

90/400

Université  
de Nancy I  
Inria Lorraine

Sc N 90/  
388 A

Centre de Recherche en  
Informatique de Nancy

# Une modélisation fonctionnelle du cortex: la Colonne Corticale.

*Aspects auditifs et visuels.*

## THÈSE

présentée et soutenue publiquement le 9 Avril 1990

pour l'obtention du

Doctorat de l'Université de Nancy I

(Spécialité Informatique)

par

Frédéric Guyot



### Composition du jury :

*Président :* J.M. PIERREL

*Rapporteurs :* M. BERTHOD  
M. CRÉHANGE

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

*Invité :* Y. BURNOD

Université  
de Nancy I  
Inria Lorraine

Centre de Recherche en  
Informatique de Nancy

# Une modélisation fonctionnelle du cortex: la Colonne Corticale.

*Aspects auditifs et visuels.*

## THÈSE

présentée et soutenue publiquement le 9 Avril 1990

pour l'obtention du

Doctorat de l'Université de Nancy I

(Spécialité Informatique)

par

Frédéric Guyot



### Composition du jury :

*Président :* J.M. PIERREL

*Rapporteurs :* M. BERTHOD  
M. CRÉHANGE

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

*Invité :* Y. BURNOD

Plus qu'un diplôme, cette thèse représente un tournant dans ma vie: fin des études et surtout début d'une période que j'espère constructive. C'est pourquoi je la dédie à tous ceux qui, de près ou de loin, ont contribué à mon évolution spirituelle: mes parents, qui m'ont toujours soutenu et encouragé, matériellement et intellectuellement, mes soeurs, qui m'ont fait découvrir la patience mais aussi la détermination, ma compagne, contrepoint féminin indispensable, qui sait moduler mes excès, et enfin mes amis qui, tout simplement, me supportent et m'enrichissent par la diversité de leurs expériences personnelles.

## Remerciements

Le mythe du chercheur solitaire ne résiste pas à la confrontation avec la réalité. Cette thèse résulte certes d'un travail personnel, mais elle n'aurait jamais pu aboutir sans le concours de certaines personnes que je tiens à remercier ici:

Mr. Jean-Marie Pierrel, professeur à l'Université de Nancy I. Parmi les premiers, Mr. Pierrel a soutenu mon approche, dans le cadre très informel des réunions de laboratoire. Ses remarques pertinentes et stimulantes m'ont accompagné durant ces trois ans. En acceptant la présidence du jury, il me témoigne son intérêt pour mes travaux; je l'en remercie vivement.

Mr. Marc Berthod, directeur de recherche à l'INRIA. Mr. Berthod a accepté de siéger dans ce jury et d'être rapporteur de ma thèse; j'ai pu tirer profit de son expérience pour améliorer la clarté de mon exposé; qu'il trouve ici mes remerciements pour l'intérêt dont il fait preuve à mon égard.

Mme Marion Créhange, professeur à l'Université de Nancy I. Mme Créhange a procédé à une relecture vraiment approfondie de mon rapport. Ses remarques, tant sur le fond que sur la forme, motivées par une soif de connaissances inextinguible, constituent à mon avis un modèle du genre et m'ont été plus qu'utiles. En me faisant l'honneur de juger mon travail, elle cautionne mes recherches; je l'en remercie sincèrement.

Mr. Jean-Pierre Aubin, professeur à l'Université de Paris Dauphine. Mr. Aubin m'a fait le grand honneur de s'intéresser à ma thèse, en siégeant dans ce jury.

Mr. Jean-Paul Haton, professeur à l'université de Nancy I. Son engagement total dans mon thème de recherche, allié à sa grande connaissance du métier et à sa disponibilité, en font un directeur de recherche inestimable. Il a su me motiver et me diriger avec un discernement qui en dit long sur ses qualités professionnelles et humaines. J'ai beaucoup apprécié la liberté

qu'il m'a laissée dans mes recherches, et la confiance qu'il m'a accordée. Qu'il trouve ici l'expression de ma plus sincère gratitude.

Mr. Yves Burnod, chargé de recherche à l'INSERM. Mr. Burnod est sans conteste le chercheur à qui je dois le plus; dès nos premières rencontres, il n'a pas hésité à passer des heures pour m'initier aux mystères de la neurobiologie. Au fil de nos longues conversations, qu'il sait rendre passionnées et captivantes, j'ai été impressionné par ses intuitions géniales, ses capacités de synthèse et l'étendue de sa culture. Je tiens à lui témoigner ma profonde reconnaissance et mon amitié.

Je tiens également à remercier tous les membres du CRIN, qui participent à l'ambiance amicale et détendue, propice à la recherche. Plus particulièrement, Martine et Nadine, sans qui tout irait de travers, Agnès, qui bichonne nos ordinateurs, Karl, qui sait (presque) tout, et bien sur Alex, avec qui se sont tissées des relations, entre autre professionnelles, où émulation et coopération ne sont pas des vains mots.

#### *Avertissement*

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

# SOMMAIRE

## Chapitre I: INTRODUCTION

### **1. Intelligence Artificielle**

- Vision
- Audition
- Raisonnement
- Bilan

### **2. Sciences Cognitives**

- Neurobiologie
- Psychologie

### **3. Coopération**

### **4. Démarche**

- Choix d'un modèle
- Plan général
- Avertissement

## Chapitre II: TROIS NIVEAUX D'ORGANISATION DU CORTEX

### 1. Introduction au système nerveux

- Le système nerveux périphérique
- Le système nerveux central
- Les voies d'information
- Trois niveaux d'organisation du cortex

### 2. Le neurone

- Le neurone vivant
- Le neurone formel
- Les réseaux à couches
- Les réseaux à satisfaction de contrainte
- Les autres modèles
- Discussion

### 3. Les aires

- Les types d'aires
- Le réseau

### 4. La colonne corticale

- Réalité biologique de la colonne
- Architecture et connectivité
- Etats d'activation
- Règles de fonctionnement
- Apprentissage
- Implantation du modèle
- Conclusion sur la modélisation

## Chapitre III: APPLICATIONS

### 1. Vision

- Introduction au problème
- Données neurobiologiques
- Le réseau
- Le fonctionnement
- Résultats

### 2. Reconnaissance de la parole

- Introduction
- Données neurobiologiques
- Données phonétiques
- Le réseau
- Fusion/division et propagation
- Résultats
- Remarques

### 3. Motricité

- Introduction au problème
- Neurobiologie
- Le réseau
- Le fonctionnement

### 4. Interprétation d'images

- Introduction au problème
- Données neurobiologiques
- Données psychologiques
- Le réseau d'interprétation d'image
- Signal, contexte, requête
- Application
- Remarques

## Chapitre IV: CONCLUSION

### 1. Critiques

- Réalité biologique
- Capacités de traitement de l'information
- Critique de la démarche
- Critiques psychologiques

### 2. Intérêt

- Comparaison aux méthodes classiques
- Comparaisons connexionnistes

### 3. Réflexions

- Theorie
- Modélisation

### 4. Perspectives

- Court terme
- Moyen terme
- Long terme

## REFERENCES

## ANNEXES

### 1. Résultats associatifs

### 2. Interprétation de clichés

### 3. Informatique

## INDEX

## 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 se situe 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 événement,

**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 "événements" 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'entreprennent 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 recoupe 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-t-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
	Diencephale	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 sensitives, 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).

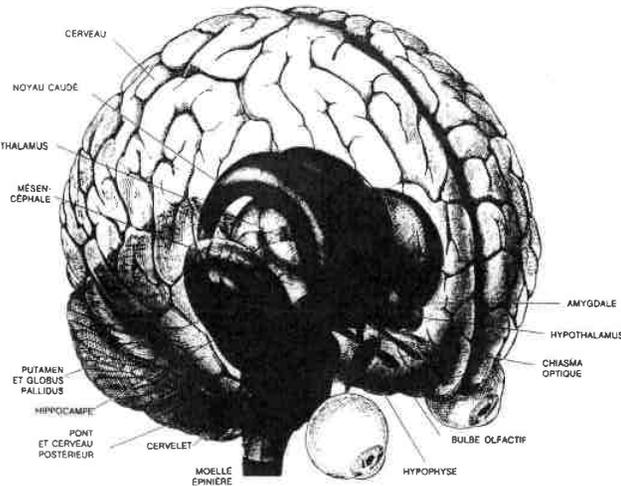


Figure 3. Le cerveau (d'après Nauta & Feirtag, 1984).

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 diencéphale 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.

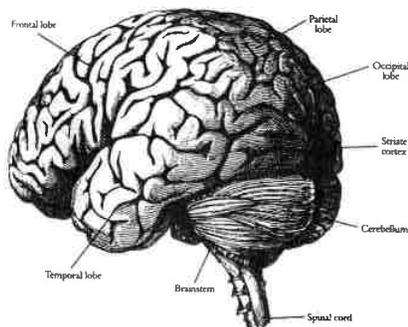


Figure 4. Régions corticales. (d'après Hubel, 1988)

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 pourrions 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).

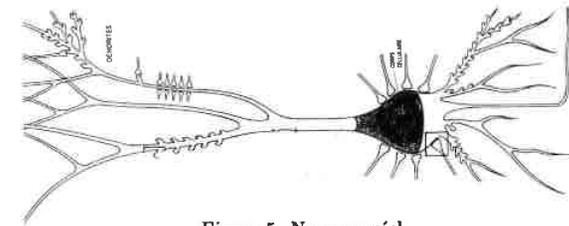


Figure 5. Neurone réel.

#### 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.

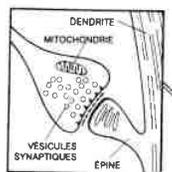


Figure 6. Synapse.

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 dues 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.

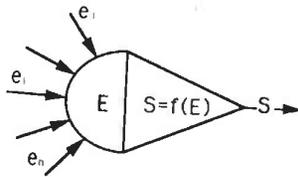


Figure 7. Neurone formel.

### 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

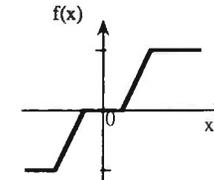


Figure 9. Fonction linéaire à seuils.

- une fonction stochastique:

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

$f(x)=0$  sinon.

- une fonction sigmoïde:

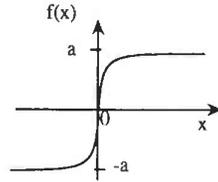


Figure 10. Fonction sigmoïde.

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

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:

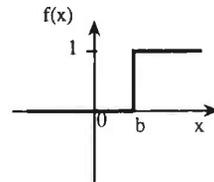


Figure 11. Fonction de MacCulloch &amp; Pitts.

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 & McClelland, 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" (Bienenstock, 1986; Kohonen, 1988b) que nous avons choisis 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. Étudions en quelques uns.

### 2.3.1 Le perceptron

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

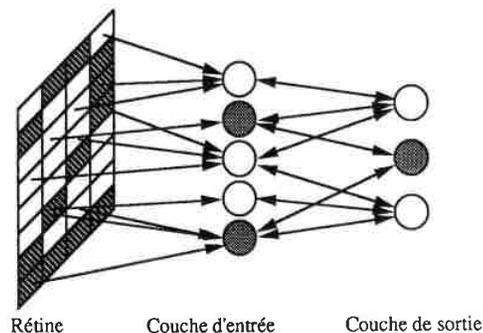


Figure 12. Perceptron.

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

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

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 apportée 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).

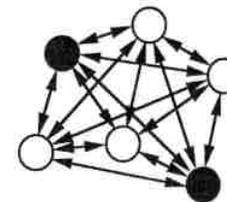


Figure 13. Réseau de Hopfield.

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, débarrassé 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.

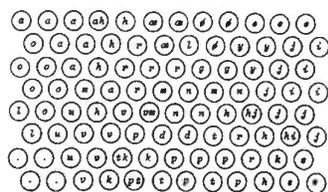


Figure 14. Carte topologique (d'après Kohonen, 1984b).

### 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...).

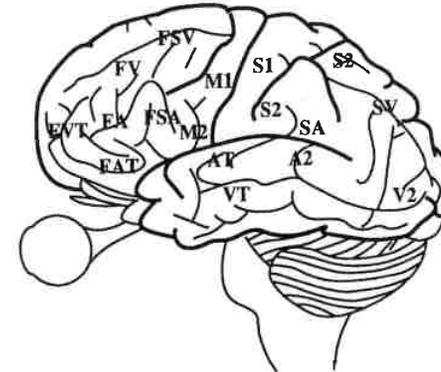


Figure 15. Les aires corticales.

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.

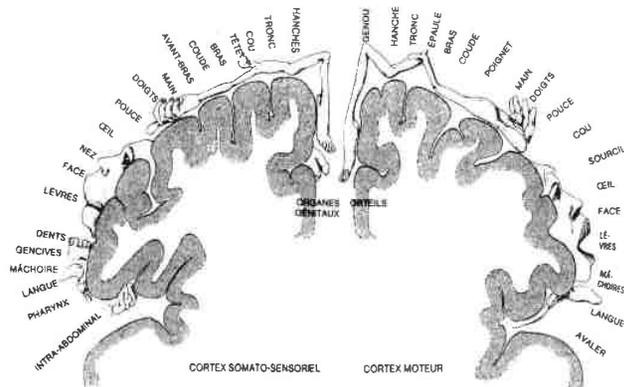


Figure 16. L'homunculus (d'après Geshwind, 1984).

### 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 sensorimotrices 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).

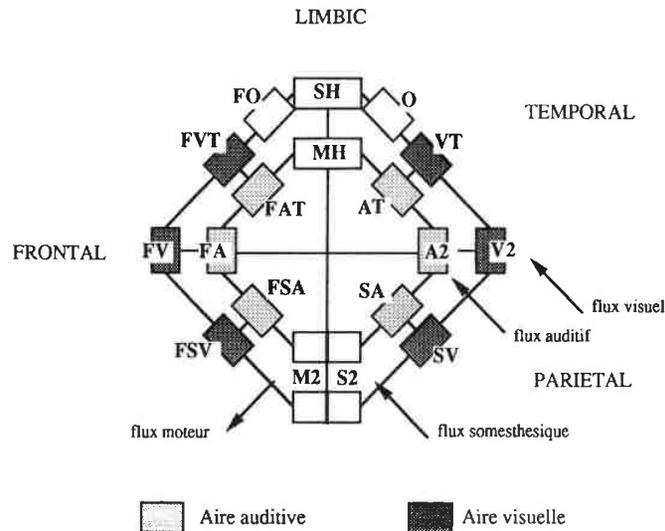


Figure 17. Forme fermée.

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.

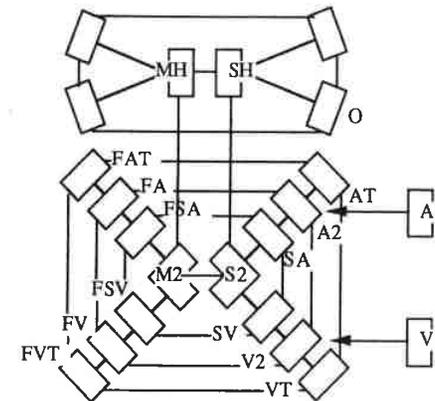


Figure 18. Forme ouverte.

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.

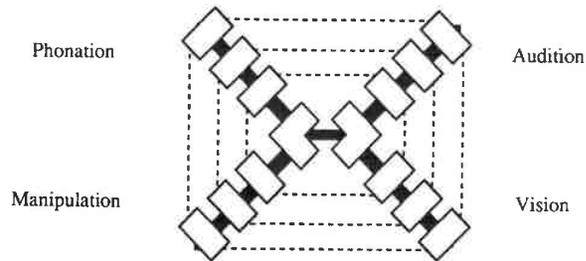


Figure 19. Axes privilégiés.

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.

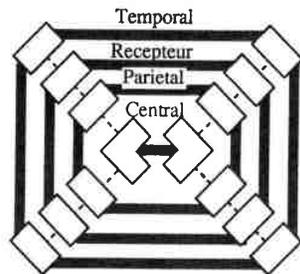


Figure 20. Circuits privilégiés.

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.

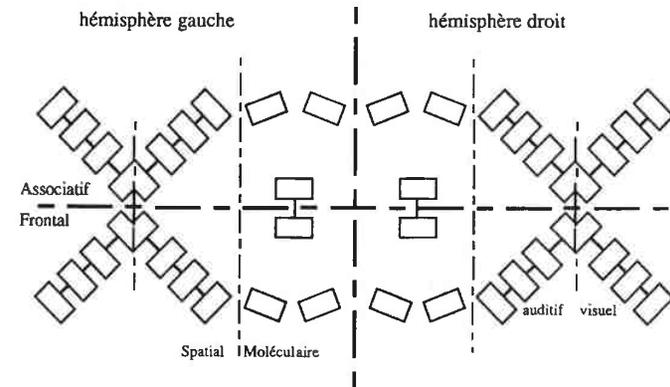


Figure 21. Symétries.

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),

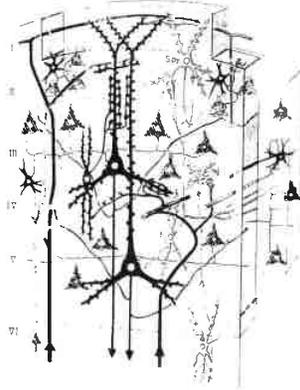


Figure 22. Cellules de la colonne (d'après Szentagothai).

qui propose une description cytotologique 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.

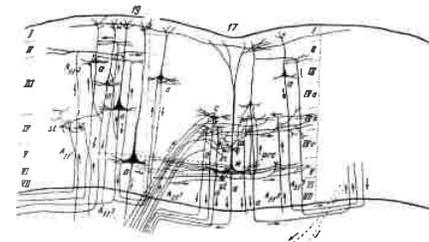


Figure 23. Connectivité (d'après Szentagothai).

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).

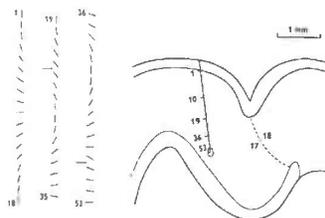


Figure 24. Pénétration oblique.

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 & McClelland, 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 interneurons, 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.

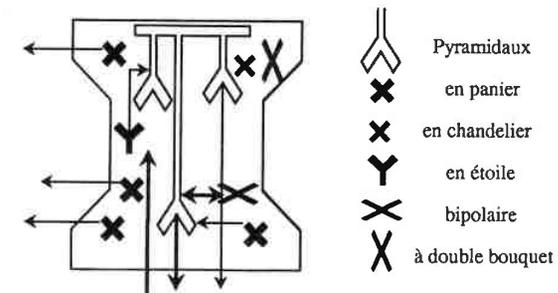


Figure 26. Modélisation des neurones de la colonne.

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:

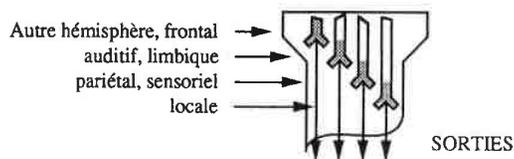


Figure 27. Hiérarchisation des entrées de la couche haute.

- 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.

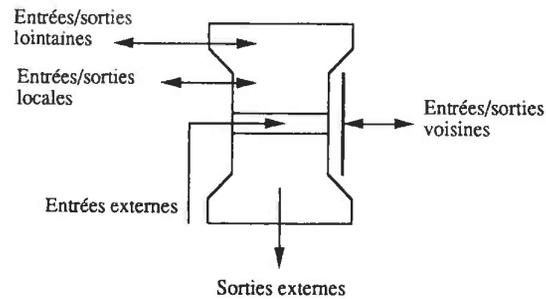


Figure 28. Entrées/sorties de la colonne corticale.

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.

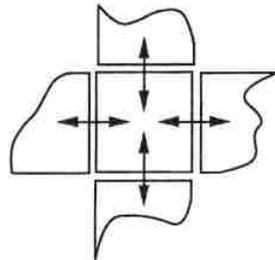


Figure 29. Liaisons voisines.

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.

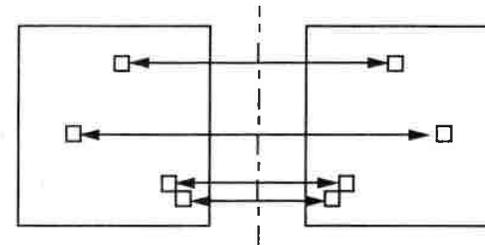


Figure 30. Connexions dans un réseau à deux aires.

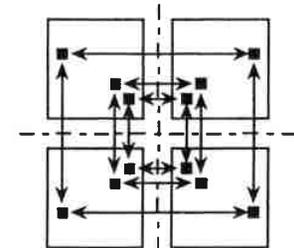


Figure 31. Connexions dans un réseau à quatre aires.

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  $2 \times 2$ ,
- soit au total, 14 connexions par unité.

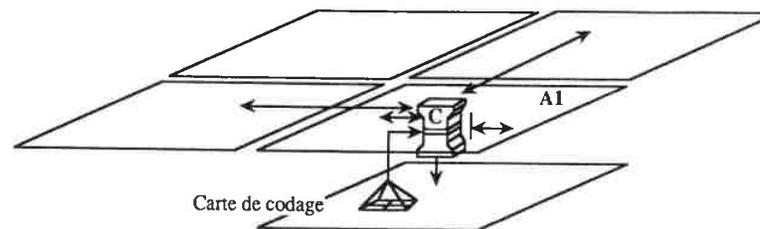


Figure 32. Résumé des connexions.

#### 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; Evars & 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<sub>i</sub> 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 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).

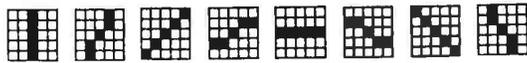


Figure 34. Masques thalamiques.

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.

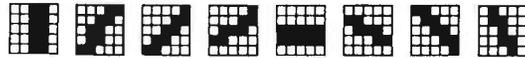


Figure 37. Masques déterminés par apprentissage.

#### 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  $P_0$  et  $P_2$  pour différents schémas:

$P_2$  mesure la probabilité que l'automate B soit dans l'état  $E_2$  avant que A ne le soit,

réciroquement,  $P_0$  mesure la probabilité d'avoir B à  $E_2$  avant A à  $E_0$ ;

$P_0$  et  $P_2$  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  $E_1$ . La Figure 38 indique, pour différentes valeurs de  $P_0$  et  $P_2$ , les conséquences sur le fonctionnement de la colonne, limitées au troisième cas de la table 33.

Coefficients		Sorties		Commentaires
$P_2$	$P_0$	Lointaine	Externe	
?	?	$E_0, E_1$	$E_0$	Aléatoire
0	1	$E_0$	$E_0$	Inhibition
$>0$	$<1$	$E_1$	$E_0$	Decl. Cond.
1	0	$E_2$	$E_2$	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 ( $E_1$ ) dans le réseau,
2. on force les réponses sur une autre aire de codage en sortie ( $E_2$ )
3. on met à jour les coefficients  $P_0$  et  $P_2$ .

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 (Ec), 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_externe;
char sortie_lointaine,
sortie_voisine,
sortie_locale,
sortie_externe;
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 *entrees\_* 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 "boîte 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 pré-traitement 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):

**creer** 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 = nbcoups2/nb.E2 et P0 = nbcoups0/nb.E2 \*/

**Si** sortie\_lointaine = E2 **alors**:

nb.E2 = nb.E2 + 1

**Pour** chaque lien\_lointain **faire**:

**Si** recep\_teur→soma→precedente.sortie\_lointaine = E2 **alors**:

recep\_teur→nbcoups2 = recep\_teur→nbcoups2 + 1

**Si** recep\_teur→soma→precedente.sortie\_lointaine = E0 **alors**:

recep\_teur→nbcoups0 = recep\_teur→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 reçue. 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: APPLICATION

Pour les quatre chapitres suivants, correspondant à quatre applications cognitives particulières, le plan sera identique: après une introduction au problème et des données expérimentales biologiques ou psychologiques, nous présenterons la structure du réseau et la caractéristique de fonctionnement qui vient enrichir le modèle. Enfin, des résultats expérimentaux viendront étayer les remarques et les interrogations soulevées par l'application.

D'autre part, dans les chapitres I et III, concernant la mémorisation associative visuelle et la motricité, nous n'aborderons que des notions indispensables pour les chapitres II et IV, les résultats étant disponibles en annexe. Une description plus détaillée de ces deux tâches est proposée par Alexandre (Alexandre 90), co-auteur de ces travaux de recherche.

## 1 VISION

### 1.1 Introduction au problème

Il s'agit dans un premier temps de valider les capacités de mémorisation associative de l'automate, pour mieux s'en détacher par la suite. Une des idées de base du connexionnisme étant le parallélisme, il est naturel que les premières applications réalisées aient porté sur la reconnaissance visuelle, mécanisme  $\infty$  combien parallèle.

Classiquement, car c'est là que les applications industrielles sont les plus nombreuses, les informaticiens se sont tournés vers la reconnaissance de caractères, tâche hétéro-associative qui consiste à étiqueter un flux visuel avec un nom de lettre. Ce problème qui reste somme toute très général, a été largement restreint, car nous ne visons pas ici la conception d'un système de lecture complet, mais seulement une démonstration des capacités associatives de notre unité; c'est pourquoi, les problèmes de localisation du texte puis des caractères dans la page imprimée ont été délibérément écartés; d'autre part, les distinctions entre les différents types d'écritures manuscrites et les aberrations liées à l'acquisition, qui impliquent l'intégration d'une variabilité insaisissable, nous ont contraints à limiter le système à la reconnaissance de caractères typographiques artificiellement bruités.

Après avoir présenté quelques données neurobiologiques sur la structuration des aires visuelles, et en avoir tiré des conséquences pour l'élaboration du réseau, nous mettrons l'accent sur les bases les plus élémentaires du fonctionnement de l'unité: la mémorisation des coefficients probabilistes, intimement liée au fonctionnement logique en feedforward ou en feedback.

### 1.2 Données neurobiologiques

Quand on étudie la structure des aires visuelles, bases expérimentales historiques de la théorie de la colonne corticale, il en émerge trois concepts, qui méritent d'être introduits:

- Le processus visuel se décompose en un certain nombre d'étapes de traitement, des capteurs (cônes et bâtonnets de la rétine) aux aires corticales les plus abstraites; l'intégration de l'information y est alors très progressive et chaque structure neuronale mise en jeu possède un rôle fonctionnel. Le thalamus, et plus particulièrement le corps genouillé latéral, effectue une extraction de contours; les aires corticales sensorielles primaire et secondaire détectent respectivement des segments orientés et des angles; l'aire associative, enfin, permet d'abstraire complètement l'information pour lui faire correspondre une réponse motrice, au travers d'une représentation distribuée.
- La connectivité entre chacune des structures qui concrétisent une étape, s'effectue localement: chaque unité d'une structure est reliée à un voisinage connexe d'une sous-structure, appelé champ récepteur. Réciproquement, dans le sens inverse, des représentations vers les données, le même schéma est conservé (figure 1).

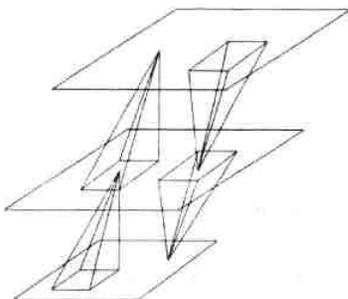


Figure 1: champs récepteurs

- Ce mode de connectivité particulier implique une certaine conservation de la topologie; deux colonnes voisines sont affectées à des portions voisines du champ visuel. En généralisant à toute une aire, il est clair que les aires sensorielles visuelles conservent une représentation globale proche de l'image réelle. Ce phénomène, connu sous le nom

de rétinotopie, est observable pour l'ensemble des aires sensorielles, auditives ou somatiques.

- Comment représenter systématiquement les orientations de petits segments et la position du stimulus dans un espace bidimensionnel comme le cortex? La nature résoud élégamment le problème, en utilisant, pour des colonnes voisines, des champs récepteurs identiques par la portion du champ visuel concernée, mais différents par les valeurs des connexions qui génèrent alors des fonctions élémentaires variées. Ces variations locales d'une dimension de représentation des données sont appelées indices secondaires, par opposition aux indices primaires, directement représentés par les deux dimensions spatiales de l'aire sensorielle. Là encore, cette particularité a été mise en évidence pour les aires sensorielles auditives, motrices et somatiques.

### 1.3 Le réseau

Ce bref aperçu neurobiologique est néanmoins suffisant pour spécifier un réseau de colonnes corticales. Le codage initial des entrées (détection de contours thalamique) et les aires impliquées (sensorielles primaire et secondaire, associative et motrice) sont les seuls paramètres dont nous avons besoin pour une application donnée. Cependant, il est clair que nous n'allons pas simuler les aires motrices impliquées dans la production de la parole ou dans l'écriture des caractères reconnus, seul moyen pour l'homme d'extérioriser sa pensée et qui nécessite un apprentissage complexe et co-opérant. Bien que, pour garder une certaine cohérence avec notre discours neurobiologique, l'aire qui fournit les réponse s'appellera aire motrice, elle n'en simulera pas la réelle fonction. Toute l'information correspondant à une image converge, de façon non distribuée, vers un ensemble de colonnes de sortie, chacune dédiée à une classe particulière. Notons au passage que cette solution d'utiliser des unités dites "grand-mères" est couramment employée en connexionnisme, et que d'autre part, les applications qui suivent y feront souvent référence.

### 1.4 Le fonctionnement

Dans cette première application, nous n'avons intégré ni le fonctionnement ni l'apprentissage sur les entrées/sorties voisines et locales (figure II.35 et II.36). Des aires sensorielles aux aires motrices, la table d'entrée/sortie de la figure II.33, définit cependant le fonctionnement des unités sous l'hypothèse d'une connectivité totale.

Par la convergence de l'aire associative vers les unités grand-mères de l'aire motrice, la mémorisation associative va s'effectuer essentiellement entre ces deux aires, les aires sensorielles servant alors à dégager des invariants du monde extérieur. Nous allons donner une interprétation logique de quelques mécanismes de fonctionnement, extraits de la table II.33, primordiaux pour l'analyse effectuée dans les aires associatives.

Soit une unité A de l'aire associative, directement connectée à une unité B de l'aire motrice; les différentes sessions d'apprentissage vont faire évoluer les valeurs des coefficients probabilistes qui rendent compte des schémas d'activité entre ces deux unités.

Dans le sens de la propagation de A vers B, en feed-forward, trois cas particuliers peuvent être mis en avant, correspondant aux variations induites par l'apprentissage sous l'effet d'un signal E1 lointain. En notant  $P2(A/B)$  [respectivement  $P0(A/B)$ ], la probabilité que l'événement "unité A dans l'état E2 [resp E0]" ait lieu avant l'événement "unité B dans l'état E2", trois interprétations des effets peuvent être données:

1.  $P2(B/A) = 1$ ; A est une condition *suffisante* pour B,  $A1 \Rightarrow B2$ ;
2.  $P0(B/A) = 1$ ; A est une condition *inhibitrice* pour B,  $A1 \Rightarrow B0$ ;
3.  $0 < P2(B/A) < 1$ ; A est une condition *sensibilisatrice* pour B,  $A1 \Rightarrow B1$ .

Réciproquement, B mémorise aussi un schéma d'activité par rapport à A; dans le sens du feed-back, deux cas sont remarquables, concernant l'effet d'un signal E1 de B sur la présence ou non d'une entrée externe pour A:

1. Si l'entrée externe de A est à E1, alors le feed-back de B suffit pour faire passer A à E1, ce qui peut se rapporter à de l'attention sélective.

2. Si l'entrée externe de A est à E0 (cas qui dépend de l'apprentissage),  $P0(A/B)=0$  définit A comme condition *nécessaire* pour B,  $A0 \Rightarrow B0$ . Remarquons que vu le séquençement de l'apprentissage,  $P2(A/B)=1$  n'est jamais réalisé, et que  $P0(A/B)=1$ , qui justifierait  $A1 \Rightarrow B0$ , n'est pas souhaitable pour des raisons de résistance au bruit.

### 1.5 Résultats

Les taux de reconnaissance des caractères en fonction du nombre de pixels inversés, ainsi que des essais sur l'écriture manuscrite, sont donnés en annexe I.

## 2 RECONNAISSANCE DE LA PAROLE

### 2.1 Introduction

La reconnaissance automatique de la parole est l'un des plus difficiles problèmes rencontrés en intelligence artificielle; elle a pour objectif la transcription d'un flux sonore en mots. Bien des modèles, mathématiques ou heuristiques, ont été proposés, mais les performances obtenues jusqu'à présent sont plutôt médiocres, même pour des applications très restreintes. Pourtant la parole est indiscutablement un phénomène bien structuré; on peut distinguer quatre niveaux d'organisation avec des constantes de temps croissantes:

1. le niveau du phonème, de l'ordre du centième de seconde,
2. le niveau du mot, de l'ordre du dixième de seconde,
3. le niveau de la phrase, de l'ordre de la seconde,
4. le niveau du discours, de l'ordre de la minute.

Pour la reconnaissance, chaque niveau d'organisation apporte un niveau de connaissance; ce qui permet d'identifier les mots en levant de plus en plus d'ambiguïtés.

Connaissances phonétiques: On connaît assez bien la structure sonore des phonèmes; la langue française n'en compte qu'une quarantaine, mais il est très difficile de les différencier, du fait de leur grande variabilité.

Connaissances lexicales: Le système fonctionne toujours avec un vocabulaire limité et si un phonème d'un mot est mal perçu, on peut le corriger par référence aux mots connus. On peut ainsi corriger "labora.oire" en "laboratoire". Il peut cependant subsister des ambiguïtés car plus le mot est court, plus il y a de possibilités ( comment corriger .on ? bon, dont, font, gond ?...)

Connaissances grammaticales: Un langage obéit à des règles de formation des phrases, utiles pour corriger d'éventuelles erreurs.

Connaissances sémantiques: Ces connaissances rendent compte du sens des mots et des relations qui existent entre eux, au niveau du discours.

### 2.1.1 Les problèmes

En considérant cette impressionnante quantité d'informations complémentaires, on peut se demander pourquoi la "machine à dicter" n'existe pas depuis 20 ans! Les connaissances lexicales et grammaticales sont facilement modélisables car on peut connaître tout le vocabulaire et toutes les règles de grammaire d'une langue. Les connaissances sémantiques sont déjà plus complexes à exprimer et demandent le développement de modèles appropriés. De plus, la quantité d'informations à stocker commence à devenir prohibitive, chaque concept pouvant être en relation avec des dizaines d'autres, sémantiquement proches. Mais le plus gros problème réside actuellement dans la reconnaissance des phonèmes (décodage acoustico-phonétique), étape préliminaire à toute reconnaissance de la parole.

Rappelons le schéma classique de la mise en correspondance associative:

- une phase de prétraitement qui, le plus souvent, consiste en une estimation spectrale du signal de la parole (FFT, ARMA...),
- une phase d'extraction d'indices acoustiques (formant, burst, friction...), sur lesquels nous reviendrons plus longuement,
- une phase de comparaison, avec des modèles prédéfinis de tous les phonèmes, par des algorithmes plus ou moins efficaces.

Ici encore, l'informaticien bute sur un problème de représentation des connaissances; en effet, les phonèmes présentent une variabilité extrême dans leurs instanciations. Pour un système de reconnaissance multi-locuteurs, les différences qui existent entre les voix des individus (accents régionaux, âge, sexe...) sont très difficiles à gérer. Même si on restreint le système à un seul locuteur, l'humeur, le ton ou la fatigue viennent modifier les caractéristiques de la voix.

Pire encore, pour un même locuteur, la prononciation d'un phonème précis dépend des phonèmes voisins (phénomène coarticulatoire).

Sous toutes ces contraintes, il est difficile de proposer un modèle universel pour chaque phonème.

Pour noircir encore un peu le tableau, on peut citer le bruit ambiant qui vient perturber l'acquisition du signal, et surtout le caractère temporel du signal de parole. En effet, si la dimension temporelle de la reconnaissance visuelle n'est guère déterminante, la reconnaissance automatique de la parole doit impérativement proposer des mécanismes associatifs qui tiennent compte des propriétés du temps.

### 2.1.2 Grandes approches du problème

Les premières générations de systèmes de reconnaissance étaient basées sur des techniques de programmation dynamique pour la mise en correspondance de schémas acoustiques [Sakoe, 71].

Plus récemment, des modèles stochastiques ont été utilisés pour construire des représentations statistiques des mots ou des phonèmes, qui prennent en compte la variabilité temporelle du signal [Levinson 83]. Ces modèles ont donné de bons résultats, particulièrement pour la reconnaissance de mots isolés. Cependant, pour des phrases entières, il semble nécessaire d'intégrer tous les niveaux de connaissance sur le langage.

Les réseaux de neurones ont déjà été utilisés avec succès pour résoudre des problèmes classiques de la reconnaissance de la parole [Lippman 87, Hinton 84, Peeling 86, Howard 88, Bengio 88, Bourlard 87, Mc Culloch 88, Waibel 88 89, Elman 88, Leboeuf 88, Kohonen 88]. Bien qu'ils n'aient pas encore démontré leur supériorité sur les modèles classiques, il n'en constituent pas moins une approche intéressante du problème:

1. Leurs grandes capacités d'apprentissage permettent de générer des modèles robustes au bruit et adaptés au flou;
2. leur structure en réseaux les rend apte à intégrer des connaissances sémantiques (réseau sémantique), par l'intermédiaire des liens entre unités;
3. ils peuvent rendre compte des trois niveaux de traitement de l'information, prétraitement (réduction du bruit, détection de la fréquence fondamentale), extraction d'indices (formants) et reconnaissance associative (phonèmes ou mots).

### 2.1.3 Notre approche

Dans ce contexte, nous nous sommes intéressés au décodage acoustico-phonétique, qui présente un double intérêt théorique: nécessité d'un apprentissage de différenciation et de généralisation, et intégration d'un traitement temporel des données.

Nous allons d'abord nous pencher sur des données neurobiologiques et linguistiques, pour extraire un codage initial pertinent. Ensuite, nous verrons plus en détail le mécanisme de fusion/division, tel qu'il a été implémenté en accord avec les caractéristiques décrites au chapitre II. Finalement, nous commenterons des résultats expérimentaux et nous soulèverons des questions et des idées relatives au modèle.

## 2.2 Données neurobiologiques

Comme nous l'avons déjà énoncé, le réseau est composé de cartes qui sont spécialisées par leurs entrées extérieures. Pour spécifier un réseau apte au décodage acoustico-phonétique, il nous faut alors déterminer le nombre de cartes impliquées, et plus spécialement, leurs entrées. Bien sûr, c'est une fois de plus le modèle humain qui va nous servir de référence.

Le traitement de l'information auditive est effectué par une chaîne de structures neuronales, du stimulus à la réponse. Grossièrement, nous pouvons dégager trois étapes d'intégration:

1. Les récepteurs auditifs sont anatomiquement situés sur la cochlée, partie de l'oreille interne enroulée en spirale. Grâce à cette forme spirale d'épaisseur décroissante, les neurones répondent sélectivement à des plages plus ou moins larges de fréquences. La principale fonction de la cochlée est donc de décomposer le signal en composantes fréquentielles [Delgutte 84].

De plus, les fibres nerveuses codent l'intensité du son jusqu'à 80 Hz, seuil de saturation. Un mécanisme d'habituation tend à faire diminuer la réponse d'une cellule au cours du temps, et réciproquement, les neurones sont très sensibles à la nouveauté [Delgutte 84] (on retrouve ces phénomènes dans le cortex visuel).

2. L'étape incontournable pour entrer dans le cortex, c'est bien sûr le thalamus, et les afférences auditives n'y échappent pas. Très globale-

ment, le thalamus (et plus particulièrement pour l'audition, le corps genouillé médian) amplifie sélectivement le signal et en rehausse le contraste [Szentagothai 70], suivant des rétroactions du cortex [Lieberman 83].

3. Le cortex auditif n'a malheureusement pas été autant étudié que le cortex visuel. Les données sur le codage initial font cruellement défaut; on peut seulement remarquer que l'axe fréquentiel est souvent représenté (figure 1).

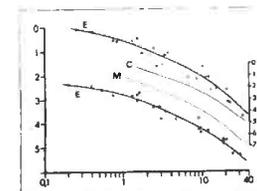


Figure 1.34  
Représentation des fréquences le long de l'axe A1 chez l'écureuil (E) et B2, chez le chat (C) et chez le singe (S).  
Présentation préférentielle en abscisse (en kHz), donnée sur le repère cortical le long d'un axe métré-couleur (pour l'écureuil) à gauche, puis le chat et le singe à droite). Toutes les courbes ont été décalées les unes par rapport aux autres en ordonnée pour le clarté de la lecture.

Figure 1: les axes fréquentiels du cortex

De nombreuses aires corticales sont impliquées dans la reconnaissance de la parole: des aires sensorielles primaires et secondaires [Goldstein 82], des aires temporales responsables des associations sémantiques [Wernicke 74] et des aires pariétales orientées vers des représentations spatio-motrices [Hecaen 83] (entre autre, production de la parole [Broca 61]).

## 2.3 Données phonétiques

La neurobiologie nous livre quelques renseignements sur les cartes impliquées et sur le prétraitement de l'information, mais elle ne donne que de très vagues idées sur l'organisation des cartes de codage initial du cortex.

Tournons-nous vers la phonétique et essayons de dégager des indices pertinents pour une discrimination associative. Les phonéticiens raisonnent très souvent sur des spectrogrammes de parole. Cette représentation de la

décomposition spectrale du signal en fonction du temps recoupe la fonction de la cochlée.

### 2.3.1 L'espace des phonèmes

Traditionnellement on structure les phonèmes dans un espace tridimensionnel, suivant le mode d'articulation, le lieu d'articulation et le voisement [Lonchamp] (illustré figure 2).

Mode:	Lieu d'articulation		
	Labiales	Dentales	Vélares
Plosives nV.	p	t	k
Plosives V	b	d	g
Nasales	m	n	gn
Fricatives nV	f	s	ch
Fricatives V	v	z	je
Glissantes	w	μ	j
Liquides		l	r

Voyelles	Antérieures ↔ Postérieures	
	Fermées	i      y      u
↑ ↓	e                      o	
	ε                      œ      )	
Ouvertes	a	

Figure 2: les classes de phonèmes

**Le mode d'articulation:** Il existe quatre façons standards de prononcer des phonèmes. Suivant les muscles mis en jeu, les phénomènes aérodynamiques ne sont pas les mêmes. On distingue ainsi les voyelles, les fricatives, les plosives et les sonnantes (nasales, glissantes, liquides).

**le lieu d'articulation:** Il module la classification suivant la localisation de l'occlusion dans le conduit vocal (labiale, dentale, vélaire). Ainsi, pour

les plosives on distingue le p de t et k, respectivement labiale, dentale et vélaire.

**le voisement:** Présence d'énergie en basse fréquence, il permet une distinction plus fine entre des phonèmes acoustiquement proches (p et b, par exemple).

### 2.3.2 Les classes de phonèmes

Nous allons maintenant passer en revue les indices révélateurs pour chaque classe de phonèmes.

#### Voyelles et formants

Les voyelles sont principalement caractérisées par la présence de maxima d'énergie stables dans le temps, pour des fréquences précises. Sur un spectrogramme, ils sont révélés par la présence de bandes noires (forte intensité) plus ou moins parallèles à l'axe du temps.

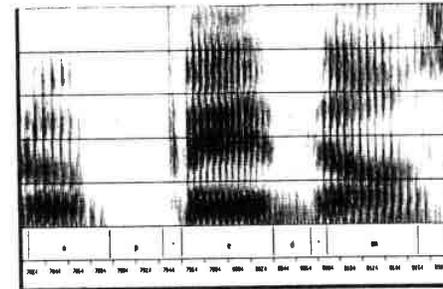


Figure 3: quelques formants

Si on examine bidimensionnellement une représentation énergie/fréquence à un instant  $t$ , on peut observer la présence de ces maxima, appelés formants. Chaque voyelle peut ainsi être caractérisée par des plages de fréquences relativement fixes pour ses formants  $F_i$ . Une représentation classique bidimensionnelle fait intervenir les valeurs des deux premiers formants de chaque voyelle (figure 4).

Réalisé sur des valeurs moyennes pour une dizaine de locuteurs masculins dans un contexte précis, ce triangle vocalique représente la localisation spatiale d'une voyelle suivant ses deux coordonnées F1, F2.

On peut comparer ce schéma avec celui des positions de la langue lors de la production des voyelles (figure 5). Les deux axes choisis représentent deux dimensions spatiales: haut/bas et avant/arrière.

La correspondance grossière qui existe entre les deux représentations est un indice intéressant pour l'élaboration d'une boucle de rétroaction externe dans le réseau.

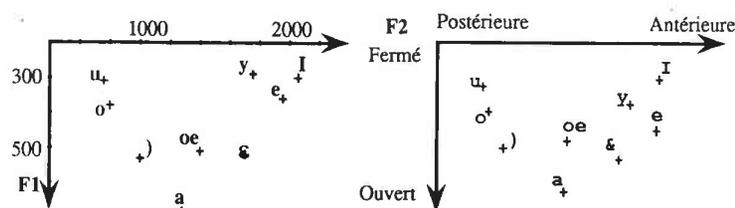


Figure 4 et 5: triangle vocalique et position de la langue

#### Les plosives

Les plosives sont caractérisées par un silence provenant de la fermeture complète du conduit vocal, suivi d'un brusque relâchement. On trouvera donc une suite d'événements acoustiques qui sont autant d'indices pour la reconnaissance:

-Le silence dure généralement plus de 50 ms,

-La barre d'explosion, onde de pression générée par le relâchement des organes articulatoires, est généralement très brève et très intense. La forme d'une coupe spectrale dépend du lieu d'articulation; l'énergie est plutôt faible, large bande et basse fréquence pour p et b; moyenne, large bande et haute fréquence pour t et d et intense, bande étroite et moyenne fréquence pour k et g.

-Les transitions formantiques, étape intermédiaire entre la production d'une plosive et d'une voyelle sont dues aux mouvements des articulatoires

vers un autre but. Ces transitions sont très révélatrices des influences coarticulatoires.

**Les fricatives** Les fricatives sont dues à des turbulences aérodynamiques provenant d'un fort resserrement du conduit vocal. Elles se caractérisent essentiellement par la présence de bruit d'énergie importante. Là encore, la coupe spectrale est révélatrice du lieu d'articulation. La figure 6 schématise la répartition de l'énergie en fréquence pour les trois types de fricatives: s et z, f et v, ch et j.

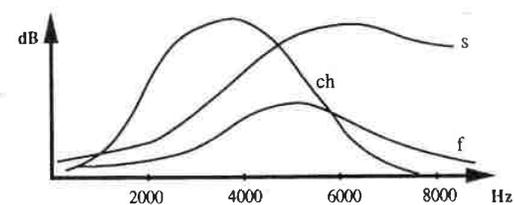


Figure 6: les fricatives

#### Les consonnes sonnantes

Les sonnantes, comme les voyelles, sont caractérisées par la présence de formants. On distingue plus finement les nasales (n, m, gn) qui font intervenir la cavité nasale par abaissement du voile du palais, les semi-voyelles (w, j), proches des voyelles, et les liquides (l et r). On peut relever deux différences significatives entre les voyelles et les sonnantes, d'une part la présence de minima locaux de très faible énergie (anti-formants), et d'autre part une variation temporelle des formants.

#### 2.3.3 La coarticulation

Outre une modélisation de la structure des phonèmes, l'étude phonétique doit rendre compte des phénomènes coarticulatoires. La coarticulation se définit comme l'influence d'un son sur un autre son voisin. Elle résulte d'un effet d'inertie mécanique, et surtout, d'une minimisation des déplacements des muscles. Son effet est d'autant plus grand que le débit est rapide.

Il est clair qu'une étude de la parole passe obligatoirement par une étude statique et dynamique des organes articulatoires. Une description

des phonèmes en terme de positions et de mouvements des lèvres et de la langue peut se baser sur des modèles de l'appareil phonatoire. Malheureusement, les modèles vibratoires proposés par l'automatique ne permettent pas de modéliser des tissus mous et déformables qui génèrent pourtant des phénomènes acoustiques particuliers.

Le meilleur espace de description des phonèmes nous semble être celui des activités musculaires, car il prend naturellement en compte les phénomènes coarticulatoires qui, à la base, ne sont que des contraintes temporelles musculaires.

## 2.4 Le réseau

### 2.4.1 Le réseau général

Pour une tâche de décodage acoustico-phonétique, nous allons développer un réseau associatif du même type que pour la reconnaissance visuelle. Les différences vont se situer au niveau du fonctionnement, puisque le mécanisme de division et le traitement temporel de l'information vont être intégrés à l'automate. Nous reviendrons plus en détail sur ces deux nouvelles propriétés, spécifions d'abord le réseau que nous allons employer. Les deux paramètres qui suffisent à décrire un réseau particulier sont, d'une part, la donnée des cartes impliquées dans le processus cognitif, et d'autre part, la donnée des cartes de codage qui vont spécifier fonctionnellement chacune des cartes. Le réseau traite le signal temporel de deux manières complémentaires: un traitement basse fréquence local pour la mise en évidence des voyelles, un traitement haute fréquence global pour les fricatives. La figure 7 illustre les étapes impliquées dans le processus de reconnaissance:

- Une étape de prétraitement, qui calcule la décomposition spectrale du signal,
- une aire primaire sensorielle qui sert de carte de codage pour l'aire secondaire,
- une aire sensorielle secondaire qui opère d'éventuels regroupements de l'information,

- une aire associative qui combine les champs auditifs et moteur; cette aire, déjà décrite pour le système visuel, va générer les représentations internes,
- une aire motrice qui simule une sortie verbale par la juxtaposition d'unités "grand mères".

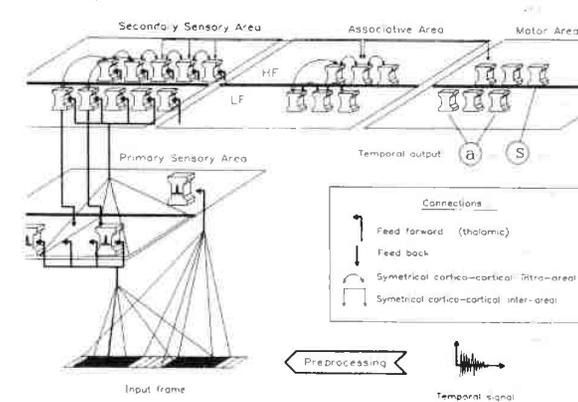


Figure 7: le réseau général (Guyot, 1989)

La connectivité est donnée pour rappel. Des liens peuvent mettre en correspondance les deux demi réseaux par même fréquence et, grâce à la mémorisation de causalité temporelle, ils permettent de tenir compte de phénomènes de coarticulation.

Nous ne reviendrons pas sur les cartes associatives et de sortie, déjà décrites dans un paragraphe précédent (III.1.3); seul le codage initial de la carte sensorielle primaire nous intéresse. Nous avons vu pour la reconnaissance visuelle directe de spectrogrammes les inconvénients soulevés par une représentation fréquence/temps. La dimension temporelle des données est maintenant prise en compte par l'unité connexionniste. Il faut donc élaborer une représentation bidimensionnelle du signal à un instant  $t$ . Trois contraintes vont nous guider:

1. Une plausibilité biologique du codage,

2. La nécessité d'une différenciation topologique entre les classes.
3. Les indices phonétiques importants.

La grande diversité des indices mis en oeuvre pour chaque classe phonétique (formants, burst, friction...), nous a amenés à réaliser des codages initiaux différents et spécifiques. Nous nous sommes surtout intéressés aux phénomènes temporellement stables (voir remarques) les voyelles et les fricatives et, dans une moindre mesure, aux plosives.

#### 2.4.2 Prétraitement commun

De toute évidence, une décomposition spectrale s'impose, justifiée tant par la fonction de la cochlée que par l'analyse des phonéticiens. Nous avons classiquement choisi d'effectuer une FFT sur le signal, qui, malgré certaines hypothèses (signal stationnaire et infini), donne des résultats corrects et rapides. Nous comptons sur la robustesse du système pour obvier aux défaillances du prétraitement. L'entrée du système à un instant  $t$  est composée d'un vecteur monodimensionnel représentant l'énergie pour chaque bande de fréquence de la FFT (voir figure 7).

#### 2.4.3 Codage des voyelles

La reconnaissance des voyelles s'effectue typiquement sur des indices formantiques. Cependant, la majorité des systèmes symboliques de codage acoustico-phonétique s'appuient sur une détermination rigoureuse des deux ou trois premiers formants; ce qui pose des problèmes d'identification exacte du numéro du formant. Nous généraliserons plutôt la notion de pics d'énergie sans leur attribuer de signification particulière.

Il est facile d'imaginer un masque d'entrée pour extraire des pics qui soit biologiquement plausible, c'est à dire, constitué d'un voisinage local avec une dualité partie centrale/partie périphérique. La largeur de bande des pics (à peu près constante) et la pente spectrale vont imposer les coefficients d'une fonction "chapeau mexicain" monodimensionnelle. Des lissages spectraux en fréquence vont permettre de réduire la taille du masque de cette fonction, les énergies ayant un écart type local moins élevé.

Comment passer d'une représentation monodimensionnelle à un instant  $t$  à une représentation bidimensionnelle ? Un codage fréquence/énergie des

pics semble peu plausible, et nous lui avons préféré, par analogie au triangle vocalique, un codage fréquence/fréquence. La figure 8 illustre un exemple de codage bidimensionnel du signal; le premier axe varie de 300 à 1300 Hz, le deuxième de 1200 à 3000 Hz. On retrouve dans cette information redondante les formants de la voyelle ( $f_1$ ,  $f_2$ ,  $f_3$ ,  $f_4$ ); des informations non significatives ( $f_{ns}$ ) peuvent aussi apparaître. Le réseau, par apprentissage, sélectionnera les indices pertinents.

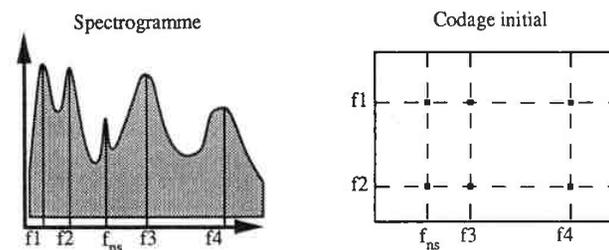


Figure 8: codage des voyelles

Nous avons vu la possibilité de définir localement des indices secondaires. Ce sont des variations locales et grossières d'une troisième dimension, les deux dimensions principales de la carte (indices primaires) étant fixées. Pour des raisons expérimentales (séparation topologique de certaines voyelles), il s'est avéré nécessaire d'ajouter l'énergie des formants au flux d'entrée. Nous avons choisi de discerner localement quatre niveaux d'énergie linéairement répartis de 50 à 80 dB.

De plus, l'automate a la possibilité de jouer sur ses états E2 (certitude) ou E1 (hypothèse). L'unité sera à E2 si elle est activée (E1 ou E2) plus de 5ms et si ses voisins sont aussi activés. Si ces conditions ne sont pas remplies mais que l'unité est quand même sollicitée, elle prendra la valeur E1, ce qui permet au système, si nécessaire, d'en tenir compte par rétroaction.

#### 2.4.4 Codage des fricatives

Les indices permettant de différencier les fricatives sont de nature moins locale que pour les voyelles. Ils rendent compte de la répartition globale de l'énergie en fréquence. C'est pourquoi le spectre est très fortement lissé en fréquence pour ne laisser apparaître que des phénomènes globaux.

On peut choisir de relever la fréquence d'un barycentre de l'énergie ou bien les limites inférieures et supérieures du bruit. Nous avons plutôt sélectionné la fréquence du maximum global de la courbe spectrale, associée aux valeurs des pentes spectrales à gauche et à droite de ce maximum. Dans un souci d'homogénéité, afin d'avoir une représentation bidimensionnelle fréquence/fréquence, la pente est exprimée par la valeur de la fréquence correspondant à l'énergie du maximum moins 10 dB, comme illustré par la figure 9.

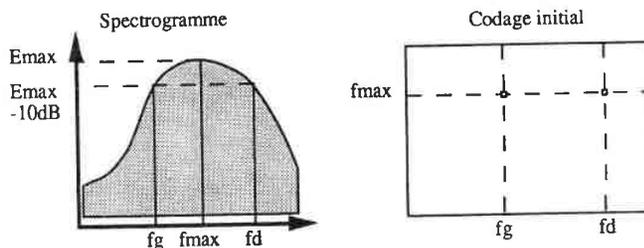


Figure 9: codage des fricatives

L'extraction de maximum globaux et de pentes spectrales peut difficilement se justifier biologiquement sans une connectivité complète et par des mécanismes d'inhibition latérale récurrente. Nous n'avons cependant pas réussi à mettre en évidence des indices pertinents calculables localement.

Ici aussi, les indices primaires ( $F_{max}$ ,  $F_g$  et  $F_d$ ) sont modulés localement par l'énergie correspondant à  $F_{max}$ . De plus, on peut associer autour des valeurs de  $F_g$  et  $F_d$  des intervalles de variation, qui permettent d'utiliser l'état  $E_1$  des automates pour une éventuelle rétroaction.

#### 2.4.5 Codage des plosives

Les plosives sont remarquables par la présence d'un silence très caractéristique d'au moins 50ms. Nous avons essayé de les différencier en ne tenant compte que de la forme spectrale de la barre d'explosion. En rehaussant les valeurs des énergies après un silence, les plosives peuvent être traitées par le même réseau que les fricatives. Cependant, la faible durée du burst, liée à une très grande variabilité coarticulatoire, ne permettent pas d'obtenir des résultats intéressants. Ainsi, même si la détection du silence est évidente, le réseau ne fournit qu'un taux de 50% de reconnaissance pour différencier les plosives entre elles.

Il serait évidemment beaucoup plus révélateur de tenir compte des transitions formantiques.

### 2.5 Fusion/division et propagation

#### 2.5.1 Rappels

Nous avons esquissé, paragraphe III.1.4, les bases du fonctionnement de l'unité; nous avons vu comment planter et utiliser les possibilités d'apprentissage probabiliste entre unités de cartes différentes par l'intermédiaire des coefficients  $P_0$  et  $P_2$ .

Cependant, la connectivité des réseaux y était implicitement totale. Nous avons défini par ailleurs deux autres mécanismes d'apprentissage entre unités voisines (paragraphe II.4.5.5) et locales (paragraphe II.4.5.4), dont les rôles sont d'assumer la connectivité locale et de permettre la généralisation et la différenciation topologique.

La notion d'intervalle d'éloignement permet d'implanter les deux apprentissages en même temps, car elle englobe les notions de localisation et d'invariance locale. Pour expliciter l'implantation réalisée, rappelons le schéma de connectivité minimale proposé paragraphe II.3.2.3.

Soit une unité  $A$  de la carte de sortie, il existe une unité  $A'$ , dans la carte associative, qui lui est directement liée par une connexion de carte à carte. Cette unité  $A'$  est aussi reliée à ses voisins dans la carte associative. D'autre part, une unité  $B'$  de la carte associative est connectée à une unité de la carte sensorielle  $B$  (figure 10).

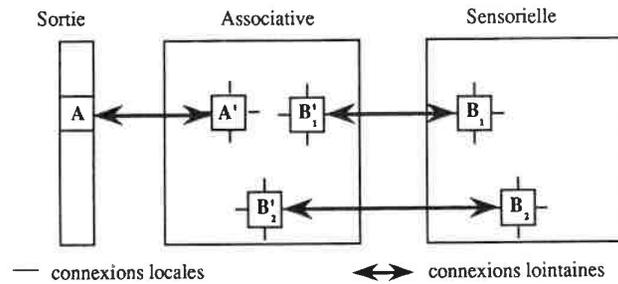


Figure 10: connectique

### 2.5.2 Implantation

Par apprentissage sur les connexions locales, la carte associative doit être en mesure de mettre en relation A' (la réponse) et les B'i qui correspondent au signal.

Dynamiquement, les messages entre les unités B'i et A' vont s'échanger par propagation de proche en proche, à travers la carte associative. Pour conserver une valeur de la provenance du signal, les messages échangés via les connexions locales, vont faire état de l'éloignement de l'émetteur par l'intermédiaire de la variable C (Cx et Cy en deux dimensions).

On peut tolérer de plus une certaine invariance topologique autour d'un indice B'i (apprentissage sur les connexions voisines). C'est pourquoi, nous parlerons plutôt de zone qui représente un voisinage local autour d'un indice et sur lequel est opéré un "ou" logique.

Une unité A' est capable de mémoriser des intervalles d'éloignement des signaux correspondant à des zones et leurs coefficients probabilistes associés; ainsi, pour le signal B'i, et le découpage en zones Zi, l'unité mémorise 3 intervalles et 3 couples (P2i, POi) associés.

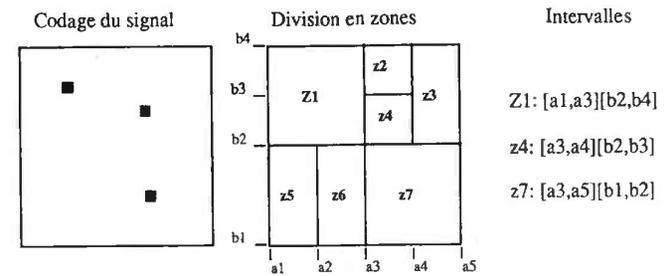


Figure 11: intervalles d'éloignement

Par ces deux mécanismes de mémorisation, l'unité est capable de créer dynamiquement des relations entre unités d'une même carte sans qu'elles aient de liens physiques. De plus, la représentation choisie, par intervalle, va permettre de mettre en place des systèmes de généralisation et de différenciation, qui vont affiner et calculer les bornes de Cx et Cy. L'idée directrice pour la mise en place de ces deux mécanismes réside dans la nécessité de la différenciation: l'objectif du réseau étant la mémorisation associative, il est rarement nécessaire de tenir compte de tous les détails du signal. C'est pourquoi, le réseau n'affine sa résolution topologique qu'en cas de confusion entre plusieurs réponses, pour converger vers une séparation des zones significatives, minimale.

L'apprentissage a donc lieu en plusieurs étapes:

- mémorisation des coefficients probabilistes P0 et P2 pour toutes les unités,
- tentative de reconnaissance d'un signal donné,
- division de certaines zones en cas de mauvaise réponse.

Le choix des zones à diviser est contraint par des critères de convergence. Nous avons vu au paragraphe III.1.4 l'importance significative des relations d'inhibition et de déclenchement inconditionnel, par contre les relations de déclenchement conditionnel ( $0 < P0 < 1, 0 < P2 < 1$ ) sont révélatrices

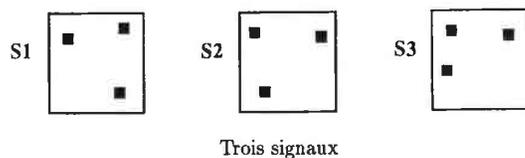
d'un indice d'entrée présent dans plusieurs signaux différents. Les unités de sortie qui ont mal répondu vont alors remettre en cause la généralisation locale effectuée par une zone en relation de déclenchement conditionnel. La division physique de la zone se concrétise par la division de l'intervalle d'éloignement et d'autres coefficients sont mémorisés, P2' P2'' et P0' P0'', respectivement significatifs des deux sous-zones.

La partition initiale la plus générale est la partition nulle, où la carte associative dans sa totalité est considérée comme intervalle d'éloignement des signaux.

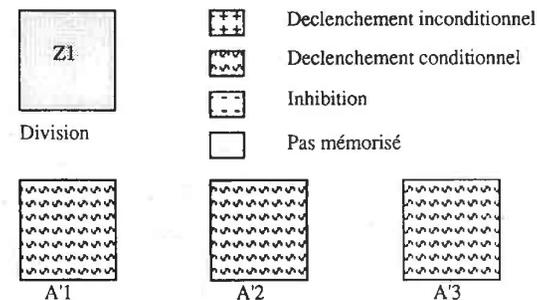
### 2.5.3 Illustration

Un exemple trivial va permettre de décrire l'apprentissage pas à pas.

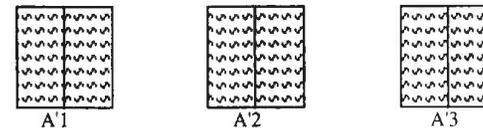
Nous voulons séparer trois signaux différents, codés respectivement par les indices suivants:



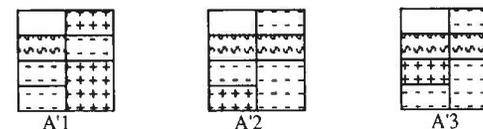
Il est commode, pour visualiser les résultats des trois apprentissages, de reconstituer l'espace de codage des signaux à partir des intervalles d'éloignement mémorisés par les symétriques A'i des réponses Ai. La figure suivante illustre ainsi la division initiale nulle et les coefficients probabilistes de type déclenchement conditionnel, résultat de la première étape d'apprentissage.



Si on essaie de reconnaître S1, les trois unités de sortie sont activées à E1; la réponse n'étant pas satisfaisante (apprentissage supervisé), les zones se divisent et mémorisent de nouveaux coefficients:



les tentatives de reconnaissance de S2, puis de S3 vont conduire aux divisions suivantes avec leurs coefficients associés.



Dès lors, la résolution est suffisante pour différencier les trois signaux.

## 2.6 Résultats

L'implantation du réseau fournit deux résultats bruts à exploiter, la résolution finale et le taux de reconnaissance. Pour les deux applications de décodage acoustico-phonétique (fricatives et voyelles) nous avons fait varier respectivement le nombre d'exemples présentés lors de l'apprentissage, et le nombre de phonèmes à reconnaître.

### 2.6.1 Les fricatives

Pour constituer le corpus d'apprentissage, un seul locuteur a prononcé 7 fois les 6 fricatives, isolées hors contexte. Par contre, la reconnaissance a été effectuée sur un corpus de 25 mots contenant 50 fricatives en contexte. Le voisement n'apparaît pas sur la carte de codage élaborée. C'est pourquoi, le réseau ne distingue que trois groupes de fricatives: les labiales, les dentales et les vélares.

Le protocole d'apprentissage est simple, le réseau apprend à reconnaître en divisant l'aire associative tant qu'il commet des erreurs de reconnaissance sur le corpus. Les figures 13 et 14 illustrent les divisions finales et les valeurs des coefficients associés suivant la même représentation que ci dessus.

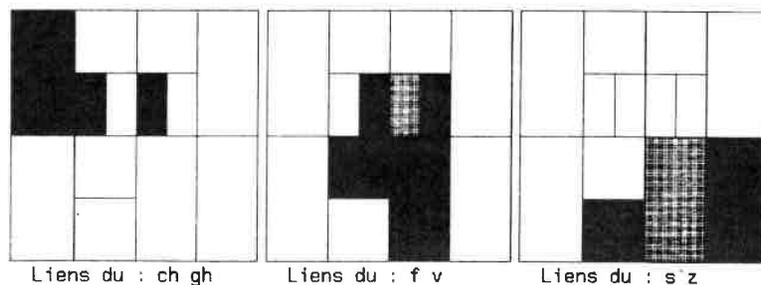


Figure 13: reconnaissance des fricatives, apprentissage sur un exemple

Plus précisément, la figure 13 a été obtenue pour un corpus d'apprentissage ne contenant qu'un seul exemple de chaque fricative, la figure 14 pour 7 exemples de chaque.

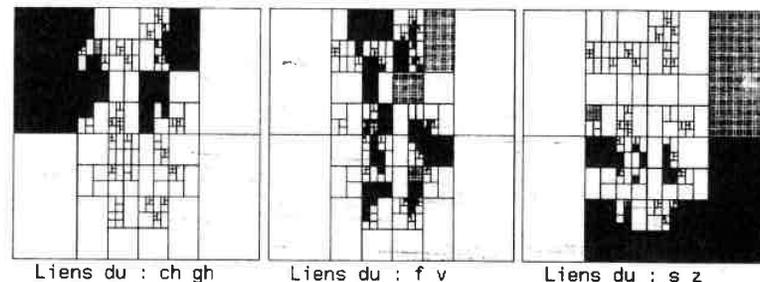


Figure 14: reconnaissance des fricatives, apprentissage sur sept exemples

En parole continue, deux critères permettent de juger des performances du réseau: il doit fournir la bonne solution quand il y a une fricative, et ne pas en proposer pour les autres phonèmes prononcés. Dans ces conditions, les tests font apparaître un taux de reconnaissance de 94% pour le premier réseau contre 96% pour le deuxième. Bien que n'ayant pas intégré de généralisation sur l'exemple, le premier réseau propose une représentation des données plus générale car peu de divisions sont nécessaires pour différencier les 6 phonèmes appris.

Les taux de reconnaissance sont très voisins. Il est clair que trop de divisions nuisent à la généralisation des données, tout en permettant cependant de différencier des cas particuliers. Une solution consiste alors à réunir les zones trop divisées en implantant la fusion des zones comme mécanisme complémentaire de la division. *Les conditions de fusion et de divisions engendreront un équilibre régulé de l'apprentissage topologique.*

On peut aussi remettre en cause le protocole d'apprentissage qui impose de reconnaître tous les exemples du corpus. Il suffit alors d'un item mal acquis pour empêcher un apprentissage optimal.

D'autres paramètres influent sur cet équilibre fusion/division qui génère une généralisation/différenciation des données. Notre propos est plus

d'étudier le concept d'apprentissage que de déterminer des valeurs précises. Nous pensons en effet, que le système doit être extrêmement robuste quant aux connaissances a priori. Sa puissance réside plus dans des mécanismes redondants et complémentaires que dans des affinages subtils et souvent instables de paramètres de fonctionnement. Notre démarche consiste alors à proposer un mécanisme, en analyser les défauts et y remédier en proposant un autre, parallèle et complémentaire.

### 2.6.2 Les voyelles

Le corpus d'apprentissage est constitué de 7 voyelles prononcées 20 fois par le même locuteur, hors contexte. Le corpus de test contient ces 7 voyelles prononcées 10 fois dans les mêmes conditions.

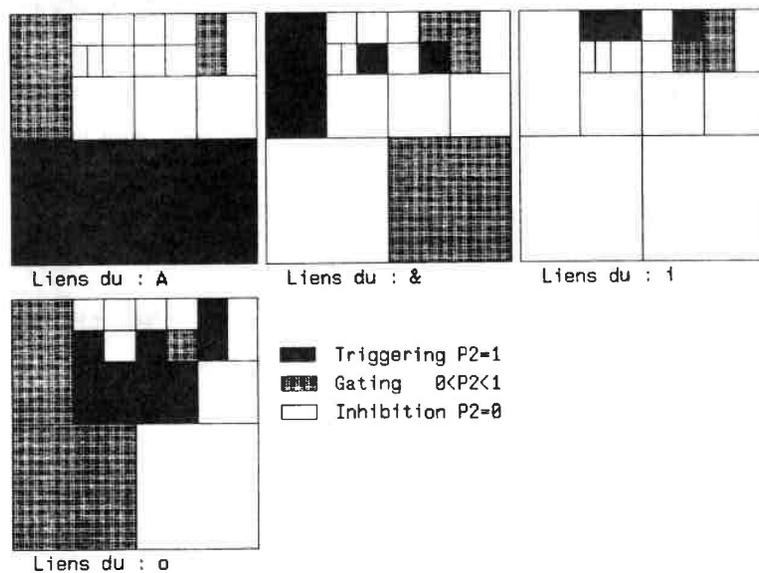


Figure 15: reconnaissance de quatre voyelles

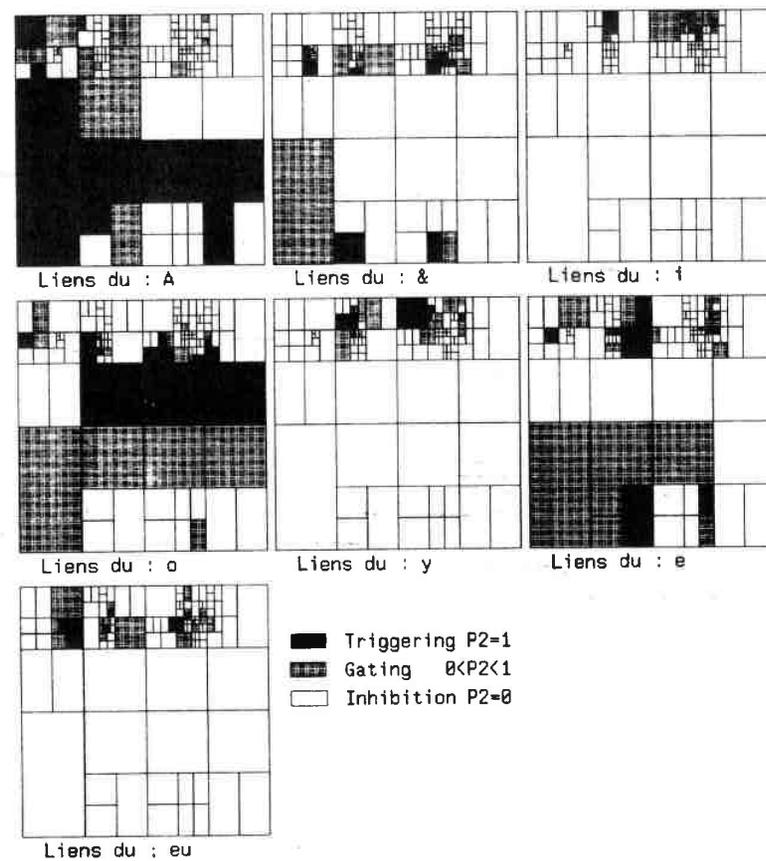


Figure 16: reconnaissance de sept voyelles

Il nous a semblé intéressant de faire varier le nombre de phonèmes à reconnaître pour mettre à l'épreuve l'apprentissage topologique. Nous avons réalisé deux simulations pour reconnaître respectivement 4 et 7 voyelles. Les figures 15 et 16 rendent compte des résolutions finales et des coefficients associés.

Très peu divisée pour quatre phonèmes, l'aire associative se différencie beaucoup plus pour sept, allant localement jusqu'à la résolution maximale possible (indice secondaire). De plus, les performances de reconnaissances chutent de 99% à 87%! Bien sûr, le choix des voyelles influe sur les performances. Les voyelles a e i o sont topologiquement très éloignées, alors que i et y, eu et & sont très proches, voire confondues. Plus finement, les 87% de reconnaissance se partagent entre 98% pour a, i, o et 76% pour eu et &.

Faut-il remettre en question un codage initial de l'information qui propose une représentation identique pour deux phonèmes différents? En comparant les coupes spectrales d'exemples particuliers, nous avons réalisé la difficulté d'établir des distinctions reproductibles, surtout en reconnaissance multi-locuteur. Il est vain de chercher à palier aux insuffisances du système en affinant le codage initial. On peut se demander quelle est la part des voyelles dans la reconnaissance globale d'un mot; intuitivement, il semble que les consonnes soient plus significatives. D'autre part, la présence des consonnes influe fortement sur les voyelles voisines. La reconnaissance des voyelles semble donc inséparable de celle des consonnes. Plus généralement, seule la prise en compte des différents niveaux de contexte peut permettre une reconnaissance globale de la parole; on ne peut pas parler de décodage acoustico-phonétique en tant que système isolé.

## 2.7 Remarques

L'application du modèle de colonne corticale à la simulation de fonctions auditives soulève de nombreux points de discussion qui, pour les plus intéressants, ont trait au traitement temporel des données par le réseau.

### 2.7.1 Les points fort du modèle

De par la dynamique de l'activité, le réseau est indépendant de toute origine temporelle de départ; il n'est pas influencé par le temps de prononciation et en parole continue, ne nécessite pas de segmentation à priori des phonèmes.

Il optimise la connectivité et généralise les données par le biais de deux mécanismes d'apprentissage complémentaires. Rappelons qu'une unité possède  $3 \times 3$  (masque d'entrée) + 4 (voisines) +  $\log_2(n)$  (carte à carte) connexions (n représente le nombre d'aires).

### 2.7.2 Entrée des données

Le prétraitement consiste en une décomposition fréquentielle, suivie d'une extraction de pics d'énergie, on pourrait remettre en cause le choix de la FFT comme analyse spectrale et proposer des techniques plus performantes.

De plus, la présence des anti-formants avec leur énergie semble nécessaire pour la reconnaissance de certains phonèmes.

Ce ne sont là, que des changements mineurs du codage. C'est plus par l'introduction de nouvelles dimensions à l'espace de représentation que le système sera amélioré.

### 2.7.3 Sortie des données

La sortie du système dans son état actuel n'est guère plausible d'un point de vue biologique. De plus, elle met de côté une propriété bien utile pour l'apprentissage: le bouclage externe de l'information. En effet, la possibilité d'entendre ce que l'on dit et sa réciproque permettent un auto-apprentissage non supervisé de la parole. De plus, le principe de recombinaison privilégiée, abordé au chapitre suivant, est sans doute appliqué chez l'homme entre les représentations musculaires et formantiques des voyelles.

### 2.7.4 Le temps

L'implantation actuelle empêche la prise en compte pourtant indispensable du contexte. Le réseau a une vision trop locale du temps. Cette première introduction de données temporelles s'est volontairement limitée aux phonèmes stables. Un léger filtrage en entrée nous assure de la stabilité des données et élimine des événements occasionnels. Ce qui empêche, d'une part de représenter les phonèmes instables, et d'autre part de tenir compte des transitions formantiques.

Il existe dans le cortex visuel des cellules sensibles au mouvement, c'est à dire à des variations topologiques temporelles. On pourrait envisager des

mécanismes similaires pour le cortex auditif, et définir des unités de codage sensibles aux variations de fréquences.

D'autre part, les phénomènes instables possèdent une propriété très intéressante pour leur représentation: les variations de fréquence sont continues. Dès lors, il est clair qu'une représentation fréquentielle fera intervenir un déplacement continu de l'activité sur une carte de codage. Cette propagation de l'activité d'unité en unité, par les connexions locales, ne nous est pas étrangère. Il semblerait tout à fait judicieux de modifier les règles de fonctionnement et d'apprentissage sur ces connexions pour pouvoir prendre en compte des événements localement successifs. C'est une voie de recherche fondamentale pour élaborer une unité plus puissante.

### 3 MOTRICITE

#### 3.1 Introduction au problème

Les perceptions auditives et visuelles constituent deux canaux cognitifs privilégiés; la poursuite de l'exploration du cortex humain passe sans conteste par l'étude des mécanismes moteurs qui orientent une grande partie de notre activité. Il est clair que les possibilités de manipulation offertes par la main, ouverture rétro-active du cortex sur le monde extérieur, sont pour beaucoup dans la perception tridimensionnelle. Nous avons choisi comme application le guidage de la main par l'oeil, qui demande une coopération étroite entre les deux types de représentation.

Il s'agit donc pour le réseau d'apprendre à effectuer des changements de repères, de passer des coordonnées rétinienne qui représentent la position de la main dans le champ visuel, à des coordonnées corporelles, en tenant compte des contraintes mécaniques des membres.

Après avoir passé en revue les aires corticales et les structures coopérantes impliquées dans la motricité, nous dégagerons un concept fondamental et riche de conséquences: les recombinaisons privilégiées. Puis nous verrons comment les connexions locales peuvent servir pour ordonnancer des tâches séquentielles, pendant indispensable du parallélisme mis en avant jusqu'à présent, directement applicable à l'organisation des contractions musculaires.

#### 3.2 Neurobiologie

Le guidage de la main par l'oeil fait intervenir au moins quatre flux d'information en entrée ou en sortie, prétraités par des structures diverses. Le corps genouillé latéral, noyau thalamique affecté à la vision, effectue un rehaussement de contraste sur l'image rétinienne en entrée, tandis qu'en sortie du cortex, le colliculus est susceptible de déclencher un mouvement oculaire en direction d'un stimulus présent dans le champ visuel. D'autre part, les flux moteurs se partagent aussi en entrées et en sorties: le thalamus code des

informations posturales en provenance de capteurs spécialisés répartis dans les différents muscles, alors que le cervelet gère les commandes motrices envoyées par le cortex, en fonction de nombreuses régulations (équilibre...).

Le réseau d'aires corticales est modelé par les connexions avec les structures extérieures au cortex: V1 et V2 reçoivent les stimuli visuels du thalamus, S2 les stimuli proprioceptifs et M2 se projette sur le cervelet. Entre les deux pôles sensori-moteurs, des aires associatives (comme SV) servent à recombinaison des différentes représentations. Nous avons déjà évoqué pour les aires auditives et visuelles une organisation sonotopique et rétinotopique, en indices primaires et secondaires. Les cortex moteur et proprioceptif ne font pas exception à la règle. Une analyse fonctionnelle de ce mode de représentation fait généralement apparaître la prépondérance d'un indice primaire sur l'autre. Ainsi, M2 et S2 s'organisent en bandes, représentant chacune une partie du corps, qui sont disposées somatotopiquement à la surface de ces aires (figure II.16: l'homoncule).

L'aire visuelles V2, grossièrement schématisée dans la première application par une répartition isotrope des champs récepteurs, fait état d'une organisation similaire en bandes qui représentent différents éloignements du centre de la rétine. La rétinotopie impose une distribution suivant des distances croissantes, de la vision centrale jusqu'à la vision la plus périphérique.

D'autre part, l'importance fonctionnelle d'une bande de codage est directement reliée à la surface qu'elle occupe. Ainsi, pour les aires motrices, la main est largement privilégiée par rapport au reste du corps.

La combinaison des flux visuel et proprioceptif dans l'aire associative SV, suit les règles de correspondance symétrique énoncées chapitre II.4.2.2; qui plus est, elle inclut des relations implicites entre la rétinotopie et la somatotopie. Le gradient "distance par rapport à l'axe visuel" représenté par une dimension de V2 est en correspondance directe avec le gradient réalisé sur l'importance fonctionnelle des parties du corps dans une dimension de M2. Ainsi et par exemple, la région qui gère les muscles des doigts est particulièrement connectée avec la vision centrale. Ce phénomène de *recombinaison privilégiée* entre différents pôles sensori-moteurs suit une logique certaine, car il est normal que des stimuli de plus en plus excentrés fassent appel à des muscles de plus en plus distants de la main (coude, épaule...). Cette prédisposition vraiment particulière optimise la longueur des chemins parcourus par l'information; on peut se demander, à juste titre,

comment la nature a su optimiser une telle représentation.

### 3.3 Le réseau

En s'en tenant à ces quelques remarques neurobiologiques et au schéma des aires, proposé chapitre II.3.2.1, le guidage de la main par l'œil fait intervenir quatre aires différentes: L'aire motrice M2 est responsable des commandes musculaires et s'organise, sur un axe principal, en bandes correspondant chacune à un muscle précis, et en déplacements relatifs sur l'autre axe.

L'aire somatique S2 code les positions absolues des différents muscles sur sa surface; l'aire visuelles V2 n'est plus une représentation fidèle de l'image d'entrée, mais fait subir aux données une transformation quasi polaire qui, à partir des coordonnées rétinienne, génère une représentation en terme d'angles et de distance par rapport à l'axe central de la vision. Enfin, l'aire associative SV permet de relier représentations visuelles et motrices.

Bien sûr, la connectivité symétrique de toutes ces aires respecte le principe de recombinaison privilégiée entre la main et la vision centrale. La figure 1 illustre ainsi la représentation des flux d'informations dans le réseau ainsi élaboré.

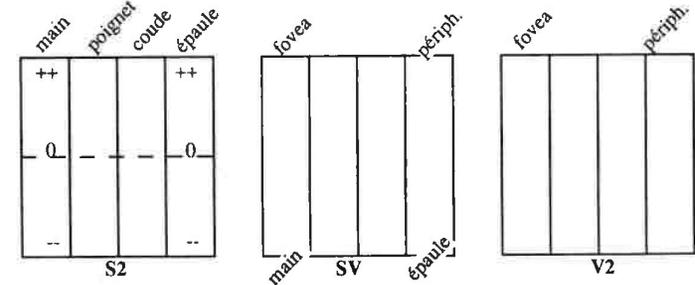


Figure 1: recombinaison privilégiée

### 3.4 Le fonctionnement

Nous avons vu au chapitre III.1, comment les coefficients probabilistes mémorisés entre unités de cartes connectées, permettent la définition d'une

logique associative; au chapitre III.2, l'utilisation des liens locaux et voisins a abouti sur des mécanismes de différenciation d'éléments pertinents et de propagation de l'information de proche en proche, permettant une connectivité locale.

Nous allons maintenant expliquer comment, en utilisant les coefficients probabilistes dans les connexions locales, des mécanismes d'ordonnement séquentiel de buts et de sous-buts peuvent émerger du fonctionnement global.

Si on fournit une interprétation orientée sur les aspects moteurs, une colonne A peut se concevoir comme un but potentiel qui connaît les sous-buts B susceptibles de favoriser sa réalisation (B2 avant A2). Récursivement, ces mêmes sous-buts ont appris leurs propres sous-buts. Ce type de mémorisation va structurer dans une aire des chemins arborescents, supports d'un séquençement temporel des actions élémentaires. Dynamiquement, lorsqu'un but quelconque est requis par le reste du réseau (des aires frontales ou limbiques), selon les tables d'entrée/sortie, cet état de recherche (E1) se propage de sous-buts en sous-buts qui, selon les conditions proposées par l'environnement extérieur, seront soit inhibés soit au contraire réalisables. Après le déclenchement du sous-but le plus lointain, les sous-buts intermédiaires vont pouvoir se réaliser les uns après les autres, dans un déclenchement en chaîne temporel, jusqu'à la réalisation du but initial.

En ce qui concerne le guidage main/œil, les buts initiaux sont typiquement de chercher à placer la main dans la région centrale de la vision. En général, les modules représentant un tel but sont activés après ceux qui se trouvent sur le passage de la main dans un champ plus périphérique, et récursivement. L'apprentissage génère ainsi des *arbres d'appels* radiaux, qui codent les mouvements possibles pour amener la main dans l'axe. Lors d'un appel, l'activité de recherche se propage jusqu'au sous-but qui représente la position courante de la main dans le champ visuel, qui, réalisable, va alors déclencher à rebours la chaîne d'actions suffisante.

Ce séquençement des actions, par apprentissage, est un premier pas vers la génération de plans d'actions, mécanisme cognitif nécessaire à la genèse du raisonnement.

## 4 INTERPRÉTATION D'IMAGES

### 4.1 Introduction au problème

Quelles différences distinguent la reconnaissance de l'interprétation d'images? Ainsi que nous l'avons vu dans les premiers paragraphes de ce chapitre (III.1 et III.2), la reconnaissance consiste à associer un nom à une image, directement, sans passer par un niveau de description "conscient". Au contraire, interpréter suppose la possibilité de donner des explications et des commentaires sur le processus qui a conduit à une solution particulière. Pour illustrer cette différence, citons, comme système d'interprétation, le diagnostic médical automatique à partir de clichés radiologiques.

#### 4.1.1 Problèmes induits par l'interprétation

Pour pouvoir expliquer son raisonnement, le système d'interprétation doit intégrer un niveau symbolique de description des images.

Pour un système connexionniste classique (reconnaissance associative), il devient alors nécessaire d'élaborer des mécanismes qui extraient des primitives symboliques à partir des valeurs des coefficients entre unités, ce qui n'est ni esthétique, ni réellement faisable.

Chez l'homme, le langage est la base même du raisonnement symbolique, car il lui permet de manipuler ses représentations internes. La communication avec un système d'interprétation suppose l'existence d'un langage commun entre l'homme et la machine, et suffisant pour décrire la démarche de l'interprétation. Il faudra donc se pencher sur la structure des images pour déterminer le vocabulaire nécessaire à une bonne description.

Plus pragmatiquement, l'interprétation d'image va différer de la reconnaissance par la taille des données. En effet, il ne s'agit plus d'identifier un flux visuel relativement limité, mais plutôt de sélectionner, parmi une quantité de données astronomique, des éléments pertinents pour le raisonnement. Ceci va mettre particulièrement l'accent sur deux des problèmes classiques liés à reconnaissance de formes: la taille des données qui doit être

impérativement réduite pour satisfaire des contraintes de place mémoire et de temps de calcul, et l'indépendance en translation qui permet au système de retrouver un indice indépendamment de sa position globale. Ici encore, c'est un problème de représentation qui est implicitement soulevé.

#### 4.1.2 Grandes approches

Depuis longtemps, la méthode la plus classique pour extraire des parcelles de connaissances d'une image, a consisté à calculer des paramètres locaux ou globaux sur toute l'image, pour les comparer ensuite à des valeurs moyennes significatives. De la couleur moyenne à l'entropie, toute une gamme de formules, plus ou moins complexes, parsèment la littérature. Bien que d'une mise au point délicate (trop de paramètres à définir), cette méthode, appliquée à des données non bruitées et stables, peut fournir des résultats honnêtes, et est encore utilisée pour des applications d'imagerie médicale [Magnin 86, Lei 86]. Peut-on déjà parler d'interprétation d'image? En s'en tenant à la définition avancée un peu plus haut, et si les "valeurs moyennes d'homogénéité locale" peuvent tenir lieu d'explication, pourquoi pas! Cependant, l'introduction de données structurelles sur les images à interpréter [Baldy 86], semble être un minimum pour mériter ce titre.

Plus actuelles, deux approches se disputent le terrain: les systèmes experts et les réseaux à propagation de contraintes. Introduisons des notions communes aux deux méthodes qui seront reprises, pour notre approche, dans les remarques de fin de chapitre. La connaissance est souvent organisée en trois classes: des traitements de bas niveaux toujours assez systématiques, des descriptions symboliques des images, et des mécanismes stratégiques qui dirigent la recherche. Trois méthodes d'analyse, dans les faits souvent combinées [Matsuyama 89, Appel 85, Nagao 85 Tsotsos 85], sont utilisées:

analyse ascendante: Des primitives de détection sont appliquées sur l'image et suivies d'une reconnaissance de formes. L'interprétation de l'ensemble des indices ne s'effectue qu'à la fin [Rosenberg 86];

analyse descendante: Le système, à partir des solutions, sélectionne des caractéristiques intéressantes, puis recherche les indices visuels correspondants [Chassery 86, Garbay 85];

analyse avec retour arrière: Cette technique, indispensable, permet de tenir compte d'éventuelles erreurs et de reprendre l'interprétation à un stade

antérieur.

Les deux approches diffèrent par la représentation des connaissances, sous forme de règles pour les systèmes experts ou sous forme de grammaire pour les systèmes à propagation de contraintes. L'aspect temporel des évaluations (chainage avant et arrière ou parallélisme synchrone et asynchrone) découle de la représentation et génère des propriétés spécifiques aux deux méthodes.

#### 4.1.3 Notre approche

Nous nous proposons de montrer comment il est possible de réaliser des réseaux de colonnes qui satisfassent aux besoins de l'interprétation, à l'aide d'un mécanisme élémentaire: la séparation entre représentation temporelle et représentation pariétale. Comme d'habitude, des notions neurobiologiques et psychologiques vont servir de base à notre réflexion. Nous proposerons ensuite une implantation dans un réseau à deux symétries (auditive-visuelle et temporelle-pariétale). Nous verrons enfin comment les résultats expérimentaux de cette application sont compatibles avec les approches théoriques et quelles remarques ils peuvent nous suggérer.

## 4.2 Données neurobiologiques

Nous n'avons pas fait état jusqu'à présent de la propriété fondamentale du système visuel humain, qui va constituer la clé de voûte de cette application. Il s'agit de la coexistence de deux chemins privilégiés de l'information de la rétine vers le cortex frontal [Whitteridge 73]. Passons en revue les différentes structures neuronales impliquées dans la perception visuelle.

Nous avons mis en évidence la notion de champs récepteurs dans le paragraphe III.1.2, les définissant comme les voisinages d'entrée sur le monde extérieur. Les observations anatomiques de la rétine [Jung 73] mettent en évidence une évolution de la taille des champs récepteurs en fonction de leur distance angulaire par rapport à l'axe central de la vision (figure 1).

On distingue ainsi deux niveaux de résolution spatiale pour l'oeil:

- La vision centrale permet une perception nette des images formées sur la rétine. Elle ne constitue cependant qu'une très faible partie du champ visuel total, entre 2 et 3 degrés.

- La vision périphérique est, quant à elle, réalisée par des cellules munies de grands champs récepteurs en entrée; elle n'est donc guère précise en résolution spatiale, mais couvre malgré tout la totalité du champ visuel.

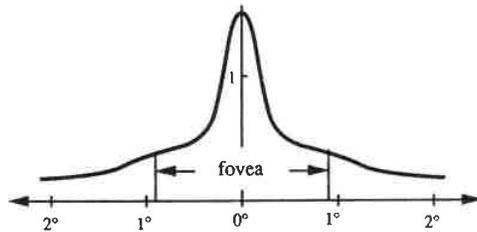


Figure 1: densité de récepteurs ( $10^5$  par  $mm^2$ )

Remarquons que, classiquement, la tache jaune qui constitue la zone de la rétine affectée à la vision centrale, est décomposée en deux parties: la fovea et la macula.

Le corps genouillé latéral, noyau thalamique affecté à la vision, ainsi que les aires sensorielles primaires et secondaires (V1 et V2), maintiennent cette distinction; la représentation de la partie centrale de la vision mobilise de plus en plus de neurones jusqu'au cortex, au détriment de la partie périphérique. Cette projection de la rétine sur le cortex sensoriel conserve la rétinotopie, mais déforme les images [Szentagothai 73b].

Le cortex sensoriel primaire (V1) possède des sorties spécifiques qui contrôlent la réception du flux visuel. En effet, les neurones pyramidaux de la couche V se projettent sur le colliculus supérieur qui commande le positionnement oculaire [Szekely 73]. Les neurones du colliculus sont activés par la présence d'un stimulus dans le champ visuel et peuvent déclencher un mouvement oculaire vers ce stimulus [Sprague 73].

Au delà des aires sensorielles, les chemins de l'information visuelle centrale et périphérique, vont se séparer. Les aires et leurs connexions réciproques vont alors former deux chemins parallèles, un chemin pariétal [Andersen 83], qui analyse les positions pour la vision périphérique et un chemin temporal [Gross 73] qui reconnaît les formes pour la vision centrale.

Le chemin pariétal est composé d'aires comme V3, dont les cellules sont spécialisées dans la reconnaissance de mouvements en trois dimensions, ou VIP plutôt orientée vers la représentation des déplacements oculaires [MacKay 73]. Le chemin temporal, quant à lui, est constitué d'aires présentant des propriétés de reconnaissance invariante, comme V4 pour la couleur ou STS et IT pour la reconnaissance des visages. Ces aires sont, de plus, directement liées au système limbique [MacLean 68] qui attribue une importance vitale aux différentes formes.

Le traitement du flux visuel est d'une part parallèle, à un niveau local pour identifier des formes élémentaires, et d'autre part séquentiel, pour reconnaître des formes plus globales par l'intermédiaire des mouvements oculaires [Otto 90].

Plus spécialisées et plus symboliques, les aires visuelles frontales organisent les explorations visuelles de scènes nouvelles, à des échelles temporelles différentes [Hecaen 83].

### 4.3 Données psychologiques

Indépendamment de l'antériorité du tout sur les parties, postulée par les gestaltistes, on ne peut nier la notion de forme. Ceux-ci la définissent comme "étant autre chose ou quelque chose de plus que la somme de ses parties" [Ehrenfels 90]. Elle fait état de propriétés qui ne résultent pas de la simple addition des propriétés de ses éléments.

Plus précisément, on distingue des formes qui gardent leurs propriétés malgré des modifications sur leurs éléments, de celles qui perdent leur caractère au moindre changement [Guillaume 79]. Cette distinction entre formes faibles et formes fortes, peut s'illustrer en déplaçant un sommet d'un triangle (forme faible) ou d'un carré (forme forte).

Essayons de percevoir objectivement (!) la figure 2:



Figure 2: Taches

Tout le monde aura sans doute reconnu deux tas de taches [Kohler 29]. Pourquoi deux? Il semble évident que si l'on diminue la distance entre les deux tas et que l'on augmente celles entre leurs éléments respectifs, l'impression de dichotomie diminue. Cet exemple, ajouté à la nécessité de distinguer des formes fortes ou faibles, montre l'importance de la distance entre éléments dans la perception des formes.

Outre la distance, il est un autre facteur qui intervient dans la perception, illustré sur la figure 3.

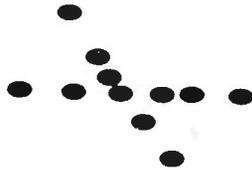


Figure 3: points

Des points rapprochés tendent, certes, à former des lignes, mais l'appartenance d'un point à une ligne vient surtout du fait qu'il en constitue le meilleur prolongement possible, qu'il en continue mieux le mouvement [Wertheimer 22]. Ce concept est connu sous le nom de loi de la bonne forme, où "la forme qui est perçue est la meilleure possible" [Wertheimer 23], bel exemple de formulation floue psychologique! Il met cependant en avant une idée intéressante: le *mouvement oculaire* contribue à la perception des formes.



Figure 4: cercles

Nous n'irons pas plus loin dans les idées empruntées à la Théorie des Formes, car elle n'est pas adaptée à une représentation du raisonnement fondée sur le besoin et l'action, telle que nous la concevons.

La figure 4 est, elle aussi, très intéressante pour comprendre les motivations qui nous ont poussés à nous intéresser au mouvement oculaire. En effet, même en l'absence de points intermédiaires entre les sommets du triangle, on en perçoit nettement les arêtes. Le balayage oculaire d'un *point de fixation* à un autre, crée l'illusion d'un contour.

La figure 5 correspond à l'enregistrement du chemin parcouru par l'oeil lors de l'exploration d'un visage [Jeannerod 74]. On remarque tout d'abord la présence de nombreux points de fixation où l'oeil s'est arrêté.

On distingue ensuite clairement des zones privilégiées dans lesquelles l'oeil est revenu de nombreuses fois. Elles correspondent à des parties du visage très contrastées et mobiles, ce qui n'est pas sans évoquer la boucle réflexe entre les aires visuelles primaires et le colliculus. Plus remarquable encore, ces indices sont pertinents pour la reconnaissance des visages, ce qui suppose de l'attention sélective *par apprentissage*.



Figure 5: visage et points de fixation

#### 4.4 Le réseau d'interprétation d'image

##### 4.4.1 Conception du réseau

A partir de ces données neurobiologiques et psychologiques, il faut spécifier un réseau dédié à l'interprétation d'images. L'idée directrice de la spécifica-

tion d'un réseau réside dans le choix des aires impliquées dans le processus et dans la détermination de leur codage d'entrée. Pour les applications précédentes (mémorisation associative), les réseaux étaient de taille restreinte: quatre cartes et un seul codage à définir. Nous sommes confrontés à une tâche cognitive plus complexe, qui nécessite une modélisation précise pour être simulée.

Interpréter une image, c'est, nous l'avons dit, être en mesure de fournir des explications sur le raisonnement qui a conduit à l'identification. Cette définition demande à être précisée:

**Les images.** Par définition, de petites perturbations locales dans la structure d'une forme forte en modifiant l'interprétation. Nous ferons l'hypothèse que la fonction du réseau se limite à différencier des formes fortes. Les formes faibles sont mises à l'écart, car leur différenciation suppose des mécanismes qui ne sont pas implantés (indépendance en taille, voir remarque III.4.7.2). *Sous cette hypothèse, une forme peut être modélisée par une succession de déplacements oculaires et de points de fixations sur des éléments locaux.*

**Les explications.** Nous supposons aussi que le réseau fournit ses explications sous la forme de noms et d'adjectifs; pour chaque élément local de la forme, il faut déterminer un nom et des adjectifs correspondant à ses diverses variantes significatives. De plus, pour rendre compte des déplacements oculaires, il est utile de définir des directions (gauche, haut...).

**Le raisonnement.** Par apprentissage, le réseau connaît des éléments locaux, les adjectifs qui les qualifient et leur structuration en forme. De plus, il possède les descriptions des variantes de la forme à identifier, en termes de différences qualitatives sur certains éléments locaux. Le raisonnement reste simpliste; il consiste à scruter les divers éléments locaux, en suivant la description des variantes.

Nous devons concevoir un réseau qui puisse nommer des éléments locaux, les qualifier, bouger le regard suivant leur structure et mémoriser une description symbolique des variantes. Partons du réseau à une symétrie, visuelle/auditive, qui permet déjà de nommer des éléments locaux. Si on développe la symétrie temporale/pariétale, les connexions réciproques entre la vision centrale et la vision périphérique vont permettre une description des formes structurelle, via les mouvements oculaires. Enfin, l'adjonction de certaines cartes frontales à ces cartes associatives, autorisera un niveau symbolique compatible avec la représentation des stratégies d'exploration.

A chaque carte définie pour cette application, va correspondre un niveau symbolique du langage développé, comme illustré figure 6.

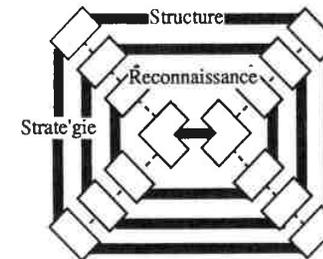


Figure 6: les cartes et leur symbolisme

#### 4.4.2 Les cartes du réseau

Passons en revue ces différentes cartes et spécifions au passage leur codage initial, en suivant la logique des symétries:

**Prétraitement des entrées visuelles.** Pour simplifier le réseau, nous n'avons pas considéré de carte visuelle primaire (V1); ce qui est justifiable, d'une part par la similitude des connexions V2/V1 et V1/monde extérieur, et d'autre part par la relative faiblesse du traitement effectué par V1 (voir paragraphe III.1.2). On intégrera les fonctions de V1 et de sa carte de codage en une seule étape. Le prétraitement consiste alors en une extraction de contours, qui simule grossièrement la fonction du corps genouillé latéral, suivie d'une extraction de segments qui simule quant à elle, la sensibilité à l'orientation de V1. Les masques d'entrée sont donnés figure 7 à titre indicatif:

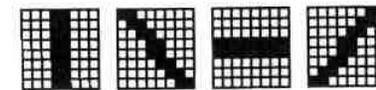


Figure 7: masques d'entrée.

La carte visuelle secondaire (V2). La carte de codage de V2 présente deux propriétés singulières:

1. Elle présente deux niveaux de résolution différents qui rendent compte de la distinction entre vision centrale et vision périphérique. Un peu comme pour l'audition (III.2.4.3), les unités responsables de la vision centrale recombinent entre elles les segments extraits par le prétraitement, ce qui, tout en correspondant à des observations neurobiologiques, génère le concept d'angle et une légère invariance en translation. Plus prosaïquement, V2 est composée de 32x32 unités juxtaposées et correspondant à une résolution d'un pixel.

Par contre, les champs récepteurs des unités de la vision périphérique sont beaucoup plus grands; ils englobent un voisinage de 16x16 pixels, ce qui ne permet pas de tenir compte des orientations. Les images de départ étant de taille 512x512, la partie périphérique de V2 sera composée de 32x32 unités (cf III.4.6.3).

La figure 8 montre d'une part, l'extraction des contours et des segments de l'image originale et, d'autre part, les activités dans les deux parties de V2 pour un point de fixation donné.

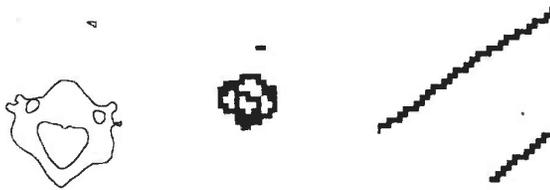


Figure 8: Codage initial dans V2

2. Les sorties de V2 vers le monde extérieur présentent elles aussi des caractéristiques bien spécifiques à la vision. En effet, la partie périphérique de V2 se projette sur une structure qui simule le colliculus

et qui commande les mouvements oculaires. Plus précisément, un stimulus visuel périphérique codé dans V2, est potentiellement capable de déclencher le mouvement des yeux vers sa position. Cette propriété fondamentale de V2, va permettre la représentation des structures des formes par des déplacements oculaires, dans les aires associatives.

L'aire temporale visuelle VT. Cette carte associative reprend le principe de division expliqué dans la section "audition" (III.2.5). Directement liée à une aire associative auditive (AT), et surtout à la partie centrale de la vision (V2), elle va élaborer une représentation centrée objet des éléments locaux qui lui sont proposés. Sa taille va varier, suivant son niveau de division, de 1 à 32x32 unités.

L'aire pariétale visuelle SV. Réciproquement, bâtie sur le même modèle mais connectée à la vision périphérique, l'aire pariétale construit une représentation centrée structure des formes globales à identifier. Ce sont les connexions entre les aires associatives pariétale et temporale qui vont permettre une représentation globale et distribuée.

L'aire auditive secondaire A2. L'objectif principal de cette application n'est pas de faire de la reconnaissance de la parole. C'est pourquoi, là encore, les aires auditives seront réduites à leur plus simple expression: une liste d'unités "grand-mères", correspondant chacune à un mot du vocabulaire du réseau.

Les aires auditives associatives AT et SA. Elles ont toutes les deux la même fonction, associer un mot et un flux visuel. Contrairement aux aires associatives visuelles, elles ne se distinguent pas par deux niveaux de résolution; c'est plutôt la fonction de la carte visuelle associative symétrique qui va les spécialiser respectivement pour la mémorisation des noms des éléments visuels, et celle des directions. Plus particulièrement, AT permet, lorsqu'un élément local a été identifié par l'intermédiaire de SA, de détecter des indices caractéristiques des variations de cet élément dans VT. C'est dans cette aire que vont se mettre en place les représentations distribuées des adjectifs qualificatifs.

Les aires frontales FV et FA. FV présente une caractéristique étonnante, vue sa position dans le cortex: elle est capable de commander des mouvements oculaires par des actions directes sur le noyau moteur. De plus, grâce à l'aire intermédiaire FA connectée à A2, elle est directement reliée aux représentations symboliques que sont les mots dans le cortex.

Ces deux aires sont impliquées dans la stratégie globale de reconnaissance.

## 4.5 Signal, contexte, requête

### 4.5.1 Rappels

Classiquement, pour des modèles de type perceptrons, on utilise des réseaux à couches simplement hiérarchisées, avec une propagation unidirectionnelle de l'information en feedforward [Fogelman 87]. Des modèles plus avancés [Waibel 89] tiennent compte de mécanismes de feedback de couches à couches (figure 9.a), qui permettent alors une remise en question d'informations de plus bas niveau. En effet, la structure même, en trois étapes de traitements linéaires, crée une hiérarchisation de l'information, des couches d'entrée aux couches de sortie via les couches cachées, qui correspond à des niveaux d'abstraction de plus en plus élevés (du moins pour les réseaux hétéroassociatifs).

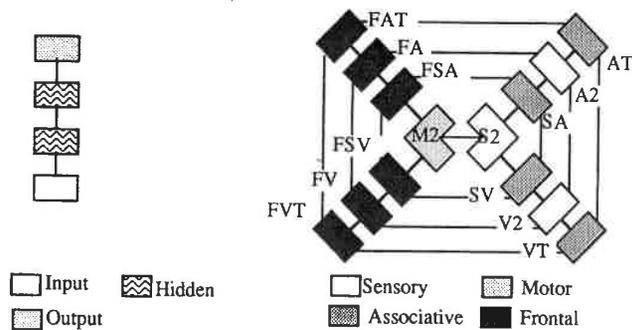


Figure 9: .a réseau classique .b notre modèle

Les réseaux de neurones classiques butent sur un problème de niveau d'organisation de l'information. En effet, s'ils possèdent des mécanismes d'apprentissage et de fonctionnement bien définis et efficaces au niveau du neurone, ils ne proposent aucune solution à la structuration des couches entre elles. Personne n'a réussi, jusqu'à présent, à corréler des informations

sur la tâche à effectuer avec le nombre de couches cachées, leur taille ou leur connectivité [Bienenstock 89].

Notre approche ne se limite pas à proposer un modèle d'unité connexionniste (chapitre II.4), car nous avons aussi défini un modèle de structuration des aires entre elles (chapitre II.3... et figure 9.b). D'inspiration biologique, ce modèle est fortement réductionniste et sera sans doute jugé simpliste par la grande majorité des neurobiologistes; pour preuve la figure 10, qui représente une modélisation des aires visuelles plus poussée qu'un simple schéma V1, V2, VT et SV.

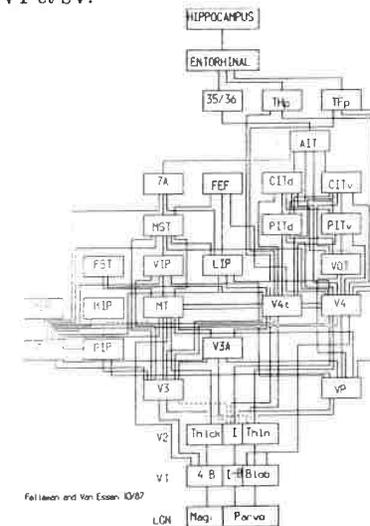


Figure 10: autre modèle des aires visuelles

Les réseaux développés jusqu'à présent (III.1, III.2, III.3), essentiellement orientés vers la mémorisation associative, étaient composés de trois cartes (sensorielle, motrice) et sans doute plus proches des modèles multicouches (figure 9.a) que du réseau des aires défini (figure 9.b). Pour dépasser ce niveau de perception, il est nécessaire de développer un réseau plus vaste, non pas au niveau de la taille des aires, mais plutôt de

leur nombre. Il est clair que l'intelligence du réseau réside dans le nombre de symétries, et donc dans le nombre de dimensions de représentation, impliquées dans un processus.

Le comportement de l'unité connexionniste demande alors à être précisé: doit-elle tenir compte, tant pour son fonctionnement que pour sa mémorisation, de la structure des aires, et donc de la provenance de ses entrées lointaines? Avant d'approfondir cette notion, revenons un peu sur les données neurobiologiques. Dans une première approximation (paragraphe II.4.2.2), les entrées d'information en provenance du cortex lui-même ont été regroupées dans une seule partie de la colonne corticale: les couches hautes. Cependant, on peut distinguer plus finement des localisations précises des afférences corticales [Jones 81b]. Dans un souci de description de ces afférences, indépendante de l'aire considérée, il est commode d'utiliser les propriétés de connectivité symétrique. Chaque aire est théoriquement reliée avec six autres, suivant les six axes de symétrie du modèle de réseau hypercubique des aires.

C'est pourquoi, on peut désigner, de façon générale, une afférence par la nature de la symétrie. Ainsi et par exemple, les liaisons de VT à AT, ou de SV à SA seront désignées sous le terme commun *d'afférences auditives/visuelles*, et suivant le même principe, les liaisons de VT à SV, sous le terme *d'afférences temporales/pariétales*.

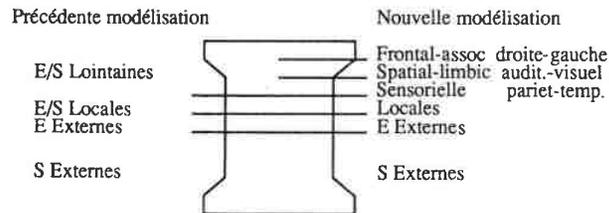


Figure 11: les sous-couches du cortex

Il est remarquable de constater que, quelle que soit l'aire, les afférences sont distribuées dans la couche haute de façon régulière, hiérarchisée et répétitive, définissant des sous-couches précises. La figure 11 illustre com-

ment, en fonction de la symétrie et donc de l'éloignement, s'organisent les entrées de la couche haute.

Il est évident que cette répartition singulière des afférences ne peut qu'avoir des répercussions sur le fonctionnement et la réponse de l'unité. Nous sommes amenés à distinguer pour l'unité, non plus quatre types d'entrées (lointaines, voisines, locales et externes), mais neuf, en décomposant les entrées lointaines en six classes. Cette distinction ne remet pas en cause la totalité de la table d'entrée/sortie proposée paragraphe II.4.4.2, mais définit plutôt un mécanisme d'intégration des entrées lointaines parallèles, différent de celui esquissé paragraphe II.4.4.3.

#### 4.5.2 Implantation

Le réseau proposé pour l'interprétation d'images se décompose fonctionnellement en deux parties: un module associatif qui réalise les traitements les plus flous et subtils de cette application et une partie frontale qui organise des processus d'exploration grossiers. Nous allons donc surtout nous intéresser à l'apprentissage et au fonctionnement dans les aires SA, VT, SV et AT, en éludant les aires frontales dont la stratégie sera imposée sans apprentissage. La figure 12 montre ainsi le sous-réseau extrait du schéma général du cortex, qui va servir de base associative.

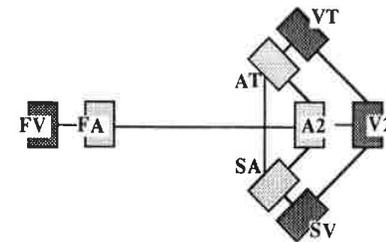


Figure 12: sous-réseau extrait

Chaque aire est connectée à trois autres aires, suivant les symétries sensorielles, pariétale/temporale et auditive/visuelle. Notre objectif est d'élaborer en fonction de ces trois types d'entrées lointaines un mécanisme

ad-hoc. C'est encore une interprétation en termes de cognition, qui va nous fournir la clé du problème. Postulons que dans une aire associative, après apprentissage, une requête R reçue, dans un contexte C, et validé par un signal S, va déclencher une action A (on pourrait ajouter pour tenir compte des stratégies frontales une priorité P, ou quelque chose d'analogue...). Il est clair que, vues les définitions fonctionnelles proposées plus haut pour chacune des aires associatives, pour une aire donnée, le contexte est représenté dans l'aire symétrique pariétale/temporale, le signal dans l'aire sensorielle et la requête dans l'aire auditive ou visuelle symétrique. Ainsi, pour l'aire pariétale visuelle, qui code les mouvements des yeux, chercher une certaine partie de la structure globale (R), alors que, localement, une certaine forme est reconnue (C) et qu'un stimulus est localisé à l'endroit présumé (S), déclenche le mouvement des yeux (A) vers celui-ci.

Nous proposons pour combiner ces trois types d'informations complémentaires, d'imposer des priorités pour l'évaluation des afférences. Il semble presque normal (!), que cette hiérarchie aille dans le sens des symétries du cortex, et donc des sous-couches de la couche haute. Ainsi, les requêtes (symétrie auditive/visuelle) sont prioritaires sur le contexte (symétrie pariétale/temporale), elles mêmes prioritaires sur les signaux. La table de la figure 13, illustre l'effet en cascade de ces différentes entrées.

	Requête	Contexte	Signal	Etat
1	Dec. Inc.	Indep.	Indep.	E2
2	Inhibition	Indep.	Indep.	E0
3	Dec. Cond.	Dec. Inc.	Indep.	E2
4		Inhibition	Indep.	E0
5		Dec. Cond.	Dec. Inc.	E2
6			Inhibition	E0
7			Dec. Cond.	E1

Figure 13: table de priorité

Illustrons un peu cette table de priorité par des exemples simples pour l'aire pariétale visuelle (SV), qui reçoit des requêtes de SA, un contexte de VT et des signaux de V2.

Ligne 1. Si à chaque fois ( $P2=1$ ) que je vois le soleil, je regarde vers le haut, alors, lorsque je cherche le soleil (R), je regarde en haut (A).

Ligne 2. Réciproquement, si à chaque fois que je vois la mer, je ne regarde jamais en l'air ( $P0=1$ ), alors lorsque je cherche la mer (R), je ne regarde pas en l'air (à tort, peut-être, mais tellement humain [Wazlawick ref]...)

Dans des cas moins extrêmes et, dans les faits, les plus courants, la localisation d'une forme locale dépend du contexte temporel, c'est à dire de la forme locale examinée (lignes 3 à 7). Par exemple, si on cherche le nez dans un visage (requête E1, mais plusieurs mouvements possibles, illustration figure 14), c'est le contexte "bouche" ou "front" qui va déterminer le mouvement à effectuer. Si à chaque fois que je vois le front puis que je regarde vers le bas, je vois le nez, alors, si je vois le front (C), et que je cherche le nez (R), je regarde vers le bas (A)...

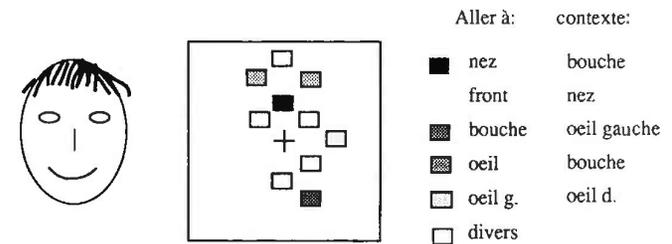


Figure 14: autre visage

Dans un cas rare, où malgré le contexte, le système ne peut pas prendre de décision, la présence ou l'absence de stimulus permet de sélectionner parmi les diverses possibilités. Toujours pour la recherche du nez, si je vois un oeil, j'ai deux mouvements possibles: en bas à gauche ou en bas à droite. C'est la présence d'un fort stimulus S (le nez dans le champ visuel périphérique) qui déclenche le bon mouvement.

Cet exemple de reconnaissance des visages, illustre l'interprétation des afférences de l'aire pariétale en fonction des types de symétrie, en terme de contexte, signal et requête. On pourrait pour chacune des aires associatives,

qui reprennent la même logique, proposer une interprétation comportementaliste. Citons, par exemple, la mémorisation des adjectifs dans l'aire AT, qui sont représentatifs d'un indice visuel (de VT), pour une forme reconnue (contexte SA).

#### 4.5.3 Apprentissage

On pourrait objecter que, pour deux recherches et contextes associés différents, le mouvement du regard peut être le même (chercher le front dans le contexte "nez" et chercher le nez dans le contexte "bouche") et que la logique définie est incapable de prendre une décision. Intuitivement, ce cas n'est pourtant pas ambigu. Ce serait ne pas tenir compte du mécanisme de division, proposé au paragraphe III.2.5, qui permet de différencier localement des représentations différentes, si nécessaire.

Il est généralement évident qu'il faut savoir marcher avant de courir, ou savoir faire des additions avant de faire des divisions. Les psychologues génétiques définissent ainsi des stades de développement des comportements humains, avec une relative précision [Piaget]. De même, nous sommes ici confrontés à un problème de maturation différenciée des aires corticales. Il est en effet impossible de mémoriser une structure entre des éléments locaux, si ceux-ci ne sont pas déjà appris et reconnus, ou bien d'apprendre des adjectifs particuliers sans connaître auparavant les objets visuels auxquels ils se rapportent. Nous proposons alors un protocole qui ordonne l'apprentissage:

1. Apprentissage des noms (maturation de VT),
2. apprentissage de la structure (SV),
3. apprentissage des adjectifs (maturation de AT et dans une moindre mesure de SA),
4. câblage imposé des stratégies (FV, FA).

## 4.6 Application

### 4.6.1 Introduction

Le projet Orfèvre initié en 85, s'était donné comme objectif le diagnostic automatique des pathologies cervicales, à partir de clichés radiologiques.

L'idée de base consistait à formaliser une sémiologie sectorielle [Wackenheim 84] et à l'implanter sous la forme d'un système expert. Des travaux sur l'extraction d'indices par des méthodes classiques et sur la modélisation des connaissances expertes, furent menés parallèlement à Nancy et Strasbourg. Cependant, le raisonnement logique, point fort des systèmes experts, n'étant pas d'une complexité affolante, seulement quelques structures à identifier, et les opérations sur le flux visuel étant par contre trop floues pour être exprimées sous forme de règles, le choix d'un système expert n'était pas des plus judicieux.

C'est le renouveau des méthodes connexionnistes qui va, en 88, relancer le projet, rebaptisé pour la circonstance Neurodile. Même si les réseaux connexionnistes classiques ne présentent pas des capacités de raisonnement logique très développées, ils sont par contre tout à fait aptes à intégrer, par apprentissage, des représentations de données floues.

### 4.6.2 Les données

Les vertèbres. L'acquisition très rigoureuse des clichés radiologiques leurs confère deux propriétés importantes:

Le scanner étant placé à une distance fixe du patient, les vertèbres étant de taille constante chez l'adulte, la dimension absolue des formes est significative et tout écart à la norme est pathologique. Il n'est ni nécessaire, ni optimal, d'élaborer un système d'interprétation indépendant de la taille des images et qui, de ce fait, nécessiterait un mécanisme de calibrage.

Le flux de particules émis par le scanner est très stable et homogène comparativement aux conditions d'éclairage en lumière visible. Les niveaux de gris du cliché sont alors significatifs du tissu; on distingue ainsi les tissus mous, les cartilages, les os et les tumeurs, dont les densités respectives sont distinctes. L'extraction des contours osseux, sur lesquels le système se repère et s'oriente, ne posera pas de difficultés.

Les éléments locaux de la forme vertébrale, sur lesquels va s'effectuer l'interprétation, sont illustrés figure 15:

Les maladies. Certaines déformations locales de ces éléments visuels sont pathologiques; nous avons sélectionné, en fonction de leur fréquence, une dizaine de maladies (voir aussi en annexe):

- Sténose: rétrécissement congénital du canal vertébral,

- arthrose: excroissance du corps vertébral dans le canal vertébral,
- maladie du japonais: corps vertébral en forme de champignon,
- métastase: corps vertébral rongé,
- hémangiome: corps vertébral granuleux,
- spina bifida: épineuse double,

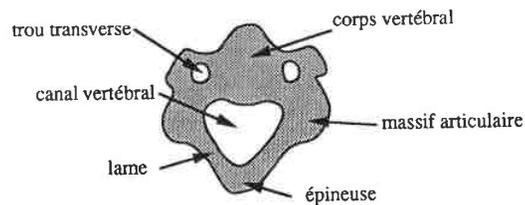


Figure 15: vertèbre et éléments

- hernie discale: débord du disque dans le canal vertébral,
- hernie de chmoll: corps vertébral troué,
- hypertrophie du trou transverse: trou transverse trop grand,
- hypertrophie du massif articulaire: massif articulaire trop grand,
- hypertrophie des lames: lames trop larges.

De ces descriptions, émergent quelques adjectifs (étroit, rongé, double, troué...) qui feront partie du vocabulaire du système. Il n'est pas nécessaire de préciser plus en détail la description; c'est par apprentissage que le réseau génère sa propre représentation des éléments pathologiques.

#### 4.6.3 Les résultats

L'apprentissage est décomposé en étapes de complexité cognitive croissante, le rôle du professeur qui le supervise devient alors crucial. Pour chaque étape qui correspond grossièrement à la maturation d'une aire, le système subit, comme pour les autres applications, deux phases d'apprentissage: présentation des données et des solutions correspondantes, qui permet la mémorisation des coefficients probabilistes, puis essais de reconnaissance, qui, en cas d'échec, conduisent à la division des zones incriminées.

Quand les résultats sont jugés suffisants, le professeur passe à l'étape suivante qui s'appuie sur les invariants dégagés antérieurement.

Apprentissage des noms. Pour la première étape, seul l'axe temporal du réseau, pour la vision centrale, est utilisé. Le professeur présente successivement tous les éléments de la forme au réseau qui apprend à les reconnaître et à les différencier. Le corpus d'apprentissage est constitué de quatre images de vertèbres saines, ce qui constitue un bon compromis entre la généralisation et la différenciation (voir paragraphe III.2.6.1). Il s'agit d'une mise en correspondance associative identique à celle mise en oeuvre pour l'audition. Testé sur des images saines, le système donne toute satisfaction, soit 90% de bonnes réponses, avec une invariance locale en translation de l'ordre du tiers de la taille de la vision centrale. Par contre, le taux de reconnaissance chute pour des images pathologiques, et se situe, suivant les maladies, entre 50 et 75%. Ce qui ne serait guère encourageant, si le système se limitait au chemin temporal.

On peut souligner les différents contrastes de résolution, en illustrant l'espace de codage du trou transverse et de la lame, figure 16.

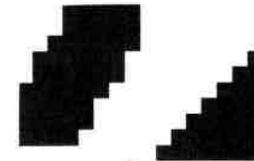


Figure 16: représentations internes

Les zones correspondant au trou transverse, sont suffisamment divisées pour qu'une "image mentale" puisse être reconstituée. Pour la lame au contraire, l'apprentissage génère un concept assez flou, qui peut éventuel-

lement être traduit par "de nombreux traits obliques, sans précision de localisation".

Apprentissage des structures. Le taux de reconnaissance sur les vertèbres saines étant jugé suffisant, l'apprentissage va porter sur le chemin pariétal, pour mémoriser des déplacements oculaires. Cette étape pourrait ne pas requérir de professeur: en promenant au hasard son regard sur l'image, le système reconnaît de temps à autre des éléments visuels; en intégrant le chemin parcouru depuis la dernière reconnaissance, il alors est capable de reconstituer la structure. Nous préférons cependant une méthode plus directe et plus rapide, qui consiste à désigner successivement tous les éléments de la forme. L'apprentissage a lieu en deux phases: mémorisation des coefficients probabilistes, puis essais de localisation d'un indice, qui conduit éventuellement à des divisions dans l'aire SV.

Vue la très grande stabilité de la forme vertébrale, les taux, non plus de reconnaissance mais de "bon déplacement vers un objectif demandé", ne peuvent être qu'excellents: pour le même corpus d'apprentissage, quatre images, et pour des essais sur des vertèbres saines et malades, ils avoisinent 95%. Mais beaucoup mieux, les transferts d'apprentissage entre les aires pariétales et temporales, lorsque les deux chemins sont couplés et se servent mutuellement de contexte, font remonter les performances de reconnaissance temporelle sur des images anormales à 90%. La figure 17 montre une reconstitution possible de la structure de la forme, à partir de tous les déplacements oculaires mémorisés.

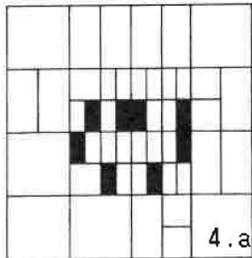


Figure 17: structure des vertèbres

câblage des stratégies. Le mécanisme de génération de plan, vu au chapitre III.3, propose une décomposition stratégique d'un problème en buts

et sous-but. Notre démarche consiste alors à câbler directement les différentes maladies avec leurs sous-but symboliques, les éléments vertébraux et les adjectifs. Par exemple, le but métastase est relié aux sous-but "corps vertébral" et "rongé". Un but initial plus général, "interpréter l'image", peut être utilisé pour gérer les différentes maladies, et sélectionner parmi les plus fréquentes, celles qui seront recherchées en premier.

Le câblage a priori n'exclut pas un affinage par apprentissage de la pertinence des indices et de l'analyse.

#### 4.7 Remarques

Cette application complexe suggère de nombreuses remarques, tant sur le comportement du réseau, que sur son évolution.

##### 4.7.1 Propriétés

Soulignons d'abord la simplicité de l'idée directrice de cette application, à savoir séparer le traitement du flux visuel en deux parties complémentaires, qui fait émerger des propriétés souhaitables.

Invariance en translation. La représentation pariétale est indépendante de toute origine. Le système se repère à partir du premier élément reconnu en balayant l'image. A un niveau global, il est donc totalement invariant en translation. Par contre, localement, cette invariance est générée par le codage initial temporel sous forme de segments orientés et d'angles.

Invariance en rotation. Elle est permise localement par la largeur de bande des détecteurs d'orientation (figure 7), et renforcée par le mécanisme de feedback. L'invariance en rotation global, supportée par les mécanismes visuels, n'est pas forcément souhaitable; chez l'homme, elle se réalise plutôt en inclinant la tête ou en tournant le support visuel.

Taille du réseau. Au lieu des 65000 unités nécessaires pour coder une image 256x256, le réseau n'en utilise, au maximum, que 2000 (2x32x32). C'est le niveau de résolution finale après apprentissage, dû au mécanisme de division, qui détermine le nombre d'unités employées, typiquement de l'ordre de 600.

Anticipation. La priorité accordée au contexte sur le signal dans les aires pariétales, permet au système de prévoir à l'avance la position d'un indice dans le champ visuel. Dans notre modèle du cortex, l'anticipation se définit

plus comme un processus temporel orienté vers les actions (sorties), tandis que l'attention sélective (paragraphe III.1.4) relève des stimuli (entrées) parallèles.

Analyse. Il est difficile d'identifier le fonctionnement du réseau à une méthode d'analyse classique. L'analyse est décomposée en un certain nombre d'étapes fonctionnelles (correspondant aux aires), qui peuvent s'organiser dans le temps séquentiellement ou parallèlement. Le raisonnement est à la fois guidé par les buts et par les données grâce à des mécanismes de feedback et d'anticipation qui se répondent mutuellement.

Représentation des connaissances. Pour la vision centrale, la connaissance est représentée sous forme de différences significatives entre les formes à discriminer. Par contre, la représentation de la structure sous forme de déplacements oculaires dans les aires pariétales, dépend des capacités d'invariance en translation des aires temporales. La stratégie d'exploration est bien pauvrement câblée dans les aires frontales. Soulignons à ce propos, le risque de confusion entre la logique émergente du système (tel que le raisonnement hypothético-déductif) et la logique de fonctionnement de l'unité.

Apprentissage. Il est vraiment fructueux de coupler et de faire coopérer les deux flux d'informations, ce qui favorise des transferts d'apprentissage, tout en ajoutant une dimension au contexte appris. Pour une configuration pariétale ou temporale mal reconnue, et en l'absence de toute autre solution, c'est le contexte symétrique qui assure la relève. En particulier, la forte structuration des images de l'application permet de localiser, et donc de reconnaître, des éléments locaux pathologiques et déformés, alors que l'apprentissage n'a eu lieu que sur des images saines.

#### 4.7.2 Idées

Sommation vectorielle. Comment se comporte le système lorsque sachant aller de A (contexte) à B (requête), et de B à C, on lui demande d'aller de A à C? Dans son état actuel, la propagation de l'état de déséquilibre (E1) induit par la requête, va déclencher séquentiellement le mouvement de A vers B, puis le mouvement de B vers C. Il est clair que le système doit apprendre à généraliser de telles chaînes d'actions et dans ce cas précis, savoir passer directement de A à C.

Sous-formes. La taille de la vision centrale n'a pas été choisie arbi-

trairement; elle est supérieure à celle du plus grand des éléments de la forme (le trou transverse). Cet artifice permet de n'avoir qu'un seul niveau de décomposition symbolique d'une forme en éléments plus simples. Plus généralement, il est indispensable de disposer de possibilités de descriptions récursives symboliques plus étendues pour prétendre à l'interprétation d'images.

Mouvement des yeux. Les vertèbres saines présentent une variabilité en taille minimale, mais certaine; les vertèbres malades peuvent subir des déformations de structure étranges. De plus, la résolution de la vision périphérique, et la plage d'invariance en translation de la vision centrale, sont du même ordre de grandeur. Un deuxième mécanisme de guidage oculaire, par la vision centrale, peut être utilement mis en place.

Formes tridimensionnelles. Le diagnostic réel a lieu sur trois radiographies qui représentent trois coupes longitudinales de la même vertèbre (annexe II). Le radiologue possède ainsi des informations sur la structure tridimensionnelle de la forme qu'il utilise pour certaines pathologies. Il nous semble pour l'instant très difficile d'intégrer naturellement (binocularité, manipulation tridimensionnelle?) des connaissances 3D, nous nous orientons plutôt vers un système "2,5D", où le contexte "hauteur de la coupe" et l'invariance de certaines structures sur les trois coupes, tiendraient lieu de troisième dimension...

Texture. On peut grossièrement différencier les adjectifs en deux classes qualifiant soit les contours (trop large, étroit...) soit les textures (granuleux, rongé...). Le système actuel n'est guère capable de percevoir les textures; l'apprentissage sur les masques d'entrée extérieure, permet, après de longs efforts, la reconnaissance des seules textures dont on peut extraire les contours. Il est beaucoup plus efficace de diversifier et de spécialiser les récepteurs visuels, ainsi que nous le suggère le système humain.

Formes fortes/faibles. La spécificité de cette application nous a permis de satisfaire à deux hypothèses pourtant bien réductionnistes, l'exclusion des formes faibles et la dépendance de la taille. La solution est commune, et consiste à implanter, non plus un intervalle d'éloignement correspondant à une localisation absolue du stimulus, mais plutôt une représentation "polaire", angles et rapports de distances. Ce mécanisme de calibrage se répercute sur l'effet des entrées locales (III.2.7.4 et III.3.4).

Transfert d'apprentissage. Nous débutons actuellement des travaux sur le mécanisme de transfert d'apprentissage qui pourrait généraliser une for-

me apprise localement en vision centrale, à une représentation globale et périphérique. Ainsi, comment le cortex fait-il le lien entre une minuscule lettre et une lettre géante? Est-ce un mécanisme acquis et sélectif? Le calibrage entre les entrées lointaines suffit-il à rendre compte de telles aptitudes?

Stratégies. La stratégie développée pour cette application reste très frustrante. Pour des raisonnements plus abstraits, du type hypothético-déductif, les aires frontales demandent à être simulées plus en détail. L'introduction de nouvelles fonctionnalités passera par la modélisation des entrées du cortex frontal, les échelles temporelles, et complétera le développement du traitement temporel de l'information.

#### 4.7.3 Symétries

Les messages. La plausibilité de l'amélioration apportée à la table d'entrée/sortie, est basée sur la distribution localisée des afférences dans les sous-couches du cortex. La provenance d'une connexion est distinguée physiquement. Pour des simulations informatiques, il est nécessaire d'ajouter un paramètre aux messages entre unités: l'aire de provenance, ou plus généralement, le type de symétrie.

Les couches basses. La projection des colonnes vers le monde extérieur, a été fortement simplifiée. Comme pour la couche haute, la position des neurones sur l'axe vertical, module des connexions vers des structures différentes. Bien que situées en dehors du cortex, les structures coopérantes participent largement au fonctionnement global et méritent sans doute une différenciation plus fine.

Gauche/droite. La symétrie développée, pariétale temporale, est essentiellement basée sur la différence de tailles des champs récepteurs visuels. Les aires auditives ne présentent pas, avant apprentissage, de différence de câblage. Quelles sont les règles qui gouvernent le choix des entrées extérieures, en fonction de la symétrie ajoutée? La symétrie gauche/droite est-elle au monde auditif, ce que la symétrie pariétal/temporal est au visuel?

Logique. Cette dernière application a mis en évidence l'interdépendance de nombreux mécanismes. Là encore, il est clair que ce système lié dépend d'une logique formelle sous-jacente. La structuration hiérarchisée des entrées lointaines, sans doute généralisable à toutes les symétries..., est un

pas décisif vers l'unification des diverses règles de fonctionnement.

## CHAPITRE IV: CONCLUSION

De toute évidence, la théorie et les applications qui viennent d'être exposées, quoique prometteuses, sont loin de constituer un aboutissement. Il est cependant temps de récapituler les avantages et les faiblesses de la modélisation, essentiellement par référence aux travaux et aux avis d'autres chercheurs. Dans une partie plus personnelle, des réflexions sur la démarche de la recherche, rendront compte de l'état d'esprit dans lequel se sont effectivement déroulés nos travaux. Bien entendu, des perspectives à plus ou moins long terme, viendront clore ce rapport et ouvrir le débat.

### 1 Critiques

Une approche tridisciplinaire telle que la nôtre, se doit d'être jugée sous une triple perspective; que les biologistes, les informaticiens et les psychologues puissent émettre un avis, même critique, valide déjà un peu la démarche.

#### 1.1 Réalité biologique

La question de la réalité biologique de la colonne corticale, restriction la plus souvent entendue, émane curieusement des informaticiens, plus particulièrement des connexionnistes. Nous avons déjà consacré un paragraphe (II.4.1) aux neurobiologistes qui ont mis en évidence cette organisation fonctionnelle des neurones, couronnée en 81 par un prix Nobel. Cependant, les raccourcis sont parfois saisissants, entre les données neurobiologiques que nous avons pris la peine de présenter avant toute modélisation, et les règles

induites qui caractérisent le modèle. Nous distinguerons deux descriptions: La description statique de la colonne prend en compte les divisions structurales en sous-couches, les localisations des afférences, les schémas de connectivité ou le réseau des aires... Basée sur des observations cytologiques précises et nombreuses, la modélisation peut, certes, être remise en cause, mais seulement sur des points de détail (nombre de connexions de sous-couches...); elle constitue une base sûre de travail. La réalité de la dynamique de l'unité, est, sans aucun doute, beaucoup plus sujette à caution. Quelques observations assez floues (Hebb, phénomène d'avalanche...) constituent les seuls points de repère biologiques pour élaborer les règles de fonctionnement ou de mémorisation. C'est seulement en modélisant mieux les neurones et leur structuration dans une colonne, que la correspondance entre un modèle de fonctionnement intuitif et la réalité pourra être établie; réciproquement, de tels travaux constitueraient une mine de renseignements précieux pour rendre le modèle plus performant...

## 1.2 Capacités de traitement de l'information.

Vu le manque de formalisme, il n'est pas question de démontrer des capacités particulières de nos réseaux, mais seulement d'obtenir une validation expérimentale sur des applications précises. Objectivement, les résultats annoncés supportent la concurrence, mais souffrent terriblement des défauts propres aux prototypes de laboratoire, domaine d'application très restreint, et mise à l'épreuve insuffisante. Si une représentation bidimensionnelle n'est pas jugée trop restrictive, on peut projeter les compétences d'un système suivant deux axes, quantitatif et qualitatif.

Une tâche cognitive étant fixée (telle la mémorisation associative), il convient de tester la quantité d'information que le réseau peut stocker. Remarquons tout de suite, que, pour des signaux artificiels reproductibles, une aire  $n \times n$  permet de coder  $2^{n \times n}$  signaux différents. Cependant, le système est construit pour l'analyse de données structurées, et ne génère pas une représentation optimale pour mémoriser des vecteurs aléatoires. Les mécanismes d'apprentissage, qui induisent une certaine généralisation, nous ferons représenter les capacités en termes de taux de reconnaissance, qui rendent compte de la quantité d'information et mesurent l'aptitude à la tâche. De plus, les taux de reconnaissance font partie du langage commun à tous les chercheurs en RFIA, et permettent de comparer les systèmes sur

la base d'un corpus semblable. C'est dans cet esprit que les résultats avancés chapitres III.1 et III.2, ont été établis, et bien qu'ils datent maintenant un peu, ils sont au moins significatifs des capacités minimales des réseaux à base de colonnes corticales.

La dimension qualitative est plus complexe à mettre en évidence et plus subjective à évaluer; il s'agit de se positionner par rapport à des stades d'analyse cognitifs. Les informaticiens proposent une classification fondée sur les connaissances représentées dans le réseau, le meilleur exemple étant celui de la reconnaissance de la parole, qui peut se décomposer en quatre stades, décodage acoustico-phonétique, reconnaissance de mots isolés, transcription de la parole continue et traduction automatique. Les psychologues sont plutôt axés sur les stades de développement de l'enfant. Si, dans les deux premières applications, nous avons plutôt recherché les performances quantitatives, les chapitres III.3 et III.4, vont plus dans le sens de l'abstraction et de l'intégration d'un plus haut niveau de raisonnement. Nous pensons en effet, qu'il est vain de développer au maximum une compétence particulière, sans tenir compte des autres aptitudes. Au contraire, le réseau doit être composé de mécanismes redondants et complémentaires, qui, isolément, ne fournissent pas des résultats quantitatifs extraordinaires, mais qui, pris dans leurs interactions réciproques parallèles et auto-correctives, transcendent leurs limitations respectives.

Sur ce, le réseau n'a guère fait preuve, pour l'instant, que de bonnes capacités de mémorisation associative, avec une prise en compte de contextes structurels et événementiels, avec une intégration de quelques aspects temporels, et avec un soupçon de stratégie de résolution.

Cependant, les résultats obtenus jusqu'à présent, laissent augurer, sous réserve de modélisation de compétence cognitives plus abstraites, des comportements globaux à la hauteur de nos espérances.

## 1.3 Critique de la démarche

Aboutissement de millions d'années d'évolution sélective sur des milliers d'espèces en parallèle, le cerveau humain peut être vu comme la convergence, temporaire mais réelle, d'un système adaptatif complexe. On peut raisonnablement postuler que le cerveau représente, dans la nature, un optimum pour le traitement de l'information sensorielle et du raisonnement flou; ce qui justifierait notre démarche. C'est cependant faire abstraction

d'un paramètre crucial, la nature ne dispose que de matériaux biologiques, avec des contraintes très fortes, pour construire son système. Elle a sans doute trouvé des solutions optimales, mais dans un cadre limité. Il est clair que l'homme est incapable de reconstituer des mécanismes subtils, comme la photosynthèse, suprême astuce pour capter l'énergie solaire. Il retourne cependant le problème en sortant du cadre étroit de la vie cellulaire et en utilisant des matériaux et des techniques artificiels.

D'autre part, non seulement nous ne construisons pas des réseaux à base de cellules vivantes, mais qui plus est, nous les simulons. L'implantation soft d'un processus permet de se libérer des contraintes imposées par le hard (thermiques, électriques, spatiales...) et offre de plus des possibilités de contrôles étendues. Ainsi, la connectivité locale et la propagation de proche associée alourdissent et ralentissent inutilement le réseau simulé.

Quelle direction prendre? Doit-on, pour élaborer une théorie du traitement de l'information par réseau, tenir des limites techniques imposées par le substrat matériel? Quelle que soit la réponse, c'est s'éloigner implicitement du modèle humain.

"Ce n'est pas en regardant les oiseaux voler que l'homme a appris à construire des avions", cette assertion résume à la fois ces objections, et en montre par sa relative valeur, leur subjectivité. La théorie de la colonne corticale se situe au niveau des prototypes à ailes battantes et autres chauve-souris à vapeur, et constitue sans nul doute une étape préliminaire fondamentale et nécessaire.

#### 1.4 Critiques psychologiques

Les psychologues modélisent les plus haut niveaux d'abstraction du cerveau, et, c'est vrai, nous sommes plus proche d'une inspiration neurobiologique. Le comportement humain semble s'organiser à partir des contraintes matérielles du corps, perçus et représentés par le cortex limbique. Le lecteur sagace, aura remarqué que, parmi les caractéristiques du réseau des aires, nous n'avons, ni étudié la symétrie Gauche/droite, ni abordé les aires limbiques. Nous en sommes convaincus, les capacités cognitives dépendent du nombre d'aires impliquées dans le réseau et tant que les réseaux ne dépasseront pas une certaine taille, les comportements émergents resteront inexploitable par le psychologue.

L'apport de la psychologie a été jusqu'à présent, sans rétroaction, notre réseau essaie de coller aux données expérimentales et ne peut fournir aucune validation de modèles de la pensée. Cependant, il est clair que pour aller de l'avant, il faudra, d'une part, mieux modéliser le fonctionnement de l'unité en interférence avec la psychologie cognitive, et d'autre part, simuler la maturation des différentes aires (psychologie génétique).

Ce n'est pas parce qu'un transistor est capable de faire un "et" logique, que le "et" entre deux conditions d'un programme est directement explicable par ce bas niveau matériel. Pourtant, les connexionnistes, et plus encore les psychologues, n'hésitent pas à attribuer localement aux neurones formels, des comportements complexes, qui révèlent plutôt de toute une chaîne de processus globaux. Quand nous introduisons les notions de contexte, sous-problème ou requêtes, ne faisons-nous pas la même erreur d'interprétation?

## 2 Intérêt

### 2.1 Comparaison aux méthodes classiques

Après cette auto-critique justifiée et nécessaire, voyons comment notre système se démarque des voies de recherches de L'intelligence artificielle classique, comme, d'ailleurs, la majorité des systèmes connexionnistes.

#### 2.1.1 Parallélisme.

La base de la théorie repose sur le fonctionnement en parallèle de toutes les unités du réseau, ce qui amène une difficulté supplémentaire pour sa conception. Comme les simulations sur ordinateurs séquentiels ne peuvent pas rendre compte de la rapidité, principal avantage du parallélisme, pourquoi persévérer dans une telle voie? Si on observe l'évolution des supercalculateurs, il est clair qu'elle s'oriente de plus en plus vers la conception de machines massivement parallèles, qui possèdent des puissances de calcul incomparables avec les machines séquentielles. Outre la technologie du silicium, les balbutiements prometteurs des systèmes optiques, éminemment parallèles ne peuvent que nous conforter dans cette voie.

L'utilisation des techniques parallèles, semble être, à moyen terme, l'unique façon de décupler les capacités.

### 2.1.2 Connaissances distribuée

. Un peu comme un hologramme, chaque élément d'un réseau connexionniste détient des parcelles de connaissance sur toute l'information. Il y a deux avantages à cela. Avec une implantation sur un ordinateur parallèle, le système est moins tributaire de défaillances locales sur des éléments, dont le travail est alors supporté par le reste du réseau. On peut, bien sûr, imaginer une implantation parallèle de connaissances non distribuées, pour des systèmes classiques, mais elle ne présentera pas cette robustesse aux pannes. Mais surtout, une représentation des connaissances distribuée, implique leur morcellement en connaissances atomiques, la prise en compte du contexte, pour tous les niveaux de connaissance entre eux, est alors implicite.

### 2.1.3 Flou

. Les données sont toujours modélisées, quel que soit le système, par des vecteurs à N dimensions (dimensions spatiales, temporelles, mais aussi plus spécifiques: température, vitesse...). Pour interpréter ces données, il faut pouvoir les différencier, classiquement, en déterminant des intervalles de variations significatifs pour chacune des dimensions. On se rend rapidement compte que cette méthode ne permet d'isoler que des hypercubes de l'espace de codage, perpendiculaires aux axes, et qu'il faut modéliser les nuages des données par des hyper-surfaces plus "souples" que les plans. Malheureusement, même dans un cas idéal de modélisation, le système peut échouer dans sa différenciation, car les nuages peuvent s'interpénétrer. Dans les systèmes symboliques sans apprentissage, le choix de l'espace de représentation et des valeurs des paramètres qui décrivent la surface, demandent une mise au point délicate et minutieuse.

Les techniques connexionnistes s'affranchissent allègrement de cette représentation explicite des connaissances floues. D'un part, certains types de réseaux déterminent eux-même les dimensions de représentation des données. Ils sont en cela très proches des techniques d'analyse vectorielles, et même, pour certaines valeurs de paramètres, n'en sont qu'une autre formulation. D'autre part, l'hyper-surface significative d'un item, est elle aussi déterminée par apprentissage. Notre approche est légèrement différente: les dimensions de représentation sont choisies a priori, par référence au sys-

tème biologique (et à notre connaissance du domaine...), et les nuages de points, non plus délimités par des hyper-plans, mais modélisés par la fusion d'hypercubes de tailles diverses, qui peuvent, par apprentissage, représenter une forme quelconque. D'autre part, les nuages peuvent s'interpénétrer, c'est alors le contexte qui détermine l'appartenance d'un point à un nuage.

### 2.1.4 Signaux dégradés

. Après la chaîne d'acquisition et de prétraitement, les données sont rarement intactes; le bruit ambiant peut générer des données non significatives, alors que l'acquisition peut faire disparaître des indices pertinents.

Notre approche du problème est soutenue par trois mécanismes: D'une part, le parallélisme implicite amoindrit fortement l'influence du bruit; tant qu'il ne représente pas une part importante du signal (paragraphe III.1.5), les performances associatives n'en sont pas affectées. D'autre part, le codage de l'information dans le réseau fait intervenir l'état E1, pour un signal mal perçu. Le mécanisme de feedback, décrit paragraphe III.1.4, permet, si besoin est, d'en modifier la perception. De plus, la prise en compte du contexte temporel, permet une certaine anticipation et une sensibilisation du réseau pour certains événements futurs qui n'en seront que mieux perçus.

## 2.2 Comparaisons connexionnistes

Toutes ces caractéristiques, intuitivement souhaitables pour des systèmes cognitifs, sont communes à la plupart des modèles de réseaux de neurones. Quelles caractéristiques démarquent la colonne corticale de tous les courants du connexionnisme confondus? L'idée essentielle étant de s'attacher à des tâches cognitives autres que la mémorisation associative, ces propriétés reparaissent à la fois les problèmes du connexionnisme classique (paragraphe II.2.6) et les principes de construction d'un automate du paragraphe II.4.6.

### 2.2.1 Connectivité locale

Pragmatiquement, il faut souligner l'importance d'une connectivité locale, indispensable pour une implantation hard, électronique ou optique, des automates de base. De plus, une connectivité réduite permet de réduire

l'explosion combinatoire qui limite les capacités quantitatives et sans doute qualitatives, des réseaux complètement interconnectés.

### 2.2.2 Intégration du temps

. De nombreux schèmes possèdent une dimension temporelle indéniable, liée soit à la nature des entrées/sorties (parole...), soit à la finalité des actions (génération de plan...). La partie associative du réseau, progressivement développée dans les applications, permet de représenter une succession d'événements par une chaîne d'unités temporellement reliées. Parallèlement, le cortex frontal est théoriquement dédié à l'organisation des stratégies de résolution à différentes échelles de temps, séquencées ou parallèles.

### 2.2.3 Unité symbolique

. Pour accéder à des stades cognitifs supérieurs, il semble naturel de modéliser et de simuler un niveau d'organisation interne du cortex plus élevé que le neurone, mais sans aller jusqu'à modéliser le comportement global de toute une aire, ce qui est un peu l'approche des systèmes multi-agents. La colonne corticale, outre sa plausibilité biologique, offre alors un support concret au développement d'une unité "symbolico-connexionniste". La définition de son rôle *fonctionnel* permet de se rattacher aux théories informatiques et psychologiques de l'intelligence, artificielle et naturelle, pour en élaborer les mécanismes internes.

### 2.2.4 Réseau des aires

. Il est clair que, si, comme nous l'annonçons, la puissance cognitive du réseau réside dans le nombre d'aires impliquées, il est indispensable de proposer un modèle d'organisation des aires entre elles. D'autre part, le développement de la structure des aires est indissociable de la modélisation de l'automate de base qui module alors ses entrées suivant l'interprétation de leur provenance. La modélisation du cortex par 6 symétries de base, permet d'éviter une hiérarchisation statique des fonctions des aires, qui privilégie obligatoirement un type d'analyse, top-down ou bottom-up.

## 3 Réflexions

### 3.1 Théorie

#### 3.1.1 Défense de la pratique

. Une théorie n'a d'intérêt que si elle supporte la critique raisonnée. Nous ne travaillons pas, pour l'instant, dans un système formel, mais nous avons toujours recherché, au travers des applications, une validation expérimentale. C'est pourquoi l'intérêt théorique de nos travaux reste largement tributaire des résultats pratiques obtenus. D'autre part, une construction théorique, même complète, ne doit pas rester figée mais au contraire évoluer pour intégrer des données nouvelles. L'expérimentation pratique, en éludant certaines voies et en guidant l'intuition, peut représenter un moteur puissant de l'évolution.

On peut se poser la question de l'utilité d'un formalisme théorique plaqué sur des idées intuitives, et qui peut n'en être qu'une reformulation verbale et lourde. Cependant, même si l'homme est relativement bien adapté au raisonnement flou, nous sommes très vite dépassé pour traiter des problèmes de logique mathématique. Si un formalisme quelconque permet de faire des prévisions sur le comportement de l'unité, de générer de nouvelles hypothèses ou de démontrer ses capacités, alors il est utile; il reste cependant difficile de juger, a priori, des compétences, et donc de l'opportunité, d'un cadre formel.

#### 3.1.2 Unification

. Faut-il essayer d'unifier la théorie pour avoir une vue plus globale, ou au contraire la morceler pour en avoir une vue plus fine?

### 3.2 Modélisation

De façon générale, les sciences expérimentales essaient de rendre cohérent un ensemble de connaissances sur un domaine, en proposant des modèles des phénomènes extérieurs (physiques, économiques, psychologiques...). L'adéquation entre le modèle et la réalité est alors l'objectif premier.

### 3.2.1 Méthodologie

La modélisation des processus physiques constitue un domaine des sciences à part entière. Les automaticiens ont développé de nombreuses techniques pour mettre en équation des systèmes complexes. Quand les mécanismes physiques du phénomène étudié sont parfaitement compris et explicites, il est possible d'établir des expressions mathématiques qui le décrivent complètement. Ce cas théorique est malheureusement bien rare. De la boîte blanche, où tous les processus sont connus, à la boîte noire, la méthodologie passe de la *modélisation* théorique à l'*identification* empirique. L'identification consiste alors à rechercher une fonction qui colle aux signaux d'entrée/sortie mesurés expérimentalement. Cependant, les données expérimentales ne sont pas toujours disponibles, et c'est plus par des approches mixtes, mêlant connaissances a priori sur les système et mesures expérimentales, que se tournent actuellement les automaticiens.

Pourrait-on entreprendre une campagne de mesures systématique sur la colonne corticale biologique? Certaines publications font déjà état de quelques observations dynamiques; cependant, les conditions expérimentales difficiles et incontournables (cellule vivante et microscopique), les effets à long terme de la mémorisation qui influent sur les sorties, et la présence d'entrées hormonales qui jouent sur le comportement dynamique, rendent sans doute cette voie impraticable...

Au contraire, l'approche "boîte blanche", basée sur la modélisation des 110 neurones de la colonne, avec leur fonctionnement, leur réseau interne de connexions et leurs mécanismes de mémorisation, semble plus facilement et rapidement applicable...

Notre approche reste encore bien intuitive, et prend ses racines dans une description floue, non pas des micro-phénomènes caractéristiques d'un automate, mais plutôt des comportements globaux attendus lors du fonctionnement en parallèle de nombre d'entre eux. Faut-il, là encore, rationaliser la démarche intellectuelle à tout prix? Quoi qu'il en soit, on peut au moins s'inspirer de deux idées clé de la modélisation: l'économie des paramètres et la construction progressive du modèle.

### 3.2.2 Parcimonie

Un modèle est décrit, entre autre, à l'aide de constantes et de paramètres, dont les valeurs peuvent être estimées à partir de données. En pratique, il est fondamental d'utiliser le moins de paramètres possible pour rendre compte d'un processus et d'autre part, la description doit rester pertinente. La modélisation est très sensible à cet équilibre entre le nombre de paramètres nécessaire et le nombre suffisant. Pour la colonne, le débat peut se rapporter au nombre de types d'entrées/sorties, qui, au fur et à mesure des applications est passé de quatre à neuf, ou bien encore à la modélisation des messages entre colonnes, progressivement enrichie, à partir des trois états, d'un éloignement puis d'un type de symétrie.

## 4 Perspectives

De toute évidence, la théorie corticale présentée est en cours de développement, et il faudra encore quelques années pour qu'elle atteigne une certaine maturité. Tout au long des applications, nous avons dégagé des voies de recherche futures qui nous semblent prometteuses; il reste à établir un programme de recherches à plus ou moins long terme.

### 4.1 Court terme

Considérant la dernière application, il est clair que les aires frontales demandent à être développées, pour intégrer dans le système des connaissances de type stratégique qui guident l'analyse, et de type grammaticale pour accéder à un niveau symbolique intéressant. Les efforts de modélisation doivent alors porter sur le réseau des aires qui conditionne le langage, et sur le fonctionnement de la colonne corticale, plus particulièrement sur les arbres d'appels, propagation d'un déséquilibre dans un réseau modelé par l'apprentissage. Pour chacune des disciplines cognitives, se dégagent des voies de recherche:

- D'un point de vue informatique, nous avons déjà mis en valeur l'approche probabiliste, par les coefficients d'apprentissage, et logique, par le mécanisme de fonctionnement. Intuitivement, il existe un lien

fort entre l'inférence et la probabilité conditionnelle. Dans un système lié, il y a forcément des connaissances a priori implicites; un cadre formel peut les mettre en évidence et permettre la déduction de nouveaux mécanismes.

- Les aires frontales, comme toutes les aires, sont caractérisées par leurs entrées externes, en l'occurrence, diverses structures qui font office d'horloges internes. Il semble que les différentes échelles temporelles qui existent dans le cortex, correspondent à des niveaux d'organisation temporelle du comportement.

Les couches hautes de la colonne recombinent les informations en provenance d'autres aires, la hiérarchisation des afférences est un paramètre déterminant pour la logique du fonctionnement.

Enfin, la neurobiologie décrit les variations cytoarchitecturales comme caractéristiques des différentes aires, n'existe-t-il pas d'autres variations qui, intégrées dans le modèle, rendraient compte des différentes interprétations des logiques frontale ou associative?

- A ce niveau de modélisation, les propriétés émergentes du réseau commencent à être intégrables dans des schémas de comportement psychologiques; réciproquement, à partir des théories génétiques, il est possible de mieux comprendre la maturation des aires du réseau. Nous avons déjà souligné l'importance d'un protocole qui, postulant une maturation progressive et successive des aires, optimise les sessions d'apprentissage et permet une construction progressive des stades cognitifs.

## 4.2 Moyen terme

L'objectif à moyen terme est d'élaborer un système complet, à six symétries, capable de simuler une grande partie des fonctions cognitives humaines. L'implantation d'un tel réseau, suppose des moyens matériels étendus; pour simuler 64 aires de taille raisonnable (quelques milliers de colonnes par aire), l'estimation de la place mémoire nécessaire varie entre 50 et 100 Mégo-octets, ce qui dépasse les capacités des ordinateurs classiques. L'utilisation d'une machine parallèle à base de transputers, chacun

de capacité suffisante pour simuler une aire, est indissociable des performances cognitives du réseau. D'autre part, les processus mis en jeu dans un tel réseau doivent être impérativement stables et convergents, seule la mise en place d'un cadre formel permettra d'établir l'équilibre des structures cognitives. A priori, la recherche de ce cadre devrait s'orienter vers la définition d'une logique trois états, probabiliste et temporelle...

En supposant maintenant l'existence d'un tel système, les perspectives d'application peuvent intéresser les trois domaines de référence:

- L'intelligence Artificielle classique différencie les méthodes de résolution des étapes de traitement de l'information, et propose différents algorithmes pour la reconnaissance de formes, l'interprétation et le raisonnement ou la génération de plans. Mettre bout à bout des systèmes différents, l'expression est en elle-même suffisamment éloquente, pose des problèmes pratiques mais aussi théoriques, qui nuisent à la circulation de l'information entre les unités de traitement, pourtant synonymes de performances remarquables, et limite les types de connaissances représentées. Ce réseau proposerait une unification possible des différentes tâches cognitives, habituellement dissociées, en définissant un atome de connaissance élémentaire suffisant pour décomposer, par apprentissage et de façon distribuée, tous les comportements qu'on peut en attendre.
- Que peuvent attendre les neurobiologistes d'un tel réseau? Il est désormais clair que les données neurobiologiques sont insuffisantes pour modéliser un automate effectif; les besoins du formalisme vont-ils nous éloigner de la réalité biologique? Les principes d'élaboration d'un automate, ajoutés aux fondements biologiques maintenant acquis, vont permettre de construire une théorie complète dont les assertions seront extrêmement intéressantes à comparer avec des données expérimentales très orientées, soit pour confirmer nos hypothèses, soit pour réorienter nos recherches.
- De possibilités quantitatives limitées, le réseau montrera un comportement complexe. Le langage semble être la clé de l'organisation de la pensée symbolique et raisonnée. Il est indiscutable, du moins pour nous, qu'un système dit intelligent, doit maîtriser les aspects majeurs de son organisation. On peut aussi penser que des conceptualisations

du temps, de la motivation ou de la symétrie, vont s'effectuer de façon distribuée dans le réseau. En montrant de tels comportements, ce dernier pourrait devenir un outil d'expérimentation et d'exploration du développement de l'intelligence...

### 4.3 Long terme

Il est très difficile de prévoir ce que pourront devenir, à long terme, de telles recherches. Les possibilités offertes par les supports matériels vont certainement conditionner les développements futurs; nous débutons, parallèlement aux travaux de modélisation, des recherches visant une implantation optique massive. Cette technique présente quelques avantages potentiellement intéressants pour des réseaux très grands. La lumière est un phénomène éminemment parallèle; de plus, si l'électronique ne joue guère que sur la tension des signaux, on peut utiliser l'intensité lumineuse, sa phase et sa polarisation. Enfin, les systèmes optiques sont implantés en trois dimensions, ce qui suggère une compacité impressionnante.

L'avenir va aussi dépendre des réelles possibilités d'une modélisation peut-être trop réductionniste, et en pratique inefficace. Comment juger des capacités développées par un futur modèle, autrement qu'en maîtrisant tous les aspects présents, et en laissant notre esprit concevoir des projections intuitives? Cette méthode, qui nous fait pencher d'un extrême à l'autre, est trop suggestive pour être reconnue.

Objectivement (!), la convergence en Intelligence Artificielle de nombreuses techniques vers des systèmes multi-agents organisés en réseaux de flux d'information, autorise des espoirs certains pour le développement de notre théorie corticale; maintenant pourra-t-on un jour comprendre et rendre compte vraiment de la complexité fascinante et mystérieuse du cerveau, le débat reste ouvert....

## REFERENCES

- 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.
- Andersen P. (1983), "Operational principles of hippocampal neurons; a summary of synaptic physiology", Neurobiology of hippocampus, W. Seifert ed., Academic Press.
- Appel R., Funk M., Dochstrasser M., Muller A.F., Pellegrini C., (1985), "Un system expert d'aide à l'analyse et à l'interpretation d'images", Proc. Cognitiva 85, Paris.
- Azencott, R. (1989), "Boltzmann machines: High-order interactions and synchronous learning" IEEE PAMI.
- Baldy R.E., Brindley G.S., Ewusi S., Mensah I., Jacobson R.R., Reveley M.A., Turner S.W., Lishman W.A.,(1986), "A fully-automated computer assisted method of CT brain scan analysis for the measurement of cerebrospinal fluid spaces and brain absorption density", Neuroradiology, 28, 109-117.
- 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.
- Bengio Y., Demori R. (1988), "Use of Neural Networks for the Recognition of Place of Articulation", Proc. ICASSP-88, 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.
- Bienestock, E. (1986), "Connectionist models of perception: a short review", Fyssen Foundation Conference in Versailles March 86.
- Bienestock, E., Dreyfus G., Personnaz L. (1989), "What's wrong with neural network?", Proc NeuroNimes 89, Nimes.
- Bourlard, H., Wellekens, C. (1987), "Multi-Layer Perceptrons and Automatic Speech Recognition", Proc. IEEE First Int. Conf. on Neural Networks, San Diego.
- Broca P., (1861), "Perte de la parole, Ramollissement chronique et destruction partielle du lobe antérieur gauche du cerveau", Bull. Soc. Anth., 2, 235-237.
- Brodman, 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.
- Changeux, J. P. (1983), "L'homme neuronal", Fayard.
- Chassery J.M., Garbay C. (1986), "Expert systems, image processing and image interpretation", Proc of 8th ICPR, 175-177.
- 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 Variable D'instance.
- Dingeon, C., Alexandre, F., Guyot, F., Haton, J. P. (1989), "Un autre apprentissage cortical: Différentier pour généraliser", Proc. NeuroNimes'89, Nimes.

- Ehrenfels C.V. (1890), "Ueber Gestaltqualitäten". *Viert. f. wiss. Phil.* 249-292.
- Elman J. L., Zipser D., (1988), "Learning the Hidden Structure of Speech", *JASA*, 83, n 4, p. 1615-1626.
- Evarts, E. V., Tanji, J. (1974), "Gating of motor cortex reflexes by prior instruction", *Brain Research*, 71, p. 479-494.
- Fogelman Soulié, F. (1987), "Le connexionnisme", Support de cours, MA-RI 87, *Cognitiva* 87.
- 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.
- 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.
- Gori, M., Bengio, Y. DeMori, R. (1989), "BPS: A learning algorithm for capturing the dynamic nature of speech", Proc. IEEE IJCNN'89, Washington.
- Garbay C., (1985), "Segmentation d'images: connaissances et stratégies d'analyse", Proc of 5ème congrès reconnaissance des formes et intelligence artificielle, 615-629.
- Geshwind, N. (1984) "Les spécialisations du cerveau humain", Le Cerveau, Belin.
- Goldstein M.H., Benson D.A., Heinz R.D., (1982), "Studies of auditory cortex in behaviorally trained monkeys", *Conditioning: representation of involved neural functions*, C.D. Woody ed., Plenum.
- Gross C.G. (1973), "Visual functions of inferotemporal cortex", *Handbook of sensory physiology*, Springer-Verlag.
- Guillaume P. (1979), *La psychologie de la forme*. Champs. Flammarion

- ed.
- 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.
- Hebb, D. O. (1949), "The organization of behaviour", Wiley, New-York.
- Heccaen H., Lanteri-Laura G., (1983), Les fonctions du cerveau. Masson.
- Hinton, G., Sejnowski, T., Ackley, D. H. (1984), "Boltzmann machines: constraint satisfaction networks that learn technical report", Carnegie Mellon University.
- 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.
- Howard I.S., Huckvale M.A. (1988), "Training Feature Detectors for Use in Automatic Speech Recognition", FASE Speech '88, Edimburgh.
- 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.
- Jacob, F. (1981), "Le jeu des possibles", Fayard.
- Jeannerod M. (1974), "Les deux mécanismes de la vision", La recherche, 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.
- Jones, E.G., (1981b), "Anatomy of cerebral cortex: columnar input/output organization", the organization of the cerebral cortex, F. Schmitt et al. eds., Cambridge, Mass., MIT Press, 199-235.
- Jung R., (1973), "Visual perception and neurophysiology", Handbook of sensory physiology, Springer-Verlag.
- 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.
- Kohler W. (1929), "Gestalt psychology", New York (Liveright), 1-403.
- 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.

- Leboeuf J., (1988), "Un Systeme Connectionniste Appliqué au Traitement Automatique de la Parole", Thèse, Université de Paris-Sud.
- LeCun, Y. (1987), "Modèles connexionnistes de l'apprentissage", Thèse de Doctorat, Université Paris VI.
- Levinson, S.E, Rabiner R.L., Sondhi M.M. (1983), "An Introduction to the Application of the Theory of Probabilistic Functions of a Markov Process to Automatic Speech Recognition", Bell Syst. Tech. J., p.1035-1073.
- Lei G., Chengke W., Xinru L., (1986), "a new method for computer recognition of small rounded pneumoconiosis opacities in chest X-rays", Proc. of 8th ICPR, 475-477.
- Lieberman, P., (1973), "On the evolution of language: a new unified view. Cognition, 2, 59-94.
- Lippmann R.P. (1987), "An Introduction to Computing with Neural Nets", IEEE ASSP Magazine, 4, n 2, p. 4-22.
- 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.
- MacCulloch, W. S., Pitts, W., (1943), "A logical calculus of the ideas immanent in nervous activity", Bulletin of Mathematical Biophysics, 5, p. 115-133.
- Mac Culloch N. A., Ainsworth W. A. (1988), "Speaker Independent Vowel Recognition Using a Mutli-Layer Perceptron", FASE Speech '88, Edimburgh.
- MacKay D.M. (1973), "Visual stability and voluntary eye movements", Handbook of sensory physiology, Springer-Verlag.
- MacLean P.D., Yokota T., Kinnard M.A. (1968), "Photically sustained on-responses of units in posterior hippocampal gyrus of awake monkey", J. Neurophysiol., 31, 870, 883.
- Magnin I.E., Cluzeau F., Odet C.L., Bremond A., (1986), "Mammographic texture analysis: an evaluation of risk for developing breast cancer", Optical Engineering, 25, 6, 780-784.

- Matsuyama T., (1989), "Expert systems for image processing: knowledge-based composition of image analysis processes", CVGIP, 48, 22-49.
- 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.
- Nagao N., (1985), "Strategies for human-like image understanding systems", Proc. Cognitiva 85, Paris.
- Nauta, W., Feirtag, M. (1984), "L'organisation du cerveau", Le Cerveau, Belin.
- Nilsson, N. J. (1986), "Probabilistic logic", Artificial Intelligence, 28, p. 71-87.
- Otto I. (1990), Proposition à INNC 90, Paris.
- Piaget, J. (1956), "Les stades de développement intellectuel de l'enfant et de l'adolescent", le problème des stades en psychologie de l'enfant, PUF, Paris.
- Piaget, J. (1975), "L'équilibration des structures cognitives, problème cen-

- tral du développement", PUF, Paris.
- Peeling S.M., Moore R.K., Tomlinson M.J. (1986), "The Multi-Layer Perceptron as a Tool for Speech Pattern Processing Research", Proc. IOA Autumn Conf. on Speech and Hearing.
- Robert, J. M. (1982), "Comprendre notre cerveau", Editions du Seuil, Paris.
- Rosenblatt, F. (1962), "A comparison of several perceptron models", in *Self Organizing Systems 1962*, ed. Yovits, Jacobi, Goldstein, Spatan Books, p. 463-484.
- Rosenberg S., Itti R., Benjelloun L., (1986), "Symbolic reasoning about myocardial scintigrams", *Eur. J. Nucl. Med.* 12, 65-68.
- Rumelhart, D. E., MacClelland, J. L. (1986), "Parallel distributed processing", MIT Press, Cambridge.
- Sakoe H., Chiba S., (1971), "A Dynamic Programming Approach to Continuous Speech Recognition", Proc. 7<sup>th</sup> Int. Cong. on Acoustics, Paper 20 C 13.
- 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.
- Sprague J.M., Berlucchi G., Rizzolatti G. (1973), "The role of the superior colliculus and pretectum in vision and visually guided behavior", *Handbook of sensory physiology*, Springer-Verlag.
- Stein, D. (1989), "Les verres de spin", *Pour la science*, 143, p. 30-37.

- Stevens, C. (1984), "Le neurone", Le Cerveau, Belin.
- Székely G. (1973), "Anatomy and synaptology of the optic tectum", *Handbook of sensory physiology*, Springer-Verlag.
- Szentagothai, J. (1970), "Glomerular synapses, complex synaptic arrangements, and their operational significance. Neurosciences, Second study program F.OBJET.Schmitt, ed., New York, Rockefeller, 427-443.
- Szentagothai, J. (1973), "Synaptology of the visual cortex", *Handbook of sensory physiology*, Springer-Verlag.
- Szentagothai, J. (1973b), "Neuronal and synaptic architecture of the lateral geniculate nucleus", *Handbook of sensory physiology*, Springer-Verlag.
- Szentagothai, J. (1975), "The "module concept" in cerebral cortex architecture", *Brain Research*, 95, p. 475-496.
- Tsotsos J., (1985), "Computer assessment of left ventricular wall motion: The ALVEN expert system", *Computer and biomedical research*, 18, 254-277.
- 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.
- Wackenheim A., (1984), "Analyse structuraliste de l'image radiologique", *Radiologie, J. CEPUR*, 284-290.
- 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.
- Waibel A. et al., (1988), "Phoneme Recognition Using Time-Delay Neural Networks", *Proc. ICASSP-88, New-York*.
- Waibel A., Sawai H., Shikano K. (1989), "Consonant recognition by modular construction of large phonemic Time Delay Neural Networks", *Proc. ICASSP-89, Glasgow*.

- Wernicke C., (1874), *Der aphasische Symptomencomplex*, Breslau, M. Cohn et Weigert.
- Wertheimer M., (1922), "Untersuchungen zur Lehre von der Gestalt", *Ps. Forsch.* 1, 47-58.
- Wertheimer M., (1923), "Untersuchungen zur Lehre von der Gestalt", *Ps. Forsch.* 4, 301-350.
- Whitteridge D., (1973), "Projection of optic pathways to the visual cortex", *Handbook of sensory physiology*, Springer-Verlag.
- Widrow, B., Hoff, M. E. (1960), "Adaptive switching circuits", 1960 IRE WESCON Convention Record, New-York, p. 96-104.
- 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.

## ANNEXES

## 1 Résultats associatifs

Le corpus est constitué d'une police de 26 caractères majuscules codés binairement dans une matrice 8x8. Pour l'apprentissage sur exemples bruités, 6, 9 et 12 pixels ont été inversés aléatoirement sur chacun des trois exemples de chaque lettre.

La reconnaissance a lieu sur les caractères originaux de plus en plus bruités. La figure 1 représente les taux de reconnaissance en fonction du bruit, pour les deux apprentissages réalisés.

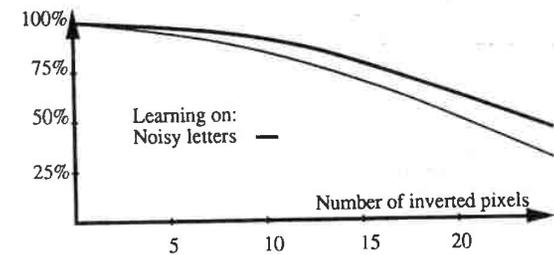


Figure 1: taux de reconnaissance

La figure 2 illustre quelques exemples de confusion pour une présence de bruit non négligeable, et particulièrement gênante.

			
Caractere	F	T	H
Reponse	E	J	N
Bruit	5	9	8
			
Caractere	B	L	I
Reponse	D	U	S
Bruit	8	7	9

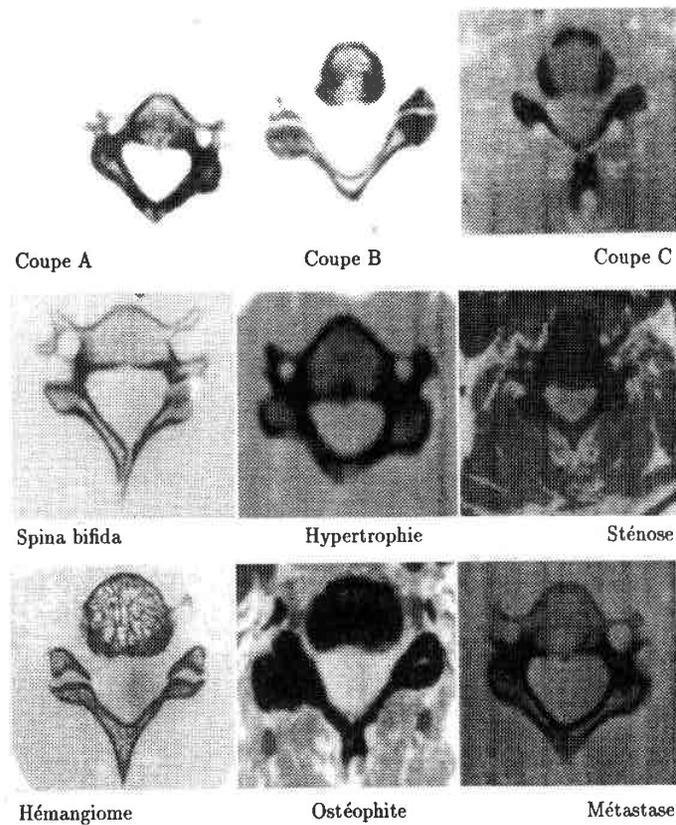
Figure 2: exemples de confusion

Enfin, le réseau a été testé sur de l'écriture manuscrite, saisie à la souris, et quelques exemples de bonnes reconnaissances sont représentés figure 3.

B O M E

Figure 3: Ecriture manuscrite

2 Interprétation de clichés radiologiques



### 3 Informatique

Malgré le caractère tridisciplinaire de nos travaux, nous soutenons avant tout une thèse d'informatique; c'est pourquoi, pour les inconditionnels de la programmation, nous avons rassemblé dans cette annexe quelques données sur l'implantation des diverses applications.

Environnement matériel: Les trois premières simulations ont été réalisées sur un SUN 3 de 8Mo de mémoire, station de travail maintenant courante et qui fournit un environnement de travail très convivial pour développer des logiciels (multi-fenêtrage, souris, bitmap...). Plus récemment, nous avons fait l'acquisition d'un SUN 4, SparcStation, bientôt up-gradé à 16Mo, écran couleur, qui a fourni des temps d'apprentissage de l'ordre de quelques minutes, et des temps de reconnaissance de l'ordre de la seconde pour la dernière application, ce qui est tout à fait acceptable. Par contre, les 8Mo actuels sont limités pour supporter nos logiciels pas vraiment optimisés et gourmands en place mémoire.

Nous avons déjà fait référence à une machine parallèle à base de transputers disponible au CRIN, qui, vraisemblablement constituera le support de futures applications de grande taille.

Langage: Les possibilités offertes par le langage C, tant du point de vue de la rapidité que de l'intégration de primitives de haut et bas niveaux, nous ont, au départ, séduits; toutes nos applications ont été écrites dans ce langage. On peut cependant mentionner les langages acteurs qui, même s'ils n'ont pas encore franchi les portes des laboratoires, permettront une programmation aisée des échanges de messages parallèles entre les unités ainsi que des héritages de propriétés nécessaires aux divers mécanismes d'apprentissage.

Structure des données: Les structures suivantes, extraites d'une application, rendent compte de la connectivité entre unités; on peut en donner une représentation graphique (figure 1), qui est à rapprocher de la modélisation tridimensionnelle, proposée par Y. Burnod, des neurones de la colonne (figure 2).

```

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;
};

```

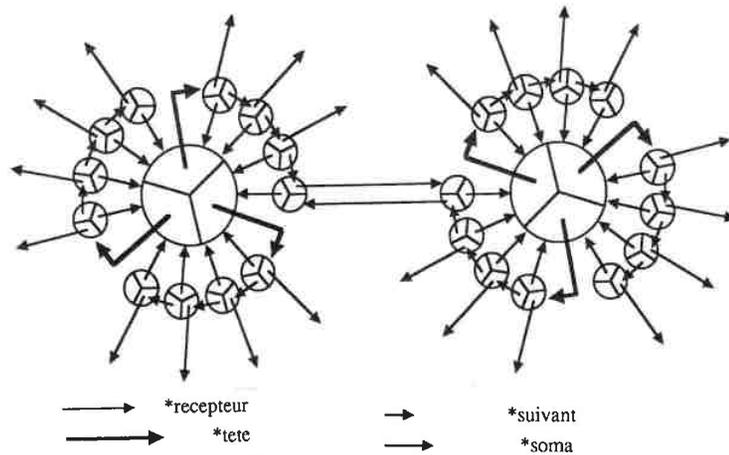


Figure 1: structure des données

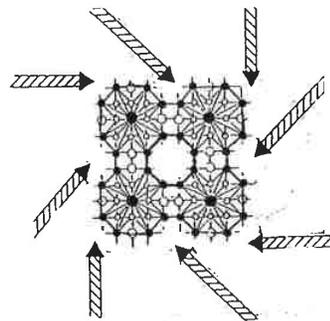


Figure 2: modèle de la colonne

## Index

- A** acquisition 3  
 aire 19, 21, 42, 50, 71  
   associative 44, 131  
   frontale 45, 131  
   motrice 42  
   réseau 45  
   sensorielle 42, 129, 131  
 analyse 144  
   ascendante 122  
   descendante 122  
 angle 87  
 apprentissage 5, 30  
   coopération 144  
   hebbien 29  
   probabiliste 89  
   protocole 107, 138  
   transfert 145  
 arbre d'appel 120  
 attention sélective. 89  
 audition 3  
 auto-association 30, 34
- B** bruit 93  
 but 71, 142  
 ordonnancement 120
- C** carte  
   de codage 62  
   de codage 100  
   topologique 38  
 cervelet 118  
 champ récepteur 87, 123, 130  
 coarticulation 92, 99  
 cochlée 94
- codage  
   initial 101, 104, 105  
 colliculus 124, 130  
 colonne 21, 52  
   sous couches 56  
 communication homme machine 121  
 condition  
   inhibitrice 89  
   nécessaire 90  
   sensibilisatrice 89  
   suffisante 89  
 connectivité 105  
   hypercubique 134  
 contour 87  
 corps genouillé latéral 87, 117  
 cortex 17  
 couche haute  
   hiérarchisation 135, 159  
 couches 30
- D** décomposition fréquentielle 94  
 diagnostic 3  
 différenciation 71, 141  
 division 107, 111
- E** écriture  
   manuscrite 90  
 évaluation  
   qualitative 150  
   quantitative 149  
 évènement 3  
 extraction 2

- F** fait 3  
 feed-back 89, 132, 154  
 feed-forward 89, 132  
 FFT 102  
 flou 93, 153  
 fonctionnalité 2  
 fonctionnement  
   logique de 89  
 formant 97, 102  
 forme  
   faible 125  
   forte 125  
   loi de la bonne 126  
   structure 142  
   tridimensionnelle 145  
 fovea 123  
 fricatives 99  
 fusion 111
- G** généralisation 71, 141  
 gestalt 125  
 grand-mère 88, 89, 101, 131
- H** hétéro-association 30, 86
- I** indice  
   primaire 88  
   secondaire 88, 103, 104  
 Intelligence Artificielle 1  
 invariant extérieur 71, 89
- L** langage 121, 128  
 latéralisation 146  
 lexique 91  
 logique 67  
   bases 146
- M** macroscopique 7, 8  
 masque 68, 102, 123, 129  
 mémorisation  
   associative 86, 92, 100  
 memorisation associative 2  
 message 146  
 microscopique 7, 8  
 modèle 2, 4, 7, 8
- N** neurone 5, 14, 20, 22, 25, 56, 64  
 formel 26  
 nouveauté 94
- O** orientation 53
- P** parallélisme 86, 152  
 pariétal 124  
 parole  
   production 88  
 phonème 91, 96  
 plosive 98  
 pluridisciplinarité 5  
 point de fixation 127  
 prétraitement 2  
 propagation 69, 71, 106  
 psychologie  
   branches 8  
   subjectivité 8
- R** raisonnement 3, 128  
 recombinaison privilégiée 63, 115, 118  
 reconnaissance 2  
   caractère 86  
   parole 91  
   phonème 92  
 représentation

- interne 108  
 interne 141, 144, 152  
 symbolique 128  
 représentation interne 128  
 réseau  
   taille 143  
 réseau de neurones 93  
 rétinotopie 87, 118  
 rétroaction 103  
   externe 98
- S** segment 87  
 sémantique 92  
 spectrogramme 95  
 stratégie 142  
 symbolique 3  
 symbolisme  
   automate 155  
   définition 121  
 symétrie hypercubique 49  
 système expert 7, 122, 139
- T** temporal 124  
 temps 24  
   anticipation 143  
   constante 91  
   contrainte 100  
   échelle 159  
   indépendance 114  
   instabilité 116  
   intégration 155  
   parole 93  
   représentation 101, 102  
   séquençement 120  
   stabilité 115  
 texture 145  
 thalamus 16, 19, 87, 95, 117, 124
- traitement  
 étape 86  
 triangle vocalique 98
- V** variabilité 3  
 vision 2  
   application 86  
   exploration 125  
   invariance 143  
   vision périphérique 124



NOM DE L'ETUDIANT : GUYOT 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° 744

LE PRESIDENT DE L'UNIVERSITE DE NANCY I



## Résumé

Trois approches scientifiques se partagent l'étude de l'esprit humain: la Psychologie, plutôt dédiée à l'étude des comportements, la Neurobiologie, intéressée par le support concret, et l'Intelligence Artificielle, focalisée sur le traitement de l'information. Cependant, les liens entre ces *sciences cognitives* sont encore bien mal perçus. L'avènement en Intelligence Artificielle des "réseaux de neurones" a pourtant suscité de grands espoirs, mais, les limitations propres à ces méthodes, ont découragé tout effort d'interprétation de haut niveau.

Actuellement, une théorie nouvelle émerge, non plus basée sur le neurone, mais postulant la *colonne corticale* comme unité élémentaire de traitement de l'information. En pratique, la mise au point et la validation de cette théorie, s'est effectuée au travers d'applications portant sur quelques fonctions cognitives importantes, la vision, l'audition, la motricité et le raisonnement interprétatif.

Emaillé de nombreuses remarques et critiques, cet essai ne doit pas être lu comme exposant une théorie complète et achevée, mais plutôt comme une tentative de compréhension tridisciplinaire des mécanismes corticaux qui en pose les bases et les principes fondamentaux.

### Mots clés :

*Connexionnisme, Cortex, Colonne corticale, Modélisation, Vision, Motricité, Interprétation d'images, Reconnaissance de la parole*