

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université de Batna

Faculté des technologies
Département de Génie Electrique
Filière : Electronique

Thème

Détection de l'Onde P dans un Signal ECG

Pour l'obtention du diplôme de
Magister en électronique
Option
Micro-onde pour Télécommunication

Par
Ibtissem HOUAMED

Jury

Président	M. Boulemden	Professeur (Université de Batna)
Rapporteur	S. Meghriche	Maitre de conférences (Université de Batna)
Examineur	S.Benabdelkader	Maitre de conférences (Université de Batna)
Examineur	M. Benslama	Professeur (Université de Constantine)
Examineur	D. Benatia	Professeur (Université de Batna)
Examineur	T.Fortaki	Professeur (Université de Batna)

17/04/2011

Dédicaces

A :

Mes chers parents

Mes frères et ma sœur

Toute ma famille

L'ensemble des enseignants

Tous mes collègues et amies sans exception

je dédie ce mémoire

Remerciements

*Je voudrais remercier **Dr. Salama MEGHRICHE** Maitre de conférences à l'université de Batna, d'avoir accepté de diriger ces travaux de recherches et qui, avec beaucoup de disponibilité, a suivi ce travail. Je la remercie pour ses conseils judicieux et pour ses qualités scientifiques et pédagogiques.*

*J'adresse ma profonde reconnaissance au **Pr. M. Boulemden**, professeur à l'Université de Batna, d'avoir accepté de présider le jury de soutenance.*

Je remercie vivement les membres du jury: monsieur M. Benslama professeur de l'université de Constantine, monsieur D. Benatia professeur de l'université de Batna, mademoiselle S. Benabdelkader Maitre de conférences de l'université de Batna, monsieur T. Fourtaki professeur de l'université de Batna qui ont accepté de juger cette thèse, l'intérêt qu'ils ont porté à mon travail et la rapidité avec laquelle ils ont lu mon manuscrit.

Je ne saurais oublier mes parents et mes sœurs, pour leur soutien moral inconditionnel, pour les aides permanentes et pour leurs encouragements tout au long de mes études.

Sans oublier les personnes avec qui j'ai eu à partager les années d'études dans l'institut d'électronique à l'université de Batna et auxquels je souhaite bon courage.

Table de matières

INTRODUCTION	1
CHAPITRE 1: L'ÉLECTROCARDIOGRAMME	3
1. SYSTÈME CARDIOVASCULAIRE	3
1.1. <i>La circulation sanguine</i>	4
1.1.1. Circulation pulmonaire	4
1.1.2. La circulation systémique	4
1.2. <i>Circulation sanguine au niveau du cœur</i>	6
1.3. <i>Cellules myocardiques</i>	7
1.4. <i>Bases Electrophysiologique</i>	7
1.4.1. Le potentiel monophasique unicellulaire	7
1.4.2. Le fonctionnement électrique du coeur.....	8
2. ELECTROCARDIOGRAPHIE	9
3. ELECTROCARDIOGRAMME.....	10
3.1. <i>Définition de l'électrocardiogramme</i>	10
3.2. <i>Dérivations</i>	10
3.2.1. Dérivations dans le plan frontal	10
a) Dérivations bipolaires.....	10
b) Dérivations unipolaires.....	10
3.2.2. Dérivations dans le plan horizontal.....	11
3.2.3. Position des électrodes	12
4. ACTIVITÉ ÉLECTRIQUE D'UN BATTEMENT CARDIAQUE NORMAL	12
5. ELECTROCARDIOGRAMME NORMAL	13
5.1. <i>Ondes Intervalles et segments</i>	13
5.1.1. Onde P.....	13
5.1.2. Intervalle PR	14
5.1.3. Onde QRS	14
5.1.4. Point J.....	14
5.1.5. Segments ST	14
5.1.6. Intervalle QT	14
5.1.7. Onde T	14
5.1.8. Onde T atriale	15
5.1.9. Onde U	15
5.2. <i>Déflexion Intrinsécoïde DI</i>	15
6. ACQUISITION ET TRAITEMENT DES SIGNAUX PHYSIOLOGIQUES	15
6.1. <i>Circuit biomédical</i>	15
6.1.1. Signal électrique.....	15
6.1.2 Tissues biologiques.....	15
6.1.3. Signal biologique	16
6.2. <i>Appareillage</i>	16
6.3. <i>Système d'acquisition de données</i>	16
6.4. <i>Chaîne d'acquisition de l'ECG</i>	17
6.5. <i>Capteurs biomédicaux</i>	18
6.5.1. Problématiques des capteurs électrophysiologiques	18
6.6. <i>Le préamplificateur</i>	19
6.7. <i>Le filtre</i>	19

6.8. Amplificateur.....	20
6.9. Echantillonnage	20
6.9.1. Echantillonneur-Bloqueur.....	20
6.9.2. Conversion analogique numérique	21
6.10. Convertisseur Analogique/Numérique.....	21
7. RÉCAPITULATION.....	23
CHAPITRE 2: QUELQUES MODÈLES DE RÉSEAUX DE NEURONES	25
I. INTELLIGENCE ARTIFICIELLE	25
1. HISTORIQUE	25
2. DEFINIR L'INTELLIGENCE ARTIFICIELLE	26
2.1. Définition	26
2.2. Différents domaines	27
2.3. Réseaux de neurones.....	27
II. RESEAUX DE NEURONES	27
1. INTRODUCTION	27
2. HISTORIQUE	28
3. NEURONE FORMEL	29
3.1. Interprétation mathématique.....	30
3.2. Fonctions d'activation.....	31
4. RÉSEAUX DE NEURONES ARTIFICIELS.....	32
4.1. Définition	32
4.2. Architecture des réseaux.....	33
4.2.1. Réseaux "FEED-FORWARD"	33
4.2.1.1. Perceptrons	33
a) Perceptron monocouche	33
b) Perceptron multicouche (MLP).....	33
4.2.1.2. Les réseaux à fonction radiale	34
4.2.2. Les réseaux "FEED-BACK"	34
4.2.2.1. Les cartes auto-organisatrices de Kohonen	34
4.2.2.2. Les réseaux de Hopfield	34
4.2.2.3. Les ART.....	34
5. RECAPITULATION	35
6. APPRENTISSAGE.....	35
6.1. Types d'apprentissage	35
6.1.1. Le mode supervisé	35
6.1.2. Le renforcement	35
6.1.3. Le mode non-supervisé (ou auto-organisationnel).....	36
6.1.4. Le mode hybride	36
6.2. Règles d'apprentissage.....	36
6.2.1. Règle de correction d'erreurs	36
6.2.2. Apprentissage de Boltzmann	36
6.2.3. Règles de Hebb	36
6.2.4. Règle d'apprentissage par compétitions.....	37
6.3. Algorithme d'apprentissage supervisé pour les RNA : la rétropropagation.....	37
6.4. Généralisation.....	38
6.4.1. Facteurs indispensables pour une bonne généralisation	39
6.4.2. Optimisation de la généralisation	39
6.4.2.1. Algorithme d'apprentissage.....	40
a) Algorithme du facteur d'apprentissage adaptatif η	40
b) Méthode de Newton	40

c) Méthode de Levenberg-Marquardt.....	41
6.4.2.2. Compromis Apprentissage Généralisation ou « early stopping ».....	41
7. RÉCAPITULATION.....	42
8. MISE EN ŒUVRE DES RÉSEAUX NEURONAUX.....	43
9. AVANTAGE DES RNA SUR LES MÉTHODES STATISTIQUES.....	44
10. APPLICATIONS	44
CHAPITRE 3: ALGORITHME DE DETECTION DE L'ONDE P	45
1. INTRODUCTION	45
1.1. Aperçus précédents des applications médicales impliquant les réseaux de neurones....	45
2. APERÇU SUR LES APPLICATIONS CARDIOVASCULAIRES.....	48
2.1. Application dans la détection du complexe QRS	48
2.2. Application dans la détection de l'infarctus du myocarde aigu (IDM)	49
2.3. Application utilisant le système de surveillance Holter.....	51
2.4. Application en détectant d'autres pathologies	53
3. ACQUISITION DES DONNÉES D'ECG.....	53
4. CONCEPTION DE LA MÉTHODE DE DÉTECTION DE L'ONDE P	53
4.1. Architecture du réseau de neurones.....	54
4.2. Apprentissage.....	55
4.3. Fonction d'apprentissage	56
4.4. Fonction de transfert.....	57
4.5. Déroulement de l'apprentissage	58
4.6. Performance et Évaluation d'exécution du l'algorithme	59
5. RÉSULTATS ET DISCUSSION	60
6. CONCLUSION.....	64
CONCLUSION.....	65

Introduction Général

Introduction

Depuis un siècle environ, EINTHOVEN a introduit en médecine l'électrocardiogramme de surface (ECG) qui constitue jusqu'aujourd'hui un des piliers essentiels du diagnostic. Au cours des dernières années, les méthodes d'analyse des pathologies cardiaques se renforcent par l'introduction des procédés d'imagerie, notamment de l'échocardiogramme. Ceci ne modifie pas l'importance et l'utilité des ECGs. C'est dans ce contexte que ce mémoire a pour but de présenter L'ECG dans son rôle diagnostique actuel et de renforcer les méthodes de traitement existantes par des nouveaux outils numériques.

Avec le développement des moyens informatiques et du calcul numérique, il devient intéressant d'intégrer une démarche d'aide au diagnostic dans un processus de calcul automatique. Le rôle du traitement du signal est primordial dans le domaine de l'électronique médicale, en cardiologie par exemple, puisqu'il réalise l'extraction des paramètres qui serviront dans des opérations de diagnostic.

Les travaux de recherches pour la classification du signal cardiaque remontent à 1957 où les premières approches étaient principalement basées sur l'utilisation de la méthode des arbres de décision et la méthode statistique. Par la suite, de nouvelles approches ont été développées telles que l'approche syntaxique, les systèmes experts, la logique floue et les réseaux de neurones. Il y a eu des centaines de citations dans la littérature biomédicale où la plupart des études décrites tentent l'identification d'un infarctus du myocarde (IDM) ou la classification des arythmies ventriculaires. Ces deux dernières se basent sur la détection et l'estimation de la morphologie du complexe QRS. Par contre on rencontre une certaine réticence dans la bibliographie en ce qui concerne les arythmies auriculaires qui dit détection de l'onde P.

Les recherches méthodologiques effectuées ces dernières années ont permis de mettre en évidence la pertinence et la supériorité des méthodes d'identification à l'aide des réseaux de neurones, par rapport aux méthodes classiques.

De plus, les réseaux de neurones se sont avérés particulièrement adaptés dans le domaine de l'automatisme. Leurs capacités d'approximation sont mises à profil pour la modélisation des systèmes. D'un autre côté, les caractéristiques dont jouissent ces réseaux, d'être utilisés pour les opérations de reconnaissance et plus généralement, en traitement de signal. Ainsi pour la classification, les réseaux de neurones ont prouvé leur efficacité et elles ne cessent de s'élargir.

L'étude que nous abordons dans ce travail vise le développement d'un logiciel de détection des paramètres de l'ECG à savoir l'onde P. Nous nous sommes concentrés essentiellement sur une approche basée sur une fragmentation du signal et qui permet de détecter cette onde en exploitant les réseaux de neurones statiques.

Cette mémoire comprend trois chapitres:

L'objectif du premier chapitre est une introduction au fonctionnement général du système cardiovasculaire, puis, de manière plus détaillée, le principe de l'électrocardiogramme (ECG). Ensuite, nous décrivons très sommairement la méthode d'acquisition numérique des ECG.

Le second chapitre est centré sur les réseaux de neurones artificiels (RNA) en général nous nous attardons sur les Perceptrons MultiCouches (PMC) pour leurs capacités à approcher les relations non-linéaires. Les propriétés des PMC en apprentissage et en généralisation sont abordées ainsi que les moyens permettant d'améliorer les performances des PMC en généralisation. Nous donnons les différents choix possibles pour les différents paramètres des modèles : architecture, fonction de transfert et algorithme d'apprentissage. Nous terminons ce chapitre en justifiant nos choix.

Le dernier chapitre du mémoire exploite l'élaboration d'un algorithme de détection d'un paramètre de diagnostic du signal dans le signal ECG qui est l'onde P.

Nous terminons cette mémoire avec nos conclusions sur l'utilité et la portée de l'utilisation de notre réseau pour la détection de l'onde P.

Chapitre 1

L'électrocardiogramme

Chapitre 1

L'électrocardiogramme

Nous présentons dans ce chapitre le fonctionnement général du système cardiovasculaire, puis, de manière plus détaillée, le principe de l'électrocardiogramme (ECG). Cette présentation se limite au strict nécessaire pour une bonne compréhension de la thèse; le lecteur intéressé par une approche médicale rigoureuse pourra se reporter aux nombreux ouvrages médicaux disponibles sur le sujet comme notamment [DeBakey et al, 1997].

1. Système cardiovasculaire

Le système cardio-vasculaire est constituée d'un vaste réseau circulatoire avec en son centre une pompe dont le rôle est de fournir la pression nécessaire à la circulation sanguine. Le système cardio-vasculaire forme un circuit fermé, de tel sorte que le sang est continuellement pompé hors du cœur par un ensemble de vaisseaux et ramené au cœur par un autre ensemble de vaisseaux. Cet ensemble vasculaire est donc constitué de deux circuits totalement distincts, qui ont tous deux leur origine et leur fin dans le cœur. Le cœur lui-même est divisé longitudinalement en deux parties ; le cœur droit et le cœur gauche.

Le sang situé au niveau du cœur droit est propulsé à travers les poumons (circulation pulmonaire) et est ramené au cœur gauche. A ce niveau, il est rejeté pour approvisionner tous les tissus (circulation systémique). Les vaisseaux qui transportent le sang du cœur aux tissus sont les artères et les vaisseaux qui le ramènent les veines. On peut donc considérer qu'il existe deux circulations distinctes : la petite circulation ou circulation pulmonaire et la grande circulation ou circulation systémique. Parallèlement, le myocarde possède sa propre circulation tissé par un fin réseau de vaisseaux, la circulation coronarienne.

1.1. La Circulation sanguine

1.1.1. Circulation pulmonaire

Le départ du sang pour les poumons prend naissance à partir du cœur droit par une grosse artère unique, l'artère pulmonaire qui se divise au-dessus du cœur en deux artères dont chacune se dirige vers un poumon. Arrivé dans les poumons, le sang pauvre en oxygène est réoxygéné au niveau des alvéoles et regagne le cœur gauche par les veines pulmonaires pour se vider dans l'oreillette gauche.

Le sang qui circule dans les veines systémiques, la moitié droite du cœur et les artères pulmonaires a un contenu relativement faible en oxygène. Quand ce sang circule dans les capillaires pulmonaires, il se charge de grandes quantités d'oxygène ; dans ces conditions, le sang contenu dans les veines pulmonaires, le cœur gauche et les artères systémiques a un contenu élevé en oxygène. Quand ce sang circule dans les capillaires des tissus et organes de l'organisme, une partie de son oxygène quitte le sang, ce qui aboutit à un abaissement du contenu en oxygène du sang veineux systémique.

Normalement la quantité de sang pompée dans les circuits pulmonaires et systémiques sont égaux, ce qui revient à dire que le cœur droit pompe la même quantité de sang que le cœur gauche. C'est seulement quand le débit sanguin subit des variations que ces volumes diffèrent l'un de l'autre, et ceci de façon momentanée.

Il est curieux de constater qu'à cet endroit précis, l'artère pulmonaire est pauvre en oxygène alors que normalement les artères sont riches en oxygène et parallèlement les veines pulmonaires riches en oxygène alors que normalement elles sont pauvres en oxygène.

1.1.2. La circulation systémique

La circulation systémique, encore appelée grande circulation prend naissance dans le cœur gauche par une seule et grosse artère, l'aorte. Lorsque le sang quitte le cœur gauche, celui-ci est riche en oxygène. L'aorte se divise ensuite en artères de plus en plus petites, les artérioles qui elles-mêmes se divisent en vaisseaux encore plus petits dans les tissus, les capillaires. A la sortie des tissus, les capillaires se lient pour former les veinules, puis en vaisseaux plus gros et moins nombreux, les veines. Le sang redevenu pauvre en oxygène par la consommation des tissus arrive alors au cœur droit par deux grosses veines ; la veine cave supérieure qui est issue des membres supérieurs et la veine cave inférieure issue des membres inférieurs. Le circulation systémique se divise donc en deux circuits : la circulation artérielle et la circulation veineuse (Figure 1.1).

Au niveau du cœur droit, la totalité du sang pompée circule dans les poumons, alors qu'une fraction seulement du débit ventriculaire gauche total circule dans un seul organe ou tissu. Ceci est dû au fait que la circulation systémique comprend de nombreuses voies parallèles.

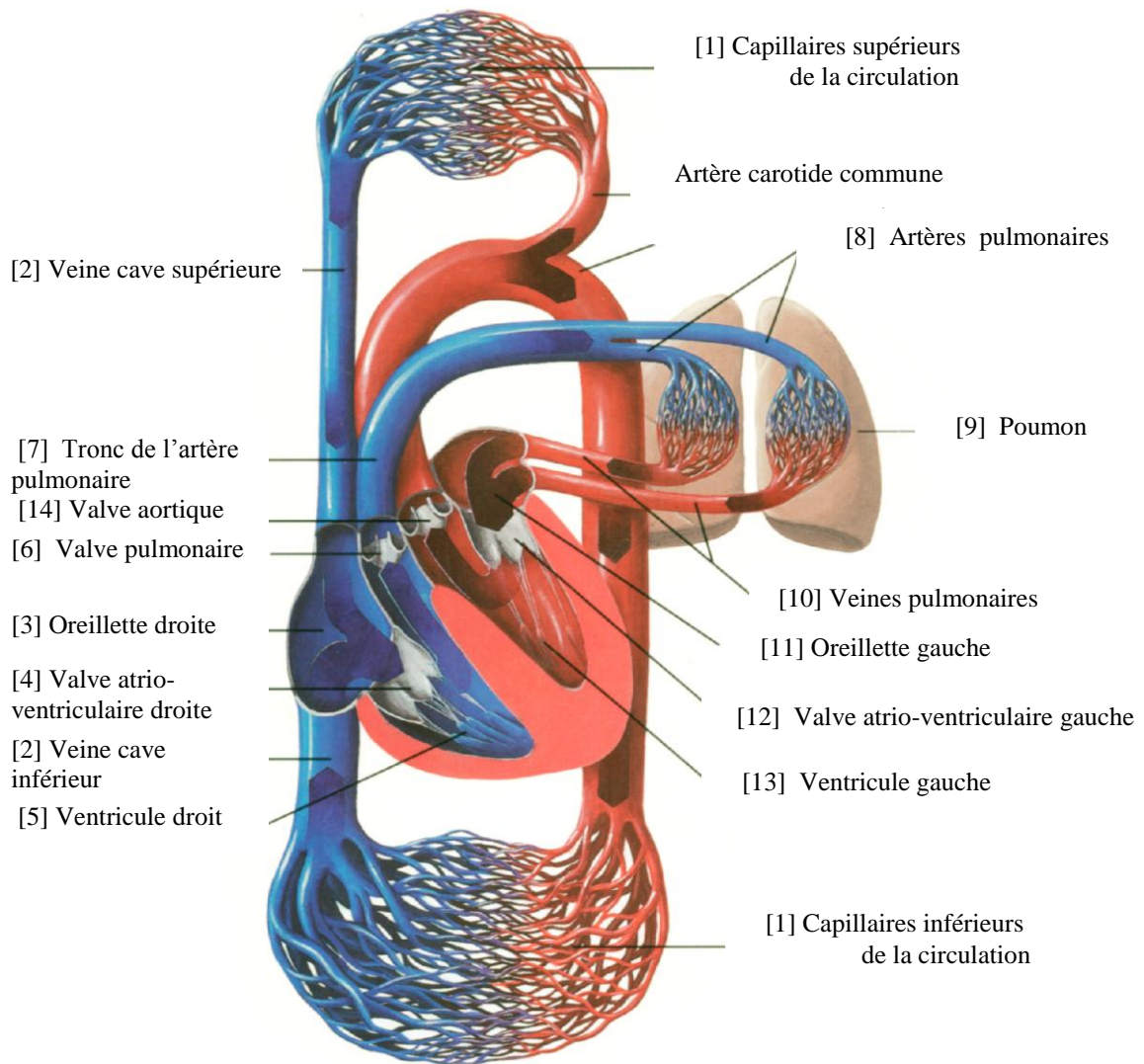


Figure 1.1. Schéma de la circulation. Le sang pauvre en oxygène est amené de la périphérie du corps [1] par l'intermédiaire de la veine cave supérieure et inférieure [2] à l'oreillette droite du cœur. Pendant la systole de l'oreillette, le sang pénètre dans le ventricule droit [5] par la valve atrio-ventriculaire droite [4]. A partir de là, le sang pendant la systole du ventricule est expulsé par la valve pulmonaire [6] dans le tronc de l'artère pulmonaire. Par l'intermédiaire des artères pulmonaires [8], le sang pénètre dans le poumon [9] où il est enrichi en oxygène inspiré. De l'anhydride carbonique se dégage. Ensuite, par l'intermédiaire des veines pulmonaires [10], le sang riche en oxygène pénètre dans l'oreillette gauche [11] et après, pendant la systole de l'oreillette, par l'intermédiaire de la valve atrio-ventriculaire gauche [12] dans le ventricule gauche [13]. Pendant la systole du ventricule, le sang est expulsé dans l'aorte [15] par la valve aortique [14]. Par l'intermédiaire du système artériel, le sang atteint les capillaires du corps [1] où il alimente les tissus en oxygène et en matières nutritives. Ensuite, le sang pauvre en oxygène retourne au cœur droit par l'intermédiaire du système veineux. Le cycle de la circulation recommence.

1.2. Circulation sanguine au niveau du cœur

Le cœur propulse le sang grâce aux contractions de son tissu musculaire appelé myocarde. Une épaisse cloison le divise en deux moitiés (cœur gauche/cœur droit), et chacune d'elles comporte deux cavités : l'oreillette et le ventricule. À chaque battement, le myocarde suit la même séquence de mouvement : le sang pauvre en oxygène arrive au cœur par la veine cave. Il y entre par l'oreillette droite, et en est chassé par sa contraction appelée systole auriculaire qui le déplace dans le ventricule droit. La systole ventriculaire (contraction des ventricules) propulse à son tour le sang du ventricule droit vers les poumons où il va se charger en oxygène. De retour au cœur par les veines pulmonaires, le sang s'accumule dans l'oreillette gauche puis, lors de la systole auriculaire, passe dans le ventricule gauche qui lors de la systole ventriculaire l'envoie vers les organes par l'artère aorte (Figure 1.2).

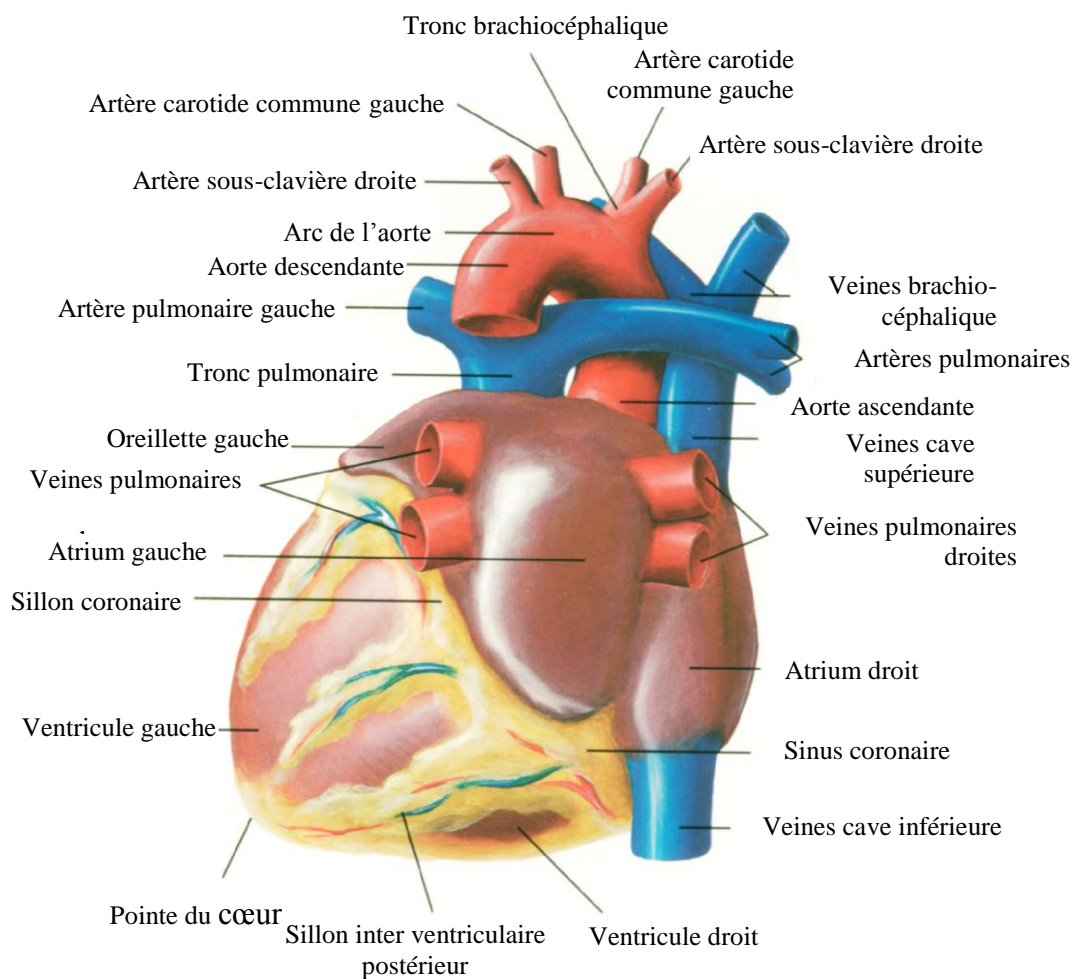


Figure 1.2. Position naturelle du cœur, vue dorsale

1.3. Cellules myocardiques

Le travail du cœur [Silbernagl et al, 1993] est dû à la contraction des cellules cardiaques, qui sont des cellules musculaires striées hautement spécialisées. A l'opposé de ce qui se passe dans les muscles communs, l'excitation électrique conduisant à la contraction se propage de cellule en cellule, ce qui assure la coordination nécessaire à un pompage efficace. Cette excitation est liée à des échanges d'ions entre les parois interne et externe de la cellule. Ce sont les ions chlorure (Cl^-), sodium (Na^+), calcium (Ca^{++}) et potassium (K^+) qui sont impliqués dans les échanges membranaires, leurs asmolarités intracellulaires et extracellulaires présentent des valeurs de potentiel électrochimiques très différentes, spécifiques de chaque ion.

1.4. Bases Electrophysiologique

1.4.1 Le potentiel monophasique unicellulaire

La cellule myocardique au repos est électropositive en surface, électronégative à l'intérieur. Si l'on place une électrode à la surface de la cellule et une autre à l'intérieur, une différence de potentiel de - 90 mV s'inscrit sur l'appareil enregistreur (Figure 1.4).

Lorsque la cellule est excitée (stimulus mécanique, chimique ou électrique), la surface devient électronégative et l'intérieur électropositif : c'est la dépolarisation. La différence de potentiel entre la surface et l'intérieur de la cellule est de + 30 mV.

Ce changement de polarité est très rapide, de l'ordre de la milliseconde; c'est la phase 0. Puis se succèdent (Figure 1.4):

- La phase de repolarisation initiale ou phase 1.
- La phase 2 de dépolarisation maintenue ou plateau.
- La phase 3 de repolarisation, la cellule redevient positive à l'extérieur, négative à l'intérieur.
- La phase 4 est une phase de polarisation stable.

Durant la phase 4 de repos, la cellule contient beaucoup de potassium et peu de sodium. Pendant la phase 0, la membrane devient très perméable au sodium, qui pénètre dans la cellule. Il y a donc un brusque mouvement d'ions responsable des variations de potentiel qui sont enregistrées par l'électrocardiogramme. Suit l'entrée du calcium et une sortie de potassium pendant la phase 2.

Durant la phase 3, le sodium et le calcium ressortent. Pendant la phase 4, une pompe sous la dépendance de membranaire rétablit la concentration de potassium intracellulaire.

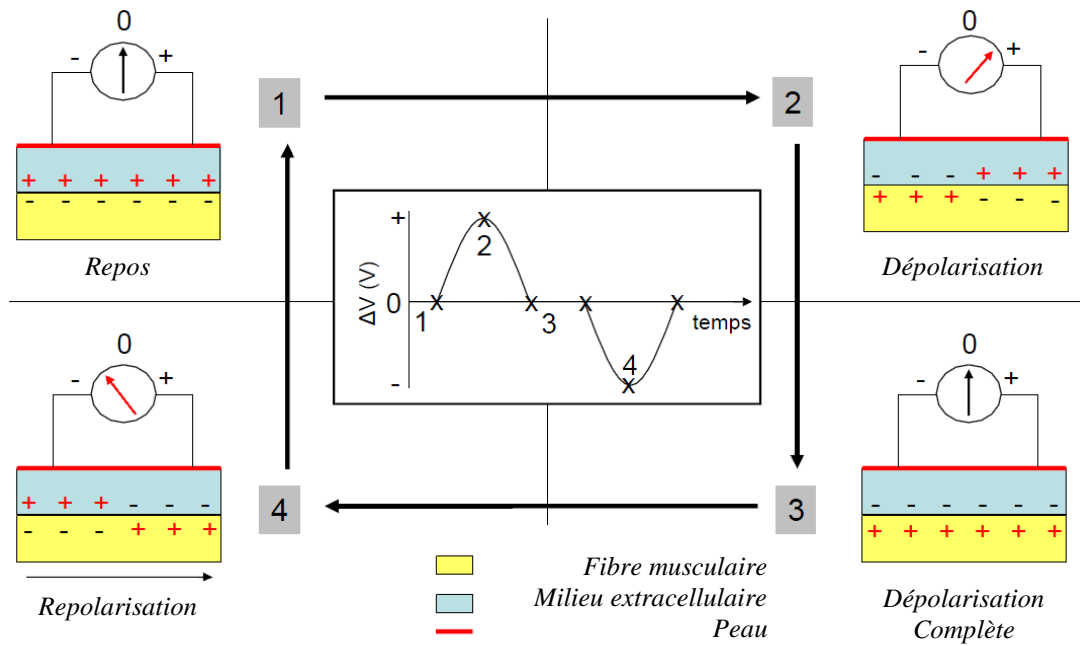


Figure 1.3. Electrophysiologie cellulaire

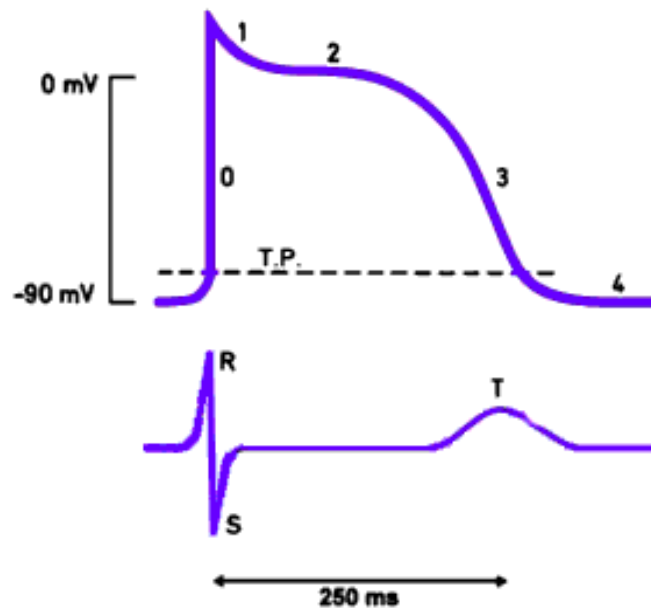


Figure 1.4. Graphe du potentiel d'action cellulaire avec ses 4 phases successives. État de base, montée rapide, plateau et descente progressive.

1.4.2. Le fonctionnement électrique du cœur

Le mécanisme essentiel du fonctionnement du cœur repose sur la contraction des fibres myocardique qui sont gérées en permanence par deux circuits électriques. Le cœur possède un système local d'autorégulation situé dans l'oreillette droite qui permet au cœur de fonctionner de façon autonome et un système central de contrôle via un centre nerveux autonome du

cerveau qui permet d'adapter le travail du cœur aux besoins de l'organisme. En fait, il faut comprendre ici que le cœur est doté d'un mécanisme autonome qui fonctionne à une fréquence moyenne pour satisfaire une activité moyenne de l'organisme. Dans la mesure où l'organisme augmente ses activités métaboliques, comme par exemple, après un repas copieux, lors d'un exercice sportif, le cœur doit augmenter sa fréquence pour satisfaire le surplus de besoins énergétiques; de même, si l'individu fait de la relaxation, son cœur doit adapter sa fréquence à cette diminution des besoins énergétiques. Comme le cerveau est le centre de contrôle capable d'être renseigné sur les activités cellulaires de l'ensemble de l'organisme, alors il faut que ce dernier soit branché sur le cœur par des voies efférentes afin de lui commander, selon le cas, d'augmenter ou de diminuer sa fréquence.

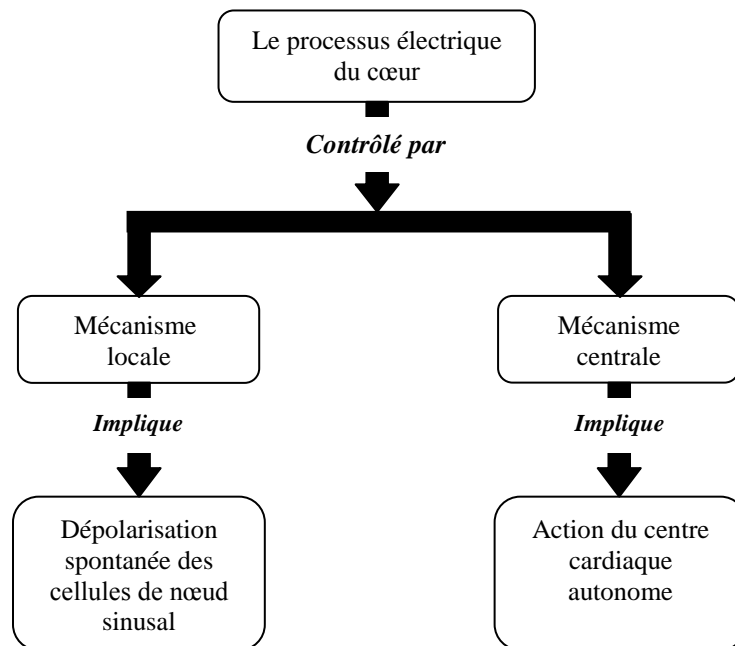


Figure 1.5. Mécanisme du cœur

2. L'électrocardiographie

L'électrocardiogramme enregistre les variations des potentiels électriques entre 2 points éloignés, à la surface du corps (ECG) : elles sont dues à la dépolarisation- repolarisation du muscle cardiaque (dérivations indirectes) selon une séquence déterminée par l'organisation fonctionnelle du tissu nodal.

Le tracé obtenu n'est pas fondamentalement différent de l'enregistrement de l'activité électrique cardiaque au moyen d'électrodes directement appliquées sur le cœur (dérivations directes) puisque le corps est un milieu conducteur. Les potentiels recueillis sont seulement plus faibles.

3. Electrocardiogramme

3.1. Définition de l'électrocardiogramme

Le tracé électrocardiographique (ECG) est une forme de visualisation des tensions électriques (de différences de potentiels électriques provoquées par la propagation de l'influx électrique au sein du cœur) qui résultent de l'excitation du cœur. Ces signaux sont obtenus à partir de points bien précis situés sur la peau par l'intermédiaire d'électrodes placées à la surface du corps (dérivations cardiaques). L'ECG exprime donc les événements électriques de l'excitation cardiaque et peut renseigner sur l'état du cœur. La fréquence de battement, la nature et la genèse du rythme, l'excitation et ses effets, ainsi que les perturbations éventuelles, que celles-ci soient d'origine anatomique ou mécanique, qu'elles concernent des altérations tissulaires ou des perturbations de la circulation sanguine sont véhiculées par le signal de l'ECG, d'où son extrême importance.

3.2. Dérivations

L'ECG s'enregistre à partir de divers points (dérivations) donnant naissance à des courbes de morphologies différentes, toujours dans l'ordre: onde P, complexe QRS et onde T. La courbe électrocardiographique va présenter des morphologies différentes selon où est placée l'électrode qui enregistre. Les endroits où sont placées les électrodes correspondent aux différentes dérivations. Les courbes de l'ECG peuvent ainsi être positives (ondes R) ou négatives (QS). Comme le cœur est un organe tridimensionnel, il est nécessaire de connaître les projections des vecteurs représentatifs de l'activité cardiaque sur tous les plans, frontaux et horizontaux, telle une photo que l'on prend à partir de plusieurs points de vue. On utilise pour faire un ECG les 12 dérivations de base: 6 dérivations sur le plan frontal et 6 sur le plan horizontal. Les dérivations frontales (dérivations de Goldberg) comprennent 3 électrodes bipolaires (DI, DII et DIII) et 3 unipolaires (aVR, aVL et aVF). Les 3 dérivations bipolaires ont tout d'abord été décrites et considérées par Einthoven comme un triangle (triangle d'Einthoven) en forme d'un circuit fermé, où selon la loi de Kirchoff $DII = DI + DIII$. Bailey en déplaçant les 3 côtés du triangle d'Einthoven au centre de ce même triangle obtenait une figure de référence appelée système triaxial de Bailey. La combinaison dans le plan frontal des dérivations bipolaires et unipolaires constituent le système hexaxial de Bailey (figure 1.6).

3.2.1. Dérivations dans le plan frontal

a) *Dérivations bipolaires*

Ces dérivations enregistrent l'activité cardiaque entre 2 points du corps. Les électrodes sont placées sur les bras droit et gauche, ainsi que sur le pied gauche. On relève ainsi les différences de potentiel entre les bras gauche et droit (DI), pied gauche et bras droit (DII) et pied gauche et bras gauches (DIII). La partie positive de (DI) correspond à 0° , celle de (DII) à $+60^\circ$ et celle de (DIII) à $+120^\circ$. Les parties négatives des 3 dérivations correspondent à $\pm 180^\circ$, -120° et -60° , respectivement.

b) *Dérivations unipolaires*

Les trois dérivations unipolaires aVR, aVL et aVF ont chacune des lignes allant d'un pôle positif à un pôle négatif en passant par le centre d'un triangle: la ligne d'aVR va de -150° (pôle positif) à $+30^\circ$ (pôle négatif), celle de aVL de -30° (pôle positif) à $+150^\circ$ (pôle négatif) et celle de aVF de $+90^\circ$ (pôle positif) à -90° (pôle négatif).

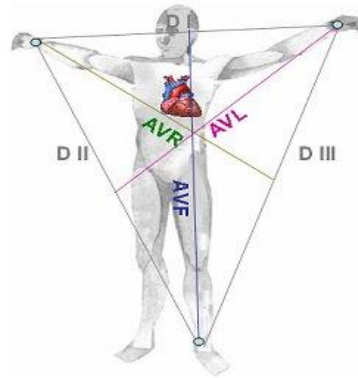


Figure 1.6. Position des électrodes des dérivations bipolaires et des dérivations unipolaires.

3.2.2. Dérivations dans le plan horizontal

Les dérivations du plan horizontal (dérivations de Wilson) constituent les dérivations unipolaires et sont au nombre de six de (V1 à V6) (figure 1.7). Ces dérivations ont une partie positive, là où est placée l'électrode exploratrice, et une négative, qui correspond à la région opposée. L'électrode exploratrice est placée sur divers points précordiaux. En plus des dérivations V1 à V6, on peut utiliser dans certaines occasions les dérivations plus droites (V1R à V4R) ou plus gauches (V7 à V9).

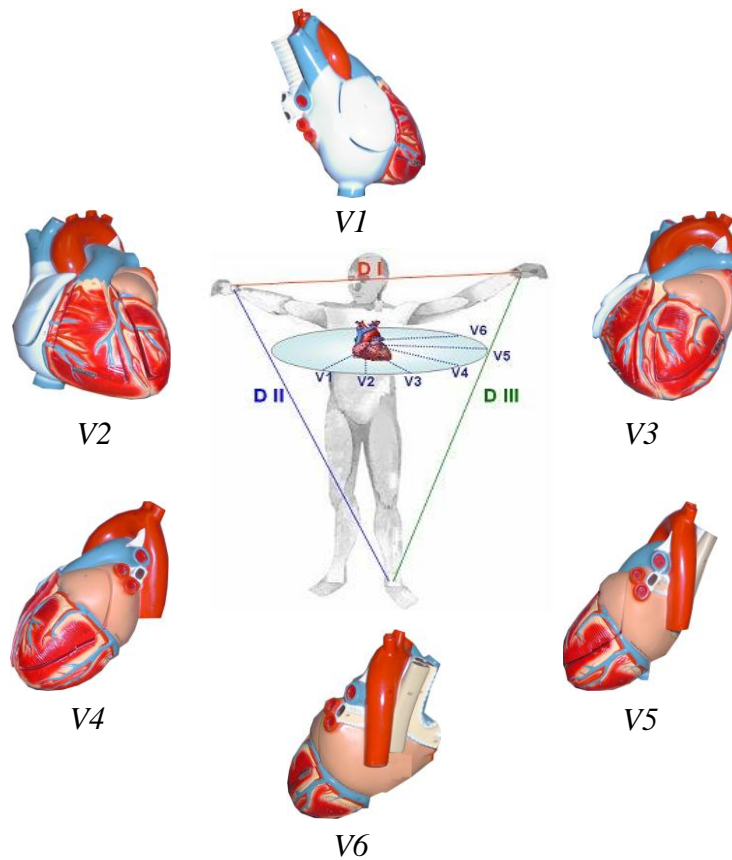


Figure 1.7. Les dérivations précordiales.

3.2.3. Position des électrodes

Les dérivations unipolaires des membres AVR, AVL et AVF utilisent le même emplacement d'électrode que les dérivations standards des membres. Quelle est la différence? Contrairement à des dérivations standards des membres, les dérivations augmentées des membres sont unipolaires. Au lieu de mesurer le potentiel électrique existant entre 2 électrodes (+ et -), IL détermine directement le point milieu. Autrement dit, les dérivations standards et unipolaires des membres ont un point en commun: elles regardent le cœur de votre patient de face (plan frontal du corps)

4. Activité électrique d'un battement cardiaque normal

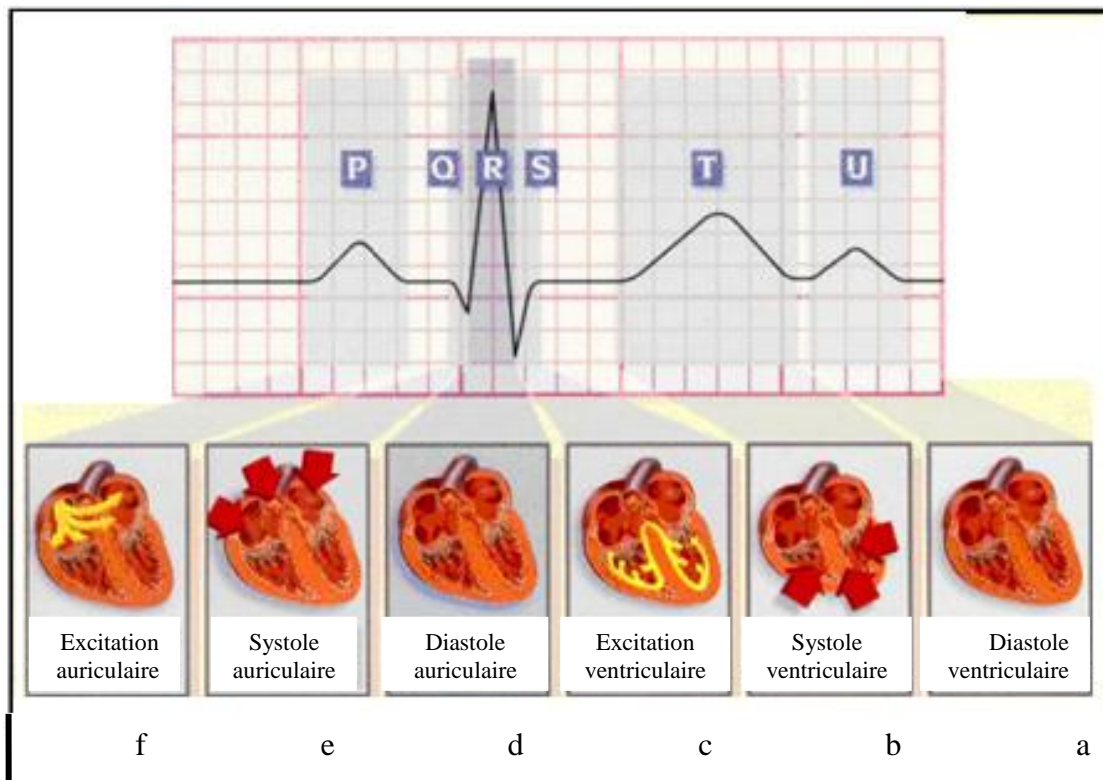


Figure 1.8. Excitation cardiaque.

Le battement cardiaque peut donc être suivi grâce à l'enregistrement en surface du signal électrique qui l'accompagne. En effet, chaque phase du battement possède une trace électrique particulière. L'impulsion initiale vient du sinus. Elle n'est pas visible sur l'ECG. L'onde électrique qui se propage ensuite dans les oreillettes, entraînant leurs contractions, laisse la trace d'une petite déflexion positive sur l'ECG : l'onde P (Figure 1.8a). L'impulsion arrive alors au nœud auriculo-ventriculaire (AV), où se produit la courte pause qui se traduit sur l'ECG par un petit segment plat PR; puis elle emprunte les voies de conduction rapide (le faisceau de His) pour entraîner la contraction des ventricules, suivie de leur repolarisation. Cette propagation de l'impulsion et la contraction brève et puissante de l'ensemble du muscle ventriculaire, dessinent sur l'ECG une succession de 3 ondes (Q, R et S) appelé complexe QRS (Figure 1.8b). L'onde Q est la première : c'est une onde dirigée vers le bas, qui n'est pas toujours visible sur le tracé ; la seconde est l'onde R : elle est de grande amplitude et dirigée

vers le haut ; la dernière est dirigée vers le bas : c'est l'onde S. C'est l'ensemble de ces trois ondes qui constitue le complexe QRS. Après chaque complexe QRS, on observe sur l'ECG une onde appelée onde T. Entre cette onde et l'onde précédente, on note une courte pause appelée le segment ST, dont l'étude est très importante pour l'identification de certaines pathologies. L'onde T traduit la phase de repolarisation des cellules constituant les ventricules; c'est un phénomène purement électrique et pendant cette phase le cœur est mécaniquement inactif (Figure 1.8c).

5. Electrocardiogramme normal

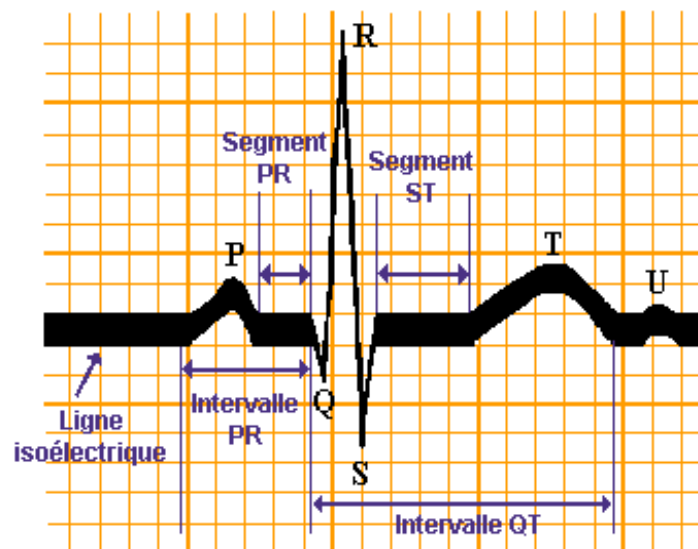


Figure 1.9. Ondes et intervalles de l'ECG.

5.1. Ondes Intervalles et Segments

Le tracé électrique comporte plusieurs accidents répétitifs appelés « ondes », et différents intervalles entre ces ondes. Les principales mesures à effectuer lors de l'analyse d'un ECG sont celles de l'onde P, de l'espace PR, du complexe QRS, du délai d'inscription de la déflexion intrinsécoïde, du point J, de l'espace QT, du segment ST et enfin de l'onde T (figure 1.9).

5.1.1. Onde P

Correspond à la dépolarisation (et la contraction) des [oreillettes](#), droite et gauche. On analyse sa morphologie (positive ou diphasique en V1 voire V2 et monophasique dans toutes les autres dérivations), sa durée (qui est de 0,08 à 0,1 seconde), son amplitude (inférieure à 2,5 mm en D2 et 2 mm en V1 et V2), son axe (déterminé de la même façon que pour l'axe des QRS, normalement situé entre 0 et 90°, généralement vers 60°) et sa synchronisation avec l'onde QRS.

5.1.2. Intervalle PR (ou PQ)

Est le temps entre le début de P et le début du QRS. Il est le témoin du temps nécessaire à la transmission de l'influx électrique du noeud sinusal des oreillettes au tissu myocardique des ventricules (conduction auriculo-ventriculaire). Sa durée normale, mesurée du début de l'onde P au début du complexe QRS est de 0,12 à 0,20 seconde. La durée de l'espace PR diminue lorsque la fréquence cardiaque augmente. Il est normalement isoélectrique.

5.1.3. Onde QRS

Appelé aussi complexe QRS qui correspond à la dépolarisation (et la contraction) des ventricules, droit et gauche. L'onde Q est la première onde négative du complexe. L'onde R est la première composante positive du complexe. L'onde S est la deuxième composante négative. Suivant la dérivation et sa forme, on parle ainsi d'aspect « QS », « RS », voire « RSR' » (pour une forme en M avec deux positivités). La forme et l'amplitude du QRS varient selon les dérivations et selon l'éventuelle pathologie du muscle cardiaque sous-jacent. Le complexe QRS a une durée normale inférieure à 0,1 seconde, le plus souvent inférieur à 0,08 s. L'axe des QRS normaux est compris entre 0 et 90°. La zone de transition correspondant à la dérivation précordiale dans laquelle les QRS sont isoélectriques est située normalement en V3 ou V4.

5.1.4. Point J

Correspond au point de transition entre le complexe QRS et le segment ST. Il est normalement isoélectrique.

5.1.5. Segment ST

Correspond au temps séparant le début de la dépolarisation ventriculaire représentée par le complexe QRS et la fin de la dépolarisation ventriculaire représentée par l'onde T. Le segment ST normal est isoélectrique du point J au début de l'onde T.

5.1.6. Intervalle QT

Mesuré du début du QRS à la fin de l'onde T correspond à l'ensemble de la dépolarisation et de la repolarisation ventriculaire (temps de systole électrique). Sa durée varie en fonction de la fréquence cardiaque, il diminue quand la fréquence cardiaque augmente et augmente quand la fréquence cardiaque diminue. Son allongement voire son raccourcissement est lié dans certaines circonstances à l'apparition d'un trouble du rythme ventriculaire complexe nommé « torsades de pointes » potentiellement mortel. Ainsi utilise-t-on le QTc (QT corrigé) qui est la mesure de l'intervalle QT corrigé par la fréquence selon la formule $QTc = QT / \text{racine carré de l'espace RR}$. L'hypoxie cardiaque et les troubles de la concentration sanguine en calcium affectent cet intervalle.

5.1.7. Onde T

Correspond à l'essentiel de la repolarisation (la relaxation) des ventricules, celle-ci commençant dès le QRS pour quelques cellules. Sa durée est de 0,20 à 0,25 secondes, l'analyse de sa durée est comprise dans l'analyse de la durée de l'intervalle QT. L'axe normal de l'onde T, calculé de la même façon que l'axe des QRS, est compris entre - 10 et 70°, souvent autour de 40°. L'onde T est normalement pointue, asymétriques et amples dans la plupart des dérivations. Elle peut être négative en V1 voire en D3 et aVF. Son amplitude

dépend généralement celle de l'onde R qui la précède, elle est comprise entre 1/8 et 2/3 de celle de l'onde R et ne dépasse pas le plus souvent 10 mm.

5.1.8. Onde T atriale

Est masquée par l'onde QRS et correspond à la repolarisation (la relaxation) des oreillettes. Celle-ci est négative.

5.1.9. Onde U

Est une petite déflexion parfois observée après l'onde T dans les dérivations précordiales V à V4. Elle est positive dans toutes les dérivations sauf en aVR, sa présence traduit une anomalie métabolique.

5.2. Déflexion Intrinsécoïde DI

Elle exprime le temps que met l'onde de dépolarisation à parcourir le myocarde de l'endocarde à l'épicarde. La DI est d'autant plus longue que le myocarde est épais. Elle se mesure au niveau des dérivations droites pour le ventricule droit et des dérivations gauches pour le ventricule gauche. La mesure de DI se fait du début de l'onde Q au sommet de l'onde R ou r, ou du dernier crochetage de R si l'onde R est crochétée.

6. Acquisition et Traitement Des Signaux Physiologiques

6.1. Circuit biomédical

6.1.1. Signal électrique

Un signal électrique est une grandeur électrique mesurable variant dans le temps ou dans l'espace et permettant de transporter une information. Il peut s'agir d'une différence de potentiel, de l'intensité d'un courant électrique qui se présente sous forme périodique, logique, analogique... pour ne citer que les plus utilisés... ou plus généralement d'une onde électromagnétique (onde radio par exemple).

Une particularité des signaux électriques est leur facilité de transmission, d'acquisition et de stockage. Le développement de l'électronique les rend également très accessibles et faciles à manipuler.

6.1.2. Tissus biologiques

Les tissus sont le niveau d'organisation intermédiaire entre les cellules et les organes. Un tissu est un ensemble de cellules pas forcément identiques mais de même origine, regroupées en ensemble fonctionnel, c'est-à-dire concourant à une même fonction. Les tissus se groupent en organes.

La membrane cellulaire agit comme un circuit électrique microscopique. Premièrement, elle agit comme un condensateur qui peut être chargé électriquement de part et d'autre puisqu'elle possède une grande surface et qu'elle est constituée de lipides (gras) très imperméables aux charges électriques (un diélectrique isolant) (Figure 1.10). De plus, comme en électronique, la membrane cellulaire se comporte comme un diélectrique de permittivité relative proche de celle des lipides formant une couche isolante séparant deux milieux conducteurs [Rigaux et al, 1996]. Afin de se rendre compte du comportement diélectrique des tissus biologiques, plusieurs modèles ont été proposés et découlent de différentes théories de

Wagner [Wagner, 1914], Debye [Debye, 1929], Fricke [Fricke, 1932], Cole [Cole, 1941] et de Maxwell [Maxwell, 1981]. Ces auteurs ont mis en place au cours de la première moitié du 20ème siècle des approches encore utilisées à ce jour.

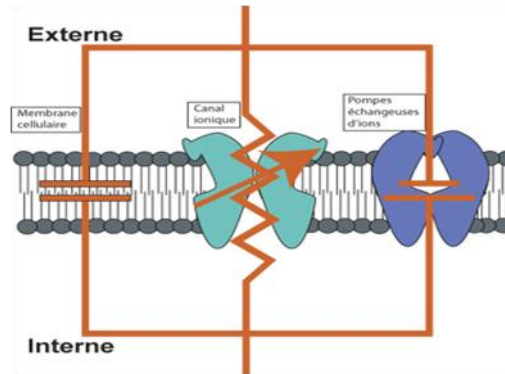


Figure 1.10. Circuit électrique de la membrane

Le circuit électronique de la membrane cellulaire des neurones. La surface de la membrane cellulaire de lipides agit comme un condensateur qui peut être chargé selon les besoins du neurone. Les canaux ioniques agissent comme des résistances variables, tandis que les pompes échangeuses d'ions agissent comme des chargeurs de batterie.

6.1.3. Signal biologique

L'électrophysiologie est la partie de la physiologie qui mesure les courants électriques des cellules. Les phénomènes électriques sont nombreux et variés dans l'organisme, en particulier dans les tissus excitables (cœur, muscle, système nerveux central), le cœur, le rein ainsi que certaines glandes.

6.2. Appareillage

Depuis le premier appareil du début du siècle, l'électrocardiographe a évolué en adoptant au fur et à mesure les technologies d'actualité. Ainsi, le signal analogique entrainent une aiguille est devenu numérique, exploité par logiciel, et transmis sur papier par l'intermédiaire d'une imprimante.

Un électrocardiographe moderne se présente sous la forme d'un appareil compact intégrant de nombreuses fonctions. En effet, celui-ci comportait un écran à cristaux liquides, une dizaine de dérivation, une imprimante, un logiciel d'exploitation des résultats, et une possibilité de stocker les enregistrements sur disque dur et disquette, ou des transmettre sur une ligne téléphonique. L'appareil pouvait aussi contrôler une épreuve d'effort.

6.3. Système d'acquisition de données

Les systèmes d'acquisition de données consistent le plus fréquemment en des interfaces avec les ordinateurs et les dispositifs d'enregistrement. Le choix des techniques d'acquisition dépend de l'environnement dans lequel le système sera utilisé [Lepoutre et al, 2001; Asch et al, 1999]

Les principales caractéristiques d'un système d'acquisition de données sont :

- La vitesse d'acquisition de données.
- La précision des mesures.

Exemple d'une chaîne de mesure numérique:

Cette dernière est munie d'une logique de commande à base de microprocesseur (Figure 1.11).

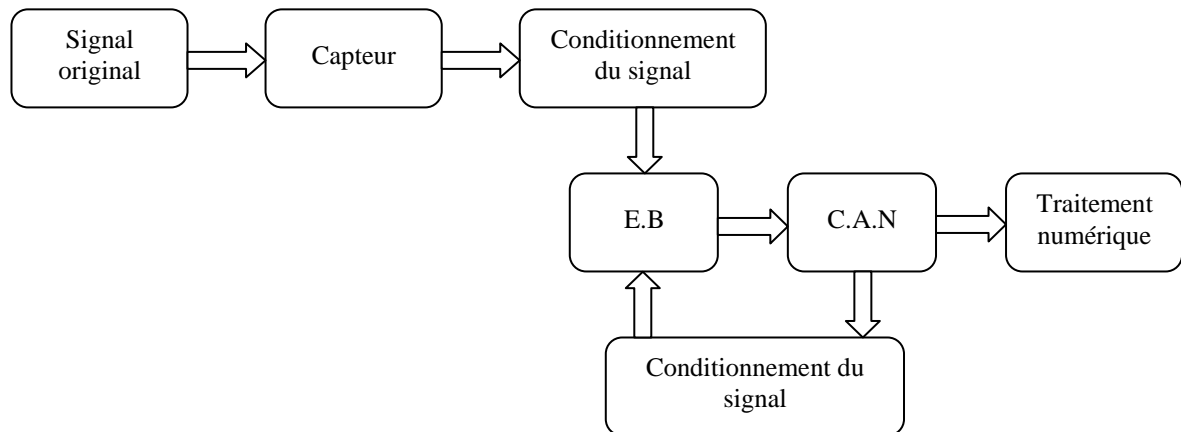


Figure 1.11. Système d'acquisition de données

6.4. Chaîne d'acquisition de l'ECG

Le traitement du signal cardiaque dépend de la chaîne d'acquisition analogique, elle représente un rapport signal sur bruit avec reproduction du signal correct et fidèle, car permettre un traitement numérique basé sur les algorithmes de détections d'anomalies et des erreurs.

Exemple d'une chaîne analogique

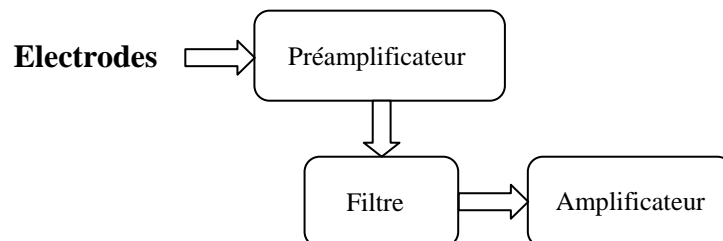


Figure 1.12. Exemple d'une chaîne analogique

Sur les appareils modernes, chaque électrode est connectée à un tampon (buffer) à haute impédance d'entrée (typiquement 100 MΩ). Le tampon supporte des hautes tensions, pour protéger les amplificateurs en cas d'électrochoc, envoyé par un défibrillateur par exemple.

Ces tampons, ou au moins tout l'étage d'entrée, sont alimentés par une alimentation isolée et stabilisée, dans le but de réduire les risques de chocs électriques. On trouve parfois une batterie pour cet usage.

Le signal de mode commun peut être utilisé pour réduire le bruit de mode commun par l'intermédiaire d'une rétroaction agissant sur la jambe droite du patient.

Sur le schéma ci-dessus, il est possible de choisir une dérivation particulière ; grâce au 'lead-Select Switch and network'. Il est à noter que la plupart des appareils actuels travaillent sur au moins dix voies en même temps.

6.5. Capteurs biomédicaux

Toute information biologique doit se présenter sous forme de signaux compréhensibles, enregistrables, et mesurables en valeurs normalisées. Le rôle joué par le capteur biomédical est l'un des aspects techniques que l'on rencontre.

Le capteur médical est constitué d'électrodes de recueil plates, métalliques, inoxydables et adhésives dont le contact électrique est assuré par une pate conductrice.

6.5.1. Problématiques des capteurs électrophysiologiques

A tout signal biologique, viennent se greffer des phénomènes accessoires qui peuvent le polluer et qui dans certain cas, le rendent méconnaissable. Ces signaux parasites sont les bruits. Un des potentiels de recueil de l'amplification et de la transmission de l'enregistrement et du traitement d'un signal est de donner au rapport signal/bruit (S/B) la valeur la plus forte possible. Ce bruit a des origines diverses :

- L'agitation thermique des électrodes.
- La présence de champs électromagnétiques due à la proximité d'émetteurs, de transformateurs, de moteurs et de générateurs d'ondes de basse ou de moyenne fréquence sont aussi à l'origine de parasites (radio, TV, radar, téléphone portable, etc.) et peuvent induire des courants dans le patient (Figure 1.13).
- Les signaux biologiques peuvent constituer du bruit les un pour les autres. Ainsi en est-il de l'électromyogramme (EMG) qui est un bruit très gênant pour l'ECG (Figure 1.14).

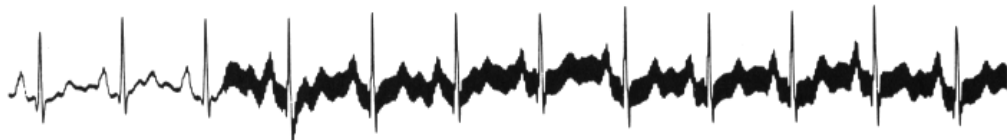


Figure.1.13. Interférences du champ magnétique à 50Hz venant du secteur.

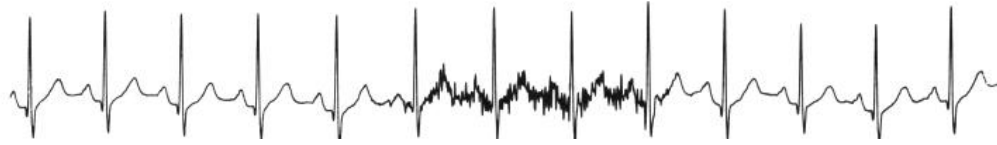


Figure.1.14. interférences d'origine électromyographique.

Les remèdes sont nombreux et varient selon la cause. Citons par exemple la mise à la terre pour évacuer les parasites du potentiel du secteur, les divers modes de blindage pour constituer des cages de faraday, utilisation des filtres analogiques et des filtres numériques [Bahoura et al, 1997].

Dans notre cas plusieurs techniques permettent d'améliorer le rapport S/B. Citons notamment:

- Amplification du signal à la source, au niveau de l'électrode, et transmission sur ligne basse impédance (insensibilité aux perturbations extérieures).
- Filtrage du signal transmis (filtre rejeteur 50 Hertz, filtre passe bas, passe haut), sachant que cela entraîne une perte d'information.
- Conversion analogique/numérique du signal à la source, permettant une transmission de type numérique et une immunité au bruit.

Réduction du bruit :

- Utilisation d'amplificateur différentiel de haute qualité.
- Utilisation de blindage efficace (câble haute qualité à blindage multiple).
- Réinjection du bruit en opposition de phase sur le blindage du câble de transmission.

Plusieurs problèmes inhérents à la conception de ces capteurs actifs, peuvent être soulignés :

- Référence : problème pour l'amplification des signaux par rapport à une référence unique.
- Alimentation : pile, batterie incorporée, téléalimentation.
- Stérilisation : étanchéité, résistance à la température.
- Marquage CE : compatibilité électromagnétique.
- Encombrement : intégration des composants, poids, taille.
- Coût : usage unique ou non.

6.6. Le préamplificateur

Le signal électrique délivré par les électrodes est très faible (quelque mV), il est en général perturbée par des bruits, un préamplificateur se révèle donc nécessaire.

L'amplificateur différentiel permet dans ce cas une bonne élimination des parasites, sans atténuation du signal utile.

6.7. Le filtre

Le signal peut être masqué par des bruits, la fonction principale d'un filtre consiste à éliminer totalement au partiellement un certain nombre d'harmoniques véhiculées par le signal mais ne faisant pas partie du signal utile.

6.8. Amplificateur

Permet d'accroître l'intensité du signal pour le rendre plus perceptible moins sensible aux bruits parasites. L'amplificateur propre à cette chaîne d'acquisition est un amplificateur d'instrumentation. Il destiner à extraire les signaux issus du capteur ayant des amplitudes faibles et de les amplifier avec un gain ajustable, précis et stable.

6.9. Echantillonnage

6.9.1. Echantillonneur-Bloqueur

L'échantillonneur bloqueur est un circuit qui permet de stocker la valeur analogique du signal (mémoire analogique) pendant un temps fixé préalablement (le pas d'échantillonnage).

On modélise l'échantillonnage à l'aide du produit du signal $S(t)$ issu du capteur par une fonction peigne noté $P(t)$ représentée par une suite d'impulsion de Dirac $\delta(t)$

$$P(t) = \sum \delta(t - kT_e)$$

Ou : T_e est la période d'échantillonnage.

Le signal échantillonné $S^*(t)$ est alors défini par :

$$S^*(t) = S(t).P(t)$$

L'échantillonneur bloqueur est constitué d'un commutateur associé à un condensateur. Lorsque le commutateur k est fermé, la tension aux bornes du condensateur C évolue avec la tension $S(t)$ (Figure 1.15). Lorsque le commutateur K est ouvert, la capacité C conserve cette tension jusqu'à ce qu'elle soit utilisée par le convertisseur analogique numérique placé en aval. les caractéristiques principales d'un échantillonneur-bloqueur sont:

- **Temps de charge**

La précision de la mesure dépend en grande partie du temps de charge de la capacité. Généralement, on veille à ce que ce temps soit le plus court possible.

- **Maintien de la tension de charge**

Durant le temps de conversion, il convient de maintenir la tension de charge aux bornes de la capacité à son niveau initial. Pour ce faire il faut, d'une part, choisir une capacité de faible coefficient de fuite et, d'autre part, faire en sorte que les circuits en amont et en aval de la capacité soient de très haute impédance.

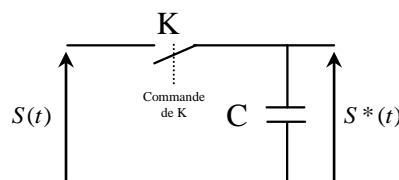
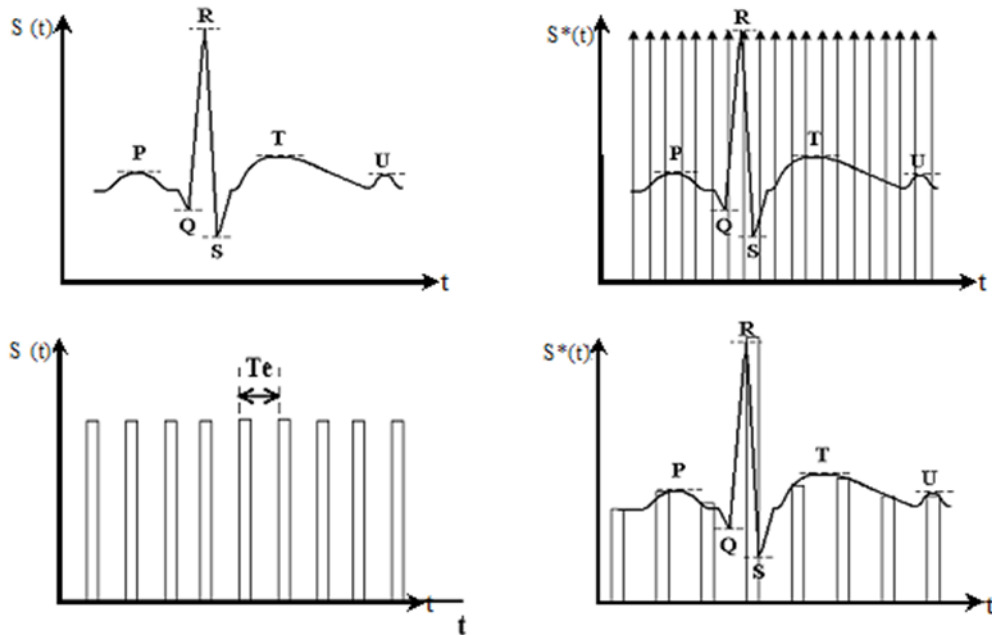


Figure.1.15. principe de fonctionnement d'un échantillonneur bloqueur.



6.9.2. Conversion analogique numérique

On désigne sous le nom de conversion analogique numérique (CAN), toute opération analogique continue dans le temps, en un nombre significatif dans un système de numérisation (décimal, binaire...).

6.10. Convertisseur Analogique/Numérique

Le convertisseur analogique/numérique [Rutkowski, 1977] effectue une conversion de la tension analogique en un signal numérique. Le signal analogique est échantillonné, quantifié puis codé en binaire. Ces trois fonctions sont assurées par le convertisseur lui-même associé à un échantillonneur bloqueur.

Le CAN est caractérisé essentiellement par sa résolution, sa précision et sa vitesse de conversion. La vitesse de conversion dépend du nombre de bits, de la technologie des éléments intégrés, de la technique de conversion ainsi que de la fréquence de l'horloge.

Parmi les très nombreuses techniques de conversion possible, nous retenons les convertisseurs parallèles, les convertisseurs à approximations successives et les convertisseurs à comptage d'impulsions.

Le principe du CAN parallèle consiste à comparer simultanément, la tension d'entrée à n niveaux de tension de référence. Ce type de convertisseur est très rapide. La fréquence de l'horloge de commande peut atteindre 100 MHz et le temps de conversion est de 500 ns pour une résolution de 12 bits.

Pour le CAN à approximation successive il existe une grande similitude entre son principe et la mesure de la tension continue à l'aide de la méthode potentiométrique. La tension d'entrée

est susceptible de varier tout au long de la conversion. Il devient donc nécessaire de figer la tension à convertir à l'aide d'un échantillonneur-bloqueur de 10 bits, la durée de conversion ne dépasse pas 25ms, la fréquence étant comprise entre 0 et 100 Hz.

La conversion par comptage d'impulsion est très simple une grand précision mais nécessite un temps de conversion excessivement long. Ce temps est généralement compris entre 0.1 et 100ms ce type de convertisseur trouve un domaine d'application dans le cas ou la grandeur physique à convertir évolue lentement.

		V1	V2	V3	V4	V5	V6	
Onde P	Durée (s)	0.06 - 0.11	0.06 - 0.11	0.06 - 0.11	0.06 - 0.11	0.06 - 0.11	0.06 - 0.11	
	Amplitude (mm)	< 2	< 2	< 2	< 2	< 2	< 2	
	Forme	P N	P N	P	P	P	P	
Complexe QRS	Durée (s)	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	
	amplitude	Q (mm)	< 25% de R	< 25% de R	< 25% de R	< 25% de R	< 3	< 3
		R (mm)	< 5	< 15	< 25	< 30	< 31	< 25
		S (mm)	< 20	< 25	< 30	< 25	< 5	< 5
	Forme	P N	P N	P N	P N	P N	P N	
Onde T	Durée (s)	0.20 - 0.25	0.20 - 0.25	0.20 - 0.25	0.20 - 0.25	0.20 - 0.25	0.20 - 0.25	
	Amplitude (mm)	> -3 et < 4	< 12	< 12	< 12	> 1 et < 8	> 1 et < 5	
	Forme	P N	P N	-	-	P	P	
Onde U	Durée (s)	0.16 - 0.25	0.16 - 0.25	0.16 - 0.25	0.16 - 0.25	0.16 - 0.25	0.16 - 0.25	
	Amplitude	F	< 2 mm	< 2 mm	F	F	F	
	Forme	N	-	-	-	-	-	
Int PR	Durée : Ad,As (s)	0.11 - 0.21	0.11 - 0.21	0.11 - 0.21	0.11 - 0.21	0.11 - 0.21	0.11 - 0.21	
	Durée : Enf (s)	0.10 - 0.17	0.10 - 0.17	0.10 - 0.17	0.10 - 0.17	0.10 - 0.17	0.10 - 0.17	
Seg ST	Durée (s)	< 0.11	< 0.11	< 0.11	< 0.11	< 0.11	< 0.11	
	Forme	I	I	I	I	I	I	
Seg PR	Durée (s)	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	0.05 - 0.10	
Int QT	Durée (s) à 60/min	0.39 ± 0.04	0.39 ± 0.04	0.39 ± 0.04	0.39 ± 0.04	0.39 ± 0.04	0.39 ± 0.04	
	à 80 bats/min	0.34 ± 0.04	0.34 ± 0.04	0.34 ± 0.04	0.34 ± 0.04	0.34 ± 0.04	0.34 ± 0.04	
	à 100 bats/min	0.31 ± 0.04	0.31 ± 0.04	0.31 ± 0.04	0.31 ± 0.04	0.31 ± 0.04	0.31 ± 0.04	
Int QU	Durée (s)	0.44 - 0.60	0.44 - 0.60	0.44 - 0.60	0.44 - 0.60	0.44 - 0.60	0.44 - 0.60	

LEGENDES DES SYMBOLES:

P : Positive.
N : Négative.
D : Diphasique.

> : Supérieur à.
< : Inférieur à.
Seg : Segment.

Int : intervalle.
mm : millimètre.
S : Seconde.
I : Isoélectrique.

F : faible.
Ad : Adulte.
As : Adolescent.

Enf : Enfant.
Bats : Battement.
min : Minute.

Chapitre 2

Quelques Modèles de Réseaux de Neurones

Chapitre 2

Quelques Modèles de Réseaux de Neurones

I. Intelligence artificielle

Le but de l'intelligence artificielle (AI) est de concevoir des systèmes capables de reproduire le comportement de l'humain dans ses activités de raisonnement.

L'IA se fixe comme but la modélisation de l'intelligence prise comme phénomène (de même que la physique, la chimie ou la biologie qui ont pour but de modéliser d'autre phénomène).

1. Historique

La préhistoire : 1945-1955

L'AI a débuté, même si elle ne portait pas encore ce nom là, après la seconde guerre mondiale. Le premier problème auquel se sont attaqués certains informaticiens était la traduction automatique. Leur erreur fut probablement de faire preuve d'un peu trop d'optimisme et surtout de croire qu'il était possible de réaliser une traduction sans qu'il soit nécessaire de comprendre le sens d'un texte. Le problème se révéla tellement ardu qu'il fut pratiquement abandonné sous sa forme initiale [Alliot et al, 1997].

Les débuts : 1955-1970

Dans les années 55-65, d'autres chercheurs s'intéressèrent à des problèmes que l'on allait commencer à appeler problèmes d'AI. La définition du mot date d'ailleurs de la conférence

tenue au Dartmouth collège (New Hampshire, USA) où il est employé pour la première fois par John McCarthy.

La spécialisation : 1970-1980

La fin des années 60 et le début des années 70 vont surtout voir l'IA se spécialiser et se théoriser.

En fin, l'IA va alors se scinder en plusieurs branches que l'on peut schématiquement résumer ainsi :

- La compréhension du langage naturel ;
- Les démonstrations automatiques de théorèmes ;
- Les jeux ;
- La résolution de problèmes généraux ;
- La résolution de problèmes experts ;
- La représentation de connaissances ;
- La perception ;
- L'apprentissage ;
- De façon générale tout problème que l'on ne sait pas résoudre informatiquement de façon classique.

Une reconnaissance : 1980-1990

L'événement majeur des années 80 est l'arrivée en force des japonais dans le domaine de l'intelligence artificielle. Le Japon lance en effet le projet cinquième génération dont le but est de développer, à la fois sur le plan matériel et sur le plan logiciel, des technologies et des techniques capables de faire de l'IA une discipline "efficace".

Le projet cinquième génération s'est officiellement terminé lors de la conférence internationale FGCS au mois de juin 1992.

2. Définir l'Intelligence artificielle

Il est extrêmement difficile de donner de l'intelligence artificielle une définition précise, tellement les domaines que recouvre le terme sont vastes et les avis divergents.

En fait, le terme même intelligence artificielle (IA) est considéré par certains auteurs comme parfaitement indu, voire contradictoire.

2.1 Définition

On désigne sous le terme d'intelligence artificielle l'élaboration de procédures automatiques de recherche de solution pour diverses classes de problèmes. Les procédures sont exprimées sous forme de programmes exécutables par les ordinateurs.

L'IA peut être envisagée selon deux points de vue complémentaires :

- L'un concerne l'étude de mécanismes de l'intelligence, l'ordinateur étant utilisé comme moyen de simulation pour tester un modèle ou une théorie ; ce point de vue relève d'une démarche cognitive.
- L'autre, concerne les efforts faits pour doter un ordinateur de capacités habituellement attribuées à l'intelligence humaine : acquisition de connaissance, prise de décision, etc.

2.2 Différents domaines

- La reconnaissance de l'écriture ;
- La reconnaissance de forme (visage, photo satellite ...) ;
- Planification
- La robotique ;
- L'apprentissage ;
- Les systèmes complexes adaptatifs.
- Les réseaux neuronaux ;

2.3 Réseaux de neurones

Un réseau de neurones est un modèle de calcul dont la conception est très schématiquement inspiré du fonctionnement de vrais neurones (humain ou non). Les réseaux de neurones sont généralement optimisés par des méthodes d'apprentissage de type statistique, si bien qu'ils sont placés d'une part dans la famille des applications statistiques, qu'ils enrichissent avec un ensemble de paradigmes permettant de générer de vastes espaces fonctionnels souples et partiellement structurés, et d'autre part dans la famille des méthodes de l'intelligence artificielle qu'ils enrichissent en permettant de prendre des décisions s'appuyant d'avantage sur la perception que sur le raisonnement logique formel.

II. Réseaux de neurones

1. Introduction

Capter une image, la numériser, la segmenter en éléments de contours, détecter un objet mobile, le reconnaître quelle que soit sa position et estimer sa profondeur. Capter le son d'une voix au milieu d'un brouhaha et du bruit ambiant et reconnaître les mots qui sont prononcés. Capter et analyser les deux cents signaux issus d'un processus industriel et en déduire si tout est conforme ou si une avarie se prépare faire le traitement d'un phénomène biologique donner un diagnostic à partir de signaux physiologique...

Voici quelques problèmes pourtant courants dans les sciences de l'informatique mais dont les solutions, encore incomplètes, impliquent de multiples efforts de recherche dans la communauté scientifique.

Malgré la constante augmentation de puissance des calculateurs, malgré les approches théoriques de plus en plus sophistiquées, un certain nombre de tâches résistent encore aux algorithmes et aux méthodes classiques de traitement des signaux et des données. Ces tâches relèvent typiquement du traitement, de très grands flots de données souvent multidimensionnelles et arrivant à des cadences élevées. Le grand nombre des données, leur variabilité, le fait qu'elles ne répondent pas à des modèles physiques connus nous laissent souvent démunis devant des tâches de caractérisation, de reconnaissance et de prise de décision.

Il y a des centaines d'exemples nous montrent à la fois combien on peut espérer de la modélisation du système nerveux mais aussi combien il sera difficile d'imaginer et de comprendre les divers aspects des problèmes de perception. Il paraît donc naturel d'essayer de comprendre comment les systèmes biologiques sont capables de telles performances, et si possible, de s'inspirer de leurs principes pour imaginer de nouveaux algorithmes ou de nouvelles machines plus efficaces que ceux dont nous disposons actuellement.

Les techniques de réseaux de neurones relèvent d'une telle approche : comprendre les principes selon lesquels les systèmes biologiques traitent l'information et s'en inspirer pour élaborer de nouvelles techniques en sciences de l'ingénieur. C'est donc une double démarche, à la fois cognitive et synthétique où le monde biologique doit être considéré comme une source de référence et de connaissance.

Les caractéristiques essentielles des réseaux de neurones réels que nous conserverons dans les modèles mathématiques étudiés, concernent le grand nombre de connexions, la non-linéarité des relations entrée-sortie et la faculté de "plasticité" ou d'adaptabilité. Ces caractéristiques, même simplifiées, leur confèrent déjà de multiples possibilités en traitement des signaux et des informations ainsi que la faculté d'apprendre à classer, à reconnaître des formes ou à réaliser des tâches complexes.

2. Historique

Des modèles de réseaux de neurones [Dreyfus et al, 2004] existent depuis longtemps. Ainsi le physiologiste viennois Sigmund Exner proposait en 1894 un modèle neuronal de la détection du mouvement par l'œil de la mouche. Cependant en 1943 W.S. McCulloch et Pitts [McCulloch et al, 1943] constituent certainement un des premiers jalons dans l'étude mathématique des réseaux de neurones formels. Dans le modèle de McCulloch et Pitts, chaque unité de traitement (ou neurone formel) calcule une somme pondérée des signaux formés par les autres unités auxquelles elle est connectée et ce résultat est comparé à une valeur de seuil. Si le seuil est franchi, le neurone est activé et cet état d'activation est propagé à travers le réseau pour les étapes de calcul ultérieures. Sous cette forme très simplifiée, les réseaux de neurones formels fonctionnent comme des systèmes itératifs discrets. A l'époque, cette idée a très vite fleuri et suscitent beaucoup d'intérêt parce qu'elle montrait que toute fonction logique pouvait être réalisée par un tel réseau de neurones formels. En 1949 le psychanalyste Donald Hebb proposait deux idées fondamentales [Hebb, 1949] :

- i) tout percept ou concept est physiquement représenté dans le cerveau par l'entrée en activité d'un ensemble de neurones (on parle de l'assemblée de Hebb) ;
- ii) deux neurones, ou deux ensembles de neurones, qui sont activés en même temps, vont finir par être «associés», de sorte que l'entrée en activité de l'un facilitera celle de l'autre. Ainsi la mise en mémoire s'effectue aux lieux d'interaction entre neurones, c'est-à-dire à la synapse (on parle de la synapse de Hebb).

Au début des années soixante, Rosenblatt [Rosenblatt, 1961], présenta le « Perceptron » qui ne différait par rapport au modèle de McCulloch et Pitts, que par sa faculté d'apprentissage basé sur la règle de Hebb, ce qui constitua une nouveauté à l'époque. Ce réseau qui était capable de prendre de bonnes décisions en classification a vite montré son inaptitude devant les problèmes non linéaires. En 1969 Papert et Minsky [Minsky et al, 1969] ont démontré clairement les limites du perceptron, et la nécessité de plusieurs couches identiques à celle du perceptron pour résoudre ces problèmes. N'ayant aucun moyen pour faire l'apprentissage de ces couches de neurones nécessaires, cette étude plongea l'idée de réseaux de neurones dans l'ombre [Simpson, 1990].

Durant les années 70, les travaux se sont orientés vers les mémoires associatives, qui sont des réseaux à apprentissage non supervisé à architecture généralement interconnectée. Anderson et Kohonen [Anderson, 1977; Kohonen, 1977] ont apporté des modèles très intéressants.

Cependant, se sont heurtés aux problèmes d'instabilité et de concepts liés aux systèmes dynamiques, que ni la théorie mathématique ni les moyens de calcul disponibles à l'époque, n'ont pu prendre en charge.

Il a fallu attendre 1982 où Hopfield [Hopfield, 1982] montra que, dans certaines conditions, le comportement dynamique des réseaux récurrents est gouverné par une fonction d'énergie décroissante le long d'une trajectoire. Ce qui implique que le réseau converge vers un minimum local de cette fonction d'énergie. Il en résulte que les réseaux récurrents peuvent être utilisés, non seulement comme mémoire associative, mais aussi pour résoudre certains problèmes d'optimisation. Le problème de stabilité des réseaux dynamiques est définitivement réglé par Cohen et Grossberg [Cohen et al, 1983]. Ce premier obstacle franchi, plusieurs travaux ont pu voir le jour.

Le plus grand événement qui est venu propulser les réseaux de neurones par la suite, est sans doute le résultat du travail d'un groupe de chercheurs [Rumelhart et al, 1986]. Ce travail a permis la résolution du vieux problème posé par Minsky et Papert, en mettant en œuvre l'algorithme d'apprentissage de « Back propagation ».

Désormais, il est possible d'effectuer l'apprentissage des réseaux de neurones statiques à plusieurs couches et de ce fait de s'attaquer aux problèmes non linéaires. A partir de là, les réseaux de neurones ont connu un essor continu. Ainsi à cet algorithme d'apprentissage, sont venus s'ajouter d'autres, qui lui ont apporté des améliorations [Trentelman et al, 1993]. Les architectures de ces réseaux ont eux aussi connu une très grande évolution donnant naissance à de nouveaux types de réseaux, parfois plus complexes, mais souvent plus adaptés à certains problèmes tels réseaux Gaussiens GPN (Gaussien Potential Network), les DBNN (Data Based Neural Network), les RBN (Radial Based Networks) [Renders, 1995], et les réseaux dynamiques.

Depuis une vingtaine d'années, les intérêts pour les RNA ont considérablement augmentés de sorte sont passé aujourd'hui du domaine de la recherche à des domaines applicatifs [Neuro Dimension, 2001].

3. Neurone formel

Le premier neurone formel est apparu en 1943. On le doit à Mac Culloch et Pitts.

Le neurone formel est donc une modélisation mathématique qui reprend les principes du fonctionnement du neurone biologique, en particulier la sommation des entrées. Sachant qu'au niveau biologique, les synapses n'ont pas toutes la même «valeur» (les connexions entre les neurones étant plus ou moins fortes), les auteurs ont donc créé un algorithme qui pondère la somme de ses entrées par des poids synaptiques (coefficients de pondération).

Voici un exemple de neurone formel :

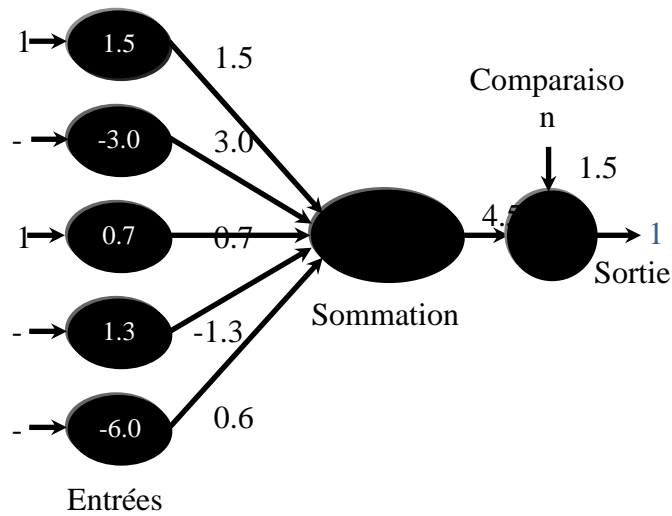


Figure 2.1. Exemple de neurone formel

3.1 Interprétation mathématique

D'un point de vue mathématique, le neurone formel peut être représenté de la manière suivante:

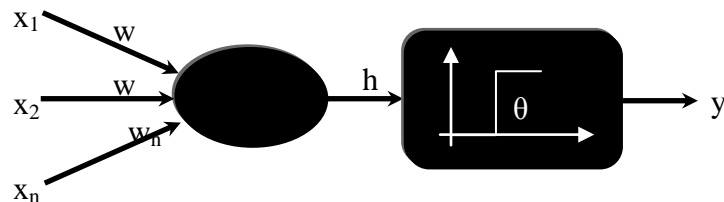


Figure 2.2. Représentation mathématique d'un neurone formel

Pour un nombre compris entre $j=1$ et un nombre quelconque n , le neurone formel va calculer la somme de ses entrées (x_1, x_2, \dots, x_n), pondérées par les poids synaptiques (w_1, \dots, w_n), et la comparer à son seuil θ . Si le résultat est supérieur au seuil, alors la valeur renvoyée est 1, sinon la valeur renvoyée est 0 [Krose et al, 1993].

D'où la formule (avec f = fonction seuil):

$$y = f\left(\sum_{j=1}^n w_j x_j - \theta\right) \tag{2.1}$$

3.2 Fonctions d'activation

La fonction d'activation, ou fonction de transfert, est une fonction qui doit renvoyer un réel proche de 1 quand les "bonnes" informations d'entrée sont données et un réel proche de 0 quand elles sont "mauvaises". On utilise généralement des fonctions à valeurs dans l'intervalle réel $[0,1]$. Quand le réel est proche de 1, on dit que l'unité (le neurone) est **active** alors que quand le réel est proche de 0, on dit que l'unité est **inactive**. Le réel en question est appelé la **sortie** du neurone et sera noté a . Si la fonction d'activation est linéaire, le réseau de neurones se réduirait à une simple fonction linéaire.

En effet, si les fonctions d'activations sont linéaires, alors le réseau est l'équivalent d'une régression multi-linéaire (méthode utilisée en statistiques).

L'utilisation du réseau de neurone est toutefois bien plus intéressante lorsque l'on utilise des fonctions d'activations non linéaires.

Il y a beaucoup de fonctions d'activations possibles, c'est à dire répondant aux critères que nous avons donnés, toutefois dans la pratique il y en a principalement deux qui sont utilisées :

Les fonctions de transfert les plus utilisées

- La fonction de Heaviside
- La fonction sigmoïde

La fonction de **Heaviside** est définie par : $\forall x \in R, g(x) = 1$ si $x \geq 0$, sinon 0

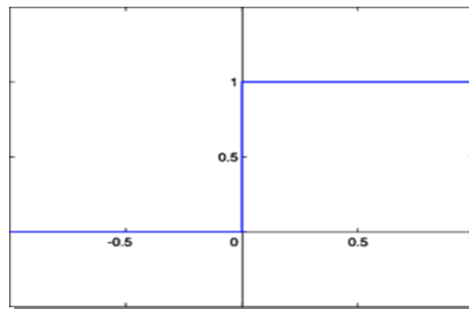


Figure 2.3. Graphe de la fonction de Heaviside

La fonction **sigmoïde** est quand à elle définie par : $\forall x \in \mathbb{R}, g(x) = \frac{1}{1 + e^{-x}}$

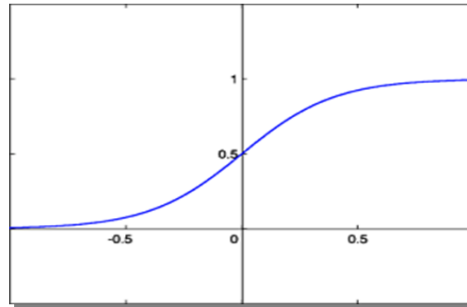


Figure 2.4. Graphe de la fonction Sigmoïde

La fonction sigmoïde présente l'avantage d'être dérivable (la fonction qui va être utilisée dans notre travail) ainsi que de donner des valeurs intermédiaires (des réels compris entre 0 et 1) par opposition à la fonction de Heaviside qui elle renvoie soit 0 soit 1. Toutefois, les deux fonctions possèdent un seuil.

4. Réseaux de neurones artificiels

4.1. Définition

Un RNA (Réseau de Neurones Artificiels) est un ensemble de neurones formels (d'unités de calcul simples, de nœuds processeurs) associés en couches (ou sous-groupes) et fonctionnant en parallèle.

Dans un réseau, chaque sous-groupe fait un traitement indépendant des autres et transmet le résultat de son analyse au sous-groupe suivant. L'information donnée au réseau va donc se propager couche par couche, de la couche d'entrée à la couche de sortie, en passant soit par aucune, une ou plusieurs couches intermédiaires (dites couches cachées). Il est à noter qu'en fonction de l'algorithme d'apprentissage, il est aussi possible d'avoir une propagation de l'information à reculons ("back propagation"). Habituellement (excepté pour les couches d'entrée et de sortie), chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante.

Les RNA ont la capacité de stocker de la connaissance empirique et de la rendre disponible à l'usage. Les habiletés de traitement (et donc la connaissance) du réseau vont être stockées dans les poids synaptiques, obtenus par des processus d'adaptation ou d'apprentissage. En ce sens, les RNA ressemblent donc au cerveau car non seulement, la connaissance est acquise au travers d'un apprentissage mais de plus, cette connaissance est stockée dans les connexions entre les entités, soit dans les poids synaptiques.

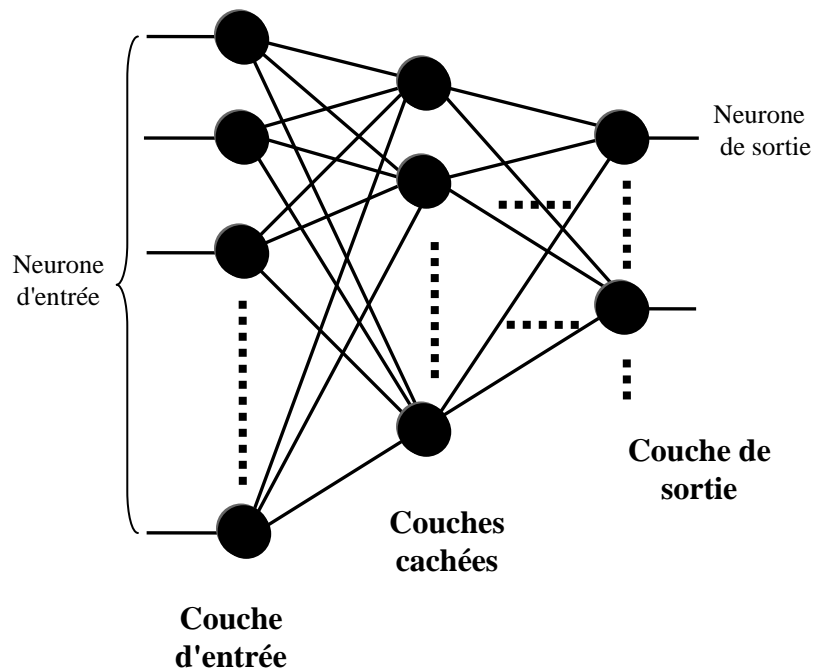


Figure 2.5. Réseau de neurone artificiel.

4.2 Architecture des réseaux

On peut classer les RNA en deux grandes catégories:

4.2.1 Réseaux "FEED-FORWARD"

Appelés aussi "réseaux de type Perceptron", ce sont des réseaux dans lesquels l'information se propage de couche en couche sans retour en arrière possible.

4.2.1.1 Perceptrons

a. Perceptron monocouche

C'est historiquement le premier RNA, c'est le Perceptron de Rosenblatt. C'est un réseau simple, puisque il ne se compose que d'une couche d'entrée et d'une couche de sortie. Il est calqué, à la base, sur le système visuel et de ce fait a été conçu dans un but premier de reconnaissance des formes. Cependant, il peut aussi être utilisé pour faire de la classification et pour résoudre des opérations logiques simples (telle "ET" ou "OU"). Sa principale limite est qu'il ne peut résoudre que des problèmes linéairement séparables. Il suit généralement un apprentissage supervisé selon la règle de correction de l'erreur [Jian-kang wu, 1993] (ou selon la règle de Hebb).

b. Perceptron multicouche (MLP)

C'est une extension du précédent, avec une ou plusieurs couches cachées entre l'entrée et la sortie. Chaque neurone dans une couche est connecté à tous les neurones de la couche précédente et de la couche suivante (excepté pour les couches d'entrée et de sortie) et il n'y a pas de connexions entre les cellules d'une même couche. Les fonctions d'activation utilisées

dans ce type de réseaux sont principalement les fonctions à seuil ou sigmoïdes. Il peut résoudre des problèmes non-linéairement séparables et des problèmes logiques plus compliqués, et notamment le fameux problème du XOR [Haykin, 1999]. Il suit aussi un apprentissage supervisé selon la règle de correction de l'erreur.

4.2.1.2 Les réseaux à fonction radiale

Ce sont les réseaux que l'on nomme aussi RBF ("Radial Basic Functions"). L'architecture est la même que pour les PMC cependant, les fonctions de base utilisées ici sont des fonctions Gaussiennes. Les RBF seront donc employés dans les mêmes types de problèmes que les PMC à savoir, en classification et en approximation de fonctions, particulièrement. L'apprentissage le plus utilisé pour les RBF est le mode hybride [Krose et al, 1993] et les règles sont soit, la règle de correction de l'erreur soit, la règle d'apprentissage par compétition.

4.2.2 Les réseaux "FEED-BACK"

Appelés aussi "réseaux récurrents", ce sont des réseaux dans lesquels il y a retour en arrière de l'information.

4.2.2.1 Les cartes auto-organisatrices de Kohonen

Ce sont des réseaux à apprentissage non-supervisé qui établissent une carte discrète, ordonnée topologiquement, en fonction de patterns d'entrée. Le réseau forme ainsi une sorte de treillis dont chaque nœud est un neurone associé à un vecteur de poids. La correspondance entre chaque vecteur de poids est calculée pour chaque entrée. Par la suite, le vecteur de poids ayant la meilleure corrélation, ainsi que certains de ses voisins, vont être modifiés afin d'augmenter encore cette corrélation [Jian-kang wu, 1993].

4.2.2.2 Les réseaux de Hopfield

Les réseaux de Hopfield sont des réseaux récurrents et entièrement connectés. Dans ce type de réseau, chaque neurone est connecté à chaque autre neurone et il n'y a aucune différenciation entre les neurones d'entrée et de sortie. Ils fonctionnent comme une mémoire associative non-linéaire et sont capables de trouver un objet stocké en fonction de représentations partielles ou bruitées. L'application principale des réseaux de Hopfield est l'entrepôt de connaissances mais aussi la résolution de problèmes d'optimisation. Le mode d'apprentissage utilisé ici est le mode non-supervisé.

4.2.2.3 Les ART

Les réseaux ART ("Adaptative Résonance Théorie") sont des réseaux à apprentissage par compétition. Le problème majeur qui se pose dans ce type de réseaux est le dilemme « stabilité/plasticité ». En effet, dans un apprentissage par compétition, rien ne garantit que les catégories formées vont rester stables. La seule possibilité, pour assurer la stabilité, serait que le coefficient d'apprentissage tende vers zéro, mais le réseau perdrait alors sa plasticité [Parizeau, 2004]. Les ART ont été conçus spécifiquement pour contourner ce problème. Dans ce genre de réseau, les vecteurs de poids ne seront adaptés que si l'entrée fournie est suffisamment proche, d'un prototype déjà connu par le réseau. On parlera alors de résonance. A l'inverse, si l'entrée s'éloigne trop des prototypes existants, une nouvelle catégorie, avec pour prototype, l'entrée qui a engendrée sa création. Il est à noter qu'il existe deux principaux types de réseaux ART : les ART-1 pour des entrées binaires et les ART-2 pour des entrées continues. Le mode d'apprentissage des ART peut être supervisé ou non.

5. Récapitulation

Les différents réseaux de neurones peuvent être classés comme dans le schéma ci dessous:

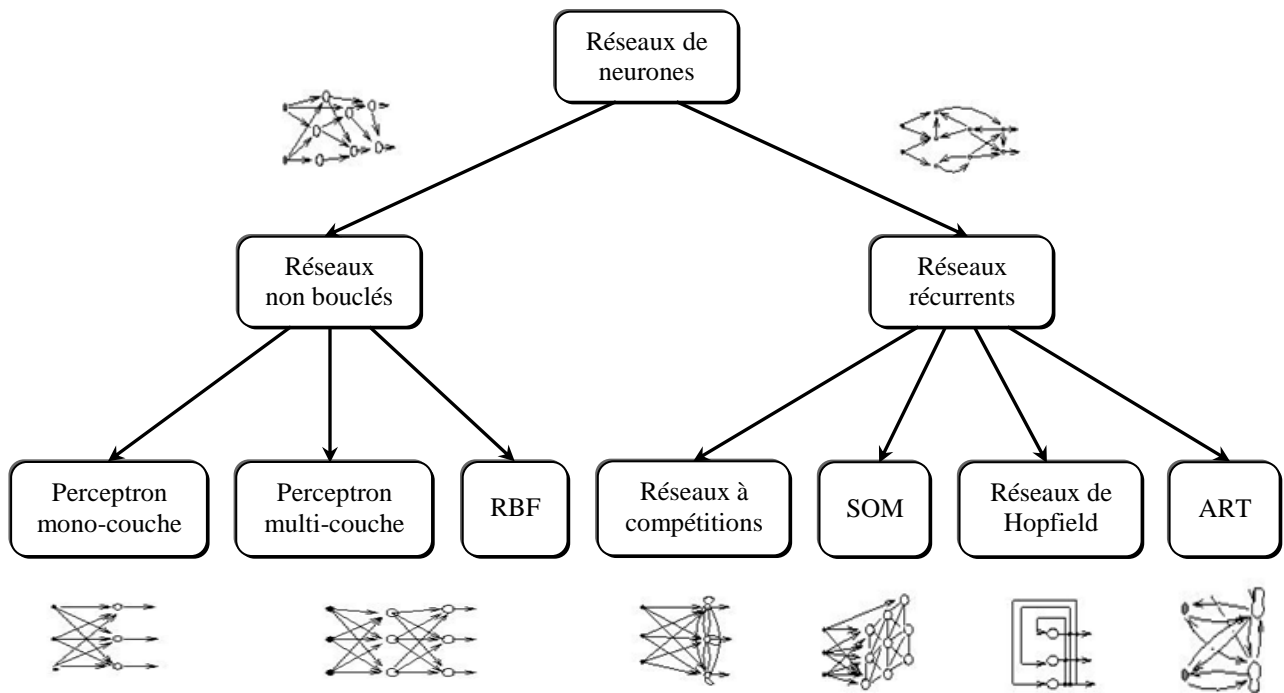


Figure 2.6. Schéma de différents réseaux de neurones.

6. Apprentissage

Pour un RNA, l'apprentissage peut être considéré comme le problème de la mise à jour des poids des connexions au sein du réseau [Huang et al, 1990], afin de réussir la tâche qui lui est demandée. L'apprentissage est la caractéristique principale des RNA et il peut se faire de différentes manières et selon différentes règles.

6.1 Types d'apprentissage

6.1.1 Le mode supervisé

Dans ce type d'apprentissage, le réseau s'adapte par comparaison entre le résultat qu'il a calculé, en fonction des entrées fournies, et la réponse attendue en sortie [Luo et al, 1989]. Ainsi, le réseau va se modifier jusqu'à ce qu'il trouve la bonne sortie, c'est-à-dire celle attendue, correspondant à une entrée donnée.

6.1.2 Le renforcement

Le renforcement est en fait une sorte d'apprentissage supervisé et certains auteurs le classe d'ailleurs, dans la catégorie des modes supervisés. Dans cette approche le réseau doit apprendre la corrélation entrée/sortie via une estimation de son erreur, c'est-à-dire du rapport échec/succès. Le réseau va donc tendre à maximiser un index de performance [R. Greiner et al, 1996] qui lui est fourni, appelé signal de renforcement. Le système étant capable ici, de savoir si la réponse qu'il fournit est correcte ou non, mais il ne connaît pas la bonne réponse.

6.1.3 Le mode non-supervisé (ou auto-organisationnel)

Dans ce cas, l'apprentissage est basé sur des probabilités. Le réseau va se modifier en fonction des régularités statistiques de l'entrée et établir des catégories, en attribuant et en optimisant une valeur de qualité, aux catégories reconnues.

6.1.4 Le mode hybride

Le mode hybride reprend en fait les deux autres approches, puisque une partie des poids va être déterminée par apprentissage supervisé et l'autre partie par apprentissage non-supervisé.

6.2 Règles d'apprentissage

6.2.1 Règle de correction d'erreurs

Cette règle s'inscrit dans le paradigme d'apprentissage supervisé, c'est -à-dire dans le cas où l'on fournit au réseau une entrée et la sortie correspondante. Si on considère y comme étant la sortie calculée par le réseau, et d la sortie désirée, le principe de cette règle est d'utiliser l'erreur $(d-y)$, afin de modifier les connexions et de diminuer ainsi l'erreur globale du système. Le réseau va donc s'adapter jusqu'à ce que y soit égal à d .

Ce Principe est notamment utilisé dans le modèle du perceptron simple.

6.2.2 Apprentissage de Boltzmann

Les réseaux de Boltzmann sont des réseaux symétriques récurrents. Ils possèdent deux sous-groupes de cellules, le premier étant relié à l'environnement (cellules dites visibles) et le second ne l'étant pas (cellules dites cachées). Cette règle d'apprentissage est de type stochastique (= qui relève partiellement du hasard) et elle consiste à ajuster les poids des connexions, de telle sorte que l'état des cellules visibles satisfasse une distribution probabiliste souhaitée.

6.2.3. Règles de Hebb

Comme je l'ai déjà dit dans l'historique, cette règle, basée sur des données biologiques, modélise le fait que si des neurones, de part et d'autre d'une synapse, sont activés de façon synchrone et répétée, la force de la connexion synaptique va aller croissant. Il est à noter ici que l'apprentissage est localisé, c'est-à-dire que la modification d'un poids synaptique w_{ij} ne dépend que de l'activation d'un neurone i et d'un autre neurone j .

6.2.4. Règle d'apprentissage par compétitions

La particularité de cette règle, c'est qu'ici l'apprentissage ne concerne qu'un seul neurone. Le principe de cet apprentissage est de regrouper les données en catégories. Les patrons similaires vont donc être rangés dans une même classe, en se basant sur les corrélations des données, et seront représentés par un seul neurone, on parle de « winner-take-all ».

Dans un réseau à compétition simple, chaque neurone de sortie est connecté aux neurones de la couche d'entrée, aux autres cellules de la couche de sortie (connexions inhibitrices) et à elle-même (connexion excitatrice) [Jian-kang wu, 1993]. La sortie va donc dépendre de la compétition entre les connexions inhibitrices et excitatrices.

6.3. Algorithme d'apprentissage supervisé pour les RNA : la rétropropagation

Quel que soit le cas de l'utilisation du réseau (approximateur de fonction ou classifieur), le processus d'apprentissage supervisé nécessite une base d'exemples composée des entrées du réseau associées aux sorties désirées. Pendant la phase d'apprentissage, les poids et biais du réseau sont ajustés pour minimiser la fonction de performance ou fonction de coût. La fonction de coût doit permettre de mesurer l'écart entre le modèle et les observations. Nous présentons ci-après la fonction d'erreur quadratique moyenne MSE qui est la plus utilisée pour l'apprentissage des RNA, mais d'autres fonctions de coût peuvent aussi être utilisées [Chauvin et al, 1995; Hagan et al, 1996]:

$$MSE = \frac{1}{N} \sum_i^N e_i^2 = \frac{1}{N} \sum_i^N (t_i - a_i)^2 \quad (2.2)$$

Où i parcourt les indices des neurones de sortie, a_i et t_i représentent respectivement l'activation mesurée et l'activation désirée pour ces neurones.

Les poids et biais du réseau sont initialisés, puis le processus itératif de l'apprentissage peut commencer à chaque itération :

- Les sorties du RNA sont calculées;
- la fonction de coût est mesurée;
- les poids et biais sont modifiés en utilisant l'algorithme de la rétropropagation;
- le critère d'arrêt est évalué.

L'arrêt de l'apprentissage est décidé lorsqu'un critère d'arrêt est atteint.

Nous présentons ici l'algorithme de rétropropagation du gradient [Chauvin et al, 1995] qui est le plus connu pour réaliser l'adaptation des réseaux multicouches. C'est à sa découverte que l'on doit le renouveau d'intérêt pour les RNA apparus au début des années 80. Il s'agit d'une méthode d'apprentissage supervisé, fondée sur la modification des poids du réseau dans le sens contraire à celui du gradient de l'erreur (Figure 2.7). La qualité d'apprentissage du réseau est évaluée par la fonction de performance décrite plus haut. A chaque itération la valeur de l'erreur quadratique MSE (équation 2.2) est renvoyée de la couche de sortie vers les couches cachées pour la mise à jour des poids et biais du réseau. L'algorithme de rétropropagation modifie les poids et les biais dans la direction de diminution de la fonction de performance (la valeur négative du gradient).

$$W_{k+1} = W_k - \eta_k G_k \quad (2.3)$$

W_k est le vecteur de poids et biais du neurone k .

η est le facteur d'apprentissage.

G est le gradient de l'erreur.

Cet algorithme, présenté ici dans sa version la plus simple, possède de nombreuses variantes. Elles correspondent généralement à l'utilisation de valeurs variables pour le facteur d'apprentissage η [Riedmiller et al, 1993], ou à l'utilisation de méthodes du deuxième ordre pour le calcul du gradient. Nous présenterons plus loin dans ce chapitre les principaux algorithmes utilisés en apprentissage des Perceptrons Multicouches [Fahlman, 1988].

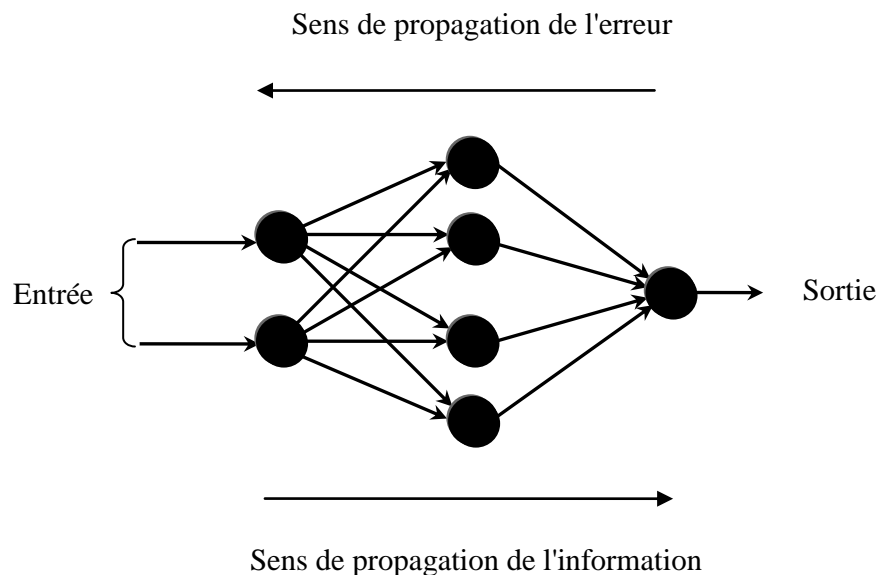


Figure 2.7. Un réseau multicouche comportant 2 neurones d'entrée, 4 neurones cachés et un neurone de sortie.

6.4. Généralisation

Pendant le processus d'apprentissage, les sorties d'un réseau non bouclé tendent à approcher les sorties fournies dans l'ensemble d'apprentissage. Mais l'intérêt principal des réseaux réside dans leur capacité de généralisation, en d'autres termes, de fournir des valeurs pour des entrées inconnues qui n'appartiennent pas à l'ensemble d'apprentissage. Pour avoir une bonne généralisation, il y a certaines règles à respecter, sans lesquelles les RNA donneront des résultats erronés. Les entrées choisies pour le réseau doivent être pertinentes et représentatives de la fonction recherchée ; nous ne pouvons pas inclure toute entrée sans lien avec les sorties fournies au risque d'augmenter artificiellement la complexité du modèle recherché.

Lorsque la séquence d'apprentissage est choisie, il faut s'assurer qu'elle soit représentative de la fonction que nous cherchons à approximer, et que les ensembles que nous essayons de prédire soient de la même classe que l'ensemble d'apprentissage.

6.4.1. Facteurs indispensables pour une bonne généralisation

Trois conditions sont nécessaires mais pas suffisantes pour obtenir une bonne généralisation:

1. Les entrées du réseau doivent contenir suffisamment d'information les reliant aux sorties désirées, pour qu'il y ait une relation mathématique reliant les entrées aux sorties avec une précision acceptable. Nous ne pouvons pas imaginer les RNA apprendre une relation non existante. Le choix des entrées pertinentes et la création d'un ensemble d'apprentissage, prennent souvent plus de temps que l'apprentissage des réseaux. Cela est particulièrement important lorsque nous essayons de modéliser des fonctions biologiques, où souvent les paramètres influençant le processus à apprendre sont difficiles à quantifier.
2. La fonction à modéliser (reliant les entrées aux sorties) doit être peu sensible. En d'autres termes, de faibles variations des valeurs d'entrées doivent, dans la plupart du temps, entraîner de faibles variations sur les sorties. Quelques réseaux de neurones sont capables d'apprendre des discontinuités à condition que la fonction à modéliser soit composée d'un nombre fini de segments continus. Cela explique l'incapacité des RNA d'apprendre certaines fonctions comme les tirages aléatoires et les algorithmes de cryptage. Pour les problèmes de classification, la condition de faible sensibilité n'est pas nécessaire. Les Perceptrons Multi-Couches avec suffisamment de neurones cachés sont des classificateurs fiables [Devroye et al, 1996].
3. L'ensemble d'apprentissage (base d'exemples) doit être suffisamment grand et représentatif de l'ensemble de l'espace des entrées et de sorties. Cette condition influence directement les deux types de généralisations possibles : l'interpolation et l'extrapolation. L'interpolation s'applique sur les cas proches des exemples d'apprentissage et l'extrapolation sur les cas loin de l'ensemble d'apprentissage. Les RNA sont capables d'une bonne interpolation. En revanche, leur performance en extrapolation reste assez médiocre. Avant d'utiliser les RNA en généralisation, leur performance en interpolation et en extrapolation doit être vérifiée.

En résumé, pour une fonction donnée, peu sensible, et pour les cas de généralisation proches des exemples d'apprentissage, la sortie du réseau sera proche des sorties correctes des exemples d'apprentissage. Si la base d'apprentissage est correctement construite, tout cas de généralisation sera proche d'un cas d'apprentissage. Sous ces conditions et avec une architecture et un algorithme d'apprentissage approprié, les RNA sont capables de généraliser correctement l'ensemble de la population.

6.4.2. Optimisation de la généralisation

Il est possible d'améliorer la qualité des réseaux de neurones en généralisation en modifiant les algorithmes d'apprentissage ou même les conditions d'arrêt du processus itératif de l'apprentissage. Nous présentons différents algorithmes d'apprentissage ainsi que la méthode « d'arrêt avant convergence » ou autrement dit la méthode du « early stopping ».

6.4.2.1. Algorithme d'apprentissage

Les algorithmes d'apprentissage itératifs utilisés dans les réseaux de neurones explorent l'espace des paramètres (poids et biais du réseau) à partir d'un point de départ fixé par les valeurs initiales de ces paramètres et selon une trajectoire qui sera fonction de l'algorithme utilisé et des données. Pour améliorer les performances des modèles neuronaux en généralisation, nous pouvons également utiliser des algorithmes basés sur la descente du gradient. Nous présentons ci-après les algorithmes les plus utilisés en apprentissage supervisé [Hagan et al, 1996].

a) Algorithme du facteur d'apprentissage adaptatif η

Il est aussi possible d'améliorer les performances du RNA en modifiant à chaque itération la valeur du facteur d'apprentissage η dans la formule modifiant les poids et les biais (équation. 3.2). Cette méthode est connue sous le nom de l'algorithme du «facteur d'apprentissage adaptatif». La performance du réseau est très sensible au facteur d'apprentissage: s'il est trop grand, l'algorithme peut osciller et devenir instable, s'il est trop petit, l'algorithme mettra beaucoup de temps avant de converger. Ce n'est pas évident de choisir une valeur avant l'apprentissage, d'autant plus que la valeur optimale du facteur d'apprentissage change pendant le processus d'apprentissage.

La performance de l'algorithme de descente du gradient peut être améliorée si on modifie le facteur d'apprentissage pendant le processus itératif. Un facteur adaptatif permet de maintenir un pas d'apprentissage le plus large possible en maintenant un apprentissage stable. A chaque itération, l'erreur d'apprentissage (e_i) (équation 2.2) est calculée. La méthode du facteur d'apprentissage adaptatif propose de modifier η au fur et à mesure de l'apprentissage en fonction de e_i selon la méthode suivante :

Si	$e_i > \alpha$	alors	
	$\eta = \eta * \delta$		$(\delta < 1)$
Sinon			
	$\eta = \eta * \delta'$		$(\delta' > 1)$
Fin Si			

b) Méthode de Newton

La méthode de Newton utilise la courbure (dérivée seconde) de la fonction de coût pour atteindre le minimum plus rapidement. La modification des paramètres est donnée par :

$$W_{k+1} = W_k - H_k^{-1} G_k \quad (2.4)$$

La direction de descente est fonction du Hessien H (matrice des dérivées secondes de la fonction de coût) et du gradient. La méthode de Newton nécessite moins d'itérations pour converger que l'algorithme de rétropropagation classique, mais malheureusement, le calcul du

Hessien est coûteux en temps pour les réseaux non bouclés. Il existe cependant une classe d'algorithmes basée sur l'algorithme de Newton appelée Quasi Newton qui utilise une approximation du Hessien pour la mise à jour des poids du réseau. Le plus connu parmi ces algorithmes est celui de Broyden, Fletcher, Goldfarb et Shanno (BFGS) [Dennis et al, 1983; Hagan et al, 1996].

c) *Méthode de Levenberg-Marquardt*

Comme les méthodes de Quasi-Newton, la méthode de Levenberg-Marquardt [More, 1977] utilise une approche du second ordre sans calculer le Hessien. Quand la fonction de coût est une somme au carré (i.e. comme le MSE équation 3.1), le Hessien peut être approximé par :

$$H = J^T J \quad (2.5)$$

Le gradient peut être décomposé sous la forme :

$$G = J^T e \quad (2.6)$$

Où J est le Jacobien qui contient les dérivées premières des erreurs du réseau

et e est le vecteur d'erreur

$$e_i = a_i - s_i \quad (2.7)$$

Le jacobien peut être calculé avec l'algorithme de rétropropagation; ce qui est beaucoup moins complexe que le calcul du Hessien.

6.4.2.2. **Compromis Apprentissage Généralisation ou « early stopping »**

Pour les réseaux de neurones artificiels, l'ensemble des données disponibles est souvent divisé en trois sous-ensembles distincts, un ensemble d'apprentissage, un ensemble de validation et un ensemble de test. La distinction entre ces sous-ensembles est cruciale, mais les termes validation et test sont souvent confondus. Nous présentons les définitions données par Ripley [Ripley, 1996] :

Ensemble d'apprentissage

Un ensemble d'exemples utilisé en apprentissage pour la modification des paramètres du réseau (poids et biais).

Ensemble de validation

Un ensemble d'exemples utilisé pour l'ajustement de l'architecture (nombre de couches et nombre de neurones cachés mais pas les poids du réseau). Cet ensemble peut aussi servir à arrêter le processus itératif de l'apprentissage lorsqu'un seuil d'erreur est atteint.

Ensemble de test

Un ensemble d'exemples utilisé pour tester la performance du réseau.

7. Récapitulation

Voici un tableau récapitulatif des différents types d'apprentissages:

<i>Paradigme</i>	<i>Règle d'apprentissage</i>	<i>Architecture</i>	<i>Algorithme</i>	<i>Tâches</i>
<i>Supervisé</i>	Correction d'erreur	Perceptron Simple ou Multi-couche	Perceptron, Rétro-Propagation, Adaline, Madaline	Classification, Approximation de fonctions, Prédiction, Contrôle
	Boltzmann	Récurrente	Apprentissage de Boltzmann	Classification
	Hebb	Multi-couche non bouclés	Analyse de discriminants linéaires	Analyse de données, Classification w
	par compétition	à compétition	LVQ	Catégorisation au sein d'une classe, Compression de données
		ART	ARTMap	Classification, Catégorisation au sein d'une classe
<i>Non supervisé</i>	Correction d'erreur	Multi-couches non bouclé	Projection de Sammon	Analyse de données
	Hebb	Non bouclé ou à compétition	analyse en composantes principales	Analyse de données, compression de données
	par compétition	à compétition	VQ	Catégorisation, Compression de données
		Cartes de Kohonen	SOM	Catégorisation, Analyse de données
		ART	ART-1, ART-2	Catégorisation
<i>Hybride</i>	Correction d'erreur et par compétition	RBF	RBF	Classification, Approximation de fonctions, Prédiction, Contrôle

8. Mise en œuvre des réseaux neuronaux

Nous allons suivre une démarche reprise par Wierenga et Kluytmans (1994) qui est composée de quatre étapes principales :

Etape 1 : fixer le nombre de couches cachées

Mis à part les couches d'entrée et de sortie, l'analyste doit décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation ; avec une couche cachée, il est capable, avec un nombre suffisant de neurones, d'approximer toute fonction continue [Hornik, 1994]. Une seconde couche cachée prend en compte les discontinuités éventuelles.

Etape 2 : déterminer le nombre de neurones par couches cachées

Chaque neurone supplémentaire permet de prendre en compte des profils spécifiques des neurones d'entrée. Un nombre plus important permet donc de mieux coller aux données présentées mais diminue la capacité de généralisation du réseau. Ici non plus il n'existe pas de règle générale mais des règles empiriques. La taille de la couche cachée doit être :

- Soit égale à celle de la couche d'entrée.
- Soit égale à 75% de celle-ci.
- Soit égale à la racine carrée du produit des nombres dans la couche d'entrée et de sortie.

Notons que le dernier choix réduit le nombre de degrés de liberté laissés au réseau, et donc la capacité d'adaptation sur l'échantillon d'apprentissage, au profit d'une plus grande stabilité. Une voie de recherche ultérieure consisterait soit à procéder à l'estimation d'un réseau comportant de nombreux neurones puis à le simplifier par l'analyse des multicollinéarités ou par une règle d'apprentissage éliminant les neurones inutiles ; soit à définir une architecture tenant compte de la structure des variables identifiée au préalable par une analyse en composantes principales.

Etape 3 : choisir la fonction d'activation

Nous considérerons la fonction logistique pour le passage de la couche d'entrée à la couche cachée. Le passage de cette dernière à la couche de sortie sera soit linéaire, soit sigmoïde (logistique) selon nos types de variables.

Etape 4 : choisir l'apprentissage

L'apprentissage par rétro-propagation nécessite la détermination du paramètre d'ajustement des poids synaptiques à chaque itération.

La détermination du critère d'arrêt est aussi cruciale dans la mesure où la convergence peut passer par des minima locaux.

9. Avantage des RNA sur les méthodes statistiques

Pourquoi alors utiliser les réseaux de neurones artificiels (le Perceptron Multicouches (PMC)) au lieu des méthodes statistiques pour l'approche des modèles non linéaires ? Il y a beaucoup de réponses à cette question [Hornik, 1994; Ripley, 1996]:

- L'utilisation d'un nombre d'entrée relativement grand.
- la plupart de l'information prédictive est contenue dans un sous-espace à faible dimension.
- capacité de représentation de n'importe quelle fonction, linéaire ou pas, simple ou complexe.
- Faculté d'apprentissage à partir d'exemples représentatifs, par « rétropropagation des erreurs ». L'apprentissage (ou construction du modèle) est automatique
- Simple à manier, beaucoup moins de travail à fournir que dans l'analyse statique classique.
- Pour l'utilisation nocive, l'idée d'apprentissage est plus simple à comprendre que les complexités des statistiques multi variables.
- Consultation rapide.
- Relations spatiales faciles à analyser et modéliser.
- Le modèle fourni est continu et dérivable et se prête bien à la visualisation, à une recherche de profils-type (par ascension du gradient de probabilité prédite), au calcul de sensibilités (par dérivation partielle), etc....

Les avantages des PMC sur les méthodes statistiques sont détaillés dans la référence [Warren, 2001].

Comme nous l'avons vu, les réseaux de neurones sont essentiellement des procédures permettant d'approcher n'importe quelle fonction linéaire ou non, à partir d'exemples de valeurs de cette fonction. La réussite de cet apprentissage suppose cependant qu'un certain nombre de conditions portant sur le choix et le nombre de ces exemples ainsi que sur l'architecture initiale du réseau, soient vérifiées.

10. Applications

Se trouvant à l'intersection de différents domaines (informatique, électronique, science cognitive, neurobiologie et même philosophie), l'étude des réseaux de neurones est une voie prometteuse de l'Intelligence Artificielle, qui a des applications dans de nombreux domaines :

- Industrie : contrôle qualité, diagnostic de panne, corrélations entre les données fournies par différents capteurs, analyse de signature ou d'écriture manuscrite...
- Finance : prévision et modélisation du marché (cours de monnaies...), sélection d'investissements, attribution de crédits...
- Télécommunications et informatique : analyse du signal, élimination du bruit, reconnaissance de formes (bruits, images, paroles), compression de données...
- Environnement : évaluation des risques, analyse chimique, prévisions et modélisation météorologiques, gestion des ressources...

Chapitre 3

Algorithme de détection de l'onde P

Chapitre 3

Algorithme de détection de l'onde P

1. Introduction

C'est seulement pendant la dernière décennie du siècle dernier que des systèmes interactifs d'aide au diagnostic ont été régulièrement employés dans la pratique clinique à une échelle significative. Parmi les méthodes qui ont abouti à de bons résultats : Les approximations des intégrales de distribution appliquée avec des méthodes numériquement robustes comme la chaîne de Markov Monté Carlo [Neal, 1996]. Ces développements ont été précurseurs d'intérêt pour les modèles graphiques [Lunn et al, 2000; Spiegelhalter et al, 1999] provenant des domaines de l'Intelligence artificielle, du modèle statistique et de l'étude assistée par l'ordinateur. Plus récemment des techniques qui impliquent les réseaux de neurones artificiels.

1.1. Aperçus précédents des applications médicales impliquant les réseaux de neurones

Les premiers tentatives d'application des réseaux de neurones au domaine médical se sont concentrées sur l'analyse des données non linéaires, le partage de la mémoire associative afin d'éviter la difficulté d'acquisition de la connaissance experte, la tolérance au bruit due à l'architecture intrinsèquement parallèle et l'adaptabilité à accommoder de nouvelles manifestations de la maladie [Reggia, 1993].

Beaucoup d'articles ont souligné la prédominance, dans la pratique clinique, d'une fonction de mémoire associative [Cross et al, 1995] adaptée à la variabilité des présentations de la maladie et des degrés de sévérité qui combinent des effets inhérents aux patients eux-mêmes et entre les patients les uns aux autres.

Cependant, technologiquement, les réseaux de neurones demeurent une partie « du sujet mystérieux de l'inférence et de la modélisation statistiques non linéaires » [Cross et al, 2000] avec l'intention de rendre ce domaine important et de développer l'accès aux statisticiens non-spécialiste. On pourrait argumenter que des principes statistiques sont importants pour compter sur l'évidence de bases fondées sur toute méthodologie de base de données, y compris les réseaux de neurones [Schwartzter, et al, 2000].

Vers 1995, il y a eu près de 1000 citations des réseaux de neurones dans la littérature biomédicale [Baxt, 1990; 1992; 1994; 1995; 1996; Baxt et White, 1995; Baxt et Skora, 1996]. La plupart des études décrites comportaient des données historiques et le plus souvent un taux d'exactitude (accuracy) sur un ensemble de données a été testé. Un bon exemple est celui de la présélection des patients qui se présentent aux unités des urgences avec des douleurs thoracique aiguës afin de tenter l'identification de ceux qui sont les plus susceptibles d'avoir fait un infarctus du myocarde aigu (IDM). Les études faisant participer 706 patients ont changé les sensibilités et les spécificités de 80 à 96% [Baxt et White, 1995], en comparaison avec une étude d'envergure des médecins qui a revendiqué une sensibilité ainsi qu'une spécificité de 88 et de 71% respectivement [Goldman et al, 1988].

Un autre aperçu de la valeur diagnostique des différentes procédures peut être dérivé en factorisant le rapport de la vraie positivité et de la fausse positivité afin de définir comment le discriminateur s'améliore sur la ligne d'estimation. Ceci définit un facteur amplifiant, entre la prédominance et la valeur prédictive positive (PPV) du procédé, donné par l'expression (3.1) ci-dessous.

$$\frac{PPV}{1-PPV} = \frac{Sensitivité}{1-Spécificité} * \frac{Pr(class)}{Pr(-class)} \quad 3.1$$

$Pr(class)$ et $Pr(-class)$ dénotent respectivement la prédominance en dehors et dans la classe diagnostique.

Afin de rendre maniable un sujet aussi vaste, l'étude des applications se consacre au domaine de la cardiologie ainsi que les fonctions cliniques de l'analyse du diagnostic, du pronostic et de la survie. Cet aperçu omet délibérément des domaines d'application importants qui mériteraient d'être étudiés. Par exemple, l'oncologie, le traitement critique et l'identification de modèle de surveillance. Cet aperçu fournit un instantané étroitement concentré sur les réseaux de neurones, à l'exclusion des méthodes hybrides telles que les systèmes neuro-flou [Szczepaniak et al, 2000].

Un aperçu de quelques publications classées par domaine médical et par fonction clinique est illustré dans les tables 3.1 et 3.2. Les domaines de l'oncologie, du soin critique et de la cardiologie sont ceux où les épreuves cliniques sont prédominantes. La plupart des études citées s'intéressent à l'extrapolation diagnostique ou pronostique avec seulement une épreuve clinique visant directement la survie.

Table 3.1. Aperçu de quelques publications classées par domaine médical et par fonction clinique.

	Oncologie	Soin critique	neurologie	Autre
Radiologie	IRM d'osteosarcoma: Glass and Reddick (1998)	Scintigraphie de perfusion pour la détection de la sténose coronaire : Goodenday et al. (1997) Le signal micro embolique du Doppler chez les patients avec des valves prothétiques du coeur : Georgiadis et al. (1995)	l'épilepsie : Bakken et al. (1999) Alzheimer : Horwitz et al. (1995)	Images myocardiques de perfusion: Lindahl et al. (2000) Détection des sténoses par écho Doppler: Smith et al. (1996)
Surveillance physio- logique		EEG en pédiatrie: Si et al. (1998)	Épreuve simple des PEVE Heinrich et al. (1999) Corrélation d'EEG et d'EMG: Gaetz et al. (1998) Potentiels évoqués dans la sclérose en plaque: Guterman et al. (1996)	Surveillance foetale pendant le travail à partir d'ECG foetal : Stock et al. (1994)
Diagnostic	cancer de la prostate: Gamito et al (2000) Finne et al. (2000) Cancer cervical : Doornewaard et al. (1999) Mango et al. (1998) Sherman et al. (1997) Cancer du sein: Naguib et al. (1996) Simpson et al. (1995) Leucémie aiguë : Kothari et al. (1996)	Hémorragie intracrânienne chez les nouveau-nés : Zernikow et al. (1998)	Détection d'embole: Kemeny et al. (1999) EEG Spontané: Anderson et al. (1998) EEG au sommeil: Baumgart-Schmitt et al. (1997, 1998) EEG Quantitatif: Winterer et al. (1998) Identification du mode de ventilation : Leon et Lorini (1997)	Appendicite : Pesonen (1997) Méthodes de référence pour des patients avec les troisièmes molaires : Goodey et al. (2000) Modèles de protéine de larme : Grus et Augustin (1999) Hémodialyse: Smith et al. (1998) moment d'ovulation: Gurgen et al. (1995)

Table 3.2. Aperçu de quelques publications classées par domaine médical et par fonction clinique.

	Oncologie	Soin critique	neurologie	Autre
Prévision de résultats	Réponse à la thérapie dans le cancer de la tête et du cou : Bryce et al. (1998) Répétition du cancer du sein chez les patientes avec nœud négatifs axillaires : McGuire et al. (1992) Myélome multiple: Bugliosi et al. (1994)	Longueur de séjours chez les nouveau-nés prématurés : Zernikow et al. (1999)		Niveaux de sang de Tracolimus : Chen et al. (1999) Croissance en calculs après lithotripsie : Michaels et al. (1998) Effet de traitement dans la schizophrénie et la dépression : Modai et al. (1996) conséquence de rupture d'anévrisme: Dombi et al. (1995)

Il y a également des domaines spécialisés liés au traitement des signaux auxquels il faudrait consacrer des études à part, tels que le traitement d'image et l'analyse en temps réel, y compris les développements dans l'analyse des composants pour la suppression d'artefact. [Lisboa et al, 2000a]. Un autre domaine d'application, la bioinformatique, est actuellement en développement en étroite corrélation avec la biomédecine [Shi et al, 2000].

2. Aperçu sur les applications cardiovasculaires

2.1. Application dans la détection du complexe QRS

L'électrocardiographie a eu un impact profond sur la pratique de la médecine. En effet, l'électrocardiogramme (ECG) contient une quantité de l'information importante qui peut être exploitée de différentes façons. L'ECG tient compte de l'analyse des aspects anatomiques et physiologiques du muscle cardiaque de totalité.

Une application des réseaux de neurones pour la discrimination entre les complexes normaux et pathologiques d'ECG est présentée dans [Conde, 1994]. La rétropropagation est employée comme algorithme d'apprentissage pour la méthode : La sortie du réseau n'utilisait pas la discrimination pour distinguer entre les complexes normaux et pathologiques, mais fait seulement l'extrait de prototype du complexe dans l'ECG analysé. Une tentative est faite pour individualiser automatiquement une morphologie pathologique du QRS, mais les battements ventriculaires étaient moins de 15% dans les ECG traités.

Dans [Dassen et al, 1992] les auteurs décrivent le développement d'un réseau de neurones conçu pour différencier l'étiologie d'un large QRS d'une tachycardie en utilisant les douze dérivations de l'ECG. De cette façon 75% de toutes les tachycardies ont été diagnostiquées correctement.

Dans le domaine du traitement des signaux en temps réel, il y a les systèmes tout à fait différents dans leur détection du complexe QRS.

Dans [Casaleggio et al, 1991], les auteurs ont développé un algorithme pour la détection des complexes QRS des signaux ambulatoires d'électrocardiogramme. Le seuil qui est obtenu à partir d'une fonction de distribution de l'amplitude du signal filtré d'ECG, change avec du temps, s'adaptant au changement des niveaux de morphologie du QRS du bruit et des artéfacts. Le seuil extrait la période qui inclut le complexe QRS et alors ce dernier est détecté pendant cette période.

2.2. Application dans la détection de l'infarctus du myocarde aigu (IDM)

Le diagnostic précoce de l'infarctus du myocarde aigu (IDM) chez les patients se présentant aux urgences avec des douleurs sévères au thorax suscite beaucoup d'intérêt pour l'application de la reconnaissance des formes. L'intention principale est d'employer les marqueurs biochimiques d'un échantillon de sang pour prévoir le résultat des mesures de protéines qui durent plusieurs heures pour être obtenus. L'intérêt quand l'IDM est détecté tôt, permet d'administrer l'anticoagulant au moment où il a le plus grand effet tout en réduisant la sévérité des dommages pouvant affecter le muscle cardiaque. Plusieurs études se concentrent sur ce thème qui, avec la radiologie, sont énumérées dans le tableau 3.1.

Dans une première étude, Ellenius et ses collaborateurs [Ellenius et al, 1997] ont suivi le diagnostic d'un patient avec un IDM mineur dès le moment de l'infarctus, en surveillant l'augmentation de la concentration des marqueurs biochimiques et en identifiant l'étape à laquelle le MLP, et chacun de trois cliniciens professionnels, confirme le diagnostic. Cette approche peu commune à l'évaluation de système a montré le modèle de détection de l'IDM et la capacité de prévoir l'étendue de l'infarctus dans le futur, simultanément avec les premières indications fournies par les experts.

Une étude à grande échelle d'interprétation automatisée des 12 dérivations de l'électrocardiogramme pour la détection de l'IDM, a été effectuée avec une multitude de patients présents dans un seul hôpital sur une période de cinq ans, comportant 1120 cas confirmés et 10.452 contrôlés. Un tracé de 20 secondes a été représenté par six mesures automatiquement produites de ST-T de chacune des 12 dérivations de l'ECG, fournissant 72 unités d'entrée d'un MLP avec une seule couche cachée [Heden et al, 1997]. Le procédé employé donne 15.5% (de 12.4 à 18.6%) d'amélioration de la sensibilité employée au-delà des critères basés sur les règles utilisées par les électrocardiographes numériques et informatisés au service des urgences.

La comparaison avec les critères basés sur les règles automatisées représente une augmentation de 50% du facteur amplifiant à partir d'environ six à neuf, alors que dans la comparaison avec les cliniciens le changement atteint 20% à un gain de probabilité d'environ 3.5. Les améliorations se détériorent à des spécificités plus élevées mais ces résultats ont néanmoins été rapportés avec enthousiasme dans les journaux médicaux [Fricker, 1997; Josefsen, 1997]. Les auteurs font remarquer que le système pourrait être d'un intérêt particulier pour des internes aux urgences, et la prochaine étape serait d'intégrer l'information contextuelle additionnelle telle que les antécédents médicaux et les résultats des enquêtes cliniques. Certaines de ces idées sont développées plus loin dans les travaux décrits ci-après.

Une étude méthodologique avancée de détection de l'IDM dans les services des urgences en utilisant les réseaux de neurones, comporte une série d'études par Baxt et collaborateurs. Des premières études qui devaient optimiser la précision des prévisions de réseau neurones [Baxt, 1992; 1994] ont été suivies d'une analyse bien soignée des différentes entrées cliniques pour avoir une bonne décision à la sortie du réseau [Dreiseitl et al, 1999], et de l'application de méthodologies rigoureuses pratiques pour l'analyse de la sensibilité [Baxt, 1995]. Il est d'un intérêt particulier d'utiliser un circuit fermé pour corriger l'erreur en effet, un biais a été utilisée au niveau de l'apprentissage des données permettait une bonne évaluation de la sensibilité.

L'analyse des entrées par le calcul des biais corrige la sensibilité chez Baxt et White [Baxt et al, 1995], classe les nouvelles variables mieux que certains indicateurs couramment utilisés par les cliniciens. Le modèle obtenu est conforme à une autre l'étude, celle des choix des variables pour la prédiction de l'IDM [Neal, 1996] comparant le LogR (multiple logistic regression), les réseaux de neurones bayésiens [Selker et al 1995] et les ensembles approximatifs. Le choix de 43 variables est convenu pour chaque approche modélisée. Seul un indicateur, sus décalage du segment ST, a été sélectionné pour toutes les méthodes. Ceci est reconnu cliniquement comme très approprié, mais trois autres variables communément employées par les cliniciens n'ont été choisies par aucun de ces modèles, à savoir antécédents de diabète, douleur de poitrine intense et durée de la douleur. Les résultats des études sont en conformité avec ces variables, suggérant qu'elles peuvent être sensibles à l'IDM.

Les études initiales ont été suivies d'une comparaison prospective entre les taux de détection par des cardiologues et le MLP pour un groupe de patients adultes présentant des douleurs thoraciques antérieures. Encore une fois, ces études se limitent à un seul hôpital [Baxt et Skora, 1996]. C'est une étude d'observation dans laquelle les cliniciens et le système de décision ont séparément examiné les mêmes patients. Elle revendique de façon discutable la plus grande augmentation dans la performance par un système du réseau de neurones comparé aux cliniciens experts, amplifiant la probabilité de prendre un diagnostic au-dessus de la prévalence.

Une épreuve multicentre plus ancienne impliquant les services des urgences de six hôpitaux a comparé trois structures parfaitement différentes en vue d'une classification, notamment l'induction régulière, LogR et le MLP, pour la prédiction de l'ischémie cardiaque aiguë (ICA), qui comprend l'IDM et l'angine de poitrine instable, à partir de huit variables [Polak et al. 1997].

Les variables représentent l'historique du patient ainsi que les signes extraits à partir d'un examen clinique et d'un électrocardiogramme. Le MLP a eu un étalonnage sensiblement plus faible que les autres méthodes. Il est également plus difficile à interpréter. On a conclu à partir de cette étude que le choix de la méthode du support de décision de base de données devrait être fait sur la base des besoins d'application spécifiques. Néanmoins, les résultats obtenus indiquent qu'une limitation importante de la performance prédictive pour ces résultats cliniques est la disponibilité de données fiables, plutôt que le besoin de développements algorithmiques.

Dans [Henrik et al, 2004] les auteurs ont employé les réseaux de neurones bayésiens comme outil pour détecter l'IDM en utilisant les 12 dérivations de l'ECG. Les six mesures suivantes, prises de chacune des 12 dérivations, ont été choisies pour une analyse meilleure utilisant les

réseaux de neurones. En outre, pour expliquer le raisonnement derrière la sortie du réseau de neurones, une méthode a été développée pour montrer les régions de l'ECG les plus pertinentes pour ce cas particulier. Les principaux ingrédients de cette approche sont : une représentation de l'ECG en utilisant des fonctions de Hermite, déjà utilisées par [Sornmo, 1981; Jane. R, 1993], et une formulation combinatoire du problème d'optimisation pour trouver les entrées importantes du réseau.

Chaque dérivation de l'ECG est exprimée comme série de fonctions de base de Hermite, et les coefficients de cette série sont employés comme entrées pour le réseau. Après l'apprentissage, un nombre restreint de coefficients pour un ECG indiqué sont perturbés dans un intervalle limité; ces perturbations sont choisies afin de maximiser le changement correspondant du diagnostic (sortie du réseau). Un ECG perturbé est reconditionné à partir des coefficients perturbés, et en comparant cet ECG perturbé à original. La possibilité de construire les ECGs perturbés à partir des entrées perturbées du réseau de neurones est l'une des vertus de la décomposition de la fonction de Hermite; aucune reconstruction n'est possible dans l'approche traditionnelle où l'extraction de dispositif se compose des mesures des amplitudes et des pentes.

L'étude de la population a été également étudiée en [Hedén et al, 1997] où une définition détaillée des critères pour le diagnostic de l'infarctus du myocarde aigu peut être trouvée. La performance était légèrement meilleure en employant des mesures traditionnelles d'amplitude et de pente du complexe STT (84.3% comparés à 83.4%). La performance employant des polynômes de Hermite était conforme également aux résultats se trouvant dans l'article de Heden et ses collaborateurs.

2.3. Application utilisant le système de surveillance Holter

Une autre application tout à fait différente est de prévoir la probabilité de développer une ischémie myocardique transitoire pendant une période de surveillance Holter ambulatoire, en utilisant des paramètres issus d'un ECG 12-lead au repos [Polak et al, 1997]. C'est l'exemple d'une étude où le MLP traité par la rétropropagation a été surpassé par une analyse linéaire de discriminants et un modèle alternatif au MLP.

Une étude de la prédiction de QT en fonction de RR utilisant les réseaux de neurones artificiels a été effectuée par [Hamidi, 1995]. Elle consistait à comparer les résultats des modèles paramétriques et neuronaux. Les performances des modèles paramétriques ont été jugées insuffisantes par l'auteur; cela est dû à la non linéarité de la relation $QT=f(RR)$. Les réseaux de neurones multicouches utilisant l'algorithme de rétropropagation dans la phase d'apprentissage, sont capables de modéliser la relation entre RR et QT. Le modèle utilisé avait pour entrée une série de RR sur une durée de 1 à 8 minutes et pour la sortie le QT correspondant au dernier RR de la série. Cette étude a mis en évidence la fiabilité de la modélisation neuronale et établit qu'un retard de 4 minutes est suffisant pour identifier la relation entre QT et RR. Ces premiers résultats ont servi de base pour un autre travail qui consiste à établir des modèles neuronaux prédictifs du comportement dynamique de l'intervalle QT en fonction des variations instantanées de l'intervalle RR et à étudier le comportement et le domaine de validité de ces modèles [El Dajani, 2002].

Table 3.3. Aperçu de quelques publications utilisant les réseaux de neurones appliqués à l'ECG.

Reference	Clinical function
<i>IDM Ellenius et al. (1997)</i>	Diagnostic précoce ou exclusion de l'IDM à partir de marqueurs biochimiques L'approche de MLP a pu fournir l'appui utile pour l'évaluation des patients en IDM suspecté
<i>IDM Baxt and White (1995)</i>	Détection d'IDM dans les départements d'urgence, à partir de 19 variables représentant l'historique du patient, les résultats cliniques et les mesures ST-T prises de l'ECG
<i>IDM Haraldsson et al. (2004)</i>	Un réseau de neurones pour détecter des signes d'infarctus du myocarde aigu dans l'ECG est employé. Les 12-dérivations d'ECG sont décomposées en fonctions de base de Hermite, et les coefficients résultants sont employés comme entrées au réseau de neurones
<i>Ischémie Polak et al. (1997)</i>	La prédiction de l'ischémie passagère pendant la surveillance par le Holter en ambulatoire des 12 dérivations de l'ECG.
<i>Ischemia Selker et al. (1995)</i>	Des indicateurs cliniques disponibles dans un délai de 10 minutes de soin au département d'urgence ont été employés pour prévoir l'IDM et l'angine de poitrine instable en temps réel
<i>Repolarisation ventriculaire R. M. EL Dajani (2001,2002)</i>	Les résultats indiquent que MLP peut approcher l'aspect non linéaire du rapport QT-RR, et modeler la réponse du comportement dynamique de $QT=f(RR)$ à différents intervalles fixes de RR
<i>BAV Meghriche et al (2006); (2007); (2008)</i>	Une performance élevée d'architecture hybride de réseaux de neurones (CNN) pour la détection de BAV en utilisant seulement une dérivation.
<i>SAO Haiying et al (2004)</i>	Décrire une plateforme principale du modèle et de visualisation de classification basée sur les réseaux neurologiques adaptatifs
<i>SAO/CHF A. Hossen (2006)</i>	L'identification entre les patients et les sujets normaux était concentrée sur les deux maladies : SAO et CHF en utilisant les ondelettes et les réseaux de neurones.
<i>Taouli sidi ahmed et Fethi berekssi-reguig (2010)</i>	Détection du Rythme Cardiaque du Signal Electrocardiogramme ECG par un seuillage Adaptatif.
<i>Mounia hendel et al. (2010)</i>	Arrhythmias Classification System based on Wavelet Transform and Bayesian RBF Neural Network
<i>Sana ktata et al (2010)</i>	Compression Robuste du Signal ECG par la Transformée en Sinus Discrète

2.4. Application en détectant d'autres pathologies

Haiying et son équipe décrivent un modèle-clé de classification ainsi qu'une plateforme de visualisation basée sur les réseaux neurologiques auto-adaptatifs [Haiying et al, 2004].

Une classification entre patients et sujets normaux s'est concentrée sur deux maladies : sommeil en apnée obstructif (SAO) et arrêt du cœur congestif (CHF) [Abdulnasir H, 2006].

Dans [Meghriche et al, 2006; 2007; 2008] la méthode développée, basée sur un réseau de neurones composé (Compound Neural Network CNN), pour classer les ECG comme normaux ou portant un bloc atrioventriculaire (BAV). Cette méthode emploie trois réseaux de neurones statiques multicouches, placés dans une configuration spécifique.

3. Acquisition des données d'ECG

Les ECG objet de la détection de l'onde P, importée de la base de données MIT et European Data Base. Ces données ont été numérisées tel que l'échantillonnage inter signal soit de l'ordre de quelques microsecondes. La plage de fréquences est en conformité avec les spécifications de l'American Heart Association (AHA) [New York Heart Association, 1953; Rantahaju et al, 1978].

Les définitions des mesures suivent les recommandations de l'action Européenne "Standards Communs pour Electrocardiographie Quantitative" CSE [CSE, 1985].

4. Conception de la méthode de Détection de l'onde P

Dans ce cadre de la réalisation de notre travail pour l'élaboration d'un algorithme de détection des paramètres de diagnostic du signal électrocardiographique à savoir l'onde P. nous procédons comme suit :

- Lecture des données à partir du fichier numérique.
- Détection du complexe QRS
- Détection de l'onde P.

L'approche adaptée est une approche qui se base sur le principe de programmation par les réseaux de neurones, à l'étape de détection de l'onde nous devons suivre les procédures illustrées par la (figure 3.1)

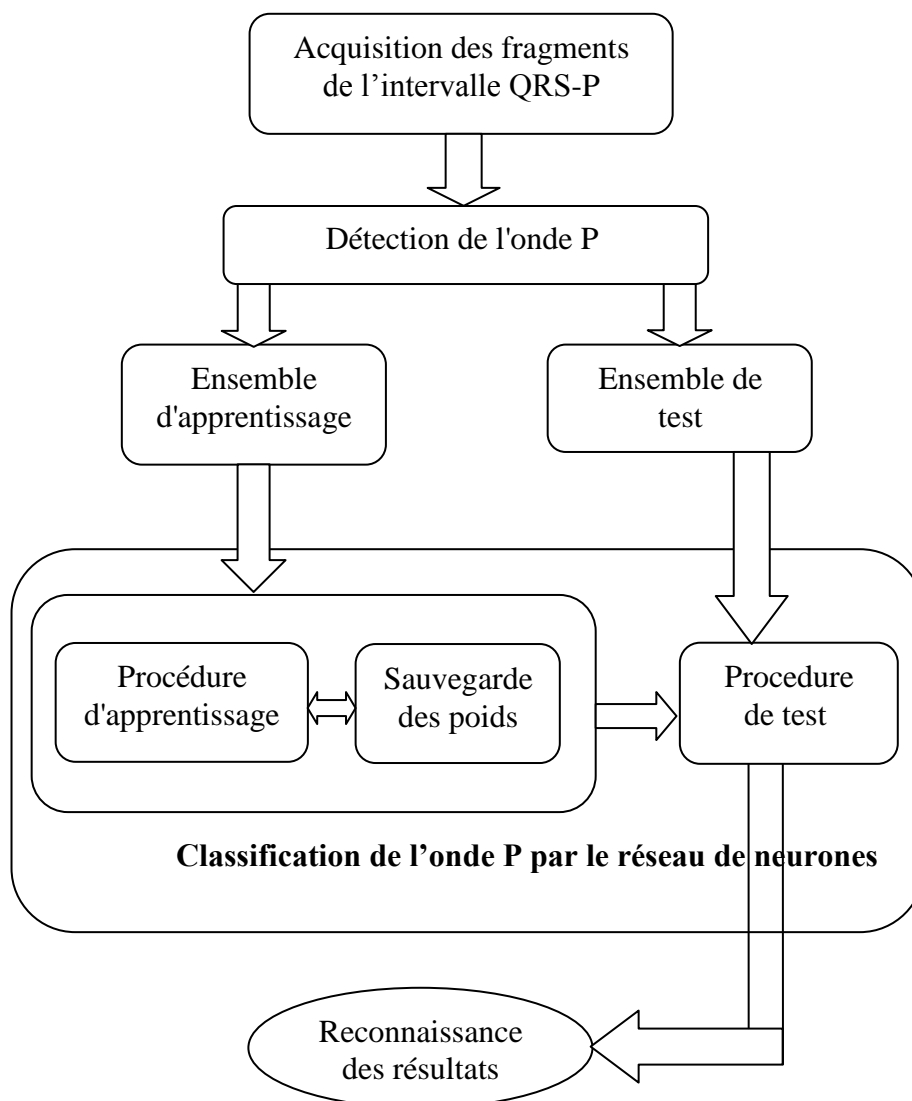


Figure 3.1 schéma global de l'algorithme.

4.1. Architecture du réseau de neurones

L'étape de détection des complexes QRS est primordiale, dans le but de localiser le début du battement et le début de l'intervalle PR. Dans cette partie du signal, on estime la présence de l'onde P qui fait l'objet de notre travail.

Pour cette application le réseau de neurone utilisé pour la détection de l'onde P est le réseau perceptron multicouche MLP (Multilayer Perceptron) possédant une couche de neurones cachée (le nombre de neurones cachés N_{cc} reste à déterminer). Les entrées du réseau sont au nombre de 14 représentés par les échantillons d'un fragment du signal ECG.

Les 14 mesures prises de chaque tracé, ont été choisies pour une analyse meilleure utilisant : Amp1 est la valeur de l'onde R qui représente l'onde la plus ample du signal. Les valeurs des entrées amp2, jusqu'à amp14 sont les amplitudes de l'intervalle PR en partant de l'onde R de

gauche à droite. Les 14 amplitudes ont été obtenues en divisant l'intervalle entre le point correspondant à l'onde R et l'extrémité qui normalement dépasse l'onde P, en quatorze parts de durée égale. Ces mesures sont obtenues à partir des enregistreurs automatisés d'ECG en utilisant leur programme de mesures. En outre, pour expliquer le raisonnement derrière la sortie du réseau de neurones, une méthode a été développée pour montrer les régions de l'ECG les plus pertinentes pour la détection de l'onde P.

Les neurones dans la couche d'entrée effectuent seulement une fonction de tampons qui distribuent les échantillons d'entrée aux neurones de la couche cachée. Chaque neurone au niveau de la couche cachée additionne ses échantillons après les avoir multipliés avec les poids des connexions respectives à partir de la couche d'entrée, et fait l'estimation de sa sortie en appliquant la fonction d'activation à la somme calculée. La sortie est représentée par un neurone unique. Les fonctions d'activation des neurones cachés et du neurone de sortie sont des sigmoïdes.

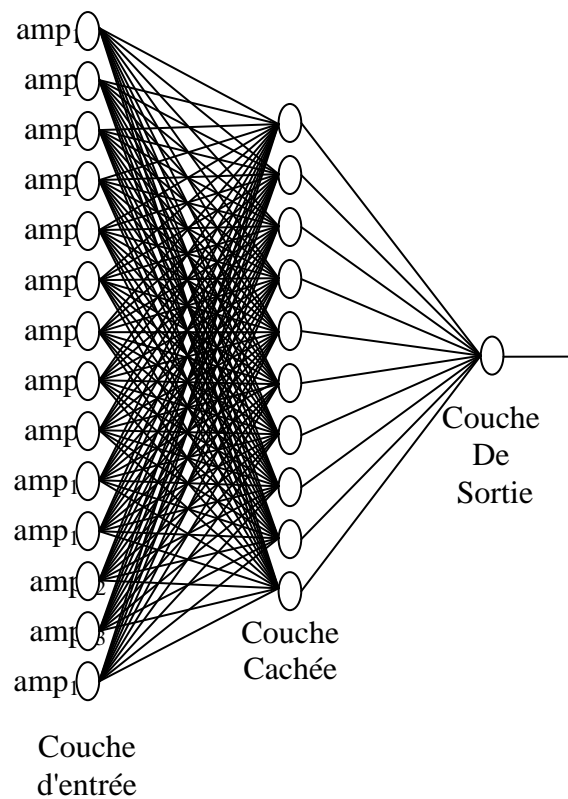


Figure 3.2 Structure Générale du réseau de neurones.

4.2. Apprentissage

L'apprentissage est défini comme étant le changement au niveau de la mémoire du réseau. Ainsi cette modification affecte les poids qui relient les neurones entre eux.

L'apprentissage a comme objectif l'amélioration des performances du réseau.

Cette phase est essentielle pour la détection d'une forme finale adaptée pour le problème.

Elle permet de trouver :

- Le nombre de couche du réseau de neurone.
- Le nombre de neurone de chaque couche du réseau de neurone.
- Le nombre d'entrée du réseau de neurone.
- Les poids du réseau de neurone.

Le mécanisme d'apprentissage diffère suivant la tâche pour laquelle ce réseau est utilisé. Nous avons utilisé l'apprentissage « supervisé » : qu'il se fait en présence d'un superviseur qui dirige le comportement du réseau en lui présentant les entrées et leurs sorties désirées.

Cet apprentissage se fait toujours par l'intermédiaire de critères à optimiser définissant la performance du réseau à chaque étape.

4.3. Fonction d'apprentissage

Les poids du réseau ont été ajustés par l'algorithme de Levenberg-Marquardt (LM) [More, 1977]. Bien que rétropropagation avec la technique de descente de gradient est un algorithme de plus grande pente, l'algorithme de Levenberg-Marquardt est une approximation de la méthode de Newton [Marquardt, 1963]. Si une fonction doit être réduite au minimum en ce qui concerne le paramètre vecteur x , alors la méthode de Newton serait:

$$MSE = \frac{1}{N} \sum_1^N E_i^2 \quad 3.2$$

Où E représente le vecteur des erreurs du réseau

$$E = output_D - output \quad 3.3$$

$output_D$ est la valeur désirée et $output$ est la sortie du réseau.

La principale motivation du choix de l'algorithme de LM repose sur la taille de la matrice du Hessien en fonction de la quantité de données de la base d'apprentissage, du coût moindre des calculs et de la garantie rapide de la convergence vers un minimum. LM entraîne le réseau de la même manière que l'algorithme de rétropropagation. Son utilisation se limite seulement aux réseaux à simple unité de sortie ainsi qu'à ceux à taille modérée (quelques centaines de poids).

L'algorithme a été conçu pour approcher le gradient du second ordre de l'apprentissage sans avoir à calculer la matrice Hessienne. La matrice Hessienne H peut être approximée comme :

$$H = J^T J + \eta I \quad 3.4$$

Et le gradient G peut être calculé comme :

$$G = J^T E \quad 3.5$$

J est la matrice de Jacobi. Elle contient les dérivées premières des erreurs du CNN en tenant compte des poids et des biais. Le Jacobien est beaucoup moins complexe que le calcul du Hessien.

L'algorithme emploie cette approximation à la matrice Hessienne dans Newton comme suit:

$$W_{k+1} = W_k - [J^T J + \eta I]^{-1} J^T E \quad 3.6$$

W_{k+1} : Valeur du vecteur poids à l'étape (k+1) après l'ajustement.

W_k : Valeur du vecteur poids à l'étape (k) avant l'ajustement.

I : Matrice identité.

η : Facteur d'apprentissage.

Le premier terme de l'équation 3.6 représente la supposition linéaire alors que le second une étape de la méthode de descente du gradient (GD).

Ainsi, quand le taux du facteur d'apprentissage est zéro, on retrouve la méthode de Newton, en utilisant la matrice Hessienne.

4.4. Fonction de transfert

Le réseau emploie la fonction sigmoïde $f(x)$ comme fonction de transfert.

$$f(x) = \frac{1}{1 + e^{-x + \text{biais}}} \quad 3.7$$

x : le vecteur d'une couche n, est défini comme :

$$x = \sum_i^N a_i W_i \quad 3.8$$

Où :

a_i : représentent les entrées et W_i : représentent les poids.

Cette fonction comprime un rang infini d'entrées en une valeur finie de sorties. Les fonctions sigmoïdes sont caractérisées par le fait que leurs pentes doivent être proches du zéro pendant que l'entrée est grande. Ainsi la sortie se trouve entre 0 et 1. Elles ont l'avantage de fournir une forme de commande de gain automatique, d'autant plus que sa dérivée est simple à calculer (figure 3.3).

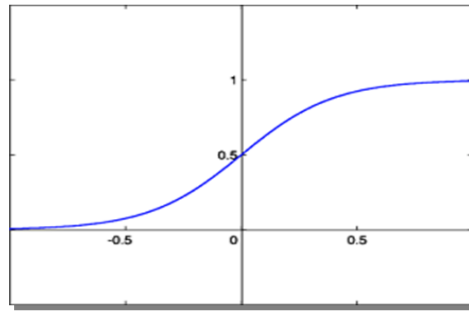


Figure 3.3. La fonction sigmoïde choisie pour le calcul des neurones.

4.5. Déroulement de l'apprentissage

L'apprentissage est réalisé par la minimisation d'une fonction de coût choisie ici comme l'erreur quadratique moyenne sur la sortie.

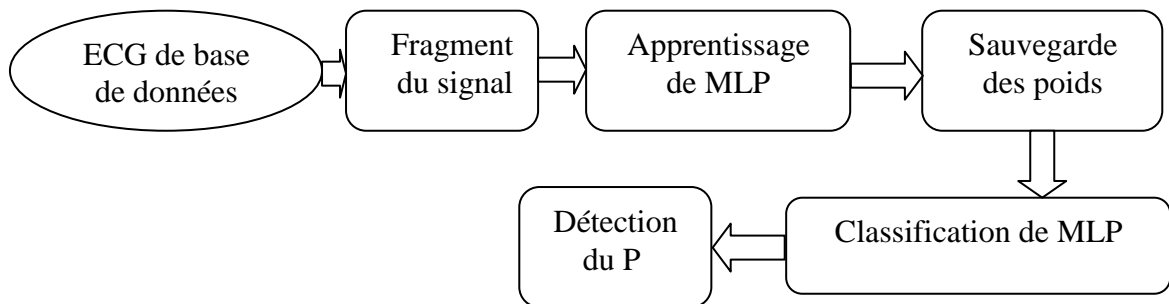


Figure 3.4. Diagramme de détection du P.

Le principe consiste à présenter au réseau le vecteur d'entrées des 14 échantillons d'un fragment du signal ECG, de procéder au calcul de la sortie par propagation à travers les couches, de la couche d'entrée vers la couche de sortie en passant par la couche cachée. Cette sortie obtenue est comparée à la sortie désirée, une erreur est alors obtenue. A partir de cette erreur, est calculée l'erreur qui est à son tour propagée de la couche de sortie vers la couche d'entrée. Cette technique permet la modification et l'ajustement des poids du réseau. L'opération est répétée pour chaque vecteur d'entrée et ce, jusqu'à ce qu'un des critères d'arrêt soit vérifié.

Notons que la phase d'apprentissage va s'arrêter sous l'une des conditions suivantes:

- L'erreur est inférieure à 10^{-25} .
- Le nombre d'itérations maximum (10^5) est atteint.

Afin de déterminer le nombre optimal de neurones cachés (N_{cc}) nous allons effectuer des apprentissages pour différentes architectures: pour chacune de ces architectures, nous effectuons 14 différentes initialisations des paramètres, ce qui conduit à 14 modèles par architecture.

Une base de validation nous permet de choisir parmi l'ensemble de ces architectures celle qui est la plus adaptée à notre problème. Pour un nombre de neurones cachés N_{cc} fixé, nous avons réalisé 14 apprentissages successifs avec des initialisations différentes des paramètres du réseau.

Nous sélectionnons ici le meilleur réseau pour N_{cc} fixé parmi les 14 initialisations des poids. Nous constatons que plus le nombre de neurones cachés est important, plus le nombre d'exemples bien classés est grand. Pour éviter le *surajustement*, nous avons choisi un N_{cc} optimal à partir de la base de validation.

4.6. Performance et Évaluation d'exécution du l'algorithme

Pour juger de la qualité d'un algorithme de ce type, cinq grandeurs sont habituellement mesurées :

- T_n (normal vrai) est le nombre d'ondes P normales identifiées en tant qu'onde P normale.
- T_{an} (anormal vrai) est le nombre d'ondes P anormales identifiées en tant qu'onde P anormale.
- F_n (normal faux) est le nombre d'ondes P anormales identifiées en tant que normal, autrement dit, c'est le nombre d'ondes P qui ont été détectées par l'algorithme alors qu'elles ne font pas partie de cette catégorie: ces erreurs peuvent correspondre à des emplacements repérés par l'algorithme alors qu'il n'y avait aucune onde caractéristique.
- F_{an} (an faux) est le nombre d'ondes P normales identifiées en tant d'onde P anormale.
- NTA (Nombre Total Analysé) : c'est le nombre de d'onde P analysés par tracé.

Afin de vérifier la performance d'un classificateur, on applique trois lois statistiques basées sur la reconnaissance des catégories mutuellement exclusives.

Ces lois sont la sensibilité, la spécificité et la précision (accuracy).

La sensibilité indique la capacité d'un classificateur à détecter des formes négatives (sujets normaux.). C'est la fraction des cas réels normaux correctement reconnus sur tous les cas réels normaux.

$$sensibilit\acute{e} = \frac{T_n}{T_n + F_{an}} 100 \quad 3.9$$

La spécificité représente la capacité d'un classificateur les cas positifs (sujets anormal) C'est la fraction des cas réels anormaux correctement reconnu sur tous les cas réels qui sont anormaux.

$$sp\acute{e}cificit\acute{e} = \frac{T_{an}}{T_{an} + F_n} 100 \quad 3.10$$

La précision (accuracy) représente la performance globale du classificateur. Elle indique le pourcentage des formes positives et négatives correctement classifiées de tous les cas.

$$accuracy = \frac{T_{an} + T_n}{T_{an} + F_n + T_n + F_{an}} 100 \quad 3.11$$

5. Résultats et discussion

La technique décrite ci-dessus a été appliquée en utilisant un grand nombre de morphologies d'ECG. Sur la base de 35% des données disponibles sont utilisées pour l'apprentissage, et de 65% pour le test, nous avons constaté que 60% des ECGs analysés étaient étiquetés pathologiques.

Chaque analyse a été effectuée en deux étapes. Dans ce qui suit nous présentons les résultats obtenus en fin d'étape.

La première étape correspond à analyser la totalité de chaque enregistrement: les annotations issues de l'algorithme sont comparées à celles des fichiers de référence et les erreurs sont comptabilisées. C'est l'étape « sans analyse du bruit ».

Par opposition, la seconde étape est l'analyse « avec analyse du bruit » : dans ce cas, on ne comptabilise pas, dans les erreurs, les annotations issues de zones trop bruitées. Ces zones sont identifiées de manière automatique pendant l'analyse. C'est cette dernière analyse qui est utilisée par la suite, car il est primordial de repérer les zones trop bruitées pour ne pas les analyser. Le tableau 3.5 présente les valeurs des trois grandeurs (NTA , F_n et F_{am}) obtenues lors de l'analyse de quelques enregistrements, ainsi que le résultat de la bonne détection de P.

<i>Edb0515</i>							
<i>intervalle 01</i>		<i>intervalle 02</i>		<i>intervalle 03</i>		<i>intervalle 04</i>	
<i>Durée (ms)</i>	<i>Amplitude (mV)</i>	<i>Durée (ms)</i>	<i>Amplitude (mV)</i>	<i>Durée (ms)</i>	<i>Amplitude (mV)</i>	<i>Durée (ms)</i>	<i>Amplitude (mV)</i>
0.28	5.1	1.016	5.105	1.732	5.145	2.444	5.125
0.26	4.11	0.996	4.065	1.712	4.035	2.424	4.085
0.24	4.03	0.976	4.035	1.692	4.02	2.404	4.02
0.22	4.02	0.956	4.03	1.672	4.005	2.384	4.005
0.2	4.005	0.936	4.02	1.652	4.005	2.364	4
0.18	4.035	0.916	4.05	1.632	4.03	2.344	4.03
0.16	4.07	0.896	4.085	1.612	4.07	2.324	4.065
0.14	4.11	0.876	4.12	1.592	4.12	2.304	4.11
0.12	4.09	0.856	4.09	1.572	4.075	2.284	4.07
0.1	4.05	0.836	4.075	1.552	4.055	2.264	4.05
0.08	4.035	0.816	4.065	1.532	4.02	2.244	4.03
0.06	4.035	0.796	4.055	1.512	4.04	2.224	4.03
0.04	4.2	0.776	4.055	1.492	4.03	2.204	4.035
0.02	4.035	0.756	4.055	1.472	4.035	2.184	4.035

Table 3.4. Exemple d'une fragmentation de quatre intervalles PR d'un signal ECG.

Table 3.5. Valeurs du traitement.

Enregistrement	NTA	F_n	F_{an}	Sensibilité
Edbe0611	776	5	4	96.66%
ECG03	831	0	1	-----
Edbe0515	1261	0	0	100%
Edbe0111	798	0	0	100%
Cudb-cu06	1289	0	1	-----
Apnea ecgx05	2116	3	0	98.90%
Apnea ecga10	2046	0	0	-----
Edbe0107	770	0	0	100%
ECG01	864	0	0	100%
Edbe0113	981	0	0	100%

TABLE 3.6: RESULTATS DES ECGs ALEATOIREMENT CHOISIS DE L'ENSEMBLE DE TEST.

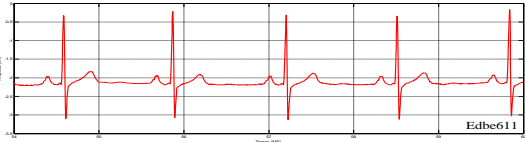
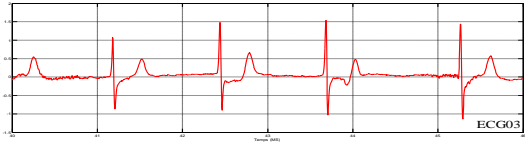
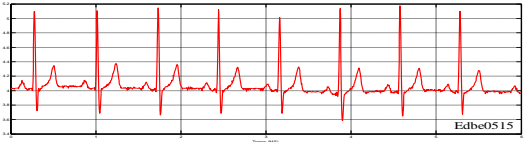
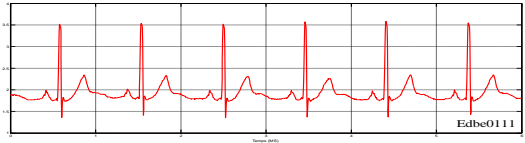
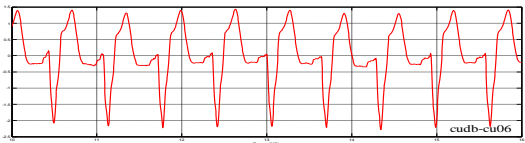
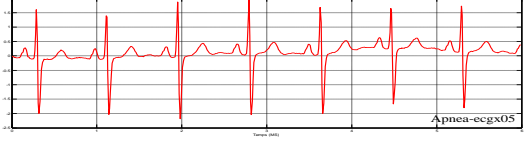
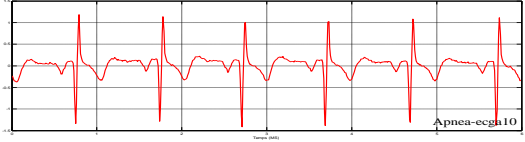
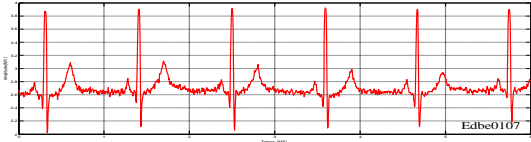
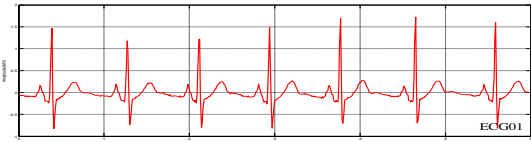
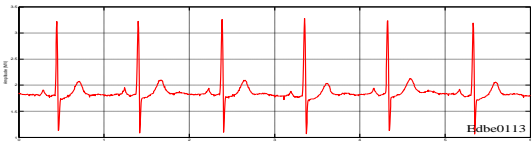
ECG	TRACÉ	Sortie Y	Cardiologue
Edbe0611		1.0000 4.3492e-006 1.0000 1.0000	Normal
ECG03		4.3111e-006 4.3113e-006 4.3115e-006	Anormal
Edbe0515		1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000	Normal
Edbe0111		1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000 1.0000	Normal
Cudb-cu06		0.0154 0.0152 0.0164 0.0164 0.0140 0.0176 0.0176	Anormal
Apnea-ecgx05		1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000 1.0000	Normal
Apnea-ecga10		3.8196e-006 3.6780e-006 3.7921e-006 3.7578e-006 3.7938e-006 3.5632e-006 3.8224e-006	Anormal

TABLE 3.7: RESULTATS DES ECGS ALEATOIREMENT CHOISIS DE L'ENSEMBLE DE TEST..

ECG	TRACÉ	Sortie Y	Cardiologue
Edbe0107		1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000 1.0000	Normal
ECG01		1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000 1.0000	Normal
Edbe0113		1.0000; 1.0000 1.0000; 1.0000 1.0000; 1.0000 1.0000	Normal

L'algorithme développé est très rapide tant dans l'apprentissage que dans l'exécution des tests. Cette performance est due principalement à l'application des réseaux de neurones et le fait qu'on fasse appel à une quantité d'information minimale.

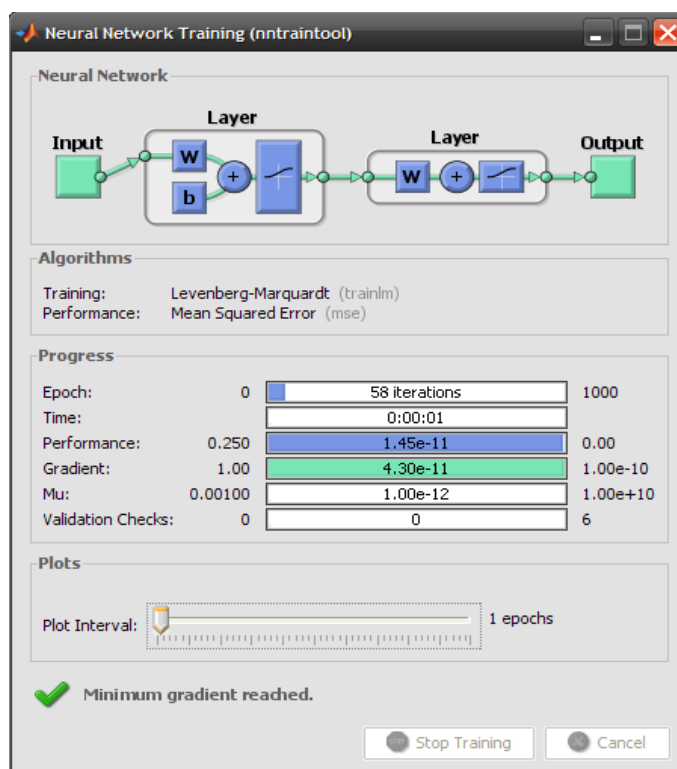


Figure 3.5 fenêtre d'apprentissage.

Les résultats montrent que le réseau se caractérise par une haute performance dans la détection de présence de l'onde P dans un ECG. Pour un seuil de 0,1 (Une valeur entre 0,1 et 1 étant la sortie désirée pour la présence de l'onde P et une valeur tombant en dehors de cette gamme, autrement dit, de 0 jusqu'à 0,1 pour la non présence de P.). Nous constatons dans des cas une sensibilité de 100%.

Cette performance est due aussi à la fragmentation du signal en petits intervalles à pas réguliers sans passer par la détection de l'onset et de l'offset de l'onde qui sont généralement non précis.

6. Conclusion

Arrivés au terme de ce dernier chapitre de thèse, nous avons pu développer un algorithme de détection de l'onde P dans un ECG. Les réseaux de neurones s'avèrent également très efficaces lorsqu'il s'agit de procéder à une classification des signaux ECG. La fragmentation du signal l'utilisation de ses fragments comme entrées du réseau ont permis de donner des résultats satisfaisants. Le temps de calcul n'est pas à proprement dit contraignant sachant que la génération des poids ne requiert que quelques secondes, la vérification séquentielle est très rapide et que le temps de détection de l'onde P est de quelques secondes aussi.

Conclusion Général

Conclusion

L'objectif assigné à ce travail consiste à mettre à la disposition du médecin un outil d'aide dans la prise de décision quant au diagnostic. Dans cette perspective, nous avons développé un programme de diagnostic automatique basé sur un réseau de neurones ayant une capacité de déterminer l'onde P du signal ECG numérique.

Cette discrimination se fait par référence à une valeur de comparaison obtenue par application de l'algorithme en question. L'application expérimentale de cet algorithme sur un grand nombre d'ECG permet une confirmation de la validité et de la fiabilité de la technique.

D'après les résultats obtenus, nous remarquons que l'apport des réseaux de neurones était de taille sur le plan de précision, de vitesse de calcul et d'architecture. Les résultats obtenus prouvent que les réseaux de neurones peuvent être utilisés pour détecter ces paramètres dans un ECG avec une grande exactitude.

Les réseaux de neurones de la présente étude peuvent être incorporés dans des programmes sur ordinateur d'interprétation d'ECG et peuvent détecter dans l'ECG au moyen des variables d'entrée de mesure la présence de pathologie.

Références Bibliographique

Références Bibliographiques

- [Abdulnasir, 2006] Abdulnasir H., "Biomedical Signals Classification Intelligent Systems", Proc. CSIT Jordan, 2006.
- [Alliot et al, 1997] Alliot J.-M. and Schiex T., Thomas Schiex, "Intelligence Artificielle & Informatique Théorique". Année p 3-11; 1997.
- [Anderson, 1997] Anderson J.A.; "Brain modeling and memory." In: X.J.R. Avula (Ed.), Proceedings of the First International Conference on Mathematical Modeling. (Invited paper). Rolla, Missouri: University of Missouri, 1977.
- [Bahoura et al, 1997] Bahoura M., Hassani M., Lee S-G., Hubin M., "Modification de la méthode de Widrow pour l'élimination de l'interférence 50Hz du signal ECG," Innov.Tech.Biol.Med.,18 (2), pp 119-127. 1997.
- [Baxt, 1990] Baxt, W. G. *Use of an artificial neural network for data analysis in clinical decision making: the diagnosis of acute coronary occlusion.* Neural Computation, 2, 480-489. 1990.
- [Baxt, 1992] Baxt, W. G. *Improving the accuracy of an artificial neural network using multiple differently trained networks.* Neural Computation, 4 (5), 772-780. 1992.
- [Baxt, 1994] Baxt, W. G.. *A neural-network trained to identify the presence of myocardial-infarction bases some decisions on clinical associations that differ from accepted clinical teaching.* Med. Dec. Making, 14 (3), 217-222. 1994.
- [Baxt, 1995] Baxt W. G. *Application of neural networks to clinical medicine.* Lancet, 346, 1135-1138; 1995.
- [Baxt et al, 1995] Baxt, W. G., White, H. *Bootstrapping confidence intervals for clinical input variable effects in a network trained to identify the presence of acute myocardial infarction.* Neural Computation, 7 (3), 624-638; 1995.
- [Baxt, 1996] Baxt, W. G. *Artificial neural network to identify acute myocardial infarction-Reply.* Lancet, 347, 551. 1996.
- [Baxt et al, 1996] Baxt, W. G., & Skora, J. *Prospective validation of artificial neural network trained to identify acute myocardial infarction.* Lancet, 347, 12-15. 1996.
- [Casaleggio et al, 1991] Casaleggio A., Morando M., Ridella S. "Neural Networks for Automatic Anomalous QRS Complex Detection," Myocardial Infarction Proc. Computer in Cardiology, pp. 553-6, 1991.
- [Chauvin et al, 1995] Chauvin, Yves et Rumethart, David E. *Back propagation: Theory architectures, and applications.* Lawrence Erlbaum Associates, 1995;

- [Cohen et al, 1983] Cohen M. A Grosberg S. *Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks*. IEEE Transactions on Systems, Man and Sybernetics, Vol.SMC-13, No.5, Sept/Oct 1983.
- [Cole et al, 1941] Cole K. S., Cole R. H., " *Dispersion and absorption in dielectrics. I. Alternating current characteristics* ", Journal of Chemical Physics, Vol. 9, pp. 341-351,1941.
- [Conde, 1994] Conde T., " *Automatic Neural Detection of Anomalies in Electrocardiogram (ECG) Signals,*" IEEE pp. 3552-58. 1994.
- [Cross et al, 1995] Cross S. S., Harrison R. F., Lee Kennedy, R. *Introduction to neural networks*. Lancet, 346, 1075-1079. 1995.
- [Cross et al, 2000] Cross S. S., Stephenson T. J., Mohammed T., Harrison R. F. *Validation of a decision support system for the cytodagnosis of fine needle aspirates of the breast using a prospectively collected dataset from multiple observers in a working clinical environment*. Cytopathology, 11 (6), 503-512. 2000.
- [CSE Working Party, 1985]The CSE Working Party- *Recommandation for Measurement Standards in Quantitative Electrocardiography*, European Heart Journal; 1985.
- [Dassen et al, 1992] Dassen W., Mulleneers R., Bleijlevens B., Den Dulk K, Rrodriguez LM, Schlapfer J., Katsivas A., wellens H., " *Determination of the Aetiology of Wide-QRS Tachycardia Using an Artificial Neural Network,*" IEEE, pp. 165-68. 1992.
- [De Bakey et al,1997] DeBakey M. E. Gotto A. M., *Le cœur en action*. Institut d'édition Sanofi-Synthelabo, 1997.
- [Debye, 1929] Debye: P., " *Polar molecules* ", Edition New York: The chemical catalog, 1929
- [Dennis et al, 1983] Dennis, J. E. Schnabel, R. B. *Numerical methods for unconstrained optimization and nonlinear equations*. Englewood Cliffs: Prentice-Hall, 354p. 1983
- [Devroye et al, 1996] Devroye L. Lugosi G. *A probabilistic theory of pattern recognition*. New York: Springer, 292 p. 1996
- [Dreiseitl et al, 1999] Dreiseitl, S., Ohno-Machado, L., Vinterbo, S. *Evaluating variable selection methods for diagnosis of myocardial infarction*. Proc. AMIA Symp. 1-2, pp 246-250. 1999.
- [Dreyfus, 2004] Dreyfus G., Martinnez J.M., Samuelides M., Gordon M.B., Badran F., Thiria S., Hérault J. Sous la direction de Gérard Dreyfus. France-Tauve, " *Réseau de Neurones Méthodes et Applications*" 11-bd de Sébastopol, Paris. Edition Eyrolles Avril 2004.
- [EL Dajani, 2002] El Dajani R. M. *Artificial neural networks for modeling and predicting the dynamical behavior of the ventricular repolarization phase*, theses INSAL 2002.
- [Ellenius et al, 1997] Ellenius, J., Groth, T., & Lindahl, B. *Neural network analysis of biochemical markers for early assessment of acute myocardial infarction*. Stud. Health Technol. Inform., 43, 382-385 1997.
- [Fahlman, 1988] Fahlman S. E. *An empirical study of learning speed in back propagation networks*. Technical Report CMU6CS-88-162, CMU.1988.
- [Fricke, 1932] Fricke H., " *The theory of electrolytic polarization* ", Phil. Mag., pp. 310-318, 1932.
- [Fricker, 1997] Fricker, J. *Artificial neural networks improve diagnosis of acute myocardial infarction*. Lancet, 350, 935. 1997.
- [Goldman al, 1988] Goldman, L., Cook, E. F., Brand, D. A., Lee, T. H., Rouan, G. W., Weisberg, M. C., *A computer protocol to predict myocardial infarction in emergency department patients with chest pain*. N. Engl. J. Med., 318, pp 797-803. 1988.

- [Greiner et al, 1996] R. Greiner , R. Isukapalli, "*Learning to select useful landmarks*", IEEE Transactions on Systems, Man, and Cybernetics-Part B, 1996
- [Hagan et al, 1996] Hagan, M. T. Demut, H. B. Beale, M. H. *Neural Network design*. Boston: PWS publishing, 344 p. 1996
- [Haiying et al, 2004] Haiying W., Azuaje F., and Norman B.. "*An Integrative and Interactive Framework for Improving Biomedical Pattern Discovery and Visualization*", IEEE Transaction on Information Technology in Biomedicine, Vol. 8, No. 1, pp. 16-27. 2004.
- [Hamidi, 1995] Hamidi S. *Analyse quantitative de l'ECG ambulatoire et etude de la dynamique spatio-temporelle de la repolarisation ventriculaire*. Thèse de doctorat en informatique et automatique appliquée. Institut National des sciences Appliquée de Lyon. 1995.
- [Haykin, 1999] Simon haykin, "*Neural networks a comprehensive foundation*" ISBN-13-908385-5 prentice hall international, Inc pp: 34-78, second edition, 1999
- [Hebb, 1949] Hebb D. O., *The Organization of Behavior : A Neuropsychological Theory*, Wiley, New York, 1949.
- [Hedén et al, 1997] Hedén, B., Öhlin, H., Rittner, R., Edenbrandt, L. *Acute myocardial infarction detected in the 12-lead ACG by artificial neural networks*. Circulation, 96 (6), pp 1798-1802 1997.
- [Hopfield, 1982] Hopfield J.J., Proc. Nat. Acad. of Sciences, vol. 79, p. 2554, 1982
- [Hornik, 1994] Hornik K. *Neural Networks: more than statistics for amateurs*. In: Dutter, R. et Grossman, W. CMPASTAT. Physica Verlag, Heidelberg. pp 223-235, 1994
- [Huang et al, 1990] Shih-chi huang , yih-fang huang, "*Learning algorithms for perceptrons using back-propagation with selective update*" pp, 1990
- [Jane et al, 1993] Jane R, Otmos S, Laguna P, Caminat P, *Adaptative hermites models for ECG data compession: performance and evaluation with automatique wave detection*. Proceed. of the computers in cardiology. IEEE computer society. 1993, pp 389-92.
- [Jian-kang, 1993] Jian-kang wu, "*Neural networks and simulation methods*", Editeur: CRC Press (Déc 1993) ISBN 0-8247-9181-9(acid-free paper)
- [Josefson, 1997] Josefson, D. *Computers beat doctors in interpreting ECGs*. BMJ, 315, pp 763-766, 1997.
- [Kohonen, 1977] Kohonen T., *Associative Memory: A System Theoretic Approach*, Springer-Verlag, Berlin, 1977.
- [krose et al, 1993] Ben J.A. krose, P.patrick van der smagt, "*An introduction to neural networks*" university of Amsterdam faculty of mathematics and computer science: pp 33-42, Fifth edition january 1993
- [Lepoutre et al, 2001] Lepoutre F. et al, *Systèmes et Microsystèmes pour la Caractérisation*, Hermès éd., Paris, vol2, pp 439-445. 2001.
- [Lisboa, 2000] Lisboa, P. J. G., Ifeachor, E. C.,& Szczepaniak, P. S. (Eds.) *Artificial neural networks in biomedicine*, New York: Springer 2000.
- [Lunn et al, 2000] Lunn, D. J., Thomas, A., Best, N. G., Spiegelhalter, D. J. WinBUGS. *A Bayesian modelling framework: concepts, structure and extensibility*. Statistics in Computing, 10, 321-333. 2000.
- [Luo et al, 1989] R. C. Luo , M. G. Kay, "*Multisensor integration and fusion in intelligent Systems*", IEEE Transactions on Systems, Man, and Cybernetics, pp: 901-931, 1989

- [Maxwell, 1981] Maxwell J. C., "A treatise on electricity and magnetism ", 2ème édition, Edition Oxford : Clarendon Press, 1881.
- [Mc Culloch et al, 1943] Mc Culloch G. et Pitts W. *A logical calculus of the ideas immanent in nervous activity*, bulletin of mathematical biophysics, V.5, pp 115-133. 1943.
- [Meghriche et al, 2006] Meghriche S., Draa A., Boulemden M., Abada A., "A Hybrid Neural Network for Detecting an AtrioVentricular Block (AVB).in ECG Signal," Proc. CSIT Jordan, pp 172-178, 2006.
- [Meghriche et al, 2007] Meghriche S., Boulemden M., "Review on the Development and Utilization of Artificial Neural Network Applied to ECG Signals Diagnosis" Proc. DCCA Jordan, pp 1066-1074, 2007.
- [Meghriche et al, 2007] Meghriche S., Draa A., Boulemden M. *A High Performance CNN Architecture for the Detection of AVB Carrying ECGs*. Asian Journal of Information Technology, Vol. 6, N° 4, pp 474-479. 2007.
- [Meghriche et al, 2008] Meghriche S., Draa A., Boulemden M. *On The Analysis of a Compound Neural Network for Detecting AtrioVentricular Heart Block (AVB) in an ECG Signal*. International Journal of Biological and Biomedical Sciences, WASET. Vol.1 N°1 pp 1-11. 2008.
- [Meghriche et al, 2008] Meghriche S., Boulemden M., Draa A., *Agreement Between Multi-Layer Perceptron and a Compound Neural Network on ECG Diagnosis of Atrioventricular Blocks*. WSEAS transaction on Biology and Biomedicine Vol. 5, N° 1, pp 12-22. 2008.
- [Meghriche et al, 2008] Meghriche S., Boulemden M., Draa A., *Two Neural Networks Architectures for Detecting AVB*. Proc. Multimedia Systems and signal processing, China, pp 193-199. 2008.
- [Minsky et al, 1969] Minsky , Papert, Perceptrons p. 231-232; 1969:
- [More, 1977] More J. J. *The Levenberg-Marquardt algorithm implementation and theory*. Ln: Watson, G. A. Numerical Analysis, Lecture Notes in Mathematics 630 Heidelberg Springer Verlag, pp. 105-116, 1977
- [Neal, 1996] Neal, R. M. *Bayesian learning for neural networks*, New York: Springer-Verlag. 1996.
- [Neuro Dimension, 2001] *Neuro Dimension. The Applications Corner*. [en ligne] disponible sur : Neuro Dimension, Inc. <http://www.nd.com/appcornr/purpose.htm> .2001
- [New York Heart Association, 1953] *New York Heart Association Subcommittee on Electrocardiography Criteria, Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Blood Vessels*. 5th ed. 1953.
- [Parizeau, 2004] Marc Parizeau, "Réseaux de neurones ", 2004
- [Polak et al, 1997] Polak, M. J., Zhou, S. H., Rautaharju, P. M., Armstrong, W. W., Chaitman, B. R. *Using automated analysis of the resting twelve-lead ECG to identify patients at risk of developing transient myocardial ischemia-an application of an adaptive logic network*. Physiol. Meas., 18 (4), pp 317-325. 1997.
- [Rantahaju, 1978] Rantahaju P. M. et al, *The Impact of Computer on Electrocardiogram Interpretation Systems*, NARDP Grants N° 410. 1978.
- [Reggia, 1993] Reggia, J. A. *Neural computation in medicine*. Artificial Intelligence in Medicine, 5, 143-157. 1993.
- [Renders et al, 1995] Renders X, De Leval N, *L'enfant au centre de la stratégie publicitaire: le phénomène du petit mannequin*. L'observateur. Rev.d'action sociale et médico-sociale, Liège, nov.1995, 1995.
- [Riedmiller et al, 1993] Riedmiller, Braun, *A direct adaptive method for faster backpropagation learning: The RPROP algorithm*, 586-591, <http://citeseer.nj.nec.com/riedmiller93direct.html> . 1993.

- [Rigaud et al, 1996] Rigaud B. Morrucci J.-P., Chauveau N. " *Bioelectrical impedance techniques in medicine. Part Bioimpedance measurement. Second section: Impedance spectrometry* ", Critical Reviews in Biomedical Engineering, Vol. 24 (4-6), p. 257-351, 1996.
- [Ripley, 1996] Ripley, B. D. *Pattern recognition neural network*. Cambridge university press. 1996.
- [Rosenblatt, 1961] Rosenblatt F. *Principales of neurodynamics: perceptron and the theory of brain mechanisms*. Washington: Spartan Books, 1961.
- [Rumelhart et al, 1986] Rumelhart, Daved E, McClelland, James L, *University of California, San Diego, and PDP Research Groupe. Parallel distributed processing explorations in the microstructure of cognition*. Cambridge, Mass: MIT Press, 1986.
- [Rutkowski, 1977] Rutkowski F., *CAN et CNA*, Electronique Applications, 3 pp 97-103, 1977.
- [Schwartzter et al, 2000] Schwartzter G., Vach W., Schumacher M. *On the misuses of artificial neural networks for pronostic and diagnostic classification in oncology*. Stat. Med., 19, 541-551, 2000.
- [Selker et al, 1995] Selker, H. P., Griffith, J. L., Patil, S., Long, W. J., D'Agostino, R. B. *A comparison of performance of mathematical predictive methods for medical diagnosis: identifying acute cardiac ischemia among emergency department patients*. J. Investig. Med., 43 (5), pp 468-476. 1995.
- [Shi et al, 2000] Shi, L. M., Fan, Y., Lee, J. K., Waltham, M., Andrews, D. T., Scherf, U., Paull, K. D., Weinstein, J. N. *Mining and visualizing large anticancer drug discovery databases*. Journal of Chemical Information and Computer Sciences, 40 (2), pp 367-379. 2000.
- [Silbernagl et al, 1993] Silbernagl S., Despopoulos A. *Atlas commenté de physiologie humaine pour étudiant et praticiens* ", Paris, janvier 1993.
- [Simpson, 1990] Simpson P. K. *Artificial Neural Systems - Foundations, Paradigms, Applications and Implementations*. Pergamon Press, 1 edition 1990.
- [Sornmo et al, 1981] Sornmo L, Borjesson P, Nygard M, Pahlm O, *A method for evaluation of QRS shape features using a mathematical model for the ECG*, IEEE Trans. Biome. Ing. pp 713-7. 1981.
- [Spiegelhalter, et al, 1999] Spiegelhalter, D. J., Myles, J. J., Jones, D. R., Abrams, K. R. *An introduction to Bayesian methods in health technology*. BMJ, 319, 508-512. 1999.
- [Szczepaniak et al, 2000] Szczepaniak, P. S., Lisboa, P. J. G., Kacprzyk, J. (Eds.) *Fuzzy systems in biomedicine*, Berlin: Springer-Verlag. 2000.
- [Trentelman, 1993] Trentelman H.L., *'Representation and learning in feedforward neural networks'* em CWI Quarterly, Vol. 6, No. 4, pp. 385-408, 1993.
- [Wagner, 1914] Wagner K. W., *"Explanation of the dielectric fatigue phenomenon on the basis of Maxwell's concept"*, Arkiv für Electrotechnik - Edition Shering H. Berlin: Springer-Verlag, 1914.
- [Warren, 2001] Warren, S. S Donoho-johnstone Benchmarks: *"Neural Net Results"* [en ligne]. Disponible sur: SAS Institute Inc., Cary, NC, USA.

Résumé

Le tracé ECG comporte plusieurs déflexions (ondes), L'analyse et l'interprétation automatique de ce signal est souvent assistée par ordinateur faisant appel à des techniques de traitement du signal et de reconnaissance de formes.

Actuellement, les réseaux de neurones bénéficient de fondements théoriques solides. Avec leurs aptitudes en classification, mémorisation, filtrage et approximation, ils sont devenus un moyen très efficace. L'opération de reconnaissance de formes constitue l'une des applications les plus connues en réseaux de neurones. Donc l'application des réseaux de neurones est bien destinée à assurer la reconnaissance et la classification des ondes et des anomalies dans un signal électrocardiographique ECG.

Le travail consiste à développer une méthode de détection de l'onde P basée sur les réseaux de neurones. Il s'agit d'utiliser de manière efficace les réseaux de neurones de type statique multicouche.

Les résultats prouvent que des réseaux de neurones peuvent être employés pour améliorer l'interprétation automatisée d'ECG pour la détection de maladies et un médecin pourrait employer cette approche comme appui de diagnostic.