortie externe vers le colliculus qui replique V2 et connaît ainsi la position du nouveau point de fixation), il n'en est pas de même pour le cervelet. Dans ce cas, la représentation est plus complexe et nécessite une place mémoire très importante pour recombinaison des états de SV et S2. Ce point est annonciateur de problèmes plus généraux: nous développons un modèle de cortex en connaissant l'importance, pour son fonctionnement, des structures coopérantes. Celles-ci devront également susciter une modélisation approfondie.

La structure en bandes de S2 et M2 nous donne des indications sur les indices primaires de ces aires.

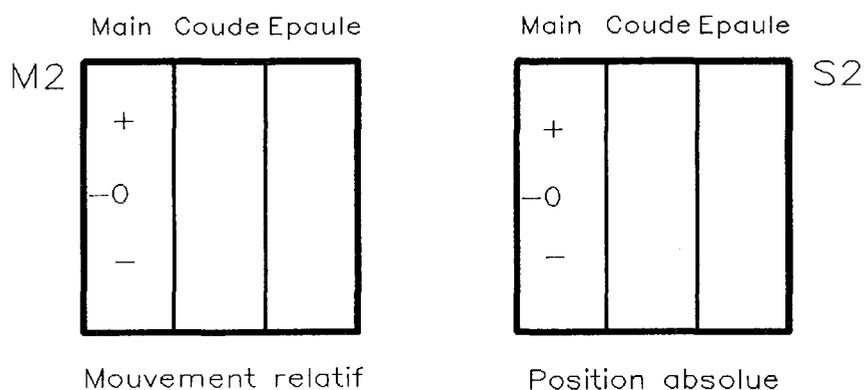


Figure 56. La dualité M2-S2.

Le premier comportera une énumération des différentes parties du corps prises en compte. Le deuxième indice primaire doit rendre compte des diffé-

rentes positions ou des différents mouvements possibles de ces articulations. Il représentera donc une base d'intensité de contractions. La dualité M2-S2 offre également, par sa représentation, un autre avantage. Nous avons ainsi accès, selon les besoins, à une information en position, soit absolue pour S2, soit relative pour M2. Ces deux points sont illustrés par la figure 56.

La représentation en bandes de V2 demande tout d'abord une explication. Dans les études visuelles précédentes, nous avons adopté les axes de l'image comme indices primaires de V2; ceux-ci ne représentaient pas la structuration en bandes de la fovea et des zones visuelles de plus en plus périphériques. Nous nous justifions de la manière suivante: cette application visuelle représentait l'axe temporel de la vision et l'image à reconnaître prenait sa place dans la partie centrale de la fovea. A cet endroit, il n'y a pas déformation de l'image et celle-ci peut donc être représentée, telle que, sur V2. Cet argument ne tient plus quand on travaille sur le champ visuel total où, cette fois, la déformation doit être représentée. Elle le sera en prenant comme indice primaire la distance au centre de la fovea. Cette déformation est illustrés par la figure 57.

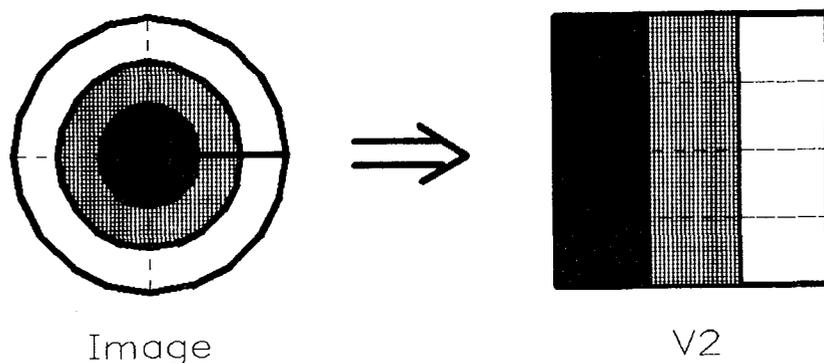


Figure 57. La déformation dans V2.

Cette figure a l'avantage de nous indiquer la nature du deuxième indice primaire: il s'agit d'une dimension angulaire prise par rapport à une référence pouvant, par exemple, correspondre à l'horizontale. On remarque que ce codage correspond en fait au passage des coordonnées cartésiennes aux coordonnées polaires. On peut également penser que la manipulation

de distances et d'angles peut se révéler plus intéressante pour un travail de guidage dans l'espace (Hinton, 1984).

L'intérêt de la proportionnalité entre surface de représentation et importance fonctionnelle est évident, tant pour le cortex que pour un programme informatique: une capacité relative supérieure pour la main ou pour la fovea pourra plus facilement s'exprimer sur une surface de représentation accrue.

La revue de la manière dont les différents indices seront représentés nous permet maintenant d'illustrer, par la figure 58, la construction de SV à partir des bandes de S2 et V2 et de la recombinaison main/fovea et des parties plus éloignées entre elles.

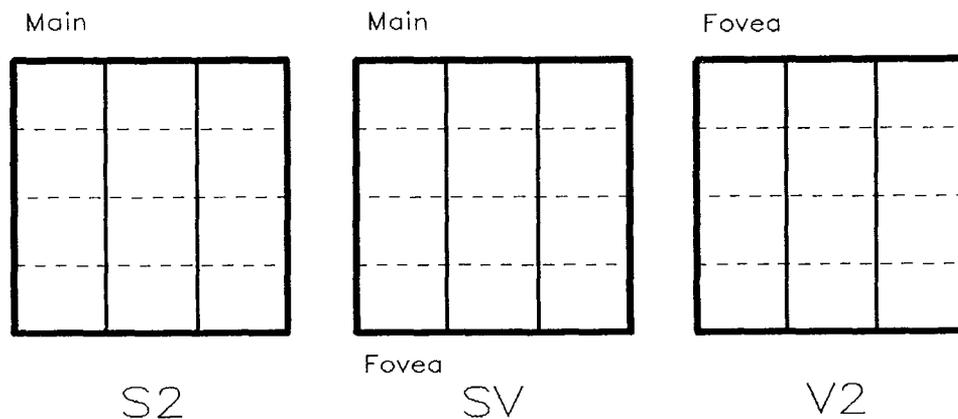


Figure 58. SV recombine S2 et V2.

D'une manière globale, on peut dire que le rôle de SV est d'effectuer un changement de repère entre les coordonnées visuelles de S2 et les coordonnées corporelles de V2 (Burnod, 1988). En effet, le guidage de la main par l'oeil ne peut pas être effectué par une simple translation entre une position rétinienne et l'intensité d'une contraction musculaire, mais par un réel changement de repère entre deux systèmes d'axes différents. Cette opération va être réalisée de manière discrète par un ensemble de relations apprises par les modules de SV. En effet, chaque module de SV teste la coactivation de ses champs récepteurs sur S2 et V2 et mémorise donc la relation entre l'image visuelle du corps reçue par V2 et la position angulaire de ses différentes parties connue par S2. Il faut également remarquer que cette mémorisation est réciproque et permettra, si le but est une image

visuelle dans V2, de connaître, via SV, la position corporelle correspondante, tout comme, si le but est une position corporelle dans S2, de guider visuellement, via SV, ce positionnement.

La recombinaison depuis la main et la fovea entre les parties de plus en plus périphériques de S2 et V2 peut prêter à discussion. En effet, la présence d'un objet dans la fovea n'implique pas sa manipulation directe, tout comme un objet vu périphériquement peut parfois être directement saisi. C'est vrai, mais cette recombinaison ne propose pas la solution unique. Il s'agit tout simplement, et tout informaticien y sera sensible, d'une prédisposition proposée, permettant statistiquement une optimisation de la longueur du chemin à parcourir entre les entrées et les sorties. Cependant, dans le cas où la recombinaison de l'aire associative ne relie pas directement le stimulus à la réponse, il y a propagation horizontale à l'aide des connexions locales à l'intérieur de l'aire associative. C'est ce mécanisme que nous proposons de détailler dans la partie suivante.

3.3 L'auto-organisation fonctionnelle

Tout au long de cette étude, nous avons mis l'accent sur la nécessité de définir des mécanismes permettant l'émergence du raisonnement. Après l'analyse des capacités de mémoire associative dans les domaines de la vision et de la parole, il est temps maintenant de présenter ces mécanismes. Certains ont déjà été vus dans la description du modèle. D'autres seront précisés ici.

Nous avons déjà annoncé les principes qui permettent l'introduction du niveau symbolique dans un réseau d'automates qui, comme les colonnes, peuvent être considérés comme des buts: il s'agit, par apprentissage de relations causales, de décomposer les buts en sous-buts validants. Ceci sera effectué, lors de l'apprentissage, de manière redondante. Lors d'une réalisation particulière où un but sera demandé, il va s'agir de chercher, parmi ses sous-buts, lesquels ne sont pas inhibés et lesquels permettent de réaliser le but de la manière la plus économique.

Nous allons donc tout d'abord procéder au rappel ou à la définition des mécanismes qui permettent ces actions. Nous nous efforcerons de le faire de manière à rester cohérents avec l'ensemble du modèle tel que nous l'avons déjà présenté. Ceci étant fait, nous placerons ces mécanismes en situation de manière à en faire une analyse moins stricte, mais plus explicite; nous

verrons comment ces mécanismes s'emboîtent et se déclenchent dans le cas du guidage de la main par l'oeil.

3.3.1 Description des mécanismes

La division

Le premier mécanisme dont nous allons parler est celui qui permet la définition, par apprentissage, des buts et des sous-buts. Nous allons voir qu'il a un certain rapport avec un phénomène dont nous avons déjà parlé lors de la modélisation, dans la partie précédente (III.2): la division; ce phénomène permet, par apprentissage, d'organiser une zone vierge en la divisant en sous-parties qui se spécialisent pour réaliser une fonction donnée sur leur champ récepteur. Ce phénomène est déjà partiellement réalisé à la naissance dans les aires sensori-motrices primaires qui sont décomposées en colonnes réalisant une fonction donnée sur un champ récepteur donné (exemple: sélectivité d'orientation dans V1). Cette organisation fonctionnelle est mise au point par apprentissage dans les aires associatives, encore partiellement vierges à la naissance.

Dans la partie III.2, nous avons présenté le phénomène de division à l'échelle de l'aire associative toute entière qui était donc complètement non spécifiée à l'origine. Dans des aires associatives telles que SV ou SA, le phénomène n'est pas aussi extrême. En effet, il y existe une spécification a priori dont nous avons déjà parlé. Il s'agit de la recombinaison par bandes des aires sensori-motrices qui définit grossièrement des correspondances entre des parties des représentations sensorielles et motrices. Ce typage initial des aires associatives a pour support les zones, ensembles de colonnes dont le but est de rendre compte de la coactivation de deux fragments de deux domaines sensori-moteurs. C'est en fait sur ces zones que le phénomène de division va faire son oeuvre d'organisation. Dans les aires sensori-motrices primaires, les colonnes voisines sont celles qui réalisent des fonctions différentes sur le même champ récepteur. On retrouve cette notion dans ces aires associatives où la division permet également de définir des fonctions différentes sur des domaines donnés. A ce titre, les colonnes d'une zone de recombinaison donnée peuvent être considérées comme voisines. Nous avons vu, dans la partie III.2, que la division permettait d'organiser l'aire associative en fonction de l'information y parvenant par des liens de type lointain. Nous allons expliquer maintenant comment

cette organisation prend également en compte les liens locaux et voir les implications très importantes de ce fait. Pour cela, commençons par décrire l'utilisation "normale" de ces liens locaux, à l'aide de la figure 59.

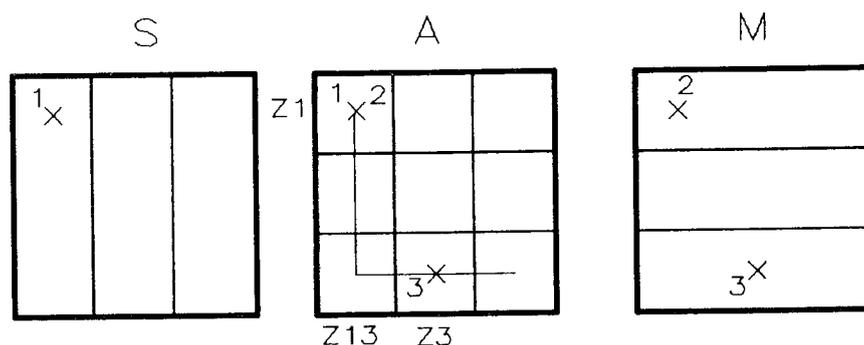


Figure 59. Utilisation des liens locaux.

Soit A une aire associative qui recombine les bandes de deux aires sensori-motrices S et M. Prenons un stimulus en 1 sur l'aire S. Deux cas peuvent alors se présenter: s'il y a un stimulus en 2 sur l'aire M, par recombinaison une partie de la zone Z1 passe à E2. Cette partie est définie par le principe de la division sur les liens lointains tels qu'on l'a vu dans la partie III.2. Par contre, si le stimulus est en 3 dans l'aire M, l'aire associative A ne peut pas recombinaison directement les stimuli 1 et 3. Ce lien va se faire grâce aux liens locaux de A selon le principe suivant. De par leurs liens lointains et la présence des stimuli 1 et 3, les zones Z1 et Z3 sont dans l'état E1 (recherche active). Dans les bandes recombinaisonées dans A, les zones développent de proche en proche des liens locaux. Selon la table 35, les états E1 de Z1 et Z3 vont se propager le long de ces bandes. Pendant cette propagation, le coefficient C des états E1 se propage également en décroissant progressivement. C représente une mesure de la distance de propagation. Ces deux E1 qui se propagent vont se rencontrer dans la zone Z13. Si ce n'est pas déjà fait, cette zone va se diviser afin de prendre en compte l'association de ces deux stimuli. Cette prise en compte passe par la mémorisation de la valeur des deux coefficients C qui parviennent dans cette zone. Cette mémorisation permet ainsi la localisation précise de l'origine de ces deux stimuli sur A et donc de leur localisation sur S et M. En

fait, au cours de l'apprentissage, les zones divisées ne vont pas se contenter d'effectuer une simple mémorisation de coefficients: si un cas proche d'un cas appris (proche par les valeurs des coefficients C) se présente, c'est la même unité qui cette fois mémorise non plus deux valeurs mais deux intervalles d'éloignement. Ce phénomène va être à la source de propriétés de généralisation. Si, plus tard, une généralisation se trouve infirmée, il y aura division de cette zone en deux sous-zones représentant chacune une partie des intervalles.

Cette mémorisation par intervalles est généralement économique car elle permet, avec une seule zone, de lier de grosses parties de la carte entre elles. De part la logique de représentation (deux zones proches ont des fonctions proches), cette liaison globale sera souvent fonctionnellement représentative d'une réelle relation associative représentant un phénomène plus macroscopique que le niveau colonnaire (par exemple, pour un intervalle de distance relativement grand, on fait toujours le même type de mouvement pour tendre la main dans une direction donnée).

Nous venons donc de donner une définition plus précise de la fonction d'une zone définie par recombinaison. D'une part, par ses liens lointains, cette zone permet d'associer par division les phénomènes issus directement des deux parties de la représentation sensori-motrice que cette zone recombine; d'autre part, le même phénomène de division permet de recombinaison des stimuli plus lointains, originaires des bandes dont cette zone est l'intersection (figure 60). Ce dernier phénomène est supporté par les liens locaux.

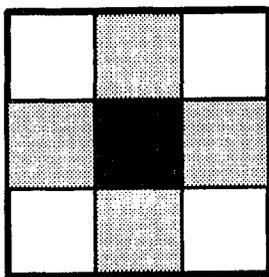


Figure 60. Liens d'une zone d'une aire associative.

Ces mécanismes permettent, par division, de dégager des zones qui de-

viennent potentiellement des buts ou des sous-buts. Après avoir vu comment la g n se de ces sous-buts est possible, nous allons voir comment ils s'organisent entre eux, en arbre d'appel.

Les arbres d'appel

Ce deuxi me m canisme a pour base les techniques d'apprentissage telles que nous les avons d j  d crites lors de la mod lisation de la colonne (II.4.5). Nous avons vu alors que, tout au long du fonctionnement du r seau, chaque colonne modifie ses liens avec une colonne selon son  tat, quand elle re oit E2 de cette colonne. Nous avons vu  galement que cette propri t  permettait d' tablir des relations causales entre les diff rents modules.

Cette m morisation est syst matique et permet donc de repr senter ces relations causales de mani re redondante. Autrement dit, une colonne en tant que but est reli e   tous ses sous-buts qui, eux-m mes en tant que buts, ont appris tous leurs sous-buts possibles. Ce type de m morisation redondante propose des liens sous forme de structure arborescente dont un exemple possible est repr sent  dans la figure 61. Cette structure va  tre le support des arbres d'appels.

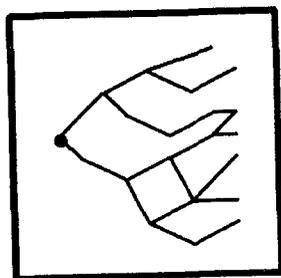


Figure 61. Arbre d'appel.

Elle va  tre utilis e de la mani re suivante: nous avons vu que la colonne effectuait des traitements sp cifiques selon l'origine de l'information. Lorsqu'une colonne re oit E1 par le r seau cortical (pas d'une aire hi rarchiquement inf rieure comme une aire sensori-motrice), cet appel n'est plus consid r  comme un stimulus devant  tre propag  pour association, mais comme une demande de r alisation du but que cette colonne repr sente.

Selon la table 35, cet appel va se propager sur les liens locaux pour appeler les sous-buts qui permettent de réaliser le but. Ces sous-buts, recevant E1, vont à leur tour tenter de se réaliser par appels de leurs sous-buts. Par cette demande, certains sous-buts seront inhibés par les conditions de l'environnement (sous-buts irréalisables); d'autres seront possibles et passeront alors à E2. Cette partie correspond à l'appel du but principal. Toujours, selon cette même table 35, les sous-buts réalisés vont transmettre leur état E2, en particulier sur les liens locaux réciproques qui les avaient excités. Les sous-buts intermédiaires seront à leur tour réalisés et ainsi de suite, récursivement, jusqu'au but initial. Cette partie correspond à l'action permettant de réaliser le but initial. Si plusieurs branches de l'action d'appel sont réalisables, c'est la première qui remontera jusqu'au but initial (prouvant ainsi une économie en termes de distance corticale) qui l'emportera.

Nous venons ainsi de voir le deuxième mécanisme permettant une organisation fonctionnelle avec un nombre de liaisons relativement réduit. Ce mécanisme, très proche de la notion de récursivité, se décompose en un appel permettant la sélection des sous-buts réalisables et en une action permettant la sélection des meilleurs sous-buts. Cependant, la description de ces mécanismes ne suffit pas à rendre compte de toute leur force. Aussi avons nous choisi d'insister sur ce point en proposant l'explication, dans le contexte précis du guidage de la main par l'oeil, de l'utilisation conjointe de ces deux mécanismes selon une procédure moins stricte mais sûrement plus explicite.

3.3.2 Les mécanismes en contexte

Nous avons déjà parlé de la décomposition du mouvement, lors d'une tâche de préhension, entre une partie balistique, commandée par le cervelet et une partie plus fine dirigée par le cortex. C'est dans cette deuxième partie que ces mécanismes s'expriment.

Tout d'abord, par le mécanisme de division, chaque zone de SV apprend certaines combinaisons correctes possibles entre la position visuelle de la main sur la rétine et la position angulaire des différentes parties du corps. Les combinaisons qu'une zone donnée mémorise sont de deux sortes: toutes celles ayant lieu dans les deux parties de la représentation sensorimotrice que cette zone recombine et toutes celles originaires des bandes

dont cette zone est l'intersection. Ainsi, d'un point de vue fonctionnel du pôle visuel vers le pôle somatique, chaque zone apprend les actions motrices qui favorisent son activation par le placement de la main dans son champ récepteur. Lors d'un exemple de guidage, nous imposons le but principal qui est la main dans la fovea (la fovea fixe la cible à atteindre). Ceci passe par la mise à E1 de tous les modules de SV représentant la main dans la fovea. Dans la réalité, ce but est imposé par les aires limbiques qui expriment la motivation.

Par apprentissage, chacun de ces modules connaît son sous-but direct qui est la colonne représentant la main dans une partie peu périphérique, proche de la fovea. En effet, en général, lors d'un mouvement, les colonnes "main dans la fovea" sont activées après les colonnes "main dans la périphérie proche" qui se trouvaient sur le passage de la main. Pour ces colonnes, leurs sous-buts validants sont situés dans une périphérie un peu plus éloignée, et ainsi de suite. La figure 62 représente le but final demandé et le sens de l'appel demandé.

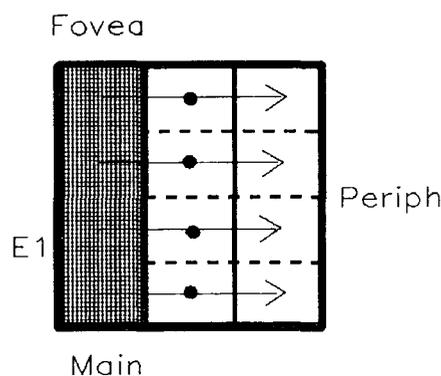


Figure 62. Appel du but.

Or, après le mouvement balistique, la main se trouve quelque part dans le champ visuel et valide donc un des sous-buts demandé: l'appel est donc couronné de succès et va déclencher, à rebours, le retour de l'action. Sur le chemin, chaque module passe à E2 et appelle les actions qui amènent la main dans son champ récepteur. En fait, ce module appelle toutes les actions motrices possibles qui peuvent amener la main dans son champ récepteur. Certains sont inhibés car impossibles (membre en bout de course), d'autres sont possibles et vont donc répondre dans S2 selon le même principe: ici aussi, il y aura sélection par la rapidité et c'est donc le mouvement

le moins coûteux en termes d'angle qui atteindra le premier le module appelant de SV et sera déclenché. De proche en proche, la main va ainsi atteindre la fovea et réaliser le but (ce qui permettra un apprentissage supplémentaire). Dans la réalité, ce processus se poursuit avec les mouvements fins des doigts permettant une préhension efficace.

Il reste plusieurs remarques à faire sur ce mécanisme. Tout d'abord, nous n'avons représenté, dans la figure 62, que les appels parallèles à l'axe des distances, c'est à dire radiaux par rapport au centre de l'oeil. Or il est possible d'atteindre la fovéa en suivant un chemin qui n'est pas radial et qui suscite un apprentissage entre colonnes pour des champs de plus en plus périphériques mais pour des directions non radiales. C'est vrai, mais ces appels, plus longs en termes de distance corticale que les appels radiaux, seront donc éliminés par la sélection de la rapidité. C'est la raison pour laquelle, lors d'une tâche de préhension, la main suit une direction rectiligne (aux phénomènes d'accélération près): c'est parce qu'elle est guidée par l'oeil. Par contre, chez les aveugles qui sont guidés par leur proprioception et non par leurs yeux, les mouvements ne sont pas rectilignes mais suivent des angles représentant une économie en terme d'enchainements musculaires. D'autre part, nous n'avons pas l'impression, lorsque nous prenons quelque chose, que notre main est guidée à ce point. C'est également vrai car, après de multiples apprentissages, certaines branches des arbres d'appel, toujours utilisées, passent dans ce que l'on appelle un déclenchement inconditionnel: il n'y a plus vérification au fur et à mesure de la progression de la main. L'appel est entièrement cortical et déclenche le mouvement d'un coup. Par contre, cette procédure reste utilisée lors des guidages très fins (exemple: manipulation d'explosifs!).

3.4 Discussion

Les mécanismes d'auto-organisation fonctionnelle que nous venons de décrire sont assurément l'aspect essentiel de cette partie. Leur application à la motricité peut être vue de deux manières différentes. L'organisation des sous-buts autour du but "main dans la fovea" se prête très simplement à la modélisation et à la description des phénomènes qui permettent la réalisation du but. Cependant, nous avons été gêné, tout au long de cette application, par une extrême difficulté pour représenter et modéliser tous les aspects moteurs. On a déjà parlé du problème du cervelet. Il existe

également beaucoup de travail autour de la représentation des muscles agonistes et antagonistes. En attendant, cette représentation des muscles en terme de base motrice ne nous satisfait guère. C'est la raison pour laquelle nous jugeons préférable de mettre l'accent ici sur les implications des mécanismes d'auto-organisation.

Nous avons vu qu'ils permettaient, dans le cadre de ce modèle, la résolution de problème avec satisfaction de contraintes. Pour nous, ils sont beaucoup plus puissants encore. En effet, nous affirmons que ces mécanismes vont permettre, dans les aires dont la fonction est plus abstraite que SV, la genèse du raisonnement. Dans les aires frontales et dans les aires associatives, la décomposition de buts en sous-buts par apprentissage va permettre, lorsqu'un problème se présentera, de sélectionner quels sous-buts sont réalisables selon le contexte externe ou interne, et de choisir, parmi ces sous-buts, lesquels sont les plus économiques selon des critères quelconques (temps, argent...). Selon nous, ce mécanisme est très voisin (sinon identique) à ce que nous appelons couramment "raisonnement".

4 Interprétation d'images

4.1 Présentation du problème

En plus d'un modèle d'unité connexionniste, nous avons proposé, dans la partie III.3, un modèle de cortex sous la forme d'un réseau des aires. La mise au point d'une application passe par l'extraction de ce réseau général de la sous-partie qui supporte la fonction considérée. En ce qui concerne les applications que nous avons jusqu'à présent décrites, les sous-parties extraites se limitaient à chaque fois à un axe sensori-moteur comprenant une partie sensorielle, une partie associative et une partie motrice. Ces axes, dont le choix était pertinent pour des fonctions de mise en correspondance, présentent une grande similitude structurelle avec les perceptrons multi-couches. Conscients que le réseau général offre, par sa structure, d'autres capacités très intéressantes, nous nous sommes employés à simuler une fonction dont le support ne se limite plus à un seul axe sensori-moteur.

La fonction choisie a pour but le diagnostic automatique de pathologies cervicales à partir de clichés radiologiques. Les besoins en terme d'aires sont d'abord visuels, pour analyser la forme des vertèbres, puis auditifs, pour définir un vocabulaire de description. Nous utilisons également des aires de la symétrie frontale, comme support à la stratégie employée lors du diagnostic. La description précise du réseau, des données traitées, des techniques employées et des résultats obtenus sont proposés dans (Guyot, 1990). Nous mettons uniquement l'accent dans cette partie sur les propriétés induites par l'utilisation d'un réseau plus complexe, en ce qui concerne les aires puis les colonnes.

4.2 Les aires

Nous avons construit ce réseau autour de trois symétries:

-Symétrie Pariétal/Temporal:

Il existe deux chemins corticaux privilégiés pour l'information visuelle. Ces chemins sont définis par deux niveaux de résolution spatiale, liés à la taille des champs récepteurs correspondants. Ces deux chemins sont anatomiquement séparés et sont déjà préparés dans la rétine par une distribution adéquate des capteurs.

D'une part, la vision centrale permet, par une résolution élevée, la reconnaissance des objets. Elle est réalisée par l'axe temporal (VT) dont on peut dire qu'il répond à la question "Quoi?". Ce traitement précis et coûteux n'est réalisé que sur une faible partie du champ visuel, son centre, ce qui montre tout l'intérêt des mouvements oculaires.

D'autre part, la vision périphérique se destine, par la taille élevée de ses champs récepteurs, à la localisation des objets. Elle est supportée par l'axe pariétal (SV) dont les liens avec les aires somato-motrices S2 et M2 permettent un premier apprentissage spatial par rapport au corps. Il est ainsi permis de dire que l'axe pariétal répond à la question "Où?". Ce traitement grossier est réalisé sur l'ensemble du champ visuel.

Ces deux axes, ajoutés à la possibilité d'effectuer par apprentissage des mouvements oculaires, permet ainsi de définir un objet (une vertèbre dans notre application) comme une séquence de points de fixation. Chaque point de fixation, réalisé par l'axe temporal, permet la reconnaissance d'une partie de l'objet; les mouvements oculaires permettent la localisation relative de ces parties. Outre son caractère descriptif, pour ne pas dire symbolique, ce mécanisme a l'avantage d'être moins coûteux en place mémoire qu'un traitement parallèle de l'image. Il permet de plus une invariance en translation immédiate et une invariance en taille relativement aisée.

-Symétrie Visuel/Auditif:

L'application considérée ne consistant plus en une mise en correspondance simple, la partie auditive du réseau aura pour utilité première le dialogue avec le réseau, afin de recueillir ses réponses, de lui poser des questions ou de le guider.

La dualité que nous venons de proposer pour le traitement visuel va symétriquement se retrouver pour le cas auditif, selon la même logique. Le chemin auditif temporel (AT) supporte le vocabulaire relatif à la dénomination des différentes parties de la vertèbre; l'aire pariétale SA recueille le vocabulaire relatif à la localisation (droite, gauche, haut, bas,...).

-Symétrie Associatif/Frontal:

Cette dernière symétrie permet de dupliquer les aires associative que nous venons de définir afin de les organiser temporellement dans les aires frontales qui deviennent ainsi le support des stratégies. Ces stratégies sont normalement apprises par essais/erreurs. En ce qui concerne l'application de diagnostic de maladies de vertèbres, cet apprentissage se résume pour le moment en la définition supervisée de stratégies extraites de l'expertise de

médecins radiologues.

La figure 63 représente le réseau que l'on a extrait du réseau général pour réaliser cette application.

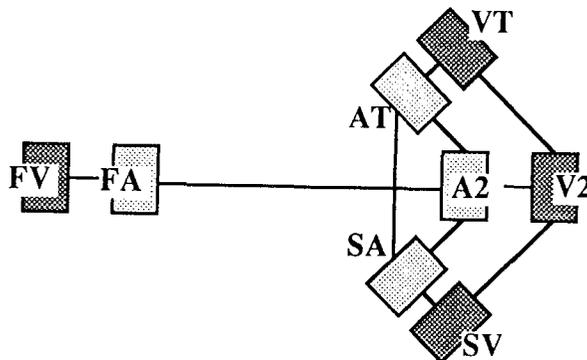


Figure 63. Réseau extrait pour l'application.

4.3 Les colonnes

Les cartes d'entrées/sorties sont définies de manière relativement similaires aux applications précédentes. Les détails de cette réalisation sont indiqués dans (Guyot, 1990). Les aires associatives utilisent le principe de division défini dans la partie III.2. En ce qui concerne la prise en compte par la colonne des informations provenant des différentes symétries, le comportement de cette unité connexionniste doit être précisé. L'étude de données neurobiologiques permet de remarquer que, à l'intérieur de la couche haute de la colonne corticale, les liens sont rangés hiérarchiquement, selon la symétrie dont ils proviennent (se reporter, dans la modélisation, à la figure 27 et au texte qui s'y rapporte). Ceci nous permet de supposer et de faire comme si la colonne pouvait également traiter sélectivement ces informations, selon leur provenance corticale. Nous ne pensons pas que ce choix conduit à une trop grande spécification du traitement puisque, et c'est là notre "garde fou", toutes les colonnes d'une même aire ont un comportement semblable.

En ce qui concerne cette application, la prise en compte de la provenance des messages échangés s'effectue comme suit: en cas de conflit sur la réponse demandée à une colonne par un message reçu, nous utilisons la hiérarchie

topologique des afférences comme une hiérarchie de priorité; autrement dit, les ordres (symétrie associatif/frontal) sont prioritaires sur les requêtes (symétrie visuel/auditif), elles mêmes prioritaires sur le contexte (symétrie pariétal/temporal), lui même prioritaire sur les signaux.

Il est certain que ces choix, qui semblent correspondre à une logique corticale, ne manquent pas d'intérêt en ce qui concerne l'interprétation psychologique de tels comportements. Pour expliquer comment un système qui fait moins confiance aux signaux qu'à ses buts internes ne se fourvoie pas, précisons qu'une certaine logique relationnelle permet une construction plutôt cohérente des voies qui supportent le fonctionnement. Il s'agit du protocole qui ordonne l'apprentissage et qui propose la maturation des aires dans un ordre inverse de celui des priorités, ce qui permet de supposer que les buts les plus abstraits qui vont guider le raisonnement seront construits en dernier sur des bases relativement fiables.

Nous avons tenté de proposer dans cette partie une première réflexion sur la manière de gérer un réseau ne se limitant pas à un ensemble de couches superposées. En dépit des difficultés qu'implique une telle gestion, nous voulons insister sur les capacités symboliques, analogiques et abstraites qui découlent d'une telle proposition.

CHAPITRE IV:

CONCLUSION

1 Bilan

Le connexionnisme a longtemps développé ses capacités de mémoire associative pour des tâches de mise en correspondance. Il tente maintenant de s'attacher au traitement de phénomènes temporels (Gori, Bengio, DeMori, 1989; Waibel, Hanazawa, Hinton, Shikano, Lang, 1989) et de déboucher sur le raisonnement. C'est à cette condition qu'il pourra répondre aux espoirs qu'il a fondé et simuler des fonctions humaines de plus haut niveau (Paillard, 87). Il existe également, pour aborder ce type de fonctions, des recherches sur la mise au point de systèmes hybrides (Memmi, 1989) mêlant l'approche connexionniste pour les tâches de bas niveau à l'approche symbolique pour les tâches abstraites. Néanmoins, la réalisation de l'interface entre ces deux modules n'est pas aisée et ces deux approches perdent certains de leurs avantages lorsqu'elles se rencontrent. De plus, l'existence d'un exemple vivant renforce la possibilité du tout connexionniste et des recherches qui en découlent. Ce n'est d'ailleurs pas pour cette raison qu'il faut rejeter en bloc l'approche symbolique qui a développé des systèmes logiques très solides, dont il est permis de s'inspirer dans le fonctionnement d'un réseau. De manière plus générale, un réseau n'est rien sans toutes les sciences, exactes ou non, qui s'intéressent de près ou de loin au domaine des neurosciences et des sciences cognitives. Elles proposent toutes des idées qui ont été prélevées à différentes échelles d'étude (moléculaire, logique, temporelle, développementale, comportementale, connectique...) et dont il est judicieux de se servir et de réinsérer à ce même niveau dans un système connexionniste.

Le modèle que nous avons présenté se déclare nouveau mais reconnaît cependant volontiers la marque de nombreuses autres approches. Du niveau moléculaire au niveau psychologique (des neurotransmetteurs aux fonctions d'adaptation), nous nous sommes toujours efforcés d'avoir à l'esprit les multiples idées issues des différents niveaux de description.

Il est indéniable que la biologie est une source d'information privilégiée de ce modèle. Ceci est d'abord dû à la paternité de Y. Burnod sur ce modèle. En fait, notre travail a consisté à comprendre le modèle tel que Y. Burnod le décrit dans son livre (ce qui ne fut pas une mince affaire...), puis, avec son aide, de revoir et d'affiner plusieurs points afin de transformer la colonne en objet informatique. Le rôle principal de la biologie est également dû au fait que, de la biologie moléculaire à l'étude des comportements, il existe, à chaque niveau, une branche de cette science qui peut nous fournir des données ou des hypothèses. Cependant, certains aspects dynamiques de la colonne restent très flous d'un point de vue biologie. Nous avons alors intégré des données psychologiques dans ce qui était le plus souvent des paramètres ou des constantes du programme: il en a été ainsi pour les constantes de temps de fonctionnement ou de transmission et pour les niveaux de priorité entre colonnes et entre aires. La psychologie se révèle également riche pour analyser le comportement d'un réseau, essayer de comprendre sa logique de fonctionnement (Vurpillot, 1972; Guillaume, 1979) ou même, plus abstraitement, pour relier notre approche à un cadre plus général (Lorenz, 1970; Girard, 1978; Hofstadter, 1985). Nous avons été ainsi frappé par l'identité qu'il existe entre les différentes symétries à intégrer progressivement et les différents stades de développement de l'enfant proposé par Piaget (Piaget, 1975).

Ce n'est d'ailleurs pas le seul rapprochement a posteriori que la mise au point et l'analyse du modèle nous ont permis de réaliser. Il est à noter que la découverte de tels points communs permet souvent, par analogie, de faire nôtre et d'utiliser tout ce qui, dans le domaine de référence, découle du point de comparaison.

Ainsi, nous avons l'habitude de considérer une colonne comme un but potentiel. Lorsque nous avons abordé le raisonnement, et décrit celui-ci en termes de buts qui, par apprentissage, se lient aux sous-buts permettant de les réaliser, la similitude avec la génération de plans (Winograd, 1983) et plus généralement, les techniques de l'intelligence artificielle (Haton & Haton, 1989) s'est imposée à nous. De même, l'analyse de la signification

des coefficients probabilistes d'apprentissage P0 et P2 nous a permis de faire un lien avec le domaine de la logique et la notion d'implication ou d'équivalence.

Les techniques classiques de reconnaissance de formes ou de parole nous ont également été d'un grand secours lors de la réalisation des applications. Bien sûr, des algorithmes classiques nous ont été utiles pour certains prétraitements, mais c'est surtout un savoir-faire et une connaissance du domaine très précieux que nous ont apportés ces approches. Il est en effet inestimable, lorsqu'on aborde une application, d'avoir une connaissance a priori des différents traitements qu'il est intéressant d'effectuer sur l'information et des difficultés particulières relatives à ce domaine (Carbonell, Fohr, Haton, 1987). C'est ainsi qu'en reconnaissance de la parole, nous avons pu utiliser les analyses cepstrales et proposer des indices secondaires permettant une bonne discrimination des voyelles.

Nous nous sommes aussi employés à suivre les développements du connexionnisme et les réalisations qu'il propose. C'est ainsi que nous nous sommes sensibilisés à la notion de mémoire associative et que nous avons mesuré les capacités de notre modèle en suivant des procédures analogues à celles de certains autres modèles. Nous affichons également, dans ce domaine qui nous touche de près, une volonté de suivre et même d'anticiper les mouvements qui, selon nous, sont un avenir possible du connexionnisme. C'est ainsi que nous mettons l'accent sur les phénomènes temporels et sur la nécessité de tenter une approche symbolique pouvant mener au raisonnement.

Toutes ces empreintes sur notre modèle des sciences cognitives et des neurosciences font alors émerger la question suivante: "Mais pourquoi ce modèle se prétend-il nouveau?". S'il l'est, c'est sans doute par la combinaison particulière qu'il propose d'idées qui ne sont pas neuves mais qui existent toujours parce qu'elles sont pertinentes et qu'elles sont sûrement des éléments du puzzle cognitif.

Evidemment, nous nous sentons redevables envers tous ces domaines pour les multiples emprunts que nous y avons fait. Cependant, nous avons l'impression d'y jouer un rôle, réciproquement, et d'apporter en retour quelques propositions susceptibles d'intéresser certains chercheurs. Ainsi, il existe des données que nous n'avons pas pu (ou peut-être pas su) trouver ailleurs. Nous avons dû alors nous inspirer de notre expérience et de notre connaissance de divers mécanismes généraux pour tenter de proposer une

solution qui nous paraisse cohérente. Le cas s'est présenté dans l'application auditive où les indices primaires et secondaires des cartes sont dûs au recoupement de données éparses (Delgutte, 1984; Knudsen, du Lac, Esterly, 1987) et à une proposition de notre part d'un mécanisme possible (Guyot, Alexandre, Haton, Burnod, 1989b; Digeon, Alexandre, Guyot, Haton, 1989). Cette application réalisée, nous avons présenté notre approche et nos résultats aux biologistes et leur avons ainsi fourni une expérimentation possible dans le cortex auditif, jusqu'alors très peu étudié. D'autre part, le côté concret de notre travail permet la validation ou, en tout cas, le renforcement de plusieurs hypothèses de mécanismes corticaux proposés par Y. Burnod. D'une manière plus concrète, nos travaux sur la motricité vont également dans le sens d'expérimentations réalisées par Georgopoulos (Georgopoulos, Lurito, Petrides, Schwartz, Massey, 1989; Kettner, Schwartz, Georgopoulos, 1988) chez le singe, sur la rotation mentale des directions pendant une tâche de préhension.

Après avoir évoqué le tissu de relations qui nous a guidé tout au long de nos recherches, nous allons maintenant résumer nos travaux dans le sens des préoccupations que nous avons, trois ans plus tôt, dans notre sujet de thèse: "L'approche que nous proposons, moins fine que l'approche connexionniste, consiste à modéliser informatiquement les fonctions cérébrales et leurs interdépendances" (Alexandre, 1987).

Ces fonctions sont représentées de manière distribuée par un ensemble d'automates dont chacun réalise une fonction élémentaire. Ces automates sont regroupés en aires et sont distribués continûment de manière à occuper potentiellement tout l'espace de représentation. L'ensemble des aires est fonctionnellement homogène et représente un hypercube de dimension 6. Une aire est donc définie par l'indication de son appartenance à chacune des 6 symétries de l'hypercube qui représentent 6 alternatives possibles pour prendre en compte les différentes modalités de l'information. L'interdépendance des fonctions est prise en compte par les liens entre les automates, dont certains résultent de l'apprentissage. Ces liens sont typés et permettent à l'automate de traiter sélectivement l'information qu'il reçoit, selon son origine.

De ces traitements, résultent par apprentissage deux phénomènes. Il y a premièrement regroupement d'automates pour rendre compte des invariants du monde extérieur. Ensuite, ces invariants s'organisent temporellement afin de rendre compte de leurs relations causales. Ces liens définis par

apprentissage sont secondés et sont guidés dans leur mise au point par un ensemble de relations a priori qui définissent grossièrement les principales voies d'information entre les différents pôles d'entrée/sortie. Ces phénomènes sont le support de mécanismes possibles pour le raisonnement. La réalisation d'une fonction d'adaptation donnée passe par la définition de l'ensemble des aires réalisant cette fonction et par la mise au point des structures coopérantes permettant l'interface entre ce réseau et le monde extérieur.

L'ensemble des applications que nous avons réalisé propose un tour d'horizon de certaines capacités du modèle et a permis un affinement progressif de celui-ci. Cependant, cette description de la molécule à la fonction d'adaptation qui nous a semblé opportune a été nécessairement trop rapide et trop restrictive. Nous avons tenté de répondre à des questions mais avons, par la même occasion, posé de nouveaux problèmes. Cette étude définit donc la base préliminaire de travaux qu'il serait bon, selon nous, de poursuivre selon des directions que nous allons maintenant proposer.

2 Perspectives

Les applications que nous avons réalisé jusqu'à présent sont principalement situées dans la partie associative (par opposition à la partie frontale). Dans cette partie, la maîtrise des relations entre les aires doit être améliorée. Ceci passe par la réalisation d'un réseau comprenant toutes les aires associatives qui nous permettra de formaliser la communication entre les aires et de ne plus les définir "au coup par coup". Cette formalisation devra pouvoir englober ultérieurement la relation avec les aires frontales, puis avec les aires d'un deuxième hémisphère, que fera apparaître la prise en compte de la symétrie droite/gauche.

Les aires frontales qui auront accès, par leurs entrées externes, à différentes échelles de temps permettront de représenter les phénomènes temporels tels que l'échéancement des actions ou le contrôle d'une action en attendant un signal de déclenchement. La définition de la symétrie frontale doit permettre de manipuler la dimension temporelle à un niveau d'abstraction que ne peut pas atteindre le simple ordonnancement généré par les relations causales entre les modules des aires associatives.

La prise en compte de la symétrie droite/gauche est envisageable pour

les fonctions nécessitant deux analyses différentes du même type d'information (en vision, analyse des champs visuels droite et gauche permettant la binocularité; en parole, différenciation des bruits de fond et de la prosodie du signal de parole). La définition de cette symétrie (comme celle de toute autre symétrie) est également intéressante car elle fournit une possibilité supplémentaire de transfert d'apprentissage qui permettra la création de liens, par analogie.

La création d'un nombre croissant d'aires doit s'accompagner de la prise en compte de leur maturation consécutive permettant un apprentissage structuré. On a vu que les colonnes traitent sélectivement les informations d'après leur provenance et rangent les informations corticales dans des sous-couches de la couche haute qui s'étagent de manière à créer une hiérarchie basée sur la distance à la symétrie (plus la symétrie est lointaine, plus elle se projette près de la surface du cortex). Or, l'étude des stades de développement de l'enfant proposée par Piaget (Piaget, 1975) nous révèle l'ordre invariable selon lequel l'enfant acquiert et intègre ses capacités intellectuelles qui prennent successivement la conduite de ses comportements. Nous avons été extrêmement frappés de constater que cet ordre, ramené à la localisation de ces fonctions dans le cortex, correspondait à la maturation progressive et ordonnée des aires prenant successivement un certain contrôle des précédentes.

On peut résumer ici les principales étapes qui surviennent lors de l'ontogénèse. Le stade de l'intelligence sensori-motrice (0-2 ans) débouche sur une connaissance pratique de l'univers environnant; il peut, selon nous, correspondre à la maîtrise des aires sensori-motrices. Le stade de préparation et d'organisation de l'intelligence opératoire concrète (2-12 ans) voit l'apparition de la fonction symbolique (débouchant sur le langage) avant l'apparition des opérations concrètes de classification, de sériation et de maîtrise spatio-temporelle de l'environnement; ces opérations pourraient résulter du mûrissement des aires associatives. Le stade de l'intelligence opératoire formelle (12-16 ans) voit apparaître le perfectionnement des opérations abstraites et peut être interprété comme l'intégration des aires frontales. Cette ébauche nous donne déjà une indication sur l'ordre dans lequel il peut être intéressant d'instruire les différentes aires d'un réseau. Il est néanmoins évident qu'un tel sujet nécessite une étude beaucoup plus approfondie.

Si la définition d'un modèle de cortex de plus en plus complet est la

suite logique de l'étude que nous avons réalisée, il est également logique de parachever les aspects extra-corticaux de notre modèle. D'une part, nous avons vu l'importance des structures coopérantes en tant qu'auxiliaires du cortex; d'autre part la simulation de celles-ci, qui ne répond plus à la logique corticale de la colonne, est à l'origine de problèmes que notre étude a révélés. Fort heureusement, ces structures sont généralement un peu moins complexes que le cortex et ont parfois déjà subi des travaux de modélisation. Le cervelet, modélisé dans (Marr, 1969), devra être pris en compte dans les aspects moteurs; de même, la modélisation de l'hippocampe proposée dans (MacNaughton & Morris, 1987) peut être un auxiliaire précieux lors de la simulation du cortex limbique permettant une meilleure réalisation des axes temporeux.

Ces perspectives doivent aussi intégrer les aspects plus techniques de notre approche informatique. Il est ainsi important de rappeler que la définition des aires frontales, puis d'un hémisphère supplémentaire, double à chaque fois la taille du réseau. Malgré sa rapidité phénoménale, notre Sun4 n'en reste pas moins une machine séquentielle qui a, de plus, une mémoire centrale limitée (quoique, nous en convenons, respectable). Il est alors indispensable de prévoir d'autres sites pour l'implantation de nos programmes. Les transputers pourraient ainsi se révéler un support privilégié. D'une part, le parallélisme est source d'un gain de temps non négligeable; d'autre part, sa structure en hypercube qui destinerait chaque aire à un transputer dédié ne manque pas d'intérêt (Germain & Sansonnet, 1989). Des réalisations plus volumineuses pourraient également être envisagée sur une Connection Machine qui reculerait encore les contraintes temporelles de nos simulations. Enfin, il n'est pas interdit de penser que, lorsque le modèle de la colonne sera bien défini et testé, nous pourrons alors envisager une implantation de cette unité en VLSI. En effet, la petitesse du nombre de ses connexions ainsi que leurs dispositions régulières permet d'espérer la possibilité d'une telle entreprise. Dans un autre domaine, il est également intéressant de garder un oeil sur les progrès de l'opto-électronique et, plus généralement, de l'optique qui suscite potentiellement de nombreux espoirs, par ses propriétés de discrimination, de parallélisme et ses capacités de miniaturisation (Bize, 1989).

Dans un autre ordre d'idées, ces développements doivent aller de pair avec le renforcement de la coopération avec les chercheurs d'autres domaines des neuro-sciences et des sciences de la cognition.

Il nous reste encore beaucoup à apprendre de leurs recherches et de leurs idées. Nous avons déjà évoqué l'intérêt d'une approche piagetienne du développement du cortex. D'une manière similaire, il peut être intéressant d'étudier le développement foetal du cerveau. Il s'agirait alors de suivre les étapes de la construction et de l'interconnexion progressive des différentes structures neuronales selon des règles génétiques générales décrivant grossièrement le plan d'ensemble, le reste de la construction étant laissé à des mécanismes de sélection locaux. L'analyse de l'ensemble de ces données peut se révéler également riche d'enseignement pour une formalisation plus générale du système nerveux.

Avec la définition de réseaux de plus en plus importants et de fonctions d'adaptation de plus en plus abstraites, la prise en compte de données psychologiques sera de plus en plus cruciale. Ces données seront d'une part nécessaires pour évaluer la valeur de certains paramètres comportementaux (constantes de temps, capacités d'apprentissage, ...); elles seront, d'autre part, une source de contrôle possible des performances de nos programmes. Enfin, mentionnons également que l'étude de théories plus fondamentales (Laborit, 1974; Levi-Strauss, 1962;...) peut nous permettre également de mieux comprendre les implications de notre modèle.

En retour, il nous semble que la mise au point de réseaux de plus en plus complets peut se révéler intéressante pour des chercheurs qui disposeraient ainsi d'un nouveau champ d'expérimentations tant pour des expérimentations biologiques que pour des études psychologiques, impossibles sur le vivant ou destructrices. En plus de ces apports réciproques, il est également intéressant de pouvoir mener des études similaires de front, dans différents domaines.

On le voit, dans ce domaine, c'est le fonds qui manque le moins. On peut néanmoins, pour un futur proche, dégager une série d'applications qu'il serait intéressant de réaliser. On peut tout d'abord penser à généraliser la dernière application afin de proposer un système visuel complet qui dégage lui-même des stratégies d'exploration d'une scène quelconque. Il serait alors fort instructif de comprendre comment ces règles sont apprises et représentées dans les aires frontales. Les dernières applications que nous avons réalisées dans le domaine auditif ne peuvent que nous donner l'envie de pousser plus loin nos investigations dans le domaine de reconnaissance de la parole. A ce propos, Y. Burnod a proposé dans (Burnod, 1988) une série d'hypothèses très alléchantes sur la représentation possible des classes de

mots (adjectifs, substantifs, verbes, ...) dans le réseau des aires corticales. Cette représentation, ajoutée aux mécanismes classiques de la colonne corticale, permet de rendre compte de l'apprentissage et de l'utilisation de nombreux aspects de la parole (grammaire, sémantique, dialogue, ...). Il nous semble enfin important de compléter les travaux sur la motricité par une meilleure prise en compte des mécanismes extra-corticaux (musculaires, cérébelleux et spinaux) permettant une application possible en robotique.

Quelqu'application que l'on choisisse, il nous semble, en résumé, primordial de mettre l'accent sur les aspects suivants:

- apprentissage progressif
- propriétés émergentes
- pluridisciplinarité
- intégration du temps
- raisonnement
- fonctions d'adaptation anthropomorphiques

3 Conclusions personnelles

Cette partie ne serait pas complète sans une mention à l'aspect humain de ces travaux. Tout d'abord, le caractère passionnant de telles recherches ne peut qu'impliquer fortement celui qui les effectue. Au cours de discussions avec mes aînés, je me suis rendu compte que la possibilité d'aborder de tels sujets avec autant de liberté venait seulement d'être offerte aux chercheurs de ma génération. L'état des connaissances pousse de toutes parts à l'éclosion de tels travaux pluridisciplinaires. J'ai tenté, tout au long de cette thèse, de garder à l'esprit le rôle qu'il m'était ainsi proposé de remplir. Cette tâche m'a été grandement facilitée par les personnes avec qui j'ai travaillé. La complémentarité des idées de F. Guyot avec les miennes nous a permis des avancées que, seuls, nous n'aurions jamais pu réaliser. Ceci nous a incité à proposer l'idée d'un cortex à 7 symétries, somme de nos deux cortex, aux capacités non pas doubles mais différentes... Dans le domaine de la biologie, notre rencontre avec de nombreux chercheurs, et en particulier avec Y. Burnod, a été d'un enrichissement considérable, non seulement au point de vue scientifique, mais aussi au point de vue humain. Je considère ainsi cette période de thèse comme une initiation réussie à mon futur métier de chercheur dans ce domaine en pleine expansion et porteur

de nombreux espoirs qu'est cette approche pluridisciplinaire du mystère humain.

ANNEXE

Aspects Informatiques

Ce rapport a pour ambition de mettre en évidence les relations entre les différents domaines des neurosciences et des sciences cognitives, que l'on peut découvrir et dont on peut se servir dans un esprit informatique. C'est la raison pour laquelle nous avons choisi de rassembler en annexe les aspects de nos travaux relatifs au matériel, au langage et à la structure des données.

Environnement matériel:

Nous avons eu accès, durant nos deux premières années de thèse, à un Sun 3-60. C'est donc sur cette station de travail, et son environnement maintenant classique (multi-fenêtrage et souris), que les trois premières applications ont été réalisées et que leur rapidité a été mesurée. Notre dernière année de thèse a vu l'apparition, dans notre bureau, d'un Sun 4 SparcStation, de 8 Mo de mémoire et avec écran couleur. Cette mémoire limitée nous fait nous intéresser maintenant, pour des applications de plus grande taille, à la machine TNode à transputers également disponible au CRIN.

Langage

Les applications ont été écrites en C en raison de la rapidité de ce langage et du large éventail de primitives dédiées, aussi bien de bas que de haut niveau, qu'il permet de spécifier. Même si ce langage n'est sûrement pas optimal pour une utilisation rationnelle d'un modèle tel que la colonne corticale, il nous a semblé très satisfaisant pour sa mise au point.

Structure des données:

Les structures suivantes, extraites d'une application, rendent compte de la connectivité entre unités.

```
    struct zone
    {
    char type;
    struct zone *fg,*fd;
    struct relation *tete1;
    struct relation *tete2;
    struct relation *tete3;
    int tempactive;
    int taille;
    char state,pstate;
    int nbc;
    char ethal,ecort;
    };
    struct relation
    {
    int nbcoups;
    int experience;
    int depot;
    struct zone *soma;
    struct relation *recepteur;
    struct relation *suivant;
    };
```

La représentation graphique que l'on peut donner des liens établis par cette structure est à rapprocher de la modélisation graphique de la colonne biologique proposée dans (Burnod, 1988).

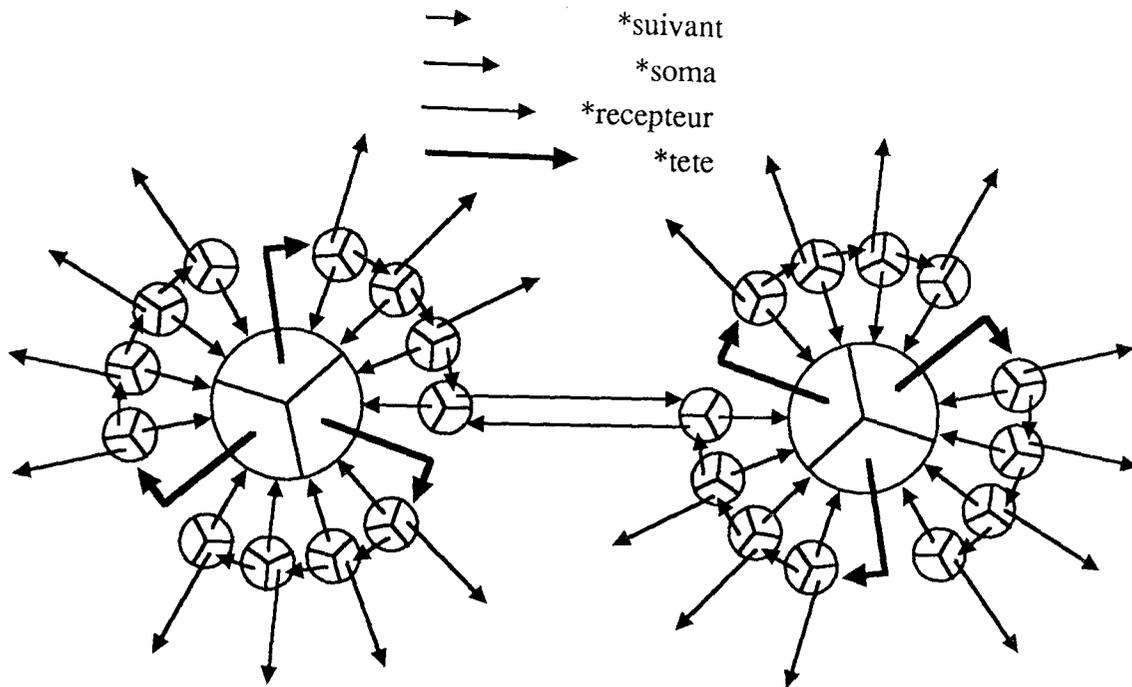


Figure 1: structure des données

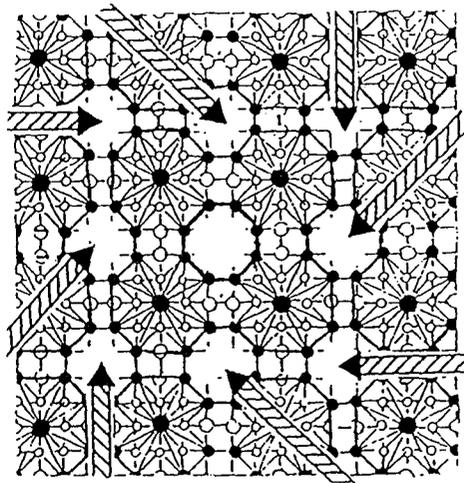


Figure 2: modèle de la colonne

BIBLIOGRAPHIE

- Alexandre, F. (1987) "Modèles connexionnistes pour l'apprentissage et le raisonnement de systèmes intelligents", Sujet de Thèse.
- Alexandre, F., Burnod, Y., Guyot, F., Haton, J. P. (1988) "La colonne corticale: nouvelle unité de base pour des reseaux multicouches", Proc. Neuro-Nimes'88, Nimes.
- Alexandre, F., Burnod, Y., Guyot, F., Haton, J. P. (1989) "La colonne corticale, unité de base pour des réseaux multicouches", C R Acad. Sci. Paris, t. 309, Série III, p. 259-264.
- Azencott, R. (1989) "Boltzmann machines: High-order interactions and synchronous learning", IEEE PAMI.
- Baird, H. S. (1988) "Feature identification for hybrid structural/statistical pattern classification", Computer vision, graphics, and image processing, 42, p. 318-333.
- Baird, H. S., Thompson, K. (1987) "Reading chess", Proceedings of workshop on computer vision, Miami Beach, FL, Nov 30-Dec 2, 87.
- Ballard, D. H. (1986) "Cortical connections and parallel processing: Structure and function", The behavioral and brain sciences 9, p. 67-120.
- Barlow, H. B. (1972) "Single units and sensation: A neuron doctrine for perceptual psychology?", Perception, 1, p. 371-394.
- Barrionuevo, G., Brown, T. H. (1983) "Associative long-term potentiation in hippocampal slices", Proceedings of the National Academy of Sciences, Neurobiology, 80, p. 7347-51.
- Bernstein, M. (1967) "Coordination and regulation of movements", Pergamon press, New-York.
- Béroule, D. (1985) "Un modèle de mémoire adaptative, dynamique et associative pour le traitement automatique de la parole", Thèse de 3ème cycle, Orsay.

- Berthod, M., Maroy, J. P. (1974) "Morphological features and sequential information in real time handprinting recognition", Second International Joint Conference on Pattern Recognition, Copenhagen, August 74.
- Bienestock, E. (1986) "Connectionist models of perception: a short review", Fyssen Foundation Conference in Versailles March 86.
- Bize, D. (1989) "Le développement de la recherche sur les réseaux neuronaux aux Etats-Unis", Rapport de fin de contrat, 20 /SGDN/STS-/VST/1.
- Boulevard, H., Wellekens, C. (1987) "Multi-Layer Perceptrons and Automatic Speech Recognition", Proc. IEEE First Int. Conf. on Neural Networks, San Diego.
- Bozinovic, R. M., Srihari, S. N. (1989) "Off-line cursive script word recognition", IEEE transactions on pattern analysis and machine intelligence, 11, 1, p. 68-82.
- Brodmann, K. (1909) "Vergleichende Lokalisationslehre der Großhirnrinde", Leipzig, J.A.Barth.
- Bullier, J. (1983) "Les cartes du cerveau", La Recherche, 14, 148, p. 1202-1214.
- Burnod, Y. (1988) "An adaptive neural network: The cerebral cortex", Masson, Paris.
- Carbonell, N., Fohr, D., Haton, J. P. (1987) "APHODEX, An Acoustic-Phonetic Decoding Expert System", Int. J. Pattern Recognition and Artificial Intelligence, 1, 2, p. 207-222.
- Changeux, J. P. (1983) "L'homme neuronal", Fayard.
- Danset, A. (1983) "Eléments de psychologie du développement", Armand Colin, Paris.
- Davalo, E., Naim, P. (1989) "Des réseaux de neurones", Eyrolles, Paris.
- Delacour, J. (1987) "Apprentissage et mémoire. Une approche neurobiologique", Masson.
- Delgutte, B. (1984) "Codage de la Parole Dans le Nerf Auditif", Thèse, Université Paris VI.
- Dingeon, C., Alexandre, F., Guyot, F., Haton, J. P. (1989) "Un autre apprentissage cortical: Différentiel pour généraliser", Proc. Neuro-Nimes'89, Nimes.
- Eccles, J. C., Ito, M., Szentagothai, J. (1967) "The cerebellum as a neuronal machine", Springer Verlag, Berlin.

- Elman, J. L., Zipser, D. (1987) "Learning the hidden structure of speech", J. Acoust. Soc. Am., p. 1615-1626.
- Evarts, E. V., Tanji, J. (1974) "Gating of motor cortex reflexes by prior instruction", Brain Research, 71, p. 479-494.
- Feldman, J. A., Ballard, D. H. (1982) "Connectionist models and their properties", Cognitive science, 6, p. 205-254.
- Fogelman Soulié, F. (1987) "Le connexionnisme", Support de cours, MARI 87, Cognitiva 87.
- Fogelman Soulie, F., Gallinari, P., Le Cun, Y., Thiria, S. (1989) "Network learning", in Machine learning vol 3 Y Kodratoff, R Michalski Eds.
- Fregnac, M., Imbert, M. (1984) "Development of neuronal selectivity in primary visual cortex of cat", Physiological Review 64.
- Fukushima, K. (1980) "Neocognitron: A self-organizing neural network for a mechanism of pattern recognition unaffected by shift in position", Biol. Cybernetics, 36, p. 193-202.
- Fuster, J. M. (1977) "Unit activity in the prefrontal cortex during delayed response performance: neuronal correlates of short-term memory", Journal of Neurophysiology, 36, p. 61-78.
- Georgopoulos, A. P., Lurito, J. T., Petrides, M., Schwartz, A. B., Massey, J. T. (1989) "Mental rotation of neuronal population vector", Science, 243, p. 234-236.
- Germain, C., Sansonnet, J. P. (1989) "L'architecture des machines massivement parallèles pour l'intelligence artificielle", 11^{èmes} journées francophones sur l'informatique, Nancy.
- Geshwind, N. (1984) "Les spécialisations du cerveau humain", Le Cerveau, Belin.
- Gilbert, C. D., Wiesel, T. N. (1981) "Laminar specialization and intracortical connections in cat primary visual cortex", In The organization of the cerebral cortex, F. Schmitt et al eds., Cambridge, Mass., MIT Press.
- Girard, R. (1978) "Des choses cachées depuis la fondation du monde", Grasset.
- Gori, M., Bengio, Y. DeMori, R. (1989) "BPS: A learning algorithm for capturing the dynamic nature of speech", Proc. IEEE IJCNN'89, Washington.
- Grossberg, S. (1988) "Nonlinear neural networks: Principles, mechanisms, and architectures", Neural Networks, 1, p. 17-61.

- Guillaume, P. (1979) "La psychologie de la forme", Flammarion.
- Guyot, F. (1990) "Une modélisation fonctionnelle du cortex: la colonne corticale. Aspects auditifs et visuels", Thèse de l'Université de Nancy I.
- Guyot, F., Alexandre, F., Haton, J.P. (1989) "Toward a continuous model of the cortical column: application to speech recognition", Proc. ICASSP-89, Glasgow.
- Guyot, F., Alexandre, F., Haton, J.P., Burnod, Y. (1989a) "The Cortical Column, a new Processing Unit for Cortex-Like Networks", Proc. CEC Workshop "From the Pixels to the Features", Elsevier.
- Guyot, F., Alexandre, F., Haton, J.P., Burnod, Y. (1989b) "A potentially powerful connectionist unit: the cortical column.", NATO Advanced Research Workshop on Neuro Computing.
- Hamburger, J. (1985) "Les miroirs multiples de la connaissance", Le Monde, 18 Mai 1985.
- Haton, J. P., Haton, M. C. (1989) "L'intelligence Artificielle", PUF, Paris.
- Hebb, D. O. (1949) "The organization of behaviour", Wiley, New-York.
- Hinton, G. (1984) "Parallel computations for controlling an arm", Journal of motor behavior, 16, 2, p. 171-194.
- Hinton, G., Sejnowski, T., Ackley, D. H. (1984) "Boltzmann machines: constraint satisfaction networks that learn technical report", Carnegie Mellon University.
- Hofstadter, D. (1985) "Godel, Escher et Bach", Interéditions.
- Hopfield, J. J. (1982) "Neural network and physical systems with emergent collective computational abilities", Proc. Natl. Acad. Sci., 79, p. 2554-2558.
- Hopfield, J. J., Tank, D. W. (1985) "Neural computation of decisions in optimisation problems", Biological Cybernetics, 52, p. 141- 152.
- Hubel, D. H. (1988) "Eye, brain, and vision", Scientific American Library, New York.
- Hubel, D. H., Wiesel, T. N. (1962) "Receptive fields binocular interaction and functional architecture in the cat's visual cortex". J. Physiol. Lond., 160, p. 106-154.
- Hubel, D. H., Wiesel, T. N. (1977) "Functional architecture of macaque monkey visual cortex", Ferrier Lecture Proc. Roy. Soc. Lond.B, 198, p. 1-59.

- Hubel, D., Wiesel, T. (1978) "Les mécanismes cérébraux de la vision", Le Cerveau, Belin.
- Imbert, M. (1983) "La neurobiologie de l'image", La Recherche, 14, 144, p. 600-613.
- Ito, M., Sakurai, M., Tongroach, P. (1982) "Climbing fiber induced depression of both mossy fibre responsiveness and glutamate sensitivity of cerebellar purkinje cells", J. Neurophysiol., 324, p. 113-134.
- Jacob, F. (1981) "Le jeu des possibles", Fayard.
- Jeannerod, M. (1974) "Les deux mécanismes de la vision", La recherche, 5, p. 41:23-32.
- Jones, E. G. (1981) "Identification and classification of intrinsic circuit elements in the neocortex", In Dynamic aspects of neocortical functions, G. M. Edelman, E. Gall, W. M. Cowan eds, John Wiley New-York.
- Kaas, J. H., Nelson, R. J., Sur, M., Merzenich, M. M. (1979) "Multiple representation of the body within the primary somatosensory cortex of primates", Science, 204, p. 521-523.
- Kettner, R. E., Schwartz, A. B., Georgopoulos, A. P. (1988) "Primate motor cortex and free arm movements to visual targets in three dimensional space", Journal of Neuroscience.
- Kienker, P. K., Sejnowski, T. J., Hinton, G. E., Schumacher, L. E. (1986) "Separating Figure from Ground with a Parallel Network", Perception, 15, p. 197-216.
- Kirkpatrick, S. Gellat, C. D., Vecchi, M. P. (1983) "Optimisation by simulated annealing", Science, 220, 4598, p. 671-680.
- Knudsen, E., du Lac, S., Esterly, S. (1987) "Computational maps in the brain", American Revue of Neuroscience, 10, p. 41-65.
- Kohonen, T. (1984a) "Self Organization and Associative Memory", Springer.
- Kohonen, T. (1984b) "Phonotopic maps. Insightful representation of phonological features for speech recognition", IEEE Pattern Recognition, Montréal.
- Kohonen, T. (1988a) "The "Neural" Phonetic Typewriter", IEEE Computer, 21, n 3, p. 25-40.
- Kohonen, T. (1988b) "An introduction to neural computing", Neural Networks, 1, p. 3-16.
- Krauth, W., Mézard, M. (1987) "Learning algorithms with optimal stability in neural networks", J. Phys., A 20, L 745.

- Laborit, H. (1974) "La nouvelle grille", Robert Laffont.
- LeCun, Y. (1987) "Modèles connexionnistes de l'apprentissage", Thèse de Doctorat, Université Paris VI.
- Levi-Strauss, C. (1962) "La pensée sauvage", Plon.
- Lorente de No, R. (1938) "The cerebral cortex: architecture, intracortical connections and motor projections", In *Physiology of the nervous system*, J.Fulton, ed., Lond., Oxford University Press, p. 291-301.
- Lorenz, K. (1970) "Comportement animal et humain", Editions du Seuil, Paris.
- MacCulloch, W. S., Pitts, W., (1943) "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, 5, p. 115-133.
- MacNaughton, B. L., Morris, R. G. M. (1987) "Hippocampal synaptic enhancement and information storage within a distributed memory system", *Trends Neurosci.*, 10, 10, p. 408-415.
- Marr, D. (1969) "A theory of cerebellar cortex", *Journal of Physiology*, 202, p. 437-470.
- Memmi, D. (1989) "Connectionism and artificial intelligence", *Proc. Neuro-Nimes'89*, Nimes.
- Merzenich, M. M., Jenkins, W. M., Middlebrooks, J. C., (1984) "Observations and hypothesis on special organisational features of the central auditory nervous system", *Dynamic aspects of neocortical functions*, G.M.Edelman, W.E.Gall, W.M.Cowan eds, John Wiley New-York.
- Minsky, M., Papert, S. (1969) "Perceptrons", MIT Press, Cambridge MA.
- Mountcastle, V. B. (1957) "Modality and topographic properties of single neurons of cat's somatic sensory cortex", *J. Neurophysiol.*, 20, p. 408-434.
- Mountcastle, V. B., Lynch, J. C., Georgopoulos, A., Sakata, H., Acuna, A. (1975) "Posterior parietal association cortex of the monkey: Command functions for operations within extra-personal space", *J. Neurophysiol.*, 38, p. 871-908.
- Mountcastle, V. B. (1978) "An organizing principle for cerebral function. The unit module and the distributed system", *The mindful brain*, MIT Press. Cambridge.
- Mountcastle, V. B., Andersen, R. A., Motter, B. C. (1981) "The influence of attentive fixation upon the excitability of the light-sensitive neurons of the posterior parietal cortex", *Journal of Neurosciences*, 1, p. 1218-

- 1235.
- Nauta, W., Feirtag, M. (1984) "L'organisation du cerveau", Le Cerveau, Belin.
- Nilsson, N. J. (1986) "Probabilistic logic", Artificial Intelligence, 28, p. 71-87.
- Oja, E. (1989) "Neural networks, principal components, and subspaces" Int. journ. of neural systems, 1, 1, p. 61-68.
- Olberg, R. M. (1981) "Parallel encoding of direction of wind, head, abdomen, and visual pattern movement by single interneurons in the dragonfly", Journal of comparative physiology, 142, p. 27-41.
- Paillard, J. (1987) "L'ordinateur et le cerveau: Un contraste saisissant", AFCET/Interfaces, 57, p. 4-9.
- Peeling, S. M., Moore, R. K., Tomlinson, M. J. (1986) "The multi-layer perceptron as a tool for speech pattern processing research", Autumn Conf on Speech and Hearing.
- Piaget, J. (1975) "L'équilibration des structures cognitives, problème central du développement", PUF, Paris.
- Poggio, T., Koch, C. (1987) "La perception visuelle du mouvement", Pour la Science, Juillet 87, p. 26-34.
- Robert, J. M. (1982) "Comprendre notre cerveau", Editions du Seuil, Paris.
- Robinson, D. L., Goldberg, M. E., Stanton, G. B. (1975) "Parietal association cortex in the primate: sensory mechanisms and behavioral modulations", J. Neurophysiol., 41, p. 910-932.
- Rosenblatt, F. (1962) "A comparison of several perceptron models", in Self Organizing Systems 1962, ed. Yovits, Jacobi, Goldstein, Spartan Books, p. 463-484.
- Rumelhart, D. E., MacClelland, J. L. (1986) "Parallel distributed processing", MIT Press, Cambridge.
- Saerens, M., Soquet, A. (1989) "Un système de controle connexionniste", Quatrièmes journées françaises de l'apprentissage.
- Sejnowski, T. J., Rosenberg, C. R. (1986) "NETTalk: a parallel network that learns to read aloud", The John Hopkins University Electrical Engineering and Computer Science Technical Report, JHU/EECS-86/01, 32p.
- Shaw, G. L., Silverman, D. J., Pearson, J. C. (1988) "Trion model of cortical organization and the search of the code of short-term memory

- and of information processing", Systems with learning and memory abilities. J. Delacour and J. C. S. Levy Eds, Elsevier, North-Holland.
- Shaw, G. L., Palm, G. (1988) "Brain Theory" World Scientific Publishing Co, Eds Shaw, Palm.
- Shafer, G., Logan, R. (1987) "Implementing Dempster's rule for hierarchical evidence", Artificial Intelligence, 33, p. 271-298.
- Stein, D. (1989) "Les verres de spin", Pour la science, 143, p. 30-37.
- Stevens, C. (1984) "Le neurone", Le Cerveau, Belin.
- Szentagothai, J. (1973) "Synaptology of the visual cortex", Handbook of sensory physiology, Springer-Verlag.
- Szentagothai, J. (1975) "The "module concept" in cerebral cortex architecture", Brain Research, 95, p. 475-496.
- Van Essen, D. C., Maunsell, J. H. R. (1983) "Hierarchical organization and functional streams in the visual cortex", Trends in neurosciences, 6, 9.
- Vincent, J. D. (1986) "Biologie des passions", Odile Jacob.
- Vurpillot, E. (1972) "Le monde visuel du jeune enfant", PUF, Paris.
- Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., Lang, K. J. (1989) "Phoneme recognition using Time Delay Neural Networks", IEEE Transactions on ASSP, 37, 3.
- Widrow, B., Hoff, M. E. (1960) "Adaptive switching circuits", 1960 IRE WESCON Convention Record, New-York, p. 96-104.
- Winograd, T. (1983) "Language as a cognitive process", Addison Wesley Publishing Company.
- Woolsey, C. N. (1958) "Organization of somatic sensory and motor areas of the cerebral cortex". In Biological and biochemical bases of behaviour, H. F. Harlow and C. N. Woolsey eds, University Wisconsin Press, Madison, p. 63-81.
- Zadeh, L. A. (1979) "Fuzzy sets and information granularity", in Advances in fuzzy sets theory and applications, p. 3-18, North-Holland, Amsterdam.
- Zadeh, L. A. (1988) "Fuzzy Logic". Computer, Avril 88, p. 83-93.
- Zeki, S., Shipp, S. (1988) "The functional logic of cortical connections", Nature 335, 311-316.

Index

- A**
acquisition 3
aire 19, 21, 42, 50, 71
 associative 44
 frontale 45
 motrice 42
 réseau 45
 sensorielle 42
apprentissage 5, 30
 hebbien 29
arbre d'appel 126
audition 3
auto-association 30, 34
auto-organisation 112
- B**
bande 117
but 71, 92, 118, 122, 126, 127, 128
- C**
carte 90, 98, 109
 de codage 62
 topologique 38
cervelet 116
champ récepteur 88, 91, 92, 93, 99,
 108, 123, 128, 131
chemin 122, 129
codage 87, 89, 92, 115
 absolu 120
 relatif 120
colliculus 116
colonne 21, 52
 sous couches 56
condition
 nécessaire 96
 suffisante 96
connexion 90
convergence 94
corps genouillé latéral 88
cortex 17
 moteur 118
 visuel 88, 89, 118
couches 30
cycle de fonctionnement 100
- D**
diagnostic 3
différenciation 71, 109
divergence 94
division 123
- E**
évènement 3
extraction 2
- F**
fait 3
feedback 87
flou 92
fonctionnalité 2
- G**
généralisation 71, 92
- H**
hétéro-association 30, 85
- I**
indice
 primaire 90, 98, 104, 117, 120
 secondaire 90, 98, 104
Intelligence Artificielle 1
intervalle 89, 125
invariant extérieur 71

- L** logique 67, 96
d'apprentissage 94
relationnelle 96
structurelle 96
- M** macroscopique 7, 8
masque 68, 88, 99
memorisation associative 2
microscopique 7, 8
modèle 2, 4, 7, 8
motricité 114
mouvement
oculaire 132
mouvement balistique 119
muscle 118
- N** nerf optique 88
neurone 5, 14, 20, 22, 25, 56, 64,
89, 116
formel 26
- O** orientation 53, 88
- P** paramètre 90
pariétal 131
pluridisciplinarité 5
prétraitement 2
pro-action 94
propagation 69, 71, 122
psychologie
branches 8
subjectivité 8
- R** raisonnement 3, 115
recombinaison privilégiée 63, 118
reconnaissance 2
de caractère 85, 97
- de la parole 103, 108
recouvrement 91
rétine 87
rétinotopie 88, 90, 118
rétro-action 86, 93, 96
- S** simulation 100
symbolique 3
symétrie hypercubique 49
système expert 7
- T** temporal 131
temps 24
thalamus 16, 19
- V** variabilité 3
vision 2

NOM DE L'ETUDIANT : ALEXANDRE Frédéric

NATURE DE LA THESE : Doctorat de l'Université de NANCY I en Informatique

VU, APPROUVE ET PERMIS D'IMPRIMER

NANCY, le 07 AVR. 1990 n°745

LE PRESIDENT DE L'UNIVERSITE DE NANCY I



100

Résumé

Depuis une dizaine d'années, la recherche sur les réseaux de neurones s'est beaucoup développée. Ce regain d'intérêt pour le connexionnisme est dû à la mise au point d'un formalisme mathématique des réseaux. Le connexionnisme a longtemps développé ses capacités de mémoire associative pour des tâches de mise en correspondance. Il tente maintenant de s'attacher au traitement de phénomènes temporels.

Nous proposons ici une nouvelle approche connexionniste, fondée sur la modélisation de la colonne corticale, architecture modulaire du cortex. Une étude approfondie du cortex nous a permis de dégager les différentes propriétés de notre modèle.

Ces propriétés sont décrites autour d'une description statique puis dynamique de la colonne. Elles débouchent sur une interprétation de l'automate colonnaire qui peut être alors considéré comme une fonction élémentaire.

De plus, l'étude de ce modèle porte également sur les propriétés émergentes de réseaux composés de tels automates. Ceci nous conduit alors à proposer la localisation de nombreuses fonctions d'adaptation humaines.

Tout au long de cette étude, nous précisons les emprunts, les coïncidences et les relations de notre modèle vis à vis d'autres domaines des neurosciences et des sciences cognitives afin d'affirmer la pluridisciplinarité de notre approche.

Des applications informatiques sont également proposées comme illustration et validation de nos propositions. Nous mettons surtout l'accent ici sur les aspects visuels et moteurs de ces applications.

Mots clés :

-Connexionnisme-Cortex-Colonne corticale-

-Vision-Motricité-

-Modélisation-