

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université Ibn Khaldoun -Tiaret-



Faculté des sciences et sciences de l'ingénieur  
Département d'Informatique

École Doctorale

Sciences et Technologies de l'Information et de la Communication

**Approche multi-agents pour le monitoring médical  
« Application cardiologique »**

**Mémoire**

Pour l'obtention du diplôme de magistère  
Option : Systèmes d'Information et de Connaissances (SIC)

par

**GOISMI Mohamed**

**Composition du jury :**

M. DAHMANI Youcef	MCA- Univ Ibn Khaldoun -TIARET-	Président
M. BESSAID Abdelhafid	Pr- Univ A Bakr Belkaid -TLEMCEN-	Examineur
M. BENABDELLAH Mohamed	Pr- Univ A Bakr Belkaid -TLEMCEN-	Examineur
M. CHADLI Abdelhafid	MAA- Univ Ibn Khaldoun -TIARET-	Invité
M. CHIKH Mohamed Amine	MCA- Univ A Bakr Belkaid -TLEMCEN-	Directeur du mémoire

**ANNEE UNIVERSITAIRE 2010/2011**



# DEDICACE

Je dédie ce travail:

A tous ceux qui me sont chers...



# REMERCIEMENTS

Tout d'abord je tiens à remercier mon DIEU LE TOUT PUISSANT pour le souffle, la patience et la force qui m'a donné pour finir ce mémoire.

Je voudrais exprimer mes forts remerciements à mon encadreur le Dr. CHIKH Mohamed Amine, pour m'avoir donné le goût de la recherche, pour m'avoir poussé et encouragé tout au long de ces dernières années et pour sa patience avec moi.

Ensuite, je tiens à remercier le membre de jury pour avoir accepté de juger ce modeste travail.

Je tiens à remercier tout le personnel de département d'informatique de l'université de Tiaret et celui de l'université de Tlemcen.

Je remercie encore, tous ceux qui ont participé de proche ou de loin pour finir ce travail.

**RESUME :**

Les maladies cardiovasculaires constituent un problème majeur de santé public, elles sont les premières causes de mortalités dans le monde.

En raison de l'ampleur du problème, divers travaux ont été mis en place afin de réduire le risque, préconisant notamment l'éducation, la prévention, et le suivi des patients à risque. Dans ce mémoire, nous avons appliqué une approche multi-agents pour le monitoring médical où nous avons proposé de mettre au point un système de classification des battements cardiaques. Ce système est basé principalement sur l'utilisation des réseaux de neurones et les systèmes experts à base d'agents intelligents, dont, le premier agent classifieur neuronal reçoit en entrée 96 paramètres temporels, et le deuxième reçoit aussi 19 paramètres morphologiques (qui sont les composantes principales retenues de l'inertie en entrée, après l'avoir compressé par la méthode d'analyse en composantes principale ACP) caractérisant un battement ECG, et dernièrement, l'agent expert qui s'occupe de la fusion des deux résultats fournis par les deux agents classifieurs. Ces agents vivent dans un environnement purement ouvert communiquant via un réseau de télécommunication. Notre système a l'objectif de faire la classification neuronale en se basant sur les données temporelles et morphologiques des ECG selon deux classes, « normale » et « anormale », qui a été validé sur des signaux de la base de données MIT\_BIH. Les résultats qualitatifs et quantitatifs obtenus démontrent l'efficacité de cette approche.

**Mots clés :**

ECG, les maladies cardio-vasculaires, Arythmies cardiaques, classification, les systèmes multi-agents, jade, réseaux de neurones, système expert, apprentissage artificielle, descripteur temporel, descripteur morphologique, ACP.

**ABSTRACT**

Cardiovascular diseases are a major public health problem; they are the leading causes of death in the world.

Because of the magnitude of the problem, various works have been implemented to reduce risk, including advocating education, prevention, and monitoring of patients at risk. In this thesis we applied a multi-agent approach for medical monitoring where we proposed to develop a classification system heartbeat. This system is primarily based on the use of neural networks, expert systems and intelligent agents, including the first neural classifier agent receives 96 temporal parameters as input, and the second has also 19 morphological parameters (which are the principal components taken of inertia as input, after being compressed by the principal components analysis method PCA) characterizing an ECG beat. Finally, the expert agent takes care of merging the two results provided by both classifier agents. These agents live in a purely open environment communicating via a telecommunication network. Our system aims to do a neuronal classification based on morphological and temporal ECG data into two classes a "normal" and "abnormal" one, which has been validated on signals from MIT\_BIH database. Qualitative and quantitative results obtained demonstrate the effectiveness of our approach.

**Keywords:**

ECG, heart disease, cardiac arrhythmias, classification, multi-agent systems, jade, neural networks, expert systems, artificial learning, temporal descriptor, morphological descriptor, PCA.

# Table des matières

<b>INTRODUCTION</b> .....	<b>14</b>
<b>I. LE MONITORING MEDICAL</b> .....	<b>20</b>
<b>I.1. Introduction</b> .....	<b>20</b>
<b>I.2. Monitoring</b> .....	<b>20</b>
I.2.1. LE MONITORING .....	20
<b>I.3. Présentation de l'organisation d'une salle de soins</b> .....	<b>22</b>
I.3.1. CARACTERISTIQUES DES DONNEES D'USIS .....	23
I.3.1.1. Haute dimensionnalité .....	23
I.3.1.2. Incertitude .....	24
I.3.1.3. Imprécision .....	24
I.3.1.4. Ambiguïté .....	25
<b>I.4. Spécificité des connaissances médicales en USIs</b> .....	<b>25</b>
I.4.1. CARACTERISTIQUES DES CONNAISSANCES MEDICALES .....	25
I.4.1.1. L'incertitude en médecine .....	25
I.4.1.2. Et le taux d'erreurs qui en découle .....	26
I.4.2. CARACTERISTIQUES DES CONNAISSANCES ASSOCIEES AUX USIS .....	26
I.4.2.1. Un nombre important d'informations à traiter .....	26
I.4.2.2. Une notion de temporalité à définir .....	26
I.4.2.3. Des comportements du patient difficilement formalisables .....	26
<b>I.5. Les systèmes de monitoring médical intelligents</b> .....	<b>27</b>
I.5.1. DES SYSTEMES DE MONITORING MEDICAL .....	28
<b>I.6. Le système CARDIOVASCULAIRE</b> .....	<b>29</b>
I.6.1. LA CIRCULATION ARTERIELLE ET VEINEUSE .....	30
I.6.1.1. Le cœur .....	30
I.6.1.1.2. Fonctionnement électrique .....	31
I.6.2. L'ELECTROCARDIOGRAPHIE .....	33
I.6.2.1. Principe de l'électrocardiogramme (ECG) .....	33
I.6.2.2. De l'ECG au Holter .....	33
I.6.3. TRACE ELECTRIQUE DU CŒUR .....	34
<b>I.7. Les pathologies cardiaques</b> .....	<b>35</b>
I.7.1. ELEMENTS FONDAMENTAUX DE LA CONDUCTION ELECTRIQUE CARDIAQUE .....	35
I.7.1.1. La conduction électrique .....	36
I.7.2. LE BATTEMENT CARDIAQUE STANDARD ET SES CARACTERISTIQUES .....	37
<b>I.8. Diagnostic à partir du rythme</b> .....	<b>39</b>
I.8.1. TROUBLES DU RYTHME CARDIAQUE .....	39
I.8.1.1. Troubles du rythme supra-ventriculaire .....	40
I.8.1.2. TROUBLES DU RYTHME VENTRICULAIRE .....	44
I.8.1.2.1. Trouble de conduction .....	46
I.8.1.2.2. Autres anomalies cardiaques .....	49
<b>I.9. Conclusion</b> .....	<b>49</b>
<b>II. LES SYSTEME MULTI-AGENTS</b> .....	<b>52</b>
<b>II.1. Introduction</b> .....	<b>52</b>
II.1.1. POSITIONNEMENT HISTORIQUE .....	52
<b>II.2. Qu'est-ce qu'un agent ?</b> .....	<b>53</b>

II.2.1 L'ARCHITECTURE D'AGENT .....	55
II.2.3 UNE COMPARAISON AVEC LES OBJETS .....	56
II.2.4 CLASSES D'AGENT .....	57
<b>II.3. Qu'est-ce qu'un système multi-agents (SMA)? .....</b>	<b>58</b>
II.3.1 LES SYSTEMES MULTI-AGENTS .....	58
II.3.2 DEFINITION: SYSTEME MULTI-AGENTS (SMA) .....	58
II.3.3 LES INTERACTIONS ET LES AGENTS .....	59
II.3.3.1 La communication.....	60
II.3.4 LES LANGAGES DE COMMUNICATION .....	61
II.3.4.1 La théorie des actes de langage .....	61
II.3.5 SEMANTIQUE DES MESSAGES.....	64
II.3.6. LA COOPERATION.....	64
<b>II.4. Outils &amp; plateformes de développement des SMA .....</b>	<b>65</b>
II.4.1 POURQUOI LA PLATEFORME JADE ? .....	66
II.4.2 LA NORME FIPA POUR LES SYSTEMES MULTI-AGENTS .....	66
<b>II.5. Domaines d'application.....</b>	<b>67</b>
II.5.1 APPLICATION DES SMA AUX TELECOMMUNICATIONS .....	67
II.5.2 LE SYSTEME ADEPT .....	67
II.5.3 TELEMEDECINE ET LA SANTE .....	68
II.5.3.1 Le système GUARDIAN .....	68
II.5.4 LES SYSTEMES D'INFORMATIONS COOPERATIFS (SIC) .....	69
<b>II.6. Conclusion.....</b>	<b>71</b>
<b>III. LES MODELES DE CLASSIFICATION .....</b>	<b>74</b>
<b>III.1. Introduction.....</b>	<b>74</b>
<b>III.2 La classification en IA.....</b>	<b>74</b>
III.2.1 CHAINE DE CLASSIFICATION D'OBJETS .....	75
III.2.1.1 Les principaux modules de classification d'objets.....	75
III.2.1.2 L'extraction des descripteurs .....	76
III.2.1.3 Procédure de résolution par apprentissage .....	76
III.2.3 FORMALISATION MATHEMATIQUE D'UN PROBLEME DE CLASSIFICATION.....	79
<b>III.3. Méthodes de classification.....</b>	<b>79</b>
III.3.1. LA LOGIQUE FLOUE EN CLASSIFICATION .....	80
III.3.1.1. Conception du classificateur flou.....	80
III.3.1.2. Avantages et Inconvénients .....	81
III.3.2. LES ARBRES DE DECISION.....	81
III.3.2.1. Exemple.....	81
III.3.2.2. Avantages et Inconvénients .....	82
III.3.3. RESEAUX BAYESIENS.....	83
III.3.3.1. Exemple.....	83
III.3.3.2. Avantages et Inconvénients .....	84
III.3.4. MACHINE A VECTEURS DE SUPPORT (SVM) .....	85
III.3.4.1. Avantages et Inconvénients .....	85
III.3.5. LES CARTE DE KOHONEN.....	86
III.3.5.1. Avantages et Inconvénients .....	87
<b>III.4. Conclusion .....</b>	<b>88</b>
<b>IV. LES RESEAUX DE NEURONES &amp; LES SYSTEMS EXPERTS.....</b>	<b>90</b>
<b>IV.1. Introduction .....</b>	<b>90</b>
<b>IV.2. Historique sur les réseaux de neurones.....</b>	<b>90</b>
<b>IV.3 Les réseaux de neurones.....</b>	<b>91</b>



IV.3.1. LE NEURONE BIOLOGIQUE .....	91
IV.3.2. LE NEURONE ARTIFICIEL ET PRINCIPE DE FONCTIONNEMENT.....	92
IV.3.3. IMPORTANCE DE LA FONCTION D'ACTIVATION .....	93
IV.3.4. RESEAU DE NEURONES .....	94
IV.3.5. ARCHITECTURE DES RESEAUX DE NEURONES.....	95
IV.3.5.1. Les réseaux de neurones non bouclés .....	95
IV.3.5.2. Les réseaux de neurones bouclés .....	96
IV.3.5.6. APPRENTISSAGE DES RESEAUX DE NEURONES.....	97
IV.3.6.1. Type d'apprentissage .....	98
IV.3.6.2. Algorithmes d'apprentissage .....	99
<b>IV.4. Modélisation à l'aide de réseaux de neurones .....</b>	<b>100</b>
IV.4.1. MODELE « BOITE NOIRE ».....	100
<b>IV.5. Conception et mise en œuvre des réseaux de neurones.....</b>	<b>101</b>
IV.5.1. DETERMINATION DES ENTREES/SORTIES DU RESEAU DE NEURONES.....	101
IV.5.2. CHOIX ET PREPARATION DES ECHANTILLONS .....	101
IV.5.3. ELABORATION DE LA STRUCTURE DU RESEAU .....	102
IV.5.4. APPRENTISSAGE .....	103
IV.5.5. VALIDATION ET TESTS.....	103
<b>IV.6. Apprentissage par les réseaux multicouches.....</b>	<b>104</b>
IV.6.1. ALGORITHME D'APPRENTISSAGE .....	105
IV.6.1.1. Algorithme de la rétropropagation du gradient.....	106
IV.6.2. PROBLEME DE L'APPRENTISSAGE.....	106
<b>IV.7. Apprentissage : point de vue de l'IAD .....</b>	<b>108</b>
<b>IV.8. Avantages et inconvénients des réseaux de neurones.....</b>	<b>109</b>
<b>IV.9. Domaines d'application des RN .....</b>	<b>110</b>
<b>IV.10. SYSTEME EXPERT .....</b>	<b>111</b>
IV.10.1. A QUOI ÇA SERT ?.....	111
IV.10.2. Composantes des systèmes experts .....	111
IV.10.2.1. BASE DE CONNAISSANCES .....	112
IV.10.2.1.1. Base de règles .....	112
IV.10.2.1.2. Base de faits.....	113
<b>IV.10.3. Moteur d'inférences.....</b>	<b>113</b>
IV.10.3.1. CHAINAGE AVANT .....	114
IV.10.3.2. CHAINAGE ARRIERE.....	114
IV.10.3.3. CHAINAGE MIXTE.....	114
<b>IV.11. Conclusion.....</b>	<b>115</b>
<b>V. CONTRIBUTION ET REALISATION.....</b>	<b>118</b>
<b>V.1. Introduction.....</b>	<b>118</b>
<b>V.2 Classification neuronale des arythmies cardiaques.....</b>	<b>118</b>
V.2.1. SELECTION DE LA BASE D'EXEMPLES .....	120
V.2.1.1. Sélection des descripteurs temporaires d'un cycle cardiaque .....	121
V.2.2. ARCHITECTURE DU RESEAU DE NEURONES .....	122
V.2.2.1. Algorithme d'apprentissage .....	122
V.2.2.2. Dimensionnement du réseau pour chaque classifieur .....	122
V.2.2.2.1 Réduction par l'ACP .....	123
V.2.3. Résultats expérimentaux des classifications .....	124
V.3.1 Architecture des agents de classification.....	127
V.3.2 Communication entre les agents du système.....	130

<b>V.4. Présentation de l'application.....</b>	<b>132</b>
V.4.1 AGENT CLASSIFIEUR NEURONAL AGENT1.....	132
V.4.1 AGENT CLASSIFIEUR NEURONAL AGENT2.....	132
V.4.1 AGENT D'ADAPTATION AGENT3.....	133
<b>V.5. Conclusion.....</b>	<b>134</b>
<b>CONCLUSION .....</b>	<b>134</b>
<b>Perspectives .....</b>	<b>135</b>
<b>Annexe A.....</b>	<b>136</b>
<b>Annexe B.....</b>	<b>141</b>
<b>Annexe C.....</b>	<b>143</b>
<b>Références bibliographiques .....</b>	<b>146</b>



## Liste des figures

FIG a :	Taux de mortalités des personnes âgées moins 70 ans .....	15
FIG 1.1 :	LE MONITORING : un ensemble de paramètres centralisés sur un écran....	22
FIG 1.2 :	Salle de soins .....	23
FIG 1.3 :	Schéma du circuit sanguin .....	30
FIG 1.4 :	Schéma fonctionnel du cœur.....	31
FIG 1.5 :	L'impulsion électrique .....	32
FIG 1.6 :	Enregistreur Syneflash.....	34
FIG 1.7 :	Éléments fondamentaux de la conduction électrique cardiaque .....	35
FIG 1.8 :	La contraction des oreillettes.....	36
Fig. 1.9 :	Dérivations bipolaires (voies I II et III) .....	36
Fig. 1.10 :	Dérivations précordiales (voies V1 V2... V6).....	37
FIG 1.11 :	Paramètres d'intérêt pour la description d'un battement .....	38
FIG 1.12 :	Tachycardie sinusale .....	40
FIG 1.13 :	Bradycardie sinusale .....	40
FIG 1.14 :	Arythmie sinusale.....	41
FIG 1.15 :	Tachycardie auriculaire.....	41
FIG 1.16 :	Emplacement des foyers ectopiques .....	42
FIG 1.17 :	Flutter auriculaire .....	42
FIG 1.18 :	Fibrillation auriculaire.....	43
FIG 1.19 :	Extrasystole auriculaire ESA.....	43
FIG 1.20 :	Tachycardie jonctionnelle.....	44
FIG 1.21 :	Battement ventriculaire prématuré.....	45
FIG 1.22 :	Tachycardie ventriculaire.....	45
FIG 1.23 :	Tachycardie ventriculaire.....	45
FIG 1.24 :	Torsade de pointe.....	46
FIG 1.25 :	Bloc auriculo-ventriculaire (B.A.V) du 1er degré.....	47
FIG 1.26 :	Blocs de branche gauches .....	48
FIG 1.27 :	Blocs de branche droits.....	48
FIG 2.1 :	Différents types d'agents .....	55
FIG 2.2 :	Architecture minimale d'un agent.....	55
FIG 2.3 :	La structure d'un message KQML.....	62
FIG 2.4 :	La structure d'un message FIPA ACL.....	63
FIG 2.5 :	Le modèle de référence pour une plate-forme multi-agents FIPA .....	66
FIG 3.1 :	Chaine de modules de classification .....	76
FIG 3.2 :	Schéma synoptique d'un classifieur supervisé.....	80
FIG 3.3 :	Arbre de décision .....	82
FIG 3.4 :	Structure de causalité.....	83
FIG 3.5 :	Intégration des connaissances d'expert et des données.....	84
FIG 3.6 :	La carte auto adaptative de Kohonen.....	87
FIG 4.1 :	Le neurone biologique.....	92
FIG 4.2 :	Neurone artificiel.....	92
FIG 4.3 :	Fonctions d'activations .....	93
FIG 4.4 :	Taxonomie des réseaux de neurones non bouclés et récurrents.....	95
FIG 4.5 :	Réseau de neurones à n entrées .....	96
FIG 4.6 :	Réseau de neurone bouclé.....	97
FIG 4.7 :	Erreur moyenne sur la base d'apprentissage .....	98
FIG 4.8 :	Diagramme schématique d'un modèle neuronal « boîte noire ».....	101
FIG 4.9 :	Organigramme de conception d'un réseau de neurones .....	104
FIG 4.10 :	Le phénomène de surapprentissage d'un RN .....	107
FIG 4.11 :	Composantes des systèmes experts .....	112
FIG 5.0 :	Schéma global de la classification .....	119
FIG 5.1 :	Battements normaux et intervalles d'intérêt.....	121
FIG 5.2 :	Variabilité des données de la base d'apprentissage.....	123
FIG 5.3 :	Architecture distribuée de l'Application.....	126
FIG 5.4 :	Architecture de l'agent classifieur AGENT1 .....	127
FIG 5.5 :	Architecture de l'agent classifieur AGENT2 .....	128

FIG 5.6 :	Architecture de l'agent d'adaptation AGENT3.....	129
FIG 5.7 :	Communication séquentielle entre les agents du système.....	131
FIG 5.8 :	Agent de classification à base des descripteurs temporels.....	132
FIG 5.9 :	Agent de classification à base des descripteurs morphologiques. ....	132
FIG 5.10 :	Agent d'adaptation. ....	132
FIG 5.11 :	SNIFFER, contrôleur des messages échangés entre les agents du système. ...	132

## Liste des tableaux

Tab 2.1 :	Approche orientée objet (AOO) versus approche orientée agent (AOA) .....	56
Tab 3.1 :	Les définitions des grandeurs VP, VN, FP et FN .....	79
Tab 3.2 :	Exemple des données .....	81
Tab 5.1 :	Les enregistrements choisis de la base de données MIT-BIH. ....	120
Tab 5.2 :	Nombre de composantes principales retenues pour chaque ECG. ....	123
Tab 5.3 :	Performances du classifieur neuronal à données temporelles.....	124
Tab 5.4 :	Performances du classifieur neuronal à données morphologiques réduite .....	124
Tab B.1 :	les 48 enregistrements de la base de données MIT-BIH.....	142
Tab C.1 :	les actes de communication fournie par Jade .....	144



## INTRODUCTION

Les maladies cardio-vasculaires sont responsables de 17 millions de décès à travers le monde. Elles sont suivies par le cancer qui tue 7,6 millions de personnes chaque année à travers le monde.

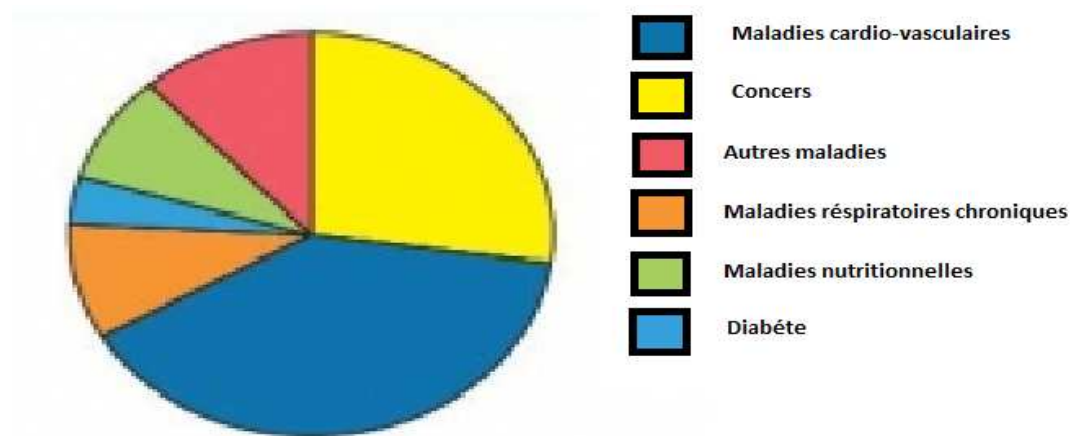
Le premier rapport de l'Organisation mondiale de la santé (Oms) sur les maladies non transmissibles (Mnt) rendu public le 27 avril 2011, révèle que les pays en voie de développement sont les plus touchés par les MNT.

« En 2008, 36,1 millions de personnes sont décédées d'affections telles que les cardiopathies, les accidents vasculaires cérébraux (Avc), les pneumopathies chroniques, les cancers et le diabète. Plus de 80% de ces décès sont survenus dans les pays à revenu faible ou moyen», indique le communiqué de l'OMS. Les maladies cardio-vasculaires et le cancer sont à l'origine de plusieurs millions de décès chaque année à travers le monde.

Les maladies cardiovasculaires sont responsables de la plupart des décès par Mnt, à savoir 17 millions de personnes chaque année ; elles sont suivies par le cancer (7,6 millions), les maladies respiratoires (4,2 millions) et le diabète (1,3 million). Ces quatre groupes de maladies représentent près de 80% de l'ensemble des décès par Mnt et ont en commun quatre facteurs de risque : le tabagisme, la sédentarité, l'usage nocif de l'alcool et une alimentation déséquilibrée, renseigne le rapport de l'Oms.

La prise en charge de ces maladies, reconnaît le directeur général de l'Oms, constitue un énorme défi pour beaucoup de pays, où la prévalence est alarmante. «Pour certains pays, il n'est pas exagéré de décrire la situation comme une catastrophe imminente, une catastrophe pour la santé, pour la société et surtout pour les économies nationales», souligne le directeur général de l'Oms, le docteur *Chan*.

La proportion des personnes en âge d'activité affectées est importante. Selon le sous-directeur général de l'Oms, le docteur *Ala Alwan*, près de 30% des personnes qui meurent de maladies non transmissibles dans les pays à faible revenu sont âgées de moins de 60 ans. Ces décès, dit-il, sont d'autant plus tragiques qu'ils sont en grande partie évitables.



**FIG a:** Taux de mortalités des personnes âgées moins 70 ans, suivants les causes de décès.

Principaux facteurs sont en cause tels que le tabagisme, l'inactivité physique, l'usage nocif de l'alcool et la mauvaise alimentation. L'influence de ces facteurs de risque comportementaux s'ajoute à d'autres causes sous-jacentes métaboliques et physiologiques;

La prise en charge des maladies cardiovasculaires a beaucoup évolué au cours de la dernière décennie, D'une part, l'accroissement des connaissances sur la physiopathologie des arythmies cardiaques et l'amélioration des techniques d'exploration électrophysiologiques permettent d'identifier les patients à risque de mort subite rythmique et de prendre des mesures préventives selon le cas rencontré et donc des millions de décès pourraient donc être évités.

Les appareils utilisés pour suivre l'évolution du cœur sont appelés systèmes de monitoring cardiaque. Ces systèmes sont composés à la fois des capteurs permettant de recueillir des données sur l'activité du cœur et des unités centrales permettant de stocker ou de visualiser ces données. L'évolution de ces systèmes entre les années 60 et 90, la multiplication du nombre de données disponibles et la relative autonomie de ces systèmes quand ils sont, par exemple, sous la forme de pacemaker cardiaque, ont conduit à l'émergence des systèmes de monitoring intelligents. Les unités centrales doivent donc maintenant, en plus des tâches de stockage et de visualisation définies précédemment, interpréter les données pour prévenir et assister le personnel médical et éventuellement, agir lorsque l'état du cœur évolue dangereusement.

L'importance de ces tâches rend la qualité, la sensibilité et la robustesse de ces systèmes primordiales. Pour augmenter la fiabilité de la détection des problèmes cardiaques, notamment en présence des bruits liés aux mouvements des patients, à la perte d'un signal, à la présence d'artefacts sur le signal etc., les systèmes de monitoring intelligents peuvent utiliser plusieurs sources de données.

Pour diagnostiquer des troubles du rythme cardiaque, les médecins ont pour habitude d'utiliser principalement l'électrocardiogramme (ECG). Lorsque l'ECG est bruité, le cerveau humain est capable, dans une certaine mesure, de retrouver les informations exactes parmi les informations bruitées ou de chercher dans d'autres sources de données (par exemple sur la morphologie de l'ECG) la confirmation de ces informations.

L'automatisation de cette recherche d'informations n'est pas triviale, notamment parce qu'il est difficile d'extraire du signal ECG toutes les informations dont se sert le médecin pour diagnostiquer des arythmies, mais aussi parce que le système de monitoring doit être capable, si l'ECG est trop bruité pour pouvoir être utilisé tel quel, de trouver l'information nécessaire dans les autres sources de données, même si elles ne sont a priori pas suffisantes pour permettre un diagnostic fiable. En outre, fournir à un système suffisamment de connaissances expertes pour automatiser le diagnostic est une tâche extrêmement coûteuse et fastidieuse.

L'analyse du rythme figure parmi les applications les plus prometteuses de l'ingénierie biomédicale : elle consiste, pour l'essentiel, à détecter les anomalies éventuelles, à identifier leur nature et à préciser un traitement approprié, d'où l'intérêt de développer un outil automatisé d'aide



au diagnostic du signal ECG. Cet objectif comporte plusieurs volets, qui sont d'ailleurs communs à tous les outils d'aide à la décision.

1. L'acquisition du signal ECG. L'instrumentation doit permettre un enregistrement fidèle de ce signal, tout en minimisant l'incidence des signaux parasites (bruits venants du secteur, de l'interface peau-électrode et des parties du corps autres que le cœur).
2. Le pré-traitement du signal ECG. L'objectif de cette étape est la transformation du signal brut en un signal interprétable par filtrage de l'ECG et la détection de ses différentes ondes.
3. La caractérisation du signal ECG. Cette étape consiste en le choix de caractéristique et paramètres permettant une description informative du signal ECG.
4. L'analyse des données et l'interprétation des résultats. Les informations obtenues lors des étapes précédentes doivent être intégrées dans des algorithmes permettant de faire la distinction entre les différents types de signaux. La classification obtenue doit avoir une signification (séparer des signaux normaux des ECGs pathologiques par exemple) et doit pouvoir être interprétée par le médecin cardiologue.

En outre, les systèmes multi-agents sont l'un des paradigmes technologiques les plus prometteurs dans le développement de systèmes logiciels distribués, ouverts et intelligents. Nous entendons par systèmes ouverts, les systèmes auxquels peuvent se rattacher d'autres systèmes. La technologie agent commence par être utilisée pour concevoir des solutions facilitant la mise en place de nouveaux concepts surtout dans le domaine médical à savoir le monitoring médical.

Les systèmes Multi agents (SMA) sont des collections d'entités intelligentes autonomes qui collaborent à la résolution commune d'un problème complexe. Ce genre de systèmes est approprié dans beaucoup de domaines médicaux, dus aux caractéristiques des problèmes dans ce secteur. Ces systèmes se positionnent au carrefour de la programmation (ce sont des logiciels), de l'intelligence artificielle (leur autonomie de décision), et des systèmes répartis (leur décentralisation). C'est-à-dire sont l'une des branches de l'intelligence artificielle distribuée (**IAD**) qui permet la résolution distribuée de problèmes complexes, par coordination d'un certain nombre d'agents.

Les agents intelligents ont l'avantage d'être facilement des entités apprenantes, qui peuvent avoir un mécanisme d'apprentissage performant, tels que les réseaux de neurones, les systèmes experts ou d'autres méthodes de l'intelligence artificielle connues, qui ont prouvés leurs performances pour la classification, la reconnaissance de forme, ...etc.

Nous sommes intéressés, dans ce mémoire, à l'application des techniques de l'intelligence artificielle classique (Les réseaux de neurones et les systèmes expert) et ceux de l'intelligence artificielle distribuée (SMA) pour la classification des battements cardiaques afin d'éviter des milliers de morts subites dans le monde.

Le but de ce mémoire est, d'une part d'appliquer des techniques permettant d'interpréter de manière automatique un signal ECG en basant sur deux voies essentielles :

- ➔ La première est d'extraire des caractéristiques temporelles telles que la fréquence et la période du temps de certaines ondes (P, Q, R, S, T, le complexe QRS, la période RR, etc..) de l'ECG, cette voie est dite Interprétation par caractéristiques temporelles.
- ➔ La deuxième consiste en premier lieu sur la morphologie du signal ECG.

Et d'autre part d'investir des avantages des systèmes multi-agents (autonomie et distribution des tâches, ...etc.) et les méthodes de l'intelligence artificielle telles que les réseaux de neurones pour faire la classification des battements cardiaques et les systèmes expert pour le couplage des résultats donnés par les PMC (Rapidité, facilité d'implémentation, fiabilité des résultats).

Le manuscrit de ce mémoire est divisé en cinq (05) chapitres :

Le premier chapitre consiste, dans la première section, à donner des notions de base pour comprendre le monitoring médical, en décrivant les propriétés des données et spécificités des connaissances médicales en unité de soin intensif, puis il est consacré, en deuxième section, sur l'anatomie et la biologie du cœur, les dérivations cardiaques pour enregistrer cette activité électrique sont ensuite présentées. Les ondes élémentaires d'un signal ECG sont décrites ainsi que les intervalles, durées, et segments que l'on mesure. Et en dernière section, le chapitre décrit brièvement les différentes arythmies et pathologiques cardiaques susceptibles d'être détectées dans un enregistrement ECG, le but, ici, est de mettre en relation certaines observations, anormales du tracé ECG avec les anomalies cardiaques les plus courantes. Il s'agit de sensibiliser le lecteur à l'importance diagnostique de l'étude des caractéristiques des ondes P, Q, R, S et T (Formes, distances, ...), au-delà d'une simple de rythme.

Le deuxième chapitre sert à donner des définitions et connaissances sur les agents intelligents et les systèmes multi-agents, nécessaires à la compréhension de notre étude. Après un bref rappel sur les outils et plateformes de développements des systèmes multi-agents (par exemple, Jade), les domaines d'application des agents intelligents, spécialement dans le domaine médical.

Le troisième chapitre définit ce qu'est un problème de classification. Parmi la variété de problèmes de classification, nous analysons les avantages et les inconvénients de quelques méthodes de classification telles que (la logique floue, les arbres de décision, les réseaux bayésiens, les machines à vecteur de support (SVM) et les cartes de kohonen).

Le quatrième chapitre permet de détailler les modèles de classification utilisées et implémentées dans ce mémoire, à savoir, les réseaux de neurones et les systèmes experts. En donnant l'historique de ces deux modèles, ces architectures, ces composants essentiels, les algorithmes d'apprentissage pour le premier et le moteur d'inférence pour le deuxième.

Le cinquième chapitre décrit notre application, le modèles neuronaux pour la classification des battements cardiaques, le modèle neuronal à base caractéristiques temporelles extraites depuis le signal ECG, et le modèle neuronal aussi à base ca caractéristiques morphologiques pour la classification des battements cardiaques. Détailler le modèle de fusion ou d'adaptation en fonction d'un système expert pour avoir la classe finale. Nous expliquons la solution de l'approche multi-agents en détaillant les agents développés dans notre application, leurs comportements, leurs architectures, leurs communications entre eux, puis on donne le schéma distribué de l'application et en fin, un scénario d'exécution de cette application.

# I. LE MONITORING MEDICAL

## I.1. Introduction

Ce chapitre a pour objectif de donner la définition du monitoring, les données médicales et ses caractéristiques dans les unités de soins intensifs et les systèmes de monitoring existant.

Ensuite on va vous présenter 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 du mémoire ; le lecteur intéressé par une approche médicale rigoureuse pourra se reporter aux nombreux ouvrages médicaux disponibles sur le sujet, comme par exemple « Le cœur en action » aux éditions Sanofi-Synthélabo [Deb, 97].

## I.2. Monitoring

### I.2.1. Le monitoring

Le terme *monitoring* est un abus de langage emprunté aux médecins pour parler de la surveillance continue de patient. En français, il faudrait parler de *monitorage*.

Le dictionnaire LARROUSSE définit le Monitoring ou monitorage comme : « surveillance médicale en continu ou à intervalles rapprochés, effectué par mesure des paramètres ou par enregistrement de phénomènes divers (contractions utérines, battements cardiaques, etc.) ».

L'activité de *monitorage* s'applique à de nombreux domaines industriels autres que la médecine tels que la surveillance de centrales nucléaires, l'industrie du pétrole, les télécommunications, etc. Le groupe de recherche ALARM [Cau et al, 98], composé de plusieurs chercheurs travaillant sur ce domaine particulier, ont proposé la définition suivante au terme de *monitorage* : « l'activité de *monitorage* d'un système dynamique peut être vue comme le suivi continu d'un système par un module de haut niveau qui analyse toutes les situations rencontrées, communique avec un opérateur humain et suggère des décisions à prendre en cas de dysfonctionnement du système. » L'activité de *monitorage* repose sur l'utilisation d'alarmes<sup>1</sup>.

Le défi pour l'opérateur humain est d'interpréter le nombre très important de ces alarmes. Dans le contexte médical le terme *monitoring* désigne à la fois des fonctions de surveillance, de diagnostic et de test, il s'applique au patient et aux systèmes qui les supportent [Pie, 98]. Nous garderons donc ce terme dans la suite de ce mémoire pour parler de *monitorage médical*.

---

<sup>1</sup> **Alarme** : indicateur discret émis par un système de surveillance à partir d'événements et censé provoquer une réaction humaine ou automatique.

On peut distinguer deux types de monitoring dans le contexte médical : le monitoring en temps réel et le monitoring hors ligne.

*Le monitoring en temps réel* est habituellement associé aux unités de soins intensifs (USI<sup>2</sup>). Ces dernières sont des unités cliniques qui accueillent les patients dont les conditions cliniques peuvent changer rapidement et pour lesquels les risques de décès sont élevés. Les soins pratiqués dans ces unités peuvent eux mêmes provoquer des changements rapides et importants de l'état physiologique du patient. Dans ces conditions, le monitoring des patients est vital.

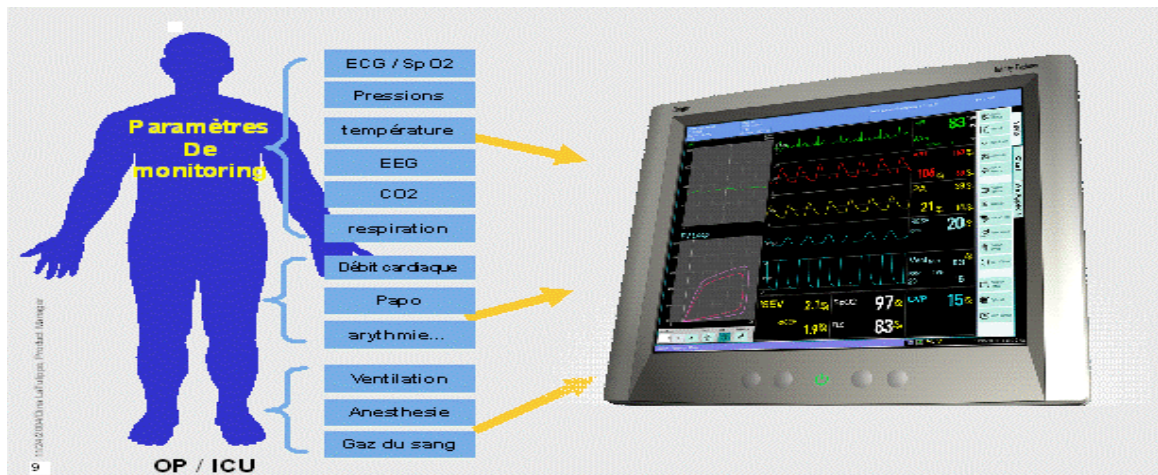
Le monitoring en USI est plus précisément défini comme l'observation et la mesure, en temps réel, de manière répétée ou continue, du patient, de ses fonctions physiologiques et des instruments qui lui permettent de se maintenir dans un état stable, pour faciliter les prises de décisions du personnel médical. Ces décisions peuvent aller de l'intervention thérapeutique proprement dite à la simple estimation de l'effet de ces interventions. Une liste détaillée des buts recherchés par les systèmes de monitoring en USI est donnée dans [Pie, 98]. Les appareils de monitoring se servent d'informations fournies par des capteurs. Ces appareils et surtout ces capteurs peuvent être invasifs<sup>3</sup> i.e. implantés dans le corps du patient (c'est le cas des cathéters artériels qui mesurent la pression sanguine artérielle, de la ventilation mécanique, etc.), semi-invasifs (comme les sondes œsophagiennes) ou non invasifs (l'électrocardiogramme, la température etc.), selon l'état du patient. Les données provenant d'un capteur invasif sont beaucoup plus précises et permettent un suivi continu de longue durée mais ces capteurs sont beaucoup moins bien tolérés par les patients et leur utilisation est plus coûteuse. Certains appareils de monitoring, comme le pacemaker cardiaque (utilisé hors USI) nécessitent un passage au bloc opératoire, ce qui augmente les risques post-opératoires pour le patient. L'intensité du monitoring, i.e. le nombre de capteurs utilisés et l'*invasivité* de ces capteurs, est fonction de la gravité de l'état du patient et des interventions qui ont été précédemment faites sur ce même patient.

Le second type de monitoring concerne les bandes enregistrées par exemple, par les moniteurs électrocardiographiques non invasifs Holter ou les Mesures Ambulatoires de Pression Artérielle (MAPA). Un Holter permet d'enregistrer l'activité électrique cardiaque de manière continue pendant 24 heures. Le Holter est un appareil de monitoring qui enregistre l'activité cardiaque du patient par le biais d'électrodes placées sur son corps. Le but d'un tel système est de s'assurer de la tolérance d'un patient à des activités de tous les jours telles que la marche, la conduite, l'ingestion de nourriture ou le sommeil. Pour permettre une interprétation *a posteriori* correcte des signaux enregistrés, le patient est tenu de répertorier lui même tous les changements d'activités effectués pendant ces 24 heures.

---

<sup>2</sup> **USI** : Unité de Soin Intensif.

<sup>3</sup> **Invasif** : se dit d'une méthode d'exploration ou de soin nécessitant une liaison de l'organisme -Larousse



**FIG 1.1 : LE MONITORING** : un ensemble de paramètres centralisés sur un écran.

### I.3. Présentation de l'organisation d'une salle de soins

De nombreux patients séjournant en Unités de Soins Intensifs (USIs) nécessitent une assistance respiratoire partielle ou totale pour prendre en charge une insuffisance respiratoire. Ce trouble peut être le résultat d'une détérioration des propriétés mécaniques du poumon mais aussi d'autres organes comme les muscles ventilatoires, le système nerveux central ou le système cardio-vasculaire. On dit qu'un patient est dépendant d'un point de vue respiratoire lorsqu'il a besoin d'une assistance pendant au moins 24 heures ou s'il a échoué lors de tests pour le replacer en respiration autonome [Tac, 02]. Plusieurs causes d'une dépendance respiratoire peuvent être répertoriées : des problèmes neurologiques (qui troublent les cycles respiratoires par exemple), des muscles respiratoires défaillants (la myopathie en est une illustration), des anomalies au niveau des échanges gazeux, des facteurs cardio-vasculaires (en particulier des dysfonctionnements au niveau du ventricule gauche), des facteurs psychologiques (comme le stress induit par le retrait de l'assistance respiratoire).

Dans tous les cas, l'assistance respiratoire justifie une surveillance importante. Pour ce faire, le patient est relié à plusieurs machines (voir FIG 1.2) qui sont, en général : un ventilateur pour assister/ suppléer la ventilation, un moniteur pour la surveillance des paramètres (visualisation des signaux, réglages de seuils) d'autres moniteurs (Scanographe, moniteur cardio-vasculaire,...), des pompes à médicament, une centrale de recueil d'informations, depuis quelques années, des progrès techniques ont permis le développement de nombreux appareils de monitoring au lit du patient. Ils se caractérisent par deux fonctions principales [Bio, 98] : une fonction d'information en temps réel sur l'état physiologique ou physiopathologique du patient, une fonction d'alarme ou d'alerte prévenant de modifications des paramètres surveillés.

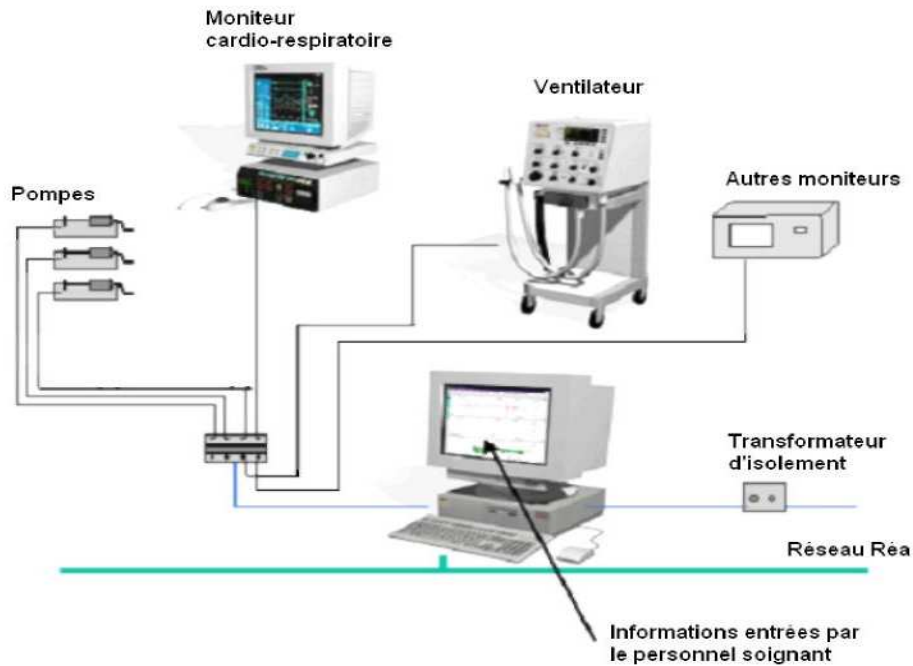


FIG 1.2 : Salle de soins [Sil, 04]

### I.3.1 Caractéristiques des données d'USIs

Les données obtenues en salle de soins ont de nombreuses caractéristiques et autant de problématiques associées. Suivant les services et les habitudes des praticiens, elles ne sont pas identiques d'un service à l'autre. Même si certains signaux physiologiques sont unanimement surveillés quelques soient le praticien et les machines utilisées comme l'électrocardiogramme par exemple, d'autres sont spécifiques aux habitudes particulières des praticiens (ou à la disponibilité de certains moniteurs). Ensuite, suivant la thérapeutique appliquée au patient, toutes les données ne sont pas disponibles continuellement ou, au contraire, d'autres sont nécessaires et de nouveaux capteurs doivent être introduits dans le système de surveillance. Hormis ces particularités, les données d'USIs ont des caractéristiques intrinsèques comme la haute dimensionnalité, l'incertitude, l'imprécision, et l'ambiguïté que nous allons détailler ci-après.

#### I.3.1.1. Haute dimensionnalité

Les données d'USIs sont temporelles, et ont une fréquence d'acquisition très élevée. Elle est de l'ordre de la mesure par seconde et la surveillance d'un patient peut durer jusqu'à plusieurs semaines ...

En USIs, la surveillance d'un patient peut impliquer la prise en compte d'un nombre considérable de variables : les paramètres physiologiques, les paramètres de réglage des machines, les données provenant des laboratoires, les médicaments... Le personnel médical se retrouve alors à devoir ingérer, analyser et éventuellement décider d'une action à entreprendre à partir d'une multitude d'informations.

Ces deux caractéristiques justifient amplement la notion de haute dimensionnalité attribuée aux données d'USIs. Le flux de données reste considérable et les algorithmes de traitement doivent en tenir compte [Sil, 04].

### I.3.1.2. Incertitude

Afin de discerner clairement les notions d'incertitude et d'imprécision, nous allons nous appuyer sur les définitions données par Isabelle Bloch dans [Blo et Maî, 94] : [...] L'incertitude est relative à la vérité d'une information, caractérisant sa conformité à la réalité.

Ainsi, la notion d'incertitude peut être rapprochée de la confiance attribuée à une information. L'incertitude se réfère au manque d'information sur l'état du monde pour déterminer si une déclaration booléenne est effectivement vraie ou fausse [Blo et Hun, 01].

Cette confiance dans les données médicales est problématique. En effet, l'incertitude est omniprésente lorsqu'il s'agit de modéliser ou du moins d'expliquer le comportement d'un patient. Elle peut être la conséquence de plusieurs facteurs :

*Le pré-traitement* : des machines assistant le patient en salle de soins (voir figure 1.2), plusieurs sortes de données sont disponibles : des signaux physiologiques et des valeurs numériques échantillonnées. En général, les valeurs numériques proviennent des signaux physiologiques comme la fréquence cardiaque de l'ECG<sup>4</sup> par exemple (voir annexe A). Ce calcul est, la plupart du temps, effectué directement par les machines au lit du patient et l'accès aux algorithmes n'est pas disponible. De plus, ces algorithmes diffèrent d'une marque à une autre [Sil, 04].

*Les variables "cachées"* : de nombreux facteurs ne peuvent être quantifiables, voire même numérisés. Un médecin, en rentrant dans la chambre du patient, peut avoir une impression ou intuition sur l'état du patient par les odeurs, la couleur de la peau... [Sil, 04].

*Les relations entre paramètres* : les liens entre les paramètres sont difficilement quantifiables et mal connus des experts mais nous nous rapprochons ici de la notion de connaissance experte que nous développerons plus tard [Sil, 04].

### I.3.1.3. Imprécision

Dans [Blo et Maî, 94], l'imprécision est attachée à une anomalie quantitative :

L'imprécision concerne le contenu de l'information et porte donc sur un défaut quantitatif de connaissance, sur une mesure [.. ..]. [.. ..] Une proposition peut être imprécise ("cet homme est grand"), incertaine ("cette lettre arrivera demain") ou à la fois imprécise et incertaine ("il pleuvra beaucoup demain").

L'imprécision se réfère au contenu de l'affirmation considérée et dépend de la "granularité" du langage utilisé (par exemple, l'affirmation "il

---

<sup>4</sup> ECG : ElectroCardioGraphie



a cinq ans" est précise seulement si l'on veut un âge en années et imprécise si l'on veut un âge en mois) [Blo et Hun, 01].

L'imprécision des données d'USIs peut provenir de plusieurs facteurs :

*Le bruit* : Par exemple, il peut suffire que le patient bouge un membre pour modifier les valeurs du paramètre acquises par un capteur.

*Des mesures imprécises* : La surveillance du patient consiste principalement en la surveillance de l'évolution de son comportement. Ainsi, le personnel soignant s'attache surtout à considérer l'évolution de chaque paramètre surveillé. Il tiendra compte de l'augmentation d'un paramètre au cours de la dernière demi-heure par exemple. Le cas de l'électrocardiogramme est frappant. En surveillance, son obtention se fait par l'intermédiaire de trois capteurs. Mais, si le médecin ou l'infirmier veut une valeur précise, il rajoutera des capteurs afin d'en avoir huit pour augmenter la précision et la qualité de la mesure. La précision de mesure peut être faible voire même très faible ; elle peut atteindre 20% pour certains paramètres [Sil, 04].

Néanmoins, les notions d'incertitude et d'imprécision sont souvent difficilement dissociables et classer les données en ces termes peut être sujet à discussion. Comme le notent les auteurs de [Blo et Hun, 01]:

#### **I.3.1.4. Ambiguïté**

Les données peuvent être interprétables de plusieurs manières. Ici, ce sont bien les données qui sont ambiguës et non l'interprétation. Nous aborderons plus tard les problèmes liés aux connaissances expertes. Par exemple, une chute de la saturation en oxygène peut correspondre à au moins deux événements : une déconnexion du capteur ou une dé-saturation. Une instabilité sur un paramètre peut correspondre à du bruit, à une angoisse du patient, à un signe clinique important ou non, à un changement de thérapeutique... Dans bien des cas, pour un même changement d'évolution ou d'état du paramètre, plusieurs interprétations sont possibles et seule une analyse *a posteriori* de tous les paramètres par les experts médicaux peut permettre de discriminer ces interprétations [Sil, 04].

### **I.4. Spécificité des connaissances médicales en USIs**

#### **I.4.1. Caractéristiques des connaissances médicales**

La difficulté de la formalisation de la connaissance médicale permet de qualifier la médecine de science "non exacte". Elle est engendrée par l'incertitude des connaissances, ce qui implique un raisonnement difficilement transcribable et reproductible. Ceci est d'autant plus vrai que le taux d'erreurs de diagnostic n'est pas négligeable (puisque'il est indissociable du raisonnement humain).

##### **I.4.1.1. L'incertitude en médecine**

La médecine est un domaine où la connaissance experte est difficilement formalisable. L'interprétation de l'état du patient est sujette à l'appréciation du médecin. En effet, bien que reposant sur des connaissances

appries, l'interprétation est affinée, améliorée par l'expérience, l'intuition que l'expert aura su développer grâce à son expérience. L'incertitude, dans de justes proportions, est omniprésente quant à l'état du patient, son évolution, sa réaction à une thérapeutique ... [Szo, 95].

#### **I.4.1.2. Et le taux d'erreurs qui en découle**

La médecine est un domaine [Mor, 02], pas à proprement parler une science exacte ; les situations sont rarement formalisables [Mor, 01].

#### **I.4.2. Caractéristiques des connaissances associées aux USIs**

Outre les difficultés inhérentes au domaine médical en général, il faut rajouter certaines caractéristiques qui sont propres au domaine des USIs : la gestion d'un nombre important d'informations disponibles, de la notion de dynamique, de l'explicitation des relations entre plusieurs variables.

##### **I.4.2.1. Un nombre important d'informations à traiter**

Le clinicien est confronté à une multitude d'informations qu'il doit intégrer, "décompiler" afin de prendre une décision. De même, cette source importante d'informations l'empêche d'explicitier des connaissances intégrant autant de paramètres. Il est même tout à fait illusoire de penser qu'il est capable d'effectuer un tel exploit [Fri et al., 02].

##### **I.4.2.2. Une notion de temporalité à définir**

Alors que les données sont acquises à une fréquence d'échantillonnage élevée, le personnel médical tient compte essentiellement de l'évolution de ces données. A partir d'une succession de données instantanées, il raisonne naturellement sur des intervalles de temps. Cependant, il ne s'agit pas d'un raisonnement prédéterminé. La difficulté du passage instantané/intervalle dans l'interprétation des données est d'une complexité importante [Dec, 91].

Tout le raisonnement nécessaire pour ce passage n'est pas objectif, difficilement reproductible. Un gros décalage existe entre la temporalité des données et le raisonnement du personnel médical. Ce dernier n'est pas capable de décrire la dynamique d'un paramètre à la seconde près, ce qui, d'ailleurs, n'est pas nécessaire dans un contexte de surveillance.

##### **I.4.2.3. Des comportements du patient difficilement formalisables**

Bien évidemment des liens entre paramètres sont connus mais ils restent tout à fait flous. Par exemple, on sait que tel paramètre est lié à un tel autre. Mais ce lien n'est pas descriptible de façon précise (voir section précédente sur la temporalité par exemple). Outre ce manque de description formelle, la synthèse de toutes les données disponibles n'est pas possible pour un médecin malgré toutes les compétences qu'il peut avoir [Fri et al., 02].

De plus, il y a une réelle difficulté à obtenir des représentations de toutes les situations des patients possibles. Malgré le fait que les données

aient une haute dimension, elles ne couvrent pas toutes les situations puisque chaque patient a des réactions différentes suivant sa morphologie, sa thérapeutique... Si on rajoute la pauvreté de la formalisation médicale, on comprend d'autant mieux l'incapacité des médecins à représenter des situations typiques de comportement du patient.

### **I.5. Les systèmes de monitoring médical intelligents**

Après la percée de technologies à base de systèmes experts à la fin des années 70 (MYCIN [Sho, 76], COMPAS [Sit *et al*, 89]) et par la suite des systèmes à base de connaissances (SBC), l'intelligence artificielle a pris une grande place dans la médecine en général et dans les systèmes de monitoring en particulier. En effet, dans un contexte d'USI par exemple, la quantité de données reflétant l'état d'un patient dépasse de beaucoup la capacité de synthèse et d'assimilation du personnel soignant. Pour éviter au personnel médical d'être submergé par les données au risque de manquer certains événements importants, les unités de soins intensifs sont équipées de *systèmes de monitoring intelligents*. En théorie, ces systèmes complexes ont pour but de collecter les données, de les traiter, de les interpréter, de faire un premier diagnostic de la situation, de prévoir l'évolution du système surveillé, de proposer des actions voire même de les exécuter (cas des systèmes en boucle fermée). En pratique, très peu de systèmes utilisés en milieu clinique disposent de toutes ces facultés. En cardiologie par exemple, le diagnostic est limité à certaines arythmies dangereuses et les systèmes n'effectuent aucune action spontanément.

*Coiera* présente dans [Coi, 93] les étapes de la construction d'un système de monitoring intelligent.

*La première étape*, au niveau du traitement du signal, consiste en le développement de nouveaux capteurs ou en la sélection des capteurs permettant de donner l'information la plus pertinente.

*La seconde étape* consiste en la validation du signal fourni par les capteurs. L'auteur propose en effet d'inclure un système de gestion des alarmes intelligent qui validerait le signal avant de générer une alarme et ainsi éviterait le problème de l'émission des fausses alarmes en cas de signal trop bruité.

*La troisième étape* est la reconnaissance de motifs caractéristiques dans le signal suivie éventuellement d'une phase de transformation des données en données qualitatives. Dans le milieu médical, cette phase est encore appelée phase *d'abstraction temporelle* car beaucoup de données manipulées font référence au temps et les symptômes d'un patient évoluent eux-mêmes au cours du temps.

*La dernière phase* est une phase d'inférence conduisant au but que se donne un système de monitoring intelligent.

Nous différencions deux types de systèmes de monitoring intelligent suivant qu'ils prennent en compte une partie acquisition automatique de la connaissance ou non.

### 1.5.1. Des systèmes de monitoring médical

Dans [Uck, 93], *Uckun* présente un état de l'art sur les systèmes de monitoring médical intelligents existant en 1992. Ces systèmes ont pour but de remédier aux défauts des systèmes de monitoring classiques sur diverses tâches telles que la gestion en temps réel d'une quantité importante de données (par exemple, par des méthodes de perception sélective consistant à se focaliser sur les données les plus critiques), la manipulation de données bruitées (en utilisant, par exemple, plusieurs sources de données), la prise en compte de l'expérience médicale sous forme de connaissances pour aider le diagnostic et la diminution des fausses alarmes [Cha, 01]. Aucun de ces systèmes n'a fait l'objet d'une réelle évaluation ni d'une utilisation clinique poussée. Parmi les systèmes référencés par ce chapitre, on peut citer *Guardian* [Lar et Hay, 98], *Ganesh* et son successeur *NéoGanesh* [Doj et al, 97] ou le système *Vie-Vent* [Mik et al, 96] de Miksch et al; ces deux derniers sont de rares exemples de systèmes de monitoring ayant été testés en milieu clinique.

*Guardian* est un système de monitoring à base de connaissances dont le but est d'assurer les tâches de monitoring et de diagnostic pour des patients venant de subir une intervention chirurgicale cardiaque. Ce système a été développé sur de nombreuses années (de 87 à 98) et inclut un simulateur de scénarios, plusieurs algorithmes de diagnostic et un algorithme de planification de tâches décrits dans un langage permettant d'étendre les fonctionnalités décrite précédemment. La base de connaissances expertes contient un très grand nombre de règles concernant jusqu'à 61 pathologies différentes couvrant les cas les plus communs de complications postopératoires liées aux interventions chirurgicales cardiaques. Il n'a pas été testé en milieu clinique mais une évaluation du système a été conduite en comparant les réactions de Guardian à des scénarios cliniques critiques à celles de personnels hospitaliers sur les mêmes scénarios dans des exercices de simulation. Les résultats prouvent que des systèmes tels que Guardian peuvent s'avérer plus efficaces et naturellement moins sensibles au stress ou à la surcharge d'événements pour évaluer l'état du patient et proposer des soins lors de situations critiques.

*Vie-Vent* est un système de monitoring à base de connaissances fonctionnant en boucle ouverte et dédié à la surveillance et à la planification d'actions pour la respiration artificielle des nouveaux nés en USI. Le système est, en particulier, équipé d'un module d'abstraction temporelle perfectionné qui permet d'unifier les représentations de données quantitatives (par exemple les paramètres du ventilateur) ou qualitatives (sexe du patient, connaissances expertes, etc.) estimées de manière continue ou discontinue (cas des mesures faites à la demande). Ce module générique est réutilisable dans de nouvelles applications. *Vie-Vent* donne de bons résultats en milieu clinique lorsque les données sont fiables (peu bruitées). Il est adapté aux domaines dans lesquels les connaissances sont incomplètes car il permet, grâce aux possibilités du module d'abstraction temporelle,

d'acquérir de nouvelles connaissances à la volée. Il n'a, en revanche, pas été conçu pour la tâche spécifique du diagnostic.

**NéoGanesh** est également un système de monitoring à base de connaissances qui contrôle, cette fois en boucle fermée, l'assistance respiratoire mécanique de patients hospitalisés en USI. Ce système fonctionne avec une architecture *distribuée* de type *multi-agent* qui contient un module d'acquisition de données, un module de classification de l'état respiratoire du patient, un module d'abstraction temporelle et un module de planification d'actions. Le module d'abstraction temporelle dont les mécanismes principaux sont l'agrégation d'événements et l'oubli de certains événements non critiques, est décrit en détail dans [Doj et Chi, 97]. Le diagnostic et la planification des actions sont basés sur plusieurs bases de connaissances orientées objet qui contiennent la description des objets qui composent l'environnement et des règles permettant d'effectuer des opérations sur ces objets. Pour des raisons de sécurité et d'éthique, l'utilisation du système de monitoring en boucle fermée est limitée. Il n'est, en particulier, dédié qu'à un mode respiratoire particulier.

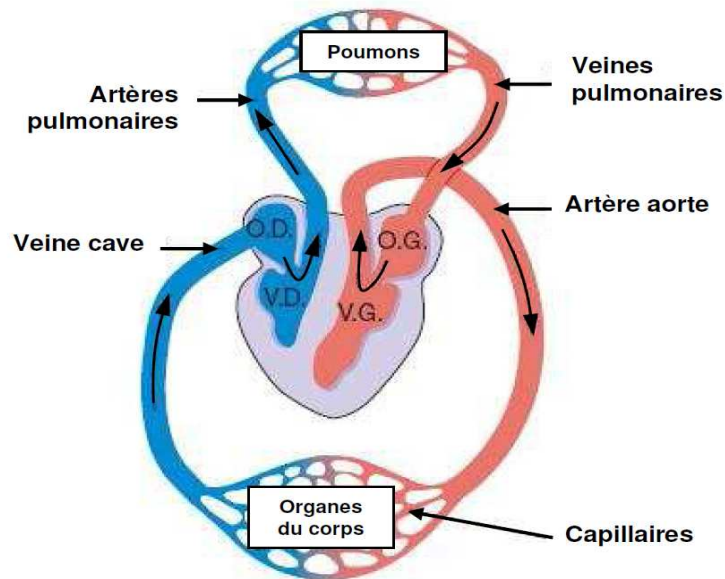
Plus récemment, le système déjà vu de **Dojat** et al. [Doj *et al*, 98] permet de reconnaître des scénarios médicaux critiques à partir d'une base de scénarios décrits de manière symbolique par des experts. Proches des chroniques de **Dousson** [Dou, 96], ces scénarios sont des réseaux de contraintes qui permettent de garder une part d'incertitude sur l'instant d'apparition des événements et ainsi permettre non seulement une reconnaissance des scénarios en ligne mais également une reconnaissance *a posteriori*.

Maintenant, dans cette session, on va vous présenter de façon détaillée le système cardiovasculaire.

## I.6. Le système CARDIOVASCULAIRE

Le système cardiovasculaire assure la circulation du sang dans l'organisme et permet ainsi son alimentation en oxygène et en nutriments. Il est composé du cœur, sorte de double pompe, qui assure la circulation dans deux réseaux complémentaires : celui des artères et celui des veines.

### I .6.1. La circulation artérielle et veineuse



**FIG 1.3:** Schéma du circuit sanguin, la couleur bleu indique que le sang est pauvre en oxygène tandis que la couleur rouge correspond au sang chargé de l'oxygène [Rém, 04].

Le réseau artériel de la grande circulation est un circuit à haute pression; il conduit le sang oxygéné à travers le corps dans des vaisseaux sanguins appelés, selon leurs tailles, *artères*, *artérioles* ou *capillaires artériels* (FIG 1.3). Ce dernier niveau est constitué de multiples petites ramifications qui facilitent le transfert de l'oxygène du sang aux organes. Le sang, devenu pauvre en oxygène, revient au cœur dans les veines, puis est envoyé par les artères pulmonaires dans la petite circulation où il est oxygéné dans les poumons. Le réseau veineux est le principal réservoir de sang : il contient environ 70% du volume total, qui est de 5 à 6 litres pour un adulte.

#### 1.6.1.1. Le cœur

Le cœur est l'élément central du système cardiovasculaire. Nous décrivons dans la suite du chapitre l'anatomie et le fonctionnement électrique d'un cœur sain.

##### I .6.1.1.1. Anatomie

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.

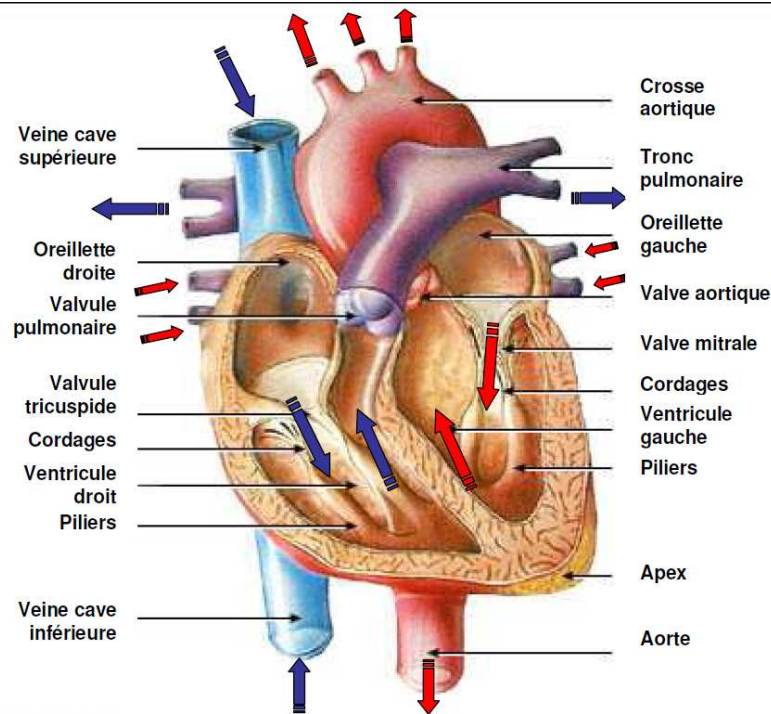
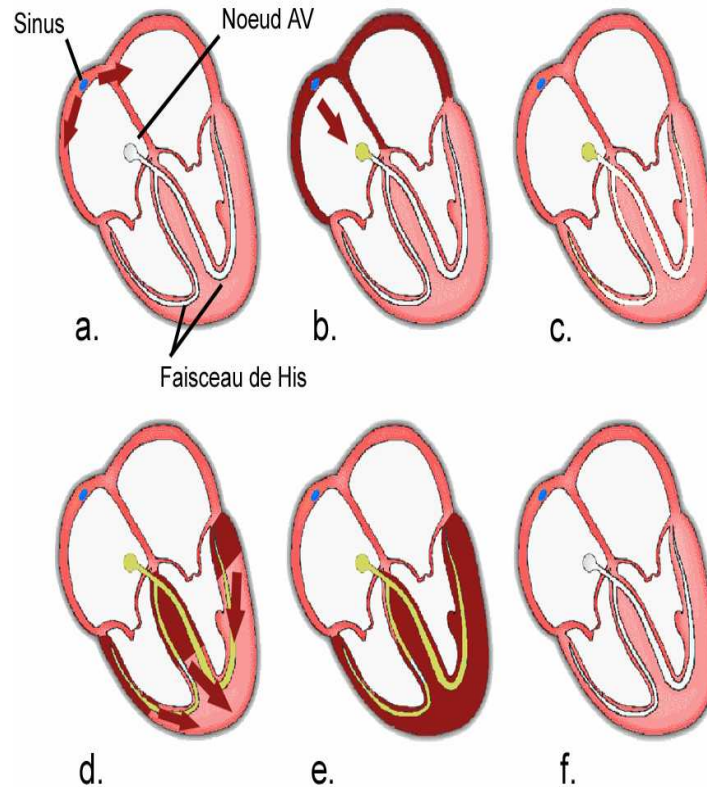


FIG 1.4: Schéma fonctionnel du cœur [Rém, 04].

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 (FIG 1.4).

#### I.6.1.1.2. Fonctionnement électrique

Comme pour tous les muscles du corps, la contraction du myocarde est provoquée par la propagation d'une impulsion électrique le long des fibres musculaires cardiaques induite par la dépolarisation des cellules musculaires. Dans le cœur, la dépolarisation prend normalement naissance dans le haut de l'oreillette droite (le sinus), et se propage ensuite dans les oreillettes, induisant la systole auriculaire (FIG 1.5) qui est suivie d'une *diastole* (décontraction du muscle).



**FIG 1.5** : L'impulsion électrique se propage dans le muscle cardiaque et induit sa contraction. Elle prend naissance dans le sinus (a) puis se propage dans les oreillettes (b) entraînant leurs contractions (systole auriculaire). L'impulsion arrive alors au nœud auriculo-ventriculaire (AV) seul point de passage électrique entre les oreillettes et les ventricules. Une courte pause est alors introduite (c) juste avant la propagation dans les fibres constituant le faisceau de His. Au passage de l'impulsion électrique (d) les ventricules se contractent à leur tour (e) (systole ventriculaire) [Rém, 04]. Après la diastole (décontraction du muscle) les cellules se repolarisent (f). Le cycle du battement cardiaque est alors terminé et le cœur est prêt pour un nouveau battement.

L'impulsion électrique arrive alors au nœud auriculo-ventriculaire (AV), seul point de passage possible pour le courant électrique entre les oreillettes et les ventricules. Là, l'impulsion électrique subit une courte pause permettant au sang de pénétrer dans les ventricules. Elle emprunte alors le faisceau de His, qui est composé de deux branches principales allant chacune dans un ventricule. Les fibres constituant ce faisceau, complétées par les fibres de Purkinje, grâce à leur conduction rapide, propagent l'impulsion électrique en plusieurs points des ventricules, et permettent ainsi une dépolarisation quasi instantanée de l'ensemble du muscle ventriculaire, malgré sa taille importante, ce qui assure une efficacité optimale dans la propulsion du sang ; cette contraction constitue la phase de systole ventriculaire. Puis suit la diastole ventriculaire (décontraction du muscle); les fibres musculaires se repolarisent et reviennent ainsi dans leur état initial.



## I.6.2. L'électrocardiographie

Ou l'art d'enregistrer l'activité électrique du cœur.

### I.6.2.1. Principe de l'électrocardiogramme (ECG)

Vers 1880, E Marey [Mar, 76] et Augustus Waller [Wal, 93] montrèrent que l'activité électrique du cœur, découverte quelques années plus tôt, pouvait être suivie à partir de la peau ; et vers 1890, Willem Einthoven réalisa le premier enregistrement cardiographique [Ein, 41].

Le courant mesuré par des électrodes sur le torse du patient mettait en mouvement un mince fil d'argent tendu entre les pôles d'un gros aimant ; ces déflexions étaient enregistrées sur du papier photographique qui se déroulait face à un rayon lumineux.

« Maintenant, nous pouvons enregistrer l'activité électrique du cœur anormale et la comparer à l'activité normale ». Ainsi naît l'électrocardiogramme à l'aube du XX<sup>ième</sup> siècle.

Le principe de l'enregistrement moderne est, à peu de chose près, celui qui fut proposé par Einthoven : grâce à deux électrodes collées à la surface de la peau, on enregistre la différence de potentiel entre deux points diamétralement opposés par rapport au cœur, ce signal étant directement corrélé au déplacement de l'impulsion électrique dans les fibres du muscle cardiaque.

### I.6.2.2. De l'ECG au Holter

En cardiologie, l'examen le plus couramment pratiqué est l'ECG 12 dériviations, où le signal électrocardiographique est visualisé selon 12 axes privilégiés : 6 axes dans le plan frontal, et 6 axes dans le plan transversal.

Sa durée peut varier de quelques secondes à une ou deux minutes ; il permet le diagnostic et la localisation précise de certaines pathologies qui laissent des traces permanentes comme, par exemple, des zones du myocarde déficientes à la suite d'un infarctus. En revanche, la courte durée de cet examen est un obstacle à la détection systématique de pathologies qui apparaissent de manière sporadique, comme certains troubles du rythme par exemple.

C'est pourquoi Norman Holter proposa, au début des années 60, un appareil « portatif » permettant d'enregistrer l'activité cardiaque pendant plusieurs heures ; cet enregistrement constitue ce que l'on appelle "l'examen Holter".

L'enregistreur Holter d'aujourd'hui se présente sous la forme d'un petit boîtier (FIG 1.6) auquel sont reliées sept électrodes, permettant ainsi l'enregistrement sur 2 ou 3 dériviations (2 frontales et 1 transversales).

Le patient se fait poser l'appareil chez un cardiologue et retourne ensuite à ses occupations habituelles. 24 heures plus tard, il revient chez le cardiologue pour se faire enlever l'appareil qui a en mémoire 24 heures d'enregistrements ECG effectués tandis que le patient a vécu une journée normale. Les résultats issus de l'analyse des quelque 100 000 battements que compte l'enregistrement permettent ainsi de diagnostiquer une plus

grande gamme de pathologies que l'ECG hospitalier. En outre, la longueur de l'enregistrement autorise par exemple le suivi du rythme cardiaque durant les phases diurne et nocturne.

L'examen Holter, examen non invasif complémentaire de l'ECG 12 dérivations, est de plus en plus prescrit ; depuis quelque temps, la communauté des cardiologues commence à envisager des enregistrements Holter de très longues durées : 1 mois par exemple, ce qui correspond à plus de 2 millions de battements enregistrés.

On comprend que l'analyse d'un tel examen n'est envisageable que parce qu'une lecture automatique des données enregistrées est aujourd'hui possible.



**FIG 1.6:** Enregistreur Syneflash® de la société *Ela Medical*. Le signal cardiaque enregistré pendant 24 heures est stocké sous forme numérique à 200Hz sur une carte à puce du type *SanDisk*®. Cette carte est ensuite directement lisible sur un ordinateur PC et l'enregistrement peut être analysé avec le logiciel *Synetec*® produit par la même société [Rém, 04].

### I.6.3. Trace électrique du cœur

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. Un œil exercé peut donc, dans la plupart des cas, différencier de manière rapide la trace d'une contraction auriculaire de la trace de contraction ventriculaire.

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 (FIG 1.8 a).

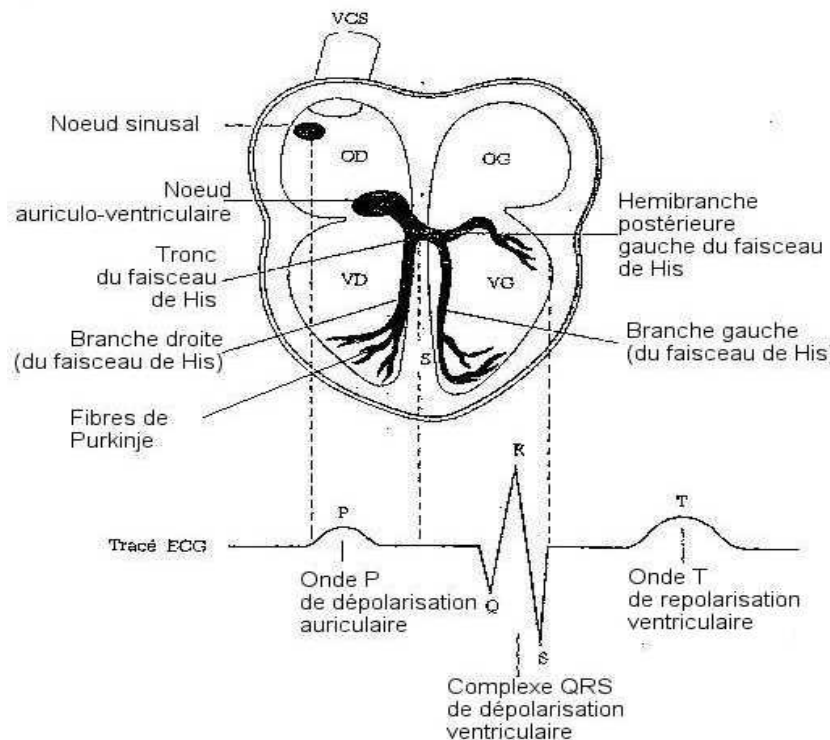
L'impulsion arrive alors au nœud auriculo-ventriculaire (AV), où se produit la courte pause qui se traduit sur l'ECG par un petit segment plat ; puis elle emprunte les voies de conductions rapides (*le faisceau de His*) pour entraîner la contraction des ventricules, suivie de leur répolarisation. Cette 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 (FIG 1.8 b). 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 [Zar, 01]. 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 (FIG 1.8 c).

## I.7. Les pathologies cardiaques

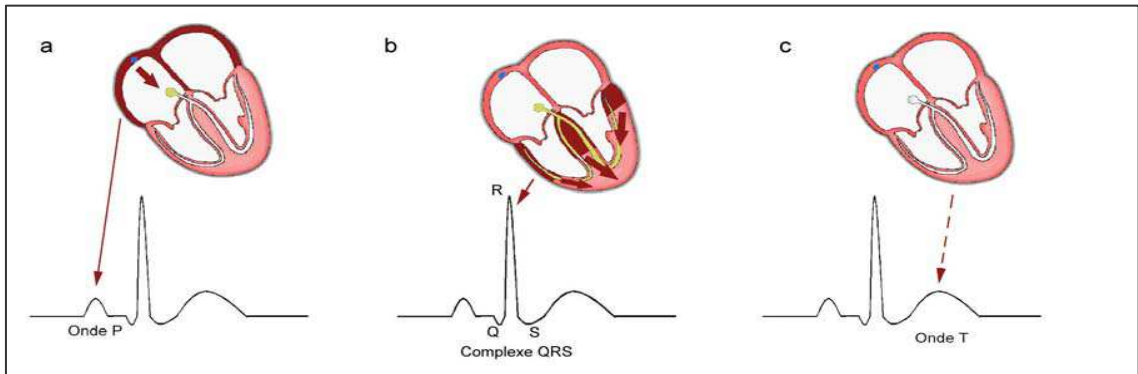
### I.7.1 Eléments fondamentaux de la conduction électrique cardiaque

Cette section est volontairement limitée aux notions utiles à la compréhension des chapitres suivants. Le lecteur intéressé pourra cependant trouver des informations complémentaires détaillées dans [Blo et Hil, 80] et de manière plus intuitive dans [Wan, 02] et [Her, 00].



**FIG 1.7 :** Eléments fondamentaux de la conduction électrique cardiaque et tracé ECG correspondant. OD, OG, VD et VG signifient respectivement oreillette gauche et droite et ventricule gauche et droit [Fro, 05].

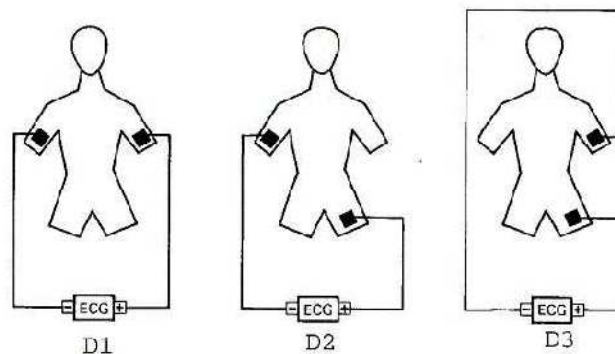
### I.7.1.1 La conduction électrique



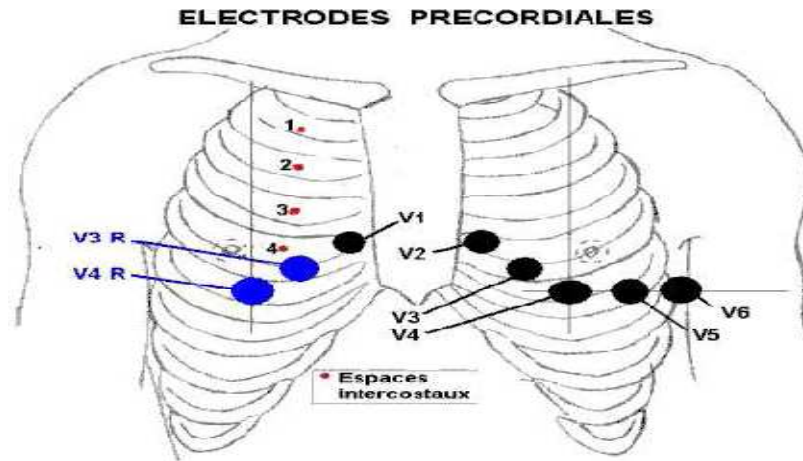
**FIG 1.8** : La contraction des oreillettes se traduit sur l'ECG par une onde positive appelée onde P (a). Elle est suivie d'une courte pause correspondant à la temporisation effectuée par le nœud (AV). La contraction brève et puissante des deux ventricules est représentée par trois ondes : l'onde Q, l'onde R et l'onde S. On parle alors de complexe QRS (b). Q est la première onde négative en début de complexe ; elle n'est pas toujours visible ; l'onde R est la deuxième onde ; elle est positive et de grande amplitude, la troisième étant l'onde S. L'onde T correspond à la repolarisation des cellules musculaires des ventricules (c). Entre cette onde et l'onde S se situe le segment ST [Rém, 04].

La contraction du muscle cardiaque (ou myocarde) a pour origine la propagation d'une onde électrique qui excite les cellules musculaires dans un ordre bien établi afin que la contraction soit la plus efficace possible. Le système de conduction électrique comprend (FIG 1.8) : le nœud sinusal, les voies spécialisées internodales, le nœud auriculo-ventriculaire, le faisceau de His, les branches gauche et droite et les fibres de Purkinje qui terminent les deux branches. Dans le cas normal, le stimulus physiologique à l'origine de l'onde électrique (et donc du battement cardiaque) provient du nœud sinusal. Ce stimulus est fourni rythmiquement à une fréquence comprise entre 60 et 100/minute. L'impulsion cardiaque initiée dans le nœud sinusal est ensuite transmise aux deux oreillettes au moyen des voies internodales qui relient le nœud sinusal au nœud auriculo-ventriculaire. L'onde se propage ensuite à travers le faisceau de His puis vers les branches gauche et droite jusqu'aux fibres de Purkinje. La propagation de l'onde se fait par une succession de dépolarisation des tissus du myocarde.

Après le passage de l'onde, la membrane des tissus se repolarise. Cette succession de dépolarisations et de repolarisations des tissus cardiaques est visible sous plusieurs angles sur un électrocardiogramme (ECG).



**Fig. 1.9** : Dérivations bipolaires (voies I II et III) [Fro, 05]



**Fig. 1.10** : Dérivations précordiales (voies V1 V2...V6) [Por, 05]

L'électrocardiogramme L'ECG reflète l'activité électrique cardiaque. L'électrocardiographie moderne s'appuie sur l'étude de douze dérivations (ou dérivations standards). Ces dernières sont obtenues en plaçant des électrodes sur le thorax près du cœur, les précordiales (6 voies de V1 à V6), ou sur les bras et les jambes (dérivations bipolaires D1 à D3 et unipolaires aVR, aVL et aVF) comme indiquées sur les Figures 1.9 et 1.10. Ces différentes dérivations permettent d'avoir plusieurs vues de la propagation électrique. Lorsqu'il n'y a aucun dysfonctionnement, ces vues sont redondantes, on se sert alors principalement des voies bipolaires I et II. L'ensemble des vues est cependant nécessaires en cas de dysfonctionnement pour localiser précisément la source du problème. Lors de la propagation de l'onde électrique dans le cœur, les électrodes permettent d'enregistrer un motif électrique composé d'une succession d'ondes caractéristiques désignées par les lettres de l'alphabet à partir de P (FIG 1.10) [Fro, 05].

L'onde P initiale représente la dépolarisation des oreillettes. L'ensemble QRS (appelé aussi complexe QRS) traduit la dépolarisation des ventricules tandis que l'onde T correspond à leur repolarisation (la repolarisation des oreillettes est masquée par celle des ventricules). L'onde U, présentée sur la Figure 1.11 est inconstante et de faible amplitude, elle a une signification encore discutée.

### **I.7.2. Le battement cardiaque standard et ses caractéristiques**

L'étude d'un enregistrement ECG est fondée sur l'analyse de quelques battements cardiaques successifs ; l'étude d'un seul battement ne fournit que peu d'indications pour la pose d'un diagnostic, mais les variations des paramètres caractéristiques de chaque battement au cours de l'enregistrement constituent une source d'information essentielle. Ces paramètres caractéristiques sont :

Les durées des ondes P, Q, R, S et T, et les amplitudes mesurées par rapport à la ligne de base (FIG 1.11). La ligne de base, est la ligne isoélectrique du cœur au repos, qui est prise comme référence pour mesurer l'amplitude des ondes : pendant l'inactivité cardiaque, le potentiel mesuré est donc normalement nul par rapport à cette référence. C'est le cas au niveau de l'intervalle entre les ondes T et P de deux battements successifs, de l'intervalle entre les ondes P et Q d'un même battement, de l'intervalle entre les ondes S et T (en l'absence de pathologie) Les distances entre ces ondes.

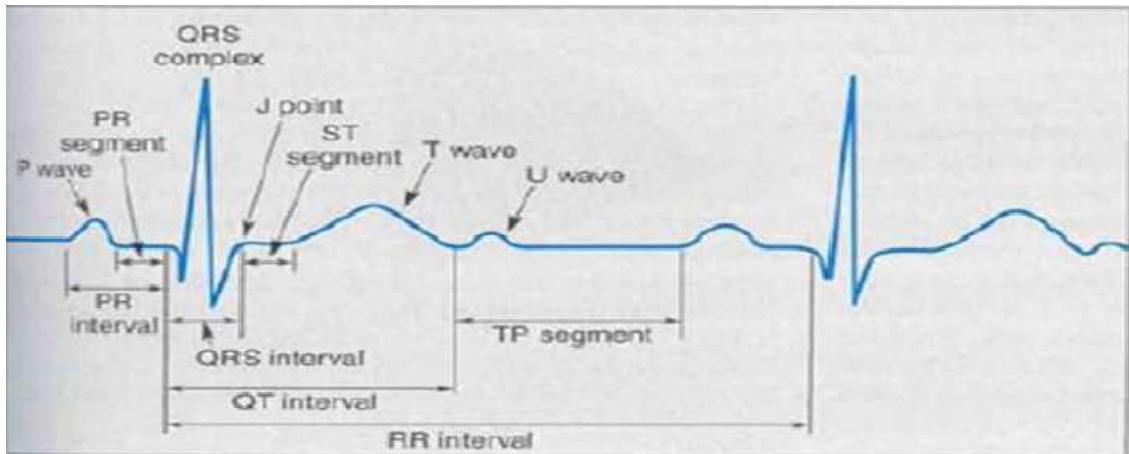


FIG 1.11: Paramètres d'intérêt pour la description d'un battement [CHA, 09].

Pour caractériser le rythme cardiaque, les cardiologues se fient à la forme générale de l'ECG. Ils portent une attention particulière à la forme des ondes citées précédemment et aux intervalles de temps qui les séparent. La FIG 1.11 répertorie les intervalles qui seront utilisés dans la suite de cette étude. Les valeurs numériques utilisées proviennent de la littérature médicale et notamment de [Ste, 99].

- l'intervalle RR (ou RR1) sépare les sommets de deux ondes R consécutives et définit le rythme ventriculaire. Un intervalle RR normal est généralement mesuré entre 600 et 1000 ms,
- l'intervalle PP (ou PP1) sépare les sommets de deux ondes P consécutives et définit le rythme auriculaire,
- l'intervalle PR (ou PR1) s'étend du début de l'onde P au début du complexe QRS. Un intervalle PR normal mesure entre 120 et 200 ms,
- l'intervalle RP (ou RP1) s'étend du début du complexe QRS jusqu'à l'onde P suivante,
- les intervalles RR2, PP2 séparent respectivement les sommets de deux QRS (resp. onde P) séparés par un QRS (resp. une onde P).

Les instants des ondes Q, S et T sont encore très difficiles à détecter automatiquement et ne seront donc pas utilisés dans cette étude. La régularité du rythme est mesurée en termes de variation de durée d'un intervalle RR à l'intervalle suivant.

Les intervalles de temps définis entre deux ondes caractéristiques de l'ECG fournissent d'importants indicateurs pour le diagnostic de maladies cardiaques car ils sont le reflet de processus physiologiques. Par exemple, l'intervalle **R-R**, représentatif de la période cardiaque, est généralement obtenu en détectant le pic de l'onde **R**. De part sa grande amplitude, la détection de cette onde est la plus accessible et nous permet souvent de segmenter l'ECG avant d'estimer toute autre onde.

L'intervalle **Q-T**, reflétant la durée de la dépolarisation et de la repolarisation ventriculaire, est l'un des intervalles les plus importants de l'ECG. En effet, sa prolongation peut être associée à des risques d'arythmie ventriculaire et de mort subite. Il est classiquement calculé entre le début du complexe **QRS** et la fin de l'onde **T**. Cependant, la fin de l'onde **T** est difficile à détecter à cause de la morphologie très variée de l'onde **T**, et surtout de la perturbation du bruit notamment sur la fin de l'onde.

Tout comme pour l'onde **T**, la détection de l'onde **P** n'est pas évidente, voire davantage complexe. Effectivement, l'onde **P** est souvent de faible amplitude et noyée dans le bruit.

L'estimation des intervalles **P-R** est donc très difficile. Alors que beaucoup d'études portent sur la détection de l'intervalle **Q-T**.

## I.8. Diagnostic à partir du rythme

Pour dresser un diagnostic à partir d'un enregistrement Holter, deux points essentiels sont à étudier : le rythme cardiaque (fréquence et régularité) et la forme des ondes (distances relatives, amplitudes, durées,...).

### I.8.1 Troubles du rythme cardiaque

La rythmologie, partie de la cardiologie qui traite des anomalies du rythme cardiaque, a fait au cours de la fin du 2<sup>ème</sup> millénaire des progrès énormes. Le nœud sinusal est le centre rythmogène ou pacemaker du cœur. Il envoie et contrôle les impulsions pour que le cœur batte de 60 à 80 fois par minute lorsque la personne est au repos.

Les arythmies sont un changement dans la fréquence des battements cardiaques qui peuvent devenir plus rapides, plus lents ou irréguliers. Beaucoup d'arythmies légères peuvent ne nécessiter aucun traitement, mais certaines formes arythmiques sont mortelles. Ainsi un cœur affaibli peut être la cause d'une arythmie, mais l'arythmie en soi peut affaiblir le cœur. Cependant, il convient également de préciser qu'un tracé d'un électrocardiogramme peut être normal et coïncider avec une arythmie particulière.

Nous classons les arythmies en fonction de leur lieu de formation et de la manière dont se propage l'excitation émise par le nœud sinusal. Si le trouble se déclare dans les oreillettes ou le nœud auriculo-ventriculaire, nous parlons d'arythmie supra-ventriculaire. S'il apparaît dans les ventricules, nous parlons d'arythmie ventriculaire. Si une arythmie accélère

le rythme cardiaque, nous parlons de tachycardie ; si elle le ralentit, nous parlons alors de bradycardie.

Les paragraphes suivants introduisent les types de troubles de rythmes et de conduction pour donner des connaissances essentielles [Jea, 94], [Gil, 82].

### I.8.1.1 Troubles du rythme supra-ventriculaire

Les troubles du rythme supra-ventriculaire sont des anomalies rythmiques trouvant leur origine en amont des ventricules, à savoir au niveau de l'oreillette (nœud sinusal ou autre origine) ou au niveau du nœud auriculo-ventriculaire (jonctionnel). Ces troubles ont des caractéristiques électrocardiographiques générales identiques (sauf pathologie particulière associée).

#### I.8.1.1.1 Anomalies rythmiques d'origine sinusale

- a- Tachycardie sinusale :** la tachycardie sinusale correspond à un rythme sinusal dont la fréquence est comprise entre 100 et 180 battements par minute ; on observe, avant chaque complexe QRS, une onde P de forme identique à celle observée lors du rythme normal (FIG 1.12).

Les causes de cette pathologie sont habituellement extracardiaques et incluent tous les facteurs de stimulation du système nerveux : surmenage, anxiété,... et l'influence de différentes substances comme l'adrénaline ou la caféine [Chi, 05].

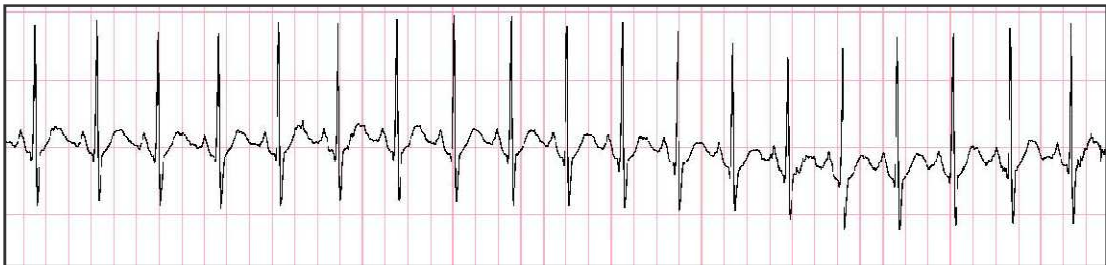


FIG 1.12 : Tachycardie sinusale [Rém, 04].

- b- Bradycardie sinusale :** outre sa fréquence basse (inférieure à 60 bpm), la bradycardie sinusale est caractérisée par la présence systématique d'une onde P avant les complexes QRS, car l'origine de la contraction ventriculaire reste la dépolarisation du sinus et des oreillettes, comme lors des battements normaux ((FIG 1.13). les causes d'une telle arythmie sont multiples et souvent extérieures au système cardio-vasculaire : traitement médicamenteux, hypothermie, urémie, etc [Chi, 05].



FIG 1.13 : Bradycardie sinusale [Rém, 04].



- c- **Arythmie sinusale** : l'arythmie sinusale est une variation de la fréquence cardiaque liée aux variations de la pression intrathoracique qui accompagne la respiration. Elle prend naissance dans le sinus, entraînant successivement les oreillettes et les ventricules avec une fréquence irrégulière. L'ECG est de morphologie normale, mais les complexes QRS sont irrégulièrement espacés, avec une variation de plus de 0,12 seconde dans la durée des intervalles PP successifs (FIG 1.14) [Chi, 05].

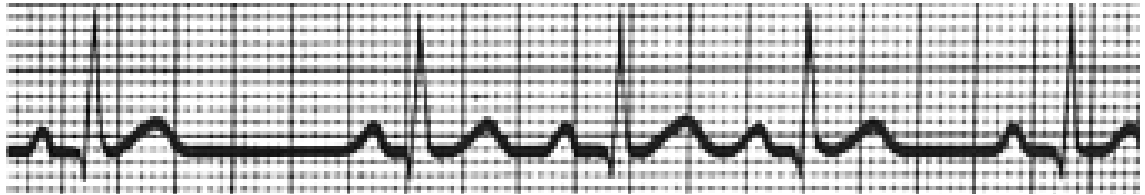


FIG 1.14 : Arythmie sinusale.

#### I.8.1.1.2 Anomalies rythmiques d'origine auriculaire

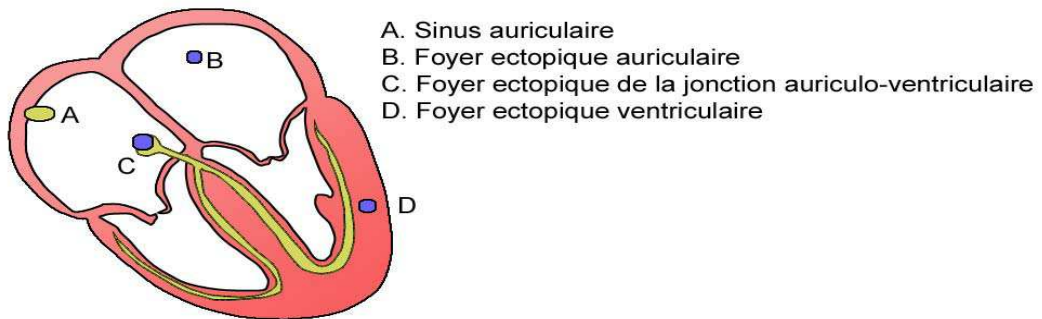
- d- **Tachycardie auriculaire** : la tachycardie auriculaire peut avoir pour origine un foyer ectopique, ou une voie qui court-circuite la voie AV, dite voie accessoire, avec réentrée par le nœud AV. Dans le cas d'un foyer ectopique, il s'agit d'un groupe de cellules situées dans les oreillettes, qui se dépolarisent spontanément et plus rapidement que le sinus, prenant ainsi sa place. La dépolarisation des oreillettes n'étant pas d'origine sinusale, la propagation de l'influx nerveux diffère de celle qui a pour origine le sinus et l'on observe une onde P de la forme inhabituelle. La fréquence typique de ce type de foyer est comprise entre 120 et 200 bpm ; en l'absence de problèmes de conduction AV, les ventricules sont entraînés au même rythme (FIG 1.15) [Chi, 05].



FIG 1.15 : Tachycardie auriculaire [Rém, 04].

- e- **Foyer ectopique** : Lorsqu'il y a un trouble de la production ou de la propagation de l'impulsion cardiaque, un certain nombre de pacemakers de réserve sont susceptibles d'engendrer des battements cardiaques, ce sont des foyers ectopiques, classés suivants leur localisation. Le rythme peut alors apparaître « régulier » si leur activité est continue (rythme d'échappement), « régulièrement irrégulier » (lorsque ces foyers n'enclenchent pas tous les battements mais se manifestent fréquemment et régulièrement, comme dans le cas des bigéminismes, par exemple) ou « irrégulier » (lorsque ces

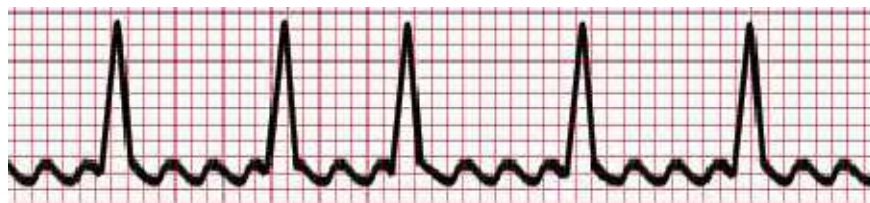
foyers prennent la commande de manière aléatoire ou lorsque leur transmission aux ventricules est aléatoire, dans le cas de foyers ou de boucles supra-ventriculaires (FIG 1.16) [Chi, 05].



**FIG 1.16** : Emplacement des foyers ectopiques.

**f- Flutter auriculaire** : le flutter auriculaire est un phénomène électrocardiographique dû à un foyer ectopique auriculaire unique, qui se traduit par des contractions régulières et coordonnées des oreillettes avec une fréquence qui peut atteindre 350 bpm. A cette fréquence, le nœud AV ne parvient pas à conduire toutes les impulsions électriques vers les ventricules et l'on observe généralement des blocs, au cours desquels seule une contraction auriculaire sur deux est suivie d'une systole ventriculaire. Le tracé ECG a les caractéristiques suivantes (FIG 1.17).

- les complexes auriculaires sont réguliers, morphologiquement anormaux : onde F de flutter avec aspect en dent de scie.
- les complexes QRS-T sont normaux [Chi, 05].



**FIG 1.17** : Flutter auriculaire.

**g- Fibrillation auriculaire** : dans cette arythmie, l'activité auriculaire n'est plus composée de battements ; soumis à plusieurs foyers ectopiques, le fonctionnement du myocarde auriculaire est totalement désorganisé. Ces mouvements anarchiques peuvent néanmoins transmettre quelques impulsions au nœud AV, qui les transmet à son tour aux ventricules et entraîne leurs contractions. Comme cette transmission revêt un caractère aléatoire, le rythme des complexes QRS est complètement irrégulier. Le tracé ECG a les caractéristiques suivantes (FIG 1.18) :

- Les complexes auriculaires sont irréguliers, morphologiquement anormaux : ondes f de fréquence 400 à 600 bpm.
- Les complexes QRS-T sont normaux mais inéquidistants [Chi, 05].

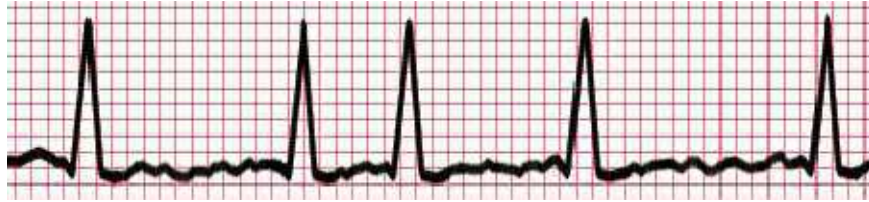


FIG 1.18 : Fibrillation auriculaire.

**k- Extrasystole auriculaire (ESA) :** c'est une systole isolée ou répétitive dont le point de départ de l'influx se situe dans l'oreillette avec une conduction et une activation normales. Sur l'ECG, elle est représentée par une onde P' de morphologie différente de l'onde P d'origine sinusale. Elle peut être visible ou confondue dans le reste du tracé ECG. La régularité de l'intervalle PP est troublée par P' : P'R peut être plus court, égal ou plus long à PR. Le QRS suivant P' est souvent identique aux QRS suivant P, mais avec souvent une pause post-extrasystolique compensatrice (FIG 1.19) [Chi, 05].

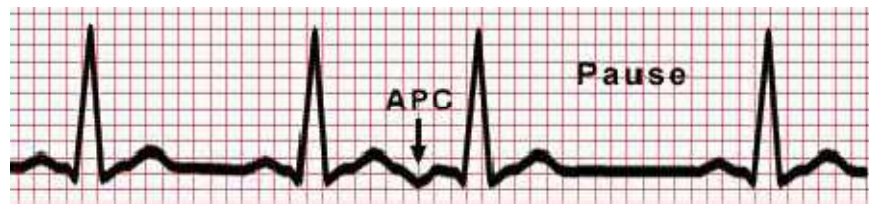


FIG 1.19: Extrasystole auriculaire ESA.

**l- Extrasystole auriculaire avec aberration ventriculaire :** c'est une extrasystole auriculaire avec un complexe QRS de morphologie différente des complexes QRS d'origine sinusale, avec un aspect de bloc de branche droit lié à un trouble intermittent, fonctionnel, de la conduction de la branche droite du faisceau de His.

**m- Battement auriculaire d'échappement :** Si le nœud sinusal est inactif, un foyer dans l'oreillette va assurer la commande du cœur et on va avoir un rythme auriculaire d'échappement avec une onde P anormal et un complexe QRS et une onde T normaux, avec une durée RR longue.

### I.8.1.1.3 Anomalies rythmiques d'origine auriculo-ventriculaire

**a- Tachycardie jonctionnelle :** En réalité, le nœud auriculo-ventriculaire lui-même n'a pas une automaticité suffisante pour être un pacemaker. Par contre, les zones directement en contact avec ce nœud (faisceau de His en aval, fibres auriculaires en amont) en ont une. La fréquence cardiaque varie entre 120 et 200 bpm. Aussi, selon l'origine de cette anomalie rythmique, on aura divers aspects

électrocardiographiques (position de l'onde P' par rapport au complexe QRS) (FIG 1.20).

- Onde P' précessive si foyer prénodal ;
- pas d'onde P' visible si foyer paranodal ;
- onde P' rétrogradé, après le complexe QRS si foyer postnodal [Chi, 05].

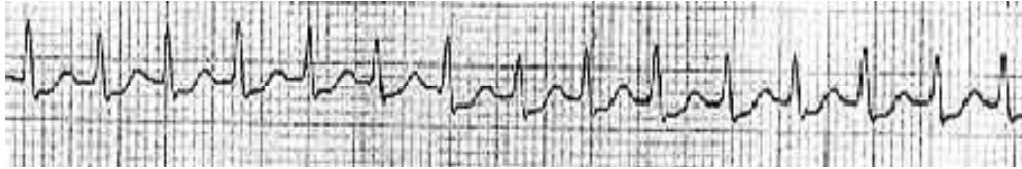


FIG 1.20: Tachycardie jonctionnelle.

**b- Battement jonctionnel d'échappement nodal** : si la région autour du nœud auriculo-ventriculaire génère l'excitation avant le nœud sinusal, le rythme s'appelle un rythme jonctionnel d'échappement. Au niveau de l'ECG, on aura un cycle cardiaque retardé sans onde P (en raison d'une absence de complexe QRS) ou une onde P inversée. La morphologie réelle de l'onde P dépend de la distance entre le foyer et le nœud auriculo-ventriculaire. L'onde de la dépolarisation sera toujours inversée parce qu'elle traverse les oreillettes dans la direction opposée comparée au rythme de sinus.

**c- Extrasystole jonctionnelle (ESJ)** : elle naît du nœud auriculo-ventriculaire et de l'activation ascendante des oreillettes, elles donnent des ondes P' négatives en D2, D3 et aVF (ou sans onde P). Les complexes QRS sont le plus souvent fins, de morphologie identique au tracé en rythme sinusal. Les extrasystoles supra-ventriculaires sont retrouvées chez des sujets normaux. Elles sont habituellement peu nombreuses (inférieures à 10 battements par heure).

### I.8.1.2 Troubles du rythme ventriculaire

Les troubles du rythme ventriculaires sont les anomalies trouvant leur origine en aval du nœud auriculo-ventriculaire (d'Aschoff-Tawara), dans les ventricules, ce qui leur donne des caractéristiques générales différentes des troubles supra-ventriculaires.

**a- Battement ventriculaire prématuré** : Contrairement aux battements normaux qui ont pour origine la dépolarisation des cellules sinusale, le BVP naît de la dépolarisation spontanée d'un petit groupe de cellules ventriculaires, appelé alors foyer ectopique ventriculaire. L'impulsion électrique créée n'emprunte pas la voie normale de conduction (faisceau de His), et se propage donc plus lentement dans les ventricules. La conduction ventriculaire ainsi étalée dans le temps perd de son efficacité. Le tracé d'un battement BVP est caractérisé par deux propriétés : l'onde R n'est pas précédée d'une onde P, puisqu'il n'y a pas eu d'activité auriculaire préalable, et la durée du complexe est supérieure à la durée d'un complexe QRS normal (FIG 1.21)[Chi, 05] .

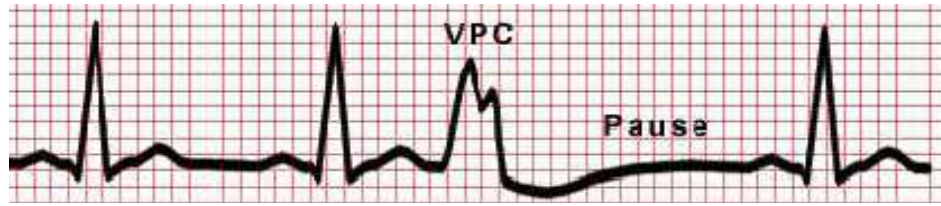


FIG 1.21 : Battement ventriculaire prématuré.

- b- **Extrasystole ventriculaire dans T (R sur T)** : les battements ectopiques ventriculaires peuvent survenir en même temps que l'onde T du battement précédent. Lors de l'installation d'un infarctus du myocarde aigu, de telles ectopies "R sur T" peuvent déclencher des arythmies ventriculaires.
- c- **Battement de fusion** : Il y'a fusion des battements ventriculaires et des battements normaux. Le complexe de fusion s'observe quand les ventricules sont activés par un influx auriculaire et un influx ventriculaires survenant de façon simultanée. Le complexe QRS résultat a donc une forme intermédiaire entre le complexe d'origine sinusale et le complexe extrasystolique.
- d- **Tachycardie ventriculaire** : la TV est régulière avec un rythme de plus de 150 bpm, des complexes QRS larges (durée supérieure ou égale à 0,12 s), avec dissociation auriculo-ventriculaire (les ventricules battent plus vite que les oreillettes). C'est donc très important de bien chercher les ondes P (FIG 1.22) [Chi, 05].



FIG 1.22 : Tachycardie ventriculaire.

- e- **Flutter ventriculaire** : le rythme ventriculaire est très rapide (supérieure à 250 bpm). Les ondes P ne sont pas visibles. Il se distingue de la fibrillation ventriculaire par l'aspect relativement organisé des QRS qui donnent un aspect sinusoïdal au tracé ECG.
- f- **Fibrillation ventriculaire** : le rythme ventriculaire est très rapide, complètement désorganisé. Les ondes P ne sont pas visibles et les QRS sont remplacées par une activité électrique anarchique, très polymorphe (FIG 1.23) [Chi, 05].



FIG 1.23 : Tachycardie ventriculaire.

- g- Battement ventriculaire d'échappement :** Si les deux nœuds de dépolarisation se ralentissent un foyer ventriculaire assure la commande du cœur ; au niveau de l'ECG on aura un complexe QRS large retardé de morphologie anormale.
- h- Torsade de pointe :** on parle de torsade de pointe devant un tracé de tachycardie ventriculaire à présentation particulière, à savoir que les complexes QRS ont une amplitude et une polarité variable, la fréquence varie entre 200 et 250 bpm. Le tracé ECG est caractérisé par un intervalle QT long (FIG 1.24) [Chi, 05].

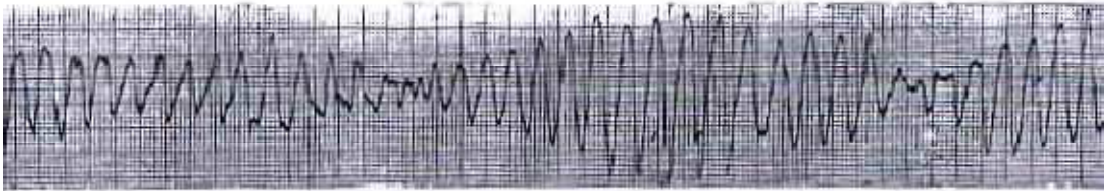


FIG 1.24 : Torsade de pointe.

### I.8.1.2.1 Trouble de conduction

Nous parlons de troubles de conduction devant toute anomalie dans le cheminement de la dépolarisation (durée de passage augmentée par un bloc). Le bloc est exemple de problèmes de conduction. C'est un défaut localisé de propagation de l'impulsion électrique dans le tissu cardiaque. Lorsqu'il est complet, c'est-à-dire que l'absence de conduction est totale, des pacemakers de réserve comme un foyer ectopique auriculaire, ou le nœud AV, peuvent prendre le relais et entraîner des bradycardies ou tachycardies. A l'inverse, lorsqu'il apparait de manière sporadique, suivant le cas, il peut se manifester par des troubles du rythme : on observe une irrégularité. Au niveau des voies de conduction rapide de l'influx nerveux dans le cœur, il existe essentiellement trois types de " blocs" classés suivant leurs localisations : le bloc sino-auriculaire (SA) situé entre le sinus et les oreillettes, le bloc auriculo-ventriculaire (AV), entre les oreillettes et les ventricules, et le bloc de branche (droit et/ ou gauche) au niveau de la transmission à chacun des deux ventricules ; ce sont surtout les premiers (blocs SA et AV) qui sont susceptibles d'entraîner des troubles du rythme. Notons que n'importe quelle partie du myocarde peut constituer un "bloc", par dégénérescence ou mort cellulaire, dans une zone infarctée par exemple, et un tel bloc peut modifier la propagation de l'influx nerveux en créant des zones de réverbération des ondes à l'origine d'éventuelles boucles, qui peuvent engendrer une arythmie.

#### I.8.1.2.1 Bloc sino-auriculaire :

Dans le cas du bloc SA, l'impulsion électrique issue du sinus n'est pas transmise aux oreillettes ; les muscles auriculaires et ventriculaires ne se contractent donc pas. Sur le tracé ECG, nous observons l'absence d'un battement là où régulièrement il devrait y en avoir un. De plus, si le bloc est bien installé, il est possible que plusieurs

impulsions du sinus ne soient pas transmises, dans un tel cas, nous observons souvent la prise de relais par un ou des pacemakers de réserve (foyer ectopique).

### I.8.1.2.2 Bloc auriculo-ventriculaire

Ce sont des blocs de conduction du courant de dépolarisation entre le nœud sinusal et les myocytes ventriculaires. Selon le type de blocage, il existe quatre types de blocs auriculo-ventriculaire (B.A.V) (FIG 1.25).

- B.A.V du 1<sup>er</sup> degré : PR allongé (supérieur à 0,12 seconde) ;
- B.A.V du 2<sup>ème</sup> degré type 1 : allongement progressif du PR jusqu'à l'onde bloquée ;
- B.A.V du 2<sup>ème</sup> degré type 2 : blocage intermittent avec passage de conduction sur 2, 3 et sur 4.
- B.A.V du 3<sup>ème</sup> degré: bloc complet, avec dissociation-auriculo-ventriculaire [Chi, 05].

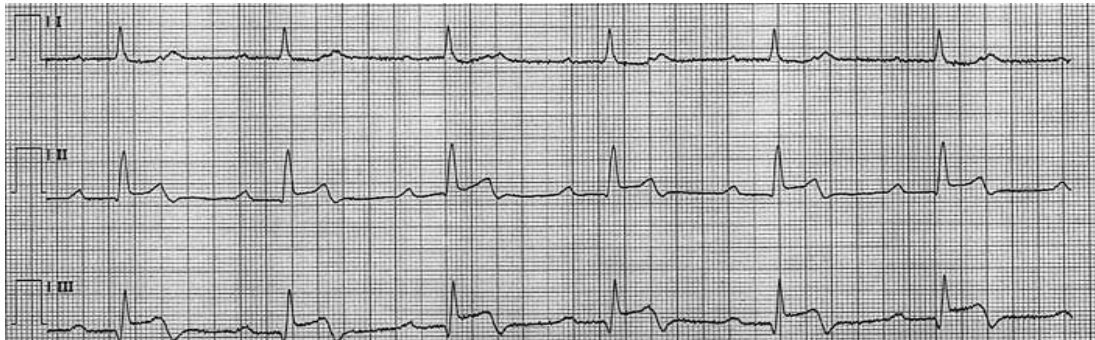


FIG 1.25 : Bloc auriculo-ventriculaire (B.A.V) du 1<sup>er</sup> degré.

### I.8.1.2.3 Bloc de branche

Un bloc de branche est un bloc de conduction situé sur une branche du faisceau de His. Sur le tracé électrocardiographique, ceci se traduit par des anomalies du complexe QRS sans atteinte de l'espace PR. Les caractéristiques générales sont l'anomalie du segment ST (donc de la repolarisation) et des complexes QRS élargis (supérieur 0,12 s).

#### I.8.1.2.3.1 Bloc de branche gauche (BBG)

Les différents signes électriques à chercher sont les suivants (FIG 1.26) :

- aspect QS en V1, V2 et V3 ;
- aspect R exclusif en D1, VL, V6 ;
- retard à l'apparition de la déflexion intrinséquoïde en V6 supérieure à 0,08 s ;
- Parfois une déviation axiale gauche modérée ;
- Troubles de la repolarisation à type d'ondes T négatives dans les précordiales gauches.

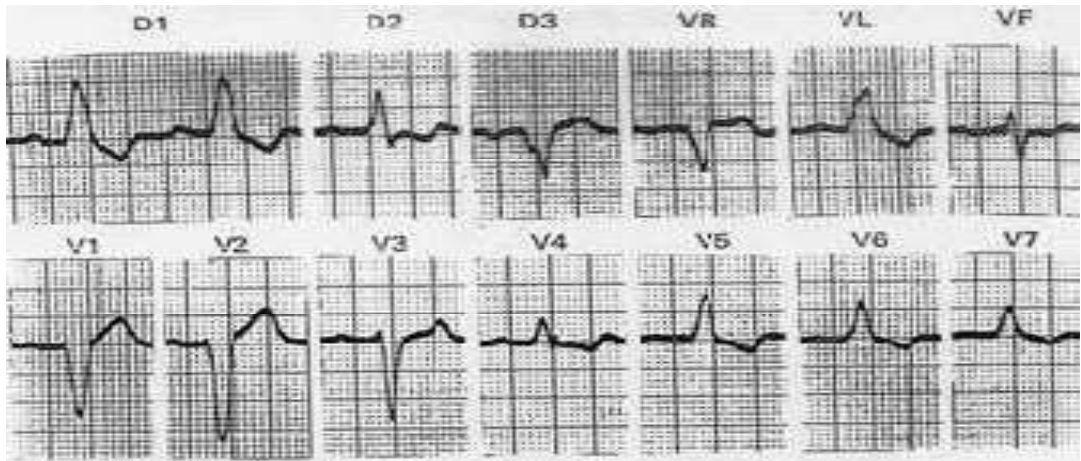


FIG 1.26 : Blocs de branche gauches [Chi, 05].

#### I.8.1.2.3.2 Bloc de branche droit (BBD)

On recherche les signes électriques suivants (FIG 1.27)

- aspect RSR' en V1 avec  $R < R'$  ;
- aspect RS en D1, VL et V6.

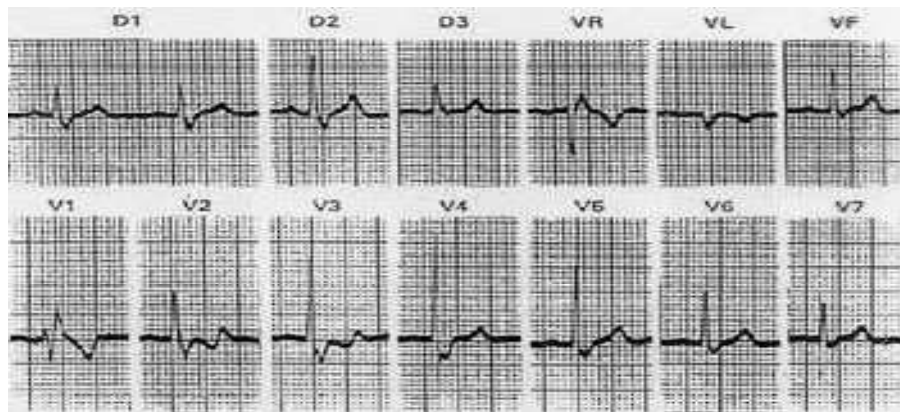


FIG 1.27 : Blocs de branche droits [Chi, 05].

#### I.8.1.2.4 Syndrome de Wolff-Parkinson-White

Ce syndrome a été décrit en 1930 par l'équipe Wolff-Parkinson-White. Il rentre dans le cadre des phénomènes de "conduction accélérée". Ce syndrome rare touche surtout les hommes (70 %), plutôt jeunes, et son pronostic est en général bénin. Il existe dans ce type de syndrome une voie de conduction anormale entre les oreillettes et les ventricules, dite une excitation précoce des ventricules : c'est le phénomène de préexcitation ventriculaire. On voit sur le tracé ECG un espace PR court (inférieur à 0,10 sec) et un complexe QRS élargi (supérieur à 0,12 sec).



### **I.8.1.2.2 Autres anomalies cardiaques**

#### **I.8.1.3.1 Infarctus du myocarde**

L'infarctus myocardique résulte en une nécrose du muscle cardiaque qui se traduit par des zones électriquement inactives. Il entraîne la disparition d'un certain nombre de vecteurs électriques et ce silence localité démasque des vecteurs opposés, modifiant l'équilibre physiologique des forces électriques de dépolarisation. On observe donc habituellement un déplacement de la boucle QRS dans une direction opposée à celle de la nécrose. A la phase aigüe de l'infarctus, au cours des premières heures, le segment ST est considérablement modifié, puis l'onde T s'inverse. Déjà à ce moment ou un peu plus tard apparaît une onde Q de nécrose. Les enregistrements électriques permettent de diagnostiquer avec précision sa localisation. Schématiquement, sur l'ECG, une onde anormale apparaît ainsi qu'une onde R très augmentée en V1 et V2.

#### **I.8.1.3.2 Hypertrophie**

L'hypertrophie est un terme utilisé pour indiquer l'augmentation des dimensions d'une cavité en épaisseur et / ou en diamètre. Elle peut être consécutive à une charge de travail excessive. C'est ainsi que l'hypertrophie ventriculaire gauche (HVG) a souvent comme origine une anomalie valvulaire aortique, une hypertension artérielle ou une myocardiopathie, tandis que l'hypertrophie ventriculaire droite (HVD) relève le plus souvent d'une anomalie valvulaire mitrale ou d'une insuffisance pulmonaire. Dans le diagnostic électrique d'une HVG, l'augmentation de l'amplitude de l'onde QRS représente un critère essentiel, mais pas toujours fiable parce que cette amplitude a une importante variable, tant interindividuelle qu'intraindividuelle, provoquée par de nombreux facteurs cardiaques et extracardiaques. Sur l'ECG, on observe une déviation à droite de l'axe de QRS, des ondes S augmentées en V5 et V6 et dans les formes importantes une grande onde R en V1 et V2 [Chi, 05].

## **I.9. Conclusion**

Le cœur est l'élément central du système cardiovasculaire qui permet l'alimentation en oxygène et en nutriments des organes. Il est principalement composé de quatre cavités : deux oreillettes et deux ventricules. À chaque battement, un influx électrique traverse le réseau de conduction du cœur et engendre les contractions successives des oreillettes et des ventricules qui envoient le sang vers les poumons ou vers le reste du corps. Comme nous le verrons dans le chapitre suivant, la propagation de cet influx électrique peut être suivie depuis l'extérieur du corps par des électrodes, collées à la surface de la peau, qui mesurent indirectement l'activité électrique des fibres musculaires cardiaques.

L'électrocardiographie est l'étude des variations de l'enregistrement de l'activité électrique des cellules cardiaques, dont dépend la contraction du

cœur. Le signal électrique enregistrable est l'électrocardiogramme (ECG). Ce signal, modifié en cas d'anomalie de la commande de l'influx électrique ou de sa propagation, de la masse globale et régionale des cellules ou de leur souffrance éventuelle, donne des renseignements importants et très utilisés en médecine.

L'étude des intervalles du tracé électrocardiographiques permet donc de mieux comprendre comment fonctionne le cœur. La suite du manuscrit présentera alors des méthodes d'extraction de certains intervalles (intervalles R-R, Q-T, P-R) puis nous tenterons d'en faire l'analyse et de tirer des conclusions, ou du moins quelques pistes, sur le fonctionnement du cœur.

Dans le chapitre suivant, nous allons vous donner un état de l'art sur les méthodes de classification, en donnant pour chacune d'elles un exemple, les avantages et les inconvénients afin de juger le choix des réseaux de neurones artificiels et les systèmes expert dans ce mémoire.



## II. LES SYSTEME MULTI-AGENTS

### II.1. Introduction

Les systèmes multi-agents (**SMA**) sont devenus l'un des courants systèmes les plus prometteurs. Le champ des applications des SMA s'est élargi avec le récent développement des réseaux de télécommunication. Le grand essor de ces derniers a fait apparaître de nouvelles opportunités de développement des SMA tels que le commerce électronique, la médecine, l'industrie...etc.

Les systèmes multi-agents sont l'un des paradigmes technologiques les plus prometteurs dans le développement de systèmes logiciels distribués, ouverts et intelligents. Nous entendons par systèmes ouverts, les systèmes auxquels peuvent se rattacher d'autres systèmes. La technologie agent commence par être utilisée pour concevoir des solutions facilitant la mise en place de nouveaux concepts surtout dans le domaine de médecine à savoir la surveillance intelligente (Le monitoring médical).

#### II.1.1 Positionnement historique

Les systèmes multi-agents se positionnent au carrefour de la programmation (ce sont des logiciels), de l'intelligence artificielle (leur autonomie de décision), et des systèmes répartis (leur décentralisation).

Historiquement, on peut replacer le concept d'agent et de système multi-agents dans l'histoire de l'intelligence artificielle et de manière duale dans l'histoire de la programmation. La notion d'agent est de fait à la base des débuts de l'intelligence artificielle (**IA**). Mais cette discipline s'est focalisée sur la modélisation des capacités intelligentes d'une unique entité pour résoudre des problèmes. Il en a résulté une première génération de programmes informatiques évolués, tels les systèmes experts.

Mais, même restreint à un domaine spécialisé (domaine expert), Un tel système expert était censé résoudre tout seul les problèmes de manière autarcique.

L'accent a donc été mis progressivement à partir de la fin des années 70 sur une résolution distribuée de problèmes complexes, par coordination d'un certain nombre d'agents, ce que l'on a alors commencé à appeler « systèmes multi-agents ». On utilise également le terme quasiment équivalent « **IA distribuée** » (avec son acronyme **IAD**, en anglais **DAI**<sup>5</sup>) pour bien montrer l'opposition à l'**IA** classique autarcique et **IA** centralisée.

Il est important pour la suite de pouvoir définir certains termes qui reviendront très souvent dans ce mémoire à savoir Agent, Système multi-agents, etc. Il est important de notifier que les définitions que nous présentons ici s'inscrivent dans le cadre de notre travail (Approche multi-

---

<sup>5</sup> **DAI**: Distributed Artificial Intelligence.

agents dans le monitoring médical) et peuvent donc avoir d'autres connotations dans d'autres domaines.

## II.2. Qu'est-ce qu'un agent ?

Il y a un nombre important d'ouvrages offrant des définitions des agents et des systèmes multi-agents. Les difficultés sont semblables en quelque sorte à celles rencontrées par les scientifiques qui ont essayé de définir la notion d'intelligence artificielle. Pourquoi a-t-il été si difficile de définir l'intelligence artificielle et pourquoi est-il si difficile de définir les systèmes d'agents, quand d'autres concepts de l'informatique, tels que l'objet et l'orienté objet, le calcul distribué, etc., n'ont pas rencontré une si grande résistance à être définis.

Une réponse possible est que la notion d'agent, ainsi que la notion d'intelligence artificielle, ont émergé des humains et de la société humaine. Il est évidemment difficile de modéliser ou de simuler le comportement spécifique humain au moyen de programmes informatiques.

Il y'a plus de 30 ans, les scientifiques ont essayé d'écrire des programmes pour mimer le comportement humain intelligent, dont le but était de créer un système artificiel ayant les mêmes capacités qu'une personne intelligente.

Il n'y a pas une définition acceptée en unanimité pour la notion d'agent. Dans ce qui suit, on présente les définitions les plus importantes.

- *Un agent est une entité qui perçoit son environnement et agit sur celui-ci* [RUS, 97].
- *Un agent est un système informatique, situé dans un environnement, et qui agit d'une façon autonome pour atteindre les objectifs pour lesquels il a été conçu* [WOO, 95a].
- *Un agent est une entité qui fonctionne continuellement et de manière autonome dans un environnement où d'autres processus se déroulent et d'autres agents existent* [SHO, 93];
- *Un agent est une entité autonome, réelle ou abstraite, qui est capable d'agir sur elle-même et sur son environnement, qui dans un univers multi-agents, peut communiquer avec d'autres agents, et dont le comportement est la conséquence de ses observations, de ses connaissances et des interactions avec les autres agents* [FER, 95].

Nous pouvons identifier deux tendances principales en comparant les définitions données ci-dessus pour les agents et les agences. Quelques chercheurs considèrent que nous pouvons définir un agent en isolation, alors que d'autres considèrent les agents principalement comme entités agissant dans une société d'autres agents, à savoir le paradigme des systèmes multi-agents.

Les deux tendances ont déjà donné des résultats. C'est vrai qu'on pense que c'est le paradigme des SMA<sup>6</sup> qui va s'imposer comme prépondérant car il est plutôt difficile de compter qu'un agent existe seulement comme une entité pour lui seul et ne rencontrera pas d'autres agents (soit artificiels ou humains) dans son environnement.

Les *agents personnels*, ou les *agents d'information*, qui ne sont pas censés principalement travailler collectivement pour résoudre des problèmes, auront certainement beaucoup à gagner s'ils agissent par interaction avec d'autres agents et bientôt, avec la diffusion large de la technologie multi-agents, ils ne pourront pas réaliser leurs tâches en isolation. Par conséquent, les chercheurs considèrent la *dimension sociale* d'un agent comme une de ses caractéristiques essentielles.

D'autres considèrent la *mobilité* en tant qu'une des caractéristiques des agents informatiques. On a des réserves sur cette opinion parce que la mobilité est un aspect relié principalement à la mise en place ou à la réalisation des agents, soient-ils logiciels (agents mobiles sur la toile) ou matériels (robots qui se déplacent dans le monde physique) et peut être comprise dans la capacité des agents d'interagir avec l'environnement.

En partant de l'ouvrage de [WOO, 95a], et des définitions citées, on peut identifier les caractéristiques suivantes pour la notion d'agent:

- *Situé* – l'agent est capable d'agir sur son environnement à partir des entrées sensorielles qu'il reçoit de ce même environnement;
- *Autonome* – l'agent est capable d'agir sans l'intervention d'un tiers (humain ou agent) et contrôle ses propres actions ainsi que son état interne;
- *Proactif* – l'agent doit exhiber un comportement proactif et opportuniste, tout en étant capable de prendre l'initiative au bon moment;
- *Capable de répondre à temps* – l'agent doit être capable de percevoir son environnement et d'élaborer une réponse dans le temps requise;
- *Social* – l'agent doit être capable d'interagir avec d'autres agents (logiciels ou humains) afin d'accomplir des tâches ou aider ces agents à accomplir les leurs.

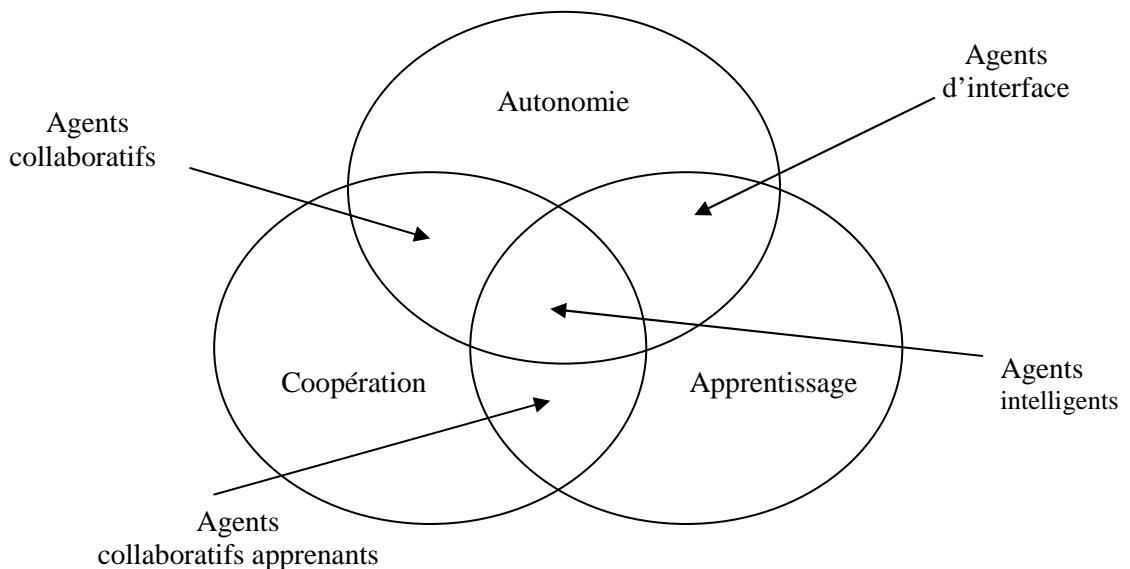
Bien que presque toutes les caractéristiques ci-dessus des agents puissent être considérées en tant que partage de quelque chose avec le comportement intelligent, les chercheurs ont essayé de donner une différence nette entre les agents logiciels (software) et les agents logiciels intelligents, glissant dans le monde des agents la différence si recherchée entre les programmes et les programmes intelligents. D'un point de vue, il est clair que, si dans la conception d'un système agent ou multi-agents, nous utilisons des méthodes et des techniques spécifiques à l'intelligence

---

<sup>6</sup> SMA: Système Multi-agents.

artificielle, alors ces agents peuvent être considérés comme étant intelligents, évidemment dans le sens de l'intelligence artificielle.

L'agent peut également posséder d'autres propriétés (de second ordre) notamment la mobilité, la faculté d'apprentissage, etc. Il est important de mentionner que celles-ci ne sont pas essentielles pour l'agent. En combinant diverses qualités propres aux agents, on peut définir différents types d'agents comme l'indique la **FIG 2.1**



**FIG. 2.1** : Différents types d'agents.

Par exemple, dans la **FIG 2.1**, les propriétés telles que la coopération, l'apprentissage et l'autonomie permettent de définir quatre types particuliers d'agents :

- ◆ les agents intelligents,
- ◆ les agents collaboratifs,
- ◆ les agents collaboratifs apprenants,
- ◆ les agents d'interface.

### II.2.1 L'architecture d'agent

La représentation d'agent en tant que boîte noire (*black box*) est considérée comme une architecture minimale, commune à tous les domaines de recherche d'IA [MUL, 96] [BRE, 98].



**FIG 2.2** : Architecture minimale d'un agent

L'architecture minimale (**FIG 2.2**) est composée de trois modules : le module de perception, le module d'action et le module de traitement intelligent. Ce dernier détermine le comportement d'agent qui peut ainsi analyser les données reçues et construire la réponse sous la forme d'action ou de message.

En général, les architectures existantes sont divisées en trois familles principales : les agents cognitifs (délibératifs), les agents réactifs et les agents hybrides.

### II.2.3 Une comparaison avec les objets

Bien des programmeurs croient encore que les agents ne sont en réalité que des simples objets, au sens attribués à ce terme dans le domaine de la programmation. En fait, un objet est une certaine modélisation d'une entité réelle ou abstraite. Cet objet encapsule un état et il est possible de modifier cet état en invoquant différentes méthodes faisant aussi partie de l'objet. Les agents sont aussi des entités informatiques qui possèdent un état interne (privé), qui sont capables d'agir et de communiquer par échanges de messages. Ce qui différencie les objets des agents, c'est que ces derniers reposent sur un concept très fort d'autonomie leur permettant d'agir de manière très flexible. La négociation, processus par lequel différents groupes arrivent à un accord, est un concept étranger au monde objet. La démarche en vue d'une négociation requiert un comportement flexible afin de pouvoir faire certaines concessions de part et d'autre.

De plus, les agents contrôlent d'une certaine façon leur état interne. Par exemple, si un objet **O1** invoque une méthode **M1** sur l'objet **O2**, cette méthode sera exécutée si aucune erreur ne survient. Cependant, si un agent **A1** effectue une requête **R1** à l'agent **A2**, c'est ce dernier qui décidera si la requête sera effectuée ou non. Toute la différence se situe dans les termes invocation et requête.

Le tableau ci-dessous (**Tab 01**) présente un récapitulatif des éléments de base de l'approche orientée objet versus l'approche orientée agent présenté par **Shoham**<sup>7</sup> [SHO, 93].

Propriétés	AOO <sup>8</sup>	AOA <sup>9</sup>
Élément de base	Objet	Agent
Paramètres définissant l'élément de base	Sans contraintes	Croyances, engagements, capacités, choix
Type d'exécution	Envoie et réception de messages	Envoie et réception de messages
Type de messages	Sans contraintes	Actes du langage
Contraintes sur les méthodes	Sans contraintes	Honnête, consistant

**Tab 2.1** : Approche orientée objet (AOO) versus approche orientée agent (AOA) [SHO, 93].

<sup>7</sup> Yoav Shoham : Professeur à Université de Stamford.

<sup>8</sup> AOO: Approche Orientée Objet.

<sup>9</sup> AOA: Approche Orientée Agent.



## II.2.4 Classes d'agent

De manière générale, il est possible de classer un agent selon les trois principaux points suivants :

### a)- Selon leur nature :

– *Agents réactifs* : Ce sont des agents qui n'ont pas une représentation explicite de leur environnement, ni de mémoire de leur passé, ni de but explicite et leur comportement est de type stimulus réponse.

– *Agents délibératifs* : Ce sont des agents qui possèdent une représentation explicite de leur environnement. Ils se basent sur la connaissance qu'ils ont de leur environnement et leur habileté à raisonner sur leurs connaissances. Les agents dirigés par leurs buts ou par une fonction d'utilité font partie de ce groupe.

– *Agents hybrides* : De façon habituelle, un agent n'est pas complètement de type réactif ni de type délibératif. C'est plutôt une combinaison de ces deux approches.

Dans certaines circonstances, un agent doit agir très rapidement alors que dans d'autres, il aura la possibilité de prendre plus de temps pour avoir un processus délibératif de meilleure qualité. Dans ce cas, les agents sont dits hybrides.

### b)- Selon l'utilisation :

– *Agents collaboratifs* : Ces agents ont des habiletés de coopération. Un regroupement de ces agents permet, entre autres, de réduire un problème complexe en sous-problèmes moins complexes, c'est le cas dans le domaine médical.

– *Agents d'interface* : Ces agents collaborent avec l'utilisateur pour effectuer certaines tâches.

– *Agents pour la recherche d'informations* : Ces agents effectuent, en premier lieu, une recherche d'informations parmi une collection de données et, en second lieu, procèdent à une analyse des informations utiles trouvées afin de découvrir de nouvelles connaissances.

– *Agents pour le commerce électronique* : La montée de l'Internet a bien entendu créé de nouvelles nécessités. Les agents issus de cette tendance permettent la promotion, la vente ainsi que l'achat de produits et de services par l'entremise des réseaux informatiques, etc.

### c)- Selon la technologie employée :

– *Agents stationnaires* : Il s'agit du cas où l'agent s'exécute toujours sur la même machine.

– *Agents mobiles* : Ces agents s'exécutent sur différentes machines en se promenant d'un hôte à l'autre. Typiquement, ils suivent ce que l'on appelle un "itinéraire".

Le plus grand avantage réside dans le fait qu'ils réduisent la charge des réseaux en se rendant directement sur la machine "utile" pour y effectuer les actions de manière locale.

D'un point de vue général, les chercheurs aimeraient que les agents soient des composants logiciels réutilisables qui permettraient d'accéder et/ou contrôler des services et des ressources informatiques. Ces agents seraient les unités de base des applications informatiques et seraient organisés en réseau dans une optique de collaboration.

L'ensemble des entités (Agents) en interaction et en vue de réaliser un but commun forme ce qu'on appelle un système multi-agents, Alors la question qui se pose est :

## II.3. QU'EST-CE QU'UN SYSTEME MULTI-AGENTS (SMA)?

### II.3.1 Les systèmes multi-agents

Les systèmes multi-agents font partie de l'Intelligence Artificielle Distribuée (**IAD**), qui, à la différence d'**IA** classique, s'intéresse aux comportements intelligents, résultant de l'activité coopérative de plusieurs agents [LAB, 93].

Bien qu'il soit difficile de donner une définition générale pour l'expression « **Système Multi-Agents "SMA"** », les différentes caractérisations proposées par [BON, 94][FER, 95] [OHA, 96], [FRA, 96], [WOO, 98b] et autres permettent de dégager des idées fondamentales communes. La définition, mentionnée ci-dessous, est conforme à la plupart de ces idées.

### II.3.2 Définition: Système multi-agents (SMA)

Le système multi-agents est considéré comme un système distribué composé d'un certain nombre d'entités autonomes<sup>10</sup> (les agents), qui travaillent selon les modes complexes d'interaction, pour réaliser leurs propres buts et par-là même atteindre l'objectif global désiré. Les agents peuvent interagir en communiquant directement entre eux ou par l'intermédiaire d'un autre agent ou en agissant sur leur environnement.

En général, dans les systèmes multi-agents, deux types d'agents sont identifiés : les agents **IAD**<sup>11</sup> et les agents **mobiles**. Les agents **IAD** sont capables de communiquer et de coopérer à l'aide de protocoles basés sur les réseaux contractuels et les actes de langage. Les agents **mobiles** peuvent se déplacer de site en site dans le réseau pour accomplir des tâches spécifiques.

<sup>10</sup> **L'autonomie** signifie que l'entité (l'agent) est capable de travailler sans l'intervention d'un humain ou des autres agents, et de contrôler ses actions ainsi que son état interne

<sup>11</sup> **IAD** : Intelligence Artificiel Distribué.

En général, ils peuvent avoir deux types de mobilité :

- ✓ L'exécution distante:  
L'agent (le programme et les données) est transféré sur le site distant, où il est exécuté.
- ✓ La migration:  
Pendant son exécution, l'agent se déplace vers un autre site afin d'accomplir progressivement sa tâche. Autrement dit, l'agent est capable de suspendre son exécution sur un site, de se déplacer vers un autre en transportant le code, les données et l'état d'exécution, et de reprendre l'exécution depuis le point de suspension.

D'après la définition du système multi-agents, proposée ci-dessus, pour qu'un système soit considéré comme SMA, il est nécessaire de satisfaire certains critères :

- Disposer d'un ensemble des agents autonomes, fonctionnant en parallèle et cherchant à satisfaire un but ;
- Les agents doivent posséder un mécanisme d'interaction de haut niveau, indépendant de problème à résoudre. Ils peuvent utiliser les langages de communication d'agents (*ACL*), par exemple, *KQML*<sup>12</sup>, *FIPA ACL*, etc. ;
- L'agent doit percevoir une partie de son environnement qui peut être le monde physique, d'autres agents, le réseau Internet, etc. Il doit répondre dans un délai acceptable.

On peut déduire, par conséquent, que la distribution modulaire d'une application ne constitue pas forcément un système multi-agents. En effet, les modules n'ont pas de buts à atteindre ou la fonction de satisfaction. Leurs mécanismes d'interactions sont de bas niveau : appels de procédures, etc. En ce qui concerne les agents, ils reçoivent des messages qui peuvent être des demandes d'exécution, mais, aussi, des requêtes d'information sur leurs capacités. Ils tentent de satisfaire les objectifs, ils disposent pour cela d'une autonomie supplémentaire par rapport aux objectifs. L'agent peut refuser une tâche donnée à cause de son manque de compétences ou de sa trop grande occupation.

### II.3.3 Les interactions et les agents

L'interaction est une notion importante dans les systèmes multi-agents. Au dire de **J.Ferber** [FER, 95], l'interaction permet d'avoir une relation dynamique entre deux ou plusieurs agents par le biais d'actions réciproques. Les situations d'interactions sont diverses (l'aide d'un robot à un autre, l'échange de données entre deux serveurs, etc.).

Pour un agent, interagir avec un autre agent constitue la source de sa puissance et l'origine de ses problèmes [FER, 96]. En effet, seulement en coopérant les uns avec les autres, les agents peuvent accomplir leurs tâches collectives. D'autre part, à cause de leurs interactions avec les autres agents, ils doivent coordonner leurs actions et résoudre des conflits.

---

<sup>12</sup> **KQML**: Knowledge and Query Manipulation Language.

Traiter le problème d'interaction revient non seulement à décrire les mécanismes permettant aux agents d'interagir, mais aussi à analyser et à concevoir les différentes formes d'interaction, utilisées par les agents pour accomplir leurs tâches et satisfaire leurs buts. Nous présentons le concept d'une de ces formes - la communication - sur laquelle repose toute l'interaction.

### II.3.3.1 La communication

La communication est essentielle dans la résolution coopérative des problèmes [LAB, 93] [FER, 95]. Elle permet de synchroniser les actions des agents et de résoudre les conflits des ressources par la négociation. D'après [BOU, 92], la communication définit l'ensemble des processus physiques et psychologiques par lesquels s'effectue l'opération de mise en relation d'un agent émetteur avec un ou plusieurs agents récepteurs, dans l'intention d'atteindre les objectifs prévus. Les processus physiques décrivent les mécanismes d'exécution des actions, par exemple, l'envoi et la réception de messages. Les processus psychologiques désignent les changements opérés par la communication sur les buts et les croyances des agents.

En général, les actions de communication entre les agents sont considérées comme les actions d'échange d'information.

Il existe plusieurs types de communication, à savoir, la communication par tableau noir, la communication par partage d'informations et la communication par envoi des messages qu'on va détailler dans ce qui suit.

#### II.3.3.1.1 La communication par envoi des messages

Ce type de communication permet aux agents d'envoyer leurs messages directement aux destinataires par les mécanismes spécifiés (les canaux ou les ports). Il existe trois types de messages : les questions, les réponses et les informations. Au niveau protocolaire, un envoi de message peut être synchrone (un agent émetteur attend la réponse de son récepteur) et asynchrone (un agent émetteur peut agir immédiatement après avoir placé son message dans une file d'attente).

La communication par l'envoi de messages peut être organisée suivant trois formes différentes :

- *La communication point à point.*

L'agent émetteur connaît l'agent destinataire et lui transmet directement son message. L'agent destinataire est le seul à recevoir le message envoyé ;

- *La distribution généralisée.*

Il s'agit d'un envoi d'un même message à tous les agents du système multi-agents. L'agent émetteur ne connaît pas forcément les destinataires du message envoyé.

- *La distribution restreinte.*

Il s'agit d'un envoi d'un même message à un certain groupe d'agents du système multi-agents. L'agent émetteur ne connaît pas forcément tous les destinataires, mais il doit être capable de les atteindre en s'appuyant, soit sur leurs caractéristiques, soit sur la notion de groupe auquel ils appartiennent.

### II.3.4 Les langages de communication

Le concept de communication permet de réaliser les échanges locaux d'information entre des agents. Le problème est alors de combiner ces actions de communication avec les compétences et les connaissances des agents, pour obtenir un comportement collectif entre ces derniers. Dans les systèmes multi-agents, la problématique susmentionnée est étudiée à travers le concept de l'interaction des agents. Cette dernière permet d'avoir une relation dynamique entre deux ou plusieurs agents par le biais d'actions réciproques [FER, 95]. La question qui se pose alors, est comment réaliser les interactions des agents ?

D'après [FIN, 97], pour que les agents puissent interagir de manière efficace, ils doivent posséder un langage de communication commun, leur permettant de se comprendre ainsi que de s'échanger des informations et des connaissances.

Avant d'aborder les langages de communication existants, nous présentons brièvement la théorie des actes de langage, considérée en intelligence artificielle comme un modèle général de communication entre les agents.

#### II.3.4.1 La théorie des actes de langage

La théorie des actes de langage (« *Speech Act Theory* »), constitue un fondement théorique de la communication, basée sur l'idée suivante : « *Lorsqu'on parle, on effectue des actions* »<sup>13</sup> [LAB, 93]. Un acte de langage définit un message, qui contient l'affirmation positive ou négative, et provoque les changements de l'environnement (voir, l'annexe C).

Chaque acte de langage comprend trois composants :

- Le composant *locutoire*, qui décrit l'expression d'un message ;
- Le composant *illocutoire*, qui définit les intentions de l'émetteur, associées implicitement au message ;
- Le composant *perlocutoire*, qui décrit les effets d'un acte de langage sur l'environnement.

<sup>13</sup> Cette théorie est introduite en 1962, par le philosophe et le linguiste anglais **J.-L.Austin** dans son livre « *How to do things with words* ».

Dans la théorie des actes de langage, les intentions des émetteurs sont identifiées en utilisant les verbes performatifs. Ces derniers sont classifiés en plusieurs catégories : **les affirmatifs** (informer), **les directifs** (ordonner), **les promissifs** (promettre), **les déclaratifs** (déclarer) et **les expressifs** (exprimer).

D'après cette théorie, chaque acte de langage multi-agents peut être décrit sous la forme d'un message, dont le type est défini par le verbe performatif, tel que « *Request* » ou « *Inform* ». Le contenu de ce message est décrit en utilisant les langages de communication (**ex:** *KQML*, *FIPA ACL*).

### II.3.4.1.1 KQML

*KQML* (« *Knowledge and Query Manipulation Language* ») [LAB, 00] [MAY, 96] est une approche basée sur les actes de langages, qui permet de réaliser les interactions des agents tenant en compte la diversité des langages de communication.

La communication est considérée comme un ensemble d'échanges de messages *KQML*. Chaque message comprend trois couches [LAB, 99] :

- La couche de *contenu*, qui spécifie le contenu réel du message d'agent ;
- La couche de *communication*, qui décrit tous les paramètres de communication de bas niveau, par exemple, l'identificateur de l'agent émetteur et celui de l'agent récepteur, l'identificateur de la communication, etc.
- La couche de *message*, qui est considéré comme le noyau de *KQML*. Sa fonction principale s'agit d'identifier le protocole de réseau, utilisé pour envoyer le message, et de déterminer le performatif, indiquant le type de ce message (**ex:** une affirmation, une requête, une commande, etc.).

Le message *KQML* est représenté sous la forme d'une liste, contenant le performatif, qui correspond au type particulier d'acte de langage (**ex:** *tell* (transférer l'information aux autres agents), *ask-one* (demander la réponse à l'agent correspondant), etc.), et les arguments associés à ce performatif (FIG 2.3).

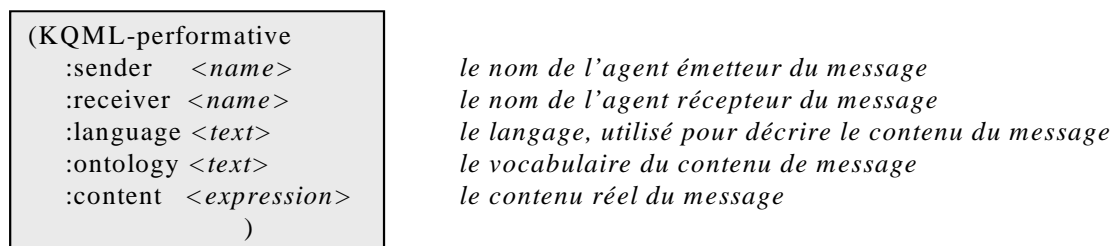


FIG 2.3. La structure d'un message **KQML**.

**Remarque:** Il est à noter que l'ordre des arguments dans une liste n'est pas important.

Le langage de *KQML* permet d'utiliser les différents protocoles de communication, notamment, *TCP/IP*, *SMTP (e-mail)*, *HTTP*, *CORBA*, etc. Parmi les travaux réalisés concernant le développement de nouveaux

protocoles, basés sur KQML, on peut mentionner les langages de coordination *COOL* [BRA, 95] et *AgentTalk* [KUW, 95].

Bien qu'il présente un grand intérêt pour les utilisateurs, *KQML* montre quelques lacunes [COH, 95] :

- La signification floue de certains performatifs (**ex**: le cas d'un performatif « dénier ») ;
- Le manque des performatifs promissifs, exprimant l'engagement, auprès d'un tiers, d'accomplir une action ;
- L'utilisation de ce langage que pour des communications isolées.

### II.3.4.1.2 FIPA ACL

Pour résoudre certains problèmes inhérents à *KQML*, *FIPA*<sup>14</sup> a proposé un nouveau langage de communication d'agents – *FIPA ACL*<sup>15</sup>.

Comme *KQML*, ce langage est fondé sur la théorie des actes de langages. Sa spécification consiste en la définition d'un ensemble des types de messages et en description de leurs pragmatiques, c'est-à-dire, les effets sur les attitudes mentaux des agents émetteurs et des agents récepteurs. La sémantique formelle de *FIPA ACL* se compose de cinq niveaux [OBR, 98] :

- Le protocole, qui décrit les règles sociales des dialogues entre les agents ;
- Les actes communicatifs (*AC*), qui définissent le type de communication entre les agents (par exemple, la demande, la confirmation, etc.) ;
- Le méta-information concernant le message (l'identification d'un agent émetteur et d'un agent récepteur, le contexte, etc.) ;
- Le langage du contenu, qui décrit la grammaire et la sémantique associée, utilisées pour exprimer le contenu d'un message ;
- L'ontologie, qui définit le vocabulaire et les significations des termes et des concepts, employés dans le contenu.

La syntaxe de *FIPA ACL* est identique à celle de *KQML*, sauf les noms de certains primitifs réservés. Le message *FIPA ACL* est représenté sous la forme d'une liste, contenant le type de l'acte communicatif (par exemple, *INFORM*, *REQUEST*), le nom de l'agent émetteur et celui de l'agent récepteur, le contenu et le contexte du message, l'ontologie à utiliser pour interpréter ce contenu, et le protocole [FIP, 00] :

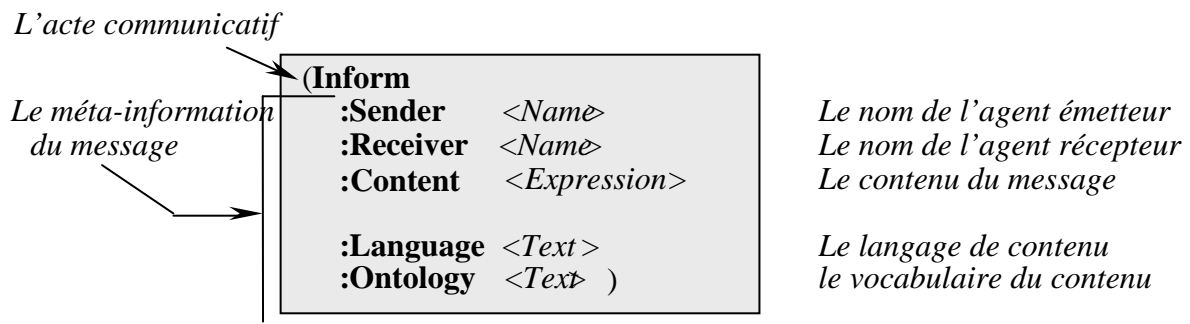


FIG 2.4. La structure d'un message FIPA ACL

<sup>14</sup> FIPA: Foundation for Intelligent Physical Agents.

<sup>15</sup> FIPA ACL: FIPA Agent Communication Language.

**FIPA ACL** peut être considéré comme l'extension de *KQML* possédant deux langages différents. Le langage externe définit la signification intentionnelle du message. Le langage interne (ou le contenu) décrit l'expression à laquelle s'appliquent les croyances, les désirs et les intentions des agents, décrites dans le primitif de communication.

Différemment de *KQML*, *FIPA ACL* est basé sur la sémantique logique de la communication. Ceci facilite la description des formats de la communication. Cependant, les agents ne possèdent pas toujours les capacités logiques nécessaires. Il est à noter que la sémantique de *FIPA ACL* est basée en grande partie sur les croyances des agents, qui peuvent être inconnus pour les autres agents.

Une autre grande différence entre *FIPA ACL* et *KQML*, concerne les actes communicatifs. *FIPA ACL* contient un ensemble d'actes communicatifs normatifs, qui peuvent être primitifs ou composés. Les nouveaux actes communicatifs ne peuvent être définis qu'en combinant les actes existants et en utilisant les opérateurs prédéfinis. Ceci permet de maintenir l'intégrité sémantique du langage.

### II.3.5 Sémantique des messages

Pour que les agents puissent comprendre les messages inter-changés, il faut qu'ils partagent un vocabulaire commun. Il s'agit de garantir que les concepts et les entités véhiculées au travers des applications ont la même signification, même si différentes applications utilisent des noms différents les référant.

Ce vocabulaire ce n'est autre qu'une ontologie qui représente une sorte d'un vocabulaire d'un domaine bien précis.

L'intérêt de l'ontologie est essentiellement de décrire le vocabulaire d'un domaine (cardiologie par exemple), elle permet la réutilisation des connaissances, en outre, elle ajoute (enrichit) une sémantique aux messages inter-changés entre les agents.

### II.3.6. La coopération

La coopération est l'une des caractéristiques fondamentales des systèmes multi-agents.

D'après [FER, 97], le problème de la coopération peut se ramener à résoudre les différents sous problèmes qui comprennent la collaboration des agents par répartition des tâches, la coordination d'actions et la résolution de conflits. **J.Ferber** présente deux points de vue sur la coopération [FER 95] [BRA, 95] :

- ✓ *Une attitude intentionnelle* d'agents qui décident de travailler ensemble. Dans ce cas, les agents coopèrent s'ils effectuent une action commune, après avoir identifié et adopté un but commun. Ce concept



présente deux inconvénients principaux : d'une part, on considère que la coopération existe, même si les agents obtiennent meilleurs résultats en travaillant individuellement ; d'autre part, d'après l'attitude intentionnelle, la possibilité pour les agents réactifs de pouvoir coopérer est supprimée ;

- ✓ *Une qualification d'une activité du groupe d'agents, observée par l'observateur*, qui interprète les comportements à partir de critères physiques et sociaux. Dans ce cas, plusieurs agents coopèrent si (1) l'ajout d'un nouvel agent permet d'accroître les performances d'un groupe, et (2) l'action d'agent sert à résoudre ou à éviter les conflits. Ces indices de coopération représentent la collaboration entre les agents ainsi que la résolution des conflits.

S'appuyant sur ces points de vue, on peut distinguer deux types de coopération : la coopération intentionnelle (ou celle d'agents cognitifs), où les agents ont l'intention de coopérer, et la coopération réactive.

**J.Ferber**<sup>16</sup> [FER, 95] précise plusieurs méthodes de coopération : le regroupement et la multiplication, la communication, la spécialisation, la collaboration par partage de tâches et de ressources, la coordination d'action ainsi que la résolution de conflit par arbitrage et la négociation. Toutes ces méthodes nécessitent d'être structurées au sein de l'organisation. Cette dernière permet de décrire la manière dont les agents sont positionnés dans un groupe ainsi que les techniques de travail coopératif efficace.

Dans la section ci-dessus on va présenter la plateforme Jade, par laquelle, on va créer et gérer les agents de notre application.

## II.4. Outils & plateformes de développement des SMA

Le meilleur moyen pour construire un système multi-agents (SMA) est d'utiliser une plate-forme multi-agents. Cette dernière est un ensemble d'outils nécessaire à la construction et à la mise en service d'agents au sein d'un environnement spécifique. Ces outils peuvent servir également à l'analyse et au test du SMA ainsi créé. Ces outils peuvent être sous la forme d'environnement de programmation (API) et d'applications permettant d'aider le développeur.

Actuellement, il existe plusieurs plates-formes développements des agents et des SMA parmi ceux, on cite : **JADE**, **MADKIT**, **AGENT**

On va étudier ci-après la plate-forme **JADE**<sup>17</sup>.

**JADE** est une plate-forme multi-agents développée en Java par **CSELT** (Groupe de recherche de Gruppo Telecom, Italie) qui a comme but la construction des systèmes multi-agents et la réalisation d'applications conformes à la norme **FIPA**<sup>18</sup> [FIP, 00]. **JADE** comprend deux composantes

<sup>16</sup> **Jacques Ferber** est un professeur de l'informatique à l'université de la Science de Montpellier. Il a mené le groupe de recherche de Myriade à Paris à partir de 1988 jusqu'en 1997. Il mène maintenant un nouveau groupe de recherche qui est également consacré aux systèmes de multi agent.

<sup>17</sup> **JADE**: Java Agent **DE**velopment Framework.

<sup>18</sup> **FIPA**: Foundation for Intelligent Physical Agents.

de base : une plate-forme agents compatible FIPA et un paquet logiciel pour le développement des agents Java.

#### II.4.1 Pourquoi la plateforme JADE ?

On a développé notre application sous la plate forme multi agent JADE qui offre les avantages suivants :

- Plate forme assez facile à mettre en place.
- Disponibilité de packages sur lesquels nous nous sommes appuyés pour développer notre application.
- Documentation claire et complète.
- Open source.

#### II.4.2 La norme FIPA pour les systèmes multi-agents

Les premiers documents de spécification de la norme FIPA, appelés spécifications [FIP, 97], établissent les règles normatives qui permettent à une société d'agents d'interopérer. Tout d'abord, les documents FIPA décrivent le modèle de référence d'une plate-forme multi-agents (FIG 2.5) où ils identifient les rôles de quelques agents clés nécessaires pour la gestion de la plate-forme, et spécifient le contenu du langage de gestion des agents et l'ontologie du langage.

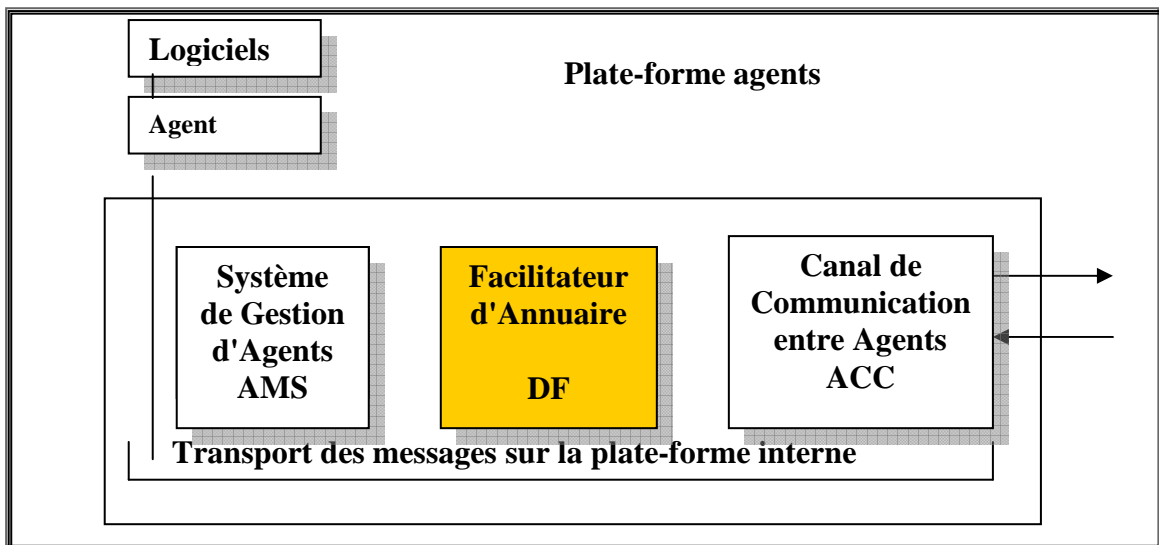


FIG 2.5 : Le modèle de référence pour une plate-forme multi-agents FIPA

Dans la FIG 2.5, on voit qu'il existe trois rôles principaux dans une plate-forme multi-agents FIPA :

- Le **Système de Gestion d'Agents** (Agent Management System - **AMS**) est l'agent qui exerce le contrôle de supervision sur l'accès à l'usage de la plate-forme; il est responsable de l'authentification des agents résidents et du contrôle d'enregistrements.
- Le **Canal de Communication entre Agents** (**ACC**<sup>19</sup>) est l'agent qui fournit la route pour les interactions de base entre les agents dans et

<sup>19</sup> ACC: Agent Communication Channel.

hors de la plate-forme; c'est la méthode de communication implicite qui offre un service fiable et précis pour le routage des messages; il doit aussi être compatible avec le protocole **IOP**<sup>20</sup> pour l'interopérabilité entre les différentes plates-formes multi-agents.

- Le **Facilitateur d'Annuaire** (Directory Facilitator - **DF**) est l'agent qui fournit un service de pages jaunes à la plate-forme multi-agents.

Le standard spécifie aussi le **Langage de Communication d'Agents** (Agent Communication Language - **ACL**), la communication des agents est basée sur l'envoi de messages. Le langage **FIPA ACL** est le langage standard des messages et impose le codage, la sémantique et la pragmatique des messages. La norme n'impose pas de mécanisme spécifique pour le transport interne de messages. Plutôt, puisque les agents différents pourraient s'exécuter sur des plates-formes différentes et utiliser technologies différentes d'interconnexion, *FIPA* spécifie que les messages transportés entre les plates-formes devraient être codés sous forme textuelle. On suppose que l'agent est en mesure de transmettre cette forme textuelle.

## II.5. DOMAINES D'APPLICATION

De nos jours, la technologie multi-agents a trouvé sa place dans les systèmes manufacturiers, les systèmes financiers, les loisirs, les télécommunications, le contrôle commande, la santé, et pas mal d'autres applications.

Dans ce qui va suivre nous n'en exposerons que quelques exemples d'applications utilisant cette technologie et nous référons le lecteur à l'article de *Jennings* [Jen, 98] pour un ensemble d'exemples plus complet.

### II.5.1 Application des SMA aux télécommunications

Ces dernières années, les télécommunications ont notamment introduit une conception de services décentralisée dans le contexte du Web, créé de nouveaux services de médiation tels que les portails et engendrés l'apparition de nombreux fournisseurs de services réseaux qui ne disposent pas de leurs propres services réseaux [Bou, 00].

L'obtention de tels services décentralisés ne peut, bien entendu, être obtenue que grâce à des logiciels pour lesquels les données et le contrôle sont forcément distribués. De ce fait, il est clair que les SMA semblent convenir aux télécommunications. C'est pourquoi les principaux acteurs de télécommunications mènent actuellement d'intenses activités de recherche sur la technologie agent, British Telecom, France Télécom, Deutch Telekom, NTT, Nortel, Siemens, etc.

### II .5.2 Le système ADEPT

Les gestionnaires de grandes compagnies effectuent des prises de

<sup>20</sup> **IOP** : Protocole de communication permettant d'intégrer et de mettre en réseau des applications de provenance diverse.

décisions en se basant sur une combinaison de jugement et d'informations provenant de plusieurs départements. Idéalement, toutes les informations pertinentes devraient être rassemblées avant qu'une décision ne soit prise. Cependant, le processus d'obtention des informations, qui sont à jour et pertinentes, est très complexe et prend énormément plus de temps. Pour cette raison, plusieurs compagnies ont cherché à développer des systèmes informatiques afin de les assister dans leur processus d'affaires.

Le système ADEPT [Mam, 94] attaque ce problème en voyant le processus d'affaires comme un ensemble d'agents qui négocient et qui offrent des services. Chaque agent représente un rôle distinct ou un département de l'entreprise et est en mesure de fournir un ou plusieurs services. Les agents qui requièrent les services d'autres agents le font par une négociation qui permet d'obtenir un coût, un délai temporel et un degré de qualité qui sont acceptables aux deux parties.

Le résultat d'une négociation terminée avec succès constitue un engagement entre les deux parties.

### II.5.3 Télémédecine et la santé

La télémédecine à base d'agents intelligents s'intéresse au problème du diagnostic médical en continu et à distance, ce type de diagnostic nous conduira à avoir une représentation permanente de l'état du patient et à la remettre à jour chaque fois que de nouvelles observations sont disponibles. Il s'agit de *monitoring*.

Le domaine de la cardiologie est considéré comme un célèbre domaine d'application par les agents intelligents, exactement pour la détection des arythmies cardiaques par le biais de l'électrocardiogramme (ECG). En ce stade, des applications ont été développées par des laboratoires de recherche tels que l'ERISA, INSA en France, d'autres sont conçues et intégrées dans des appareils portatifs, en vue de les commercialiser.

Généralement, ces agents sont dotés par un moteur de raisonnement basé sur une ou un mariage entre les méthodes de l'intelligence artificiel (IA) telles que les réseaux des neurones, les réseaux bayésiens, les chaînes de Markov, les systèmes experts et la logique floue.

Dans notre cas, on s'intéresse au monitoring médical à base d'agents intelligents, dont le comportement de classification (Classification behaviour) s'effectue par un réseau neurones multicouches (Perceptron multicouches), par contre le comportement d'adaptation se fait par un système expert fonctionnant en chainage avant.

Passons maintenant à la présentation des systèmes d'aide au diagnostic médical à base d'agents intelligents ou des systèmes réalisés pour le domaine médical aussi en se basant sur les systèmes multi-agents.

#### II.5.3.1 Le système GUARDIAN

Le système GUARDIAN [Fro, 05] a pour but de gérer les soins aux patients d'une unité de soins intensifs chirurgicale. Les principales

motivations de ce système sont premièrement, le modèle des soins d'un patient dans une unité de soins intensifs est essentiellement celui d'une équipe, où un ensemble d'experts dans des domaines distincts coopèrent pour organiser les soins des patients; deuxièmement, le facteur le plus important pour donner de bons soins au patients est le partage d'informations entre les membres de l'équipe de soins critiques. Particulièrement, les médecins spécialistes n'ont pas l'opportunité de superviser l'état d'un patient minute par minute; cette tâche revient aux infirmières qui, quant à elles, ne possèdent pas les connaissances nécessaires à l'interprétation des données qu'elles rassemblent.

Le système GUARDIAN répartit donc le suivi des patients à un certain nombre d'agents de trois types différents. Les agents *perception/action* sont responsables de l'interface entre GUARDIAN et le monde environnant, établissant la relation entre les données des capteurs et une représentation symbolique que le système pourra utiliser, et traduisant les requêtes d'action du système en commandes pour les effecteurs. Les agents en charge du raisonnement sont responsables d'organiser le processus de prise de décision du système. Finalement, les agents en charge du contrôle (il n'y en a habituellement qu'un seul) assurent le contrôle de haut niveau du système.

#### II. 5.4 Les systèmes d'informations coopératifs (SIC)

Les SIC sont généralement caractérisés par la grande variété et le grand nombre de sources d'informations. Ces sources d'informations sont hétérogènes et distribuées soit sur un réseau local (Intranet) soit sur l'Internet. De tels systèmes doivent être capables d'exécuter principalement les tâches suivantes :

- la découverte des sources : trouver la bonne source de données pour l'interroger;
- la recherche d'informations : identifier les informations non structurées et semi- structurées;
- le filtrage des informations : analyser les données et éliminer celles qui sont inutiles;
- la fusion de l'information : regrouper les informations d'une manière significative.

Le système multi-agents « Warren » pourrait constituer un exemple spécifique de l'utilisation des agents dans ce type d'application. C'est un système d'agents intelligents pour l'aide des users dans la gestion des portefeuilles [CHA, 02]. Ce système combine les données du marché financier, les rapports financiers, les modèles techniques et les rapports analytiques avec les prix courants des actions des compagnies. Toutes ces informations sont déjà disponibles sur le Web; "Warren" ne fait que les intégrer via des agents spécialisés, les agents d'informations et ensuite les présenter aux usagers. Pour ce faire, "Warren" dispose de six agents ressources, deux agents de tâches et un agent utilisateur pour chaque usager. L'agent utilisateur affiche (via le web) les informations financières de son usager, lui permettant de faire des simulations d'achat

et de vente des actions. Il affiche également les prix courants des actions et les nouvelles informations du marché financier. Le même agent permet également d'accéder aux rapports produits par les deux agents de tâches. Ces deux agents fournissent d'une part, une intégration graphique des prix et des nouvelles concernant les actions et, d'autre part, une analyse fondamentale des actions en tenant compte de leurs historiques. Les agents d'informations accèdent à différentes sources d'informations, comme les pages Web, les nouvelles de "Clarinet et Dow-Jones", les rapports financiers électroniques de "SEC Edgar" ainsi que d'autres rapports sous un format texte. "Warren" n'est qu'un exemple et il existe actuellement plusieurs autres systèmes qui touchent à ce genre d'application. Parmi ces applications, nous pouvons citer :

- ✓ Infosleuth : C'est un système multi-agents pour la recherche coopérative d'informations dans des bases de données distribuées. Ce système a été appliqué aux domaines médicaux [Nod, 98].
- ✓ NetSA (pour "Networked Software Agents") : C'est un système proche de Infosleuth et dédié aux environnements riches en informations [Tro, 98], [Côt, 99].
- ✓ UMDL : C'est un système d'informations coopératif pour la recherche des documents dans une librairie digitale [Vid, 98].

À titre d'exemple, le système NETSA est un système multi-agents coopératif, développé à l'université Laval [Tro, 98], [Côt, 99] est destiné aux environnements riches en informations. Ce système comporte plusieurs types d'agents :

- ❖ Un agent *utilisateur* en charge de la cueillette et du filtrage des informations provenant et allant vers l'utilisateur;
- ❖ Un agent *courtier* servant de répertoire pour les agents qui évoluent au sein de NETSA;
- ❖ Des agents *ressources* reliés chacun à une ressource d'informations et pouvant rapatrier et mettre à jour les données;
- ❖ Un agent d'*exécution* en charge de la décomposition des tâches et du suivi du déroulement d'exécution des différentes sous-tâches;
- ❖ Un agent *ontologie* en charge du maintien de la cohérence des concepts utilisés par les agents.

Les agents et les systèmes multi-agents sont utilisés dans plusieurs domaines d'applications, comme par exemple : la gestion des réseaux, la recherche d'informations, le commerce électronique et la planification des tâches. Pour une grande partie de ces applications, les agents sont utilisés dans les systèmes d'informations coopératifs ou comme assistants personnels dans certains travaux. Ce type d'applications nécessite des recherches approfondies et il convient en particulier :

- ❖ de pousser l'ingénierie de la construction des systèmes à agents collaboratifs. Il faut concevoir plus de méthodes et d'outils pour faciliter leurs implantations [Smt, 94];
- ❖ de maîtriser et de bien identifier la coordination entre les agents.

Il s'agit en particulier d'établir une théorie claire et formelle pour cette coordination [Ros, 94];

- ❖ d'assurer les critères de stabilité, de clarté et de performance pour de tels systèmes;
- ❖ de trouver des techniques qui permettent l'évaluation, la vérification et la validation de ces systèmes.

Il est à signaler que NetSA (Networked Software Agents), l'architecture sur laquelle nous avons travaillé, fait elle aussi partie des systèmes d'informations coopératifs.

## II.6. CONCLUSION

Nous avons vu, tout au long de ce chapitre que la technologie agent et multi-agents n'est pas un concept voué à rester sur les tablettes des laboratoires de recherche puisque plusieurs exemples d'applications existent déjà.

Les personnes qui ont développé des SMA vous diront toutefois qu'il est difficile de concevoir et de bâtir un système multi-agents. En effet, la construction de ce type de système comportent toutes les difficultés inhérentes aux systèmes répartis, auxquelles s'ajoute le caractère flexible et sophistiqué des interactions entre agents. À cela s'ajoute le fait que les concepteurs de SMA font de nos jours face à deux difficultés majeures. Tout d'abord, l'absence de méthodologie systématique qui permettrait de spécifier et de structurer une application multi-agents. Ensuite, le manque d'outils commerciaux pour bâtir des SMA.

Une des forces importantes derrière la croissance rapide que connaissent aujourd'hui les systèmes multi-agents est l'Internet, où la population d'agents est sans cesse croissante. Ces agents devront savoir collaborer afin d'atteindre les objectifs de leurs concepteurs. Dans ce type d'environnement, les agents rencontrent deux défis majeurs : ils doivent être en mesure de se rencontrer et interopérer. Une première solution à ce problème est l'introduction d'agents intermédiaires (*brokers*, *facilitateurs*, etc.) [Tro, 98],[Côt, 99].

Ces agents intermédiaires ont pour principales fonctions : (1) d'associer au mieux les besoins des utilisateurs et les services des fournisseurs; (2) d'unifier et probablement traiter les réponses des fournisseurs pour produire un résultat approprié; (3) aviser régulièrement les utilisateurs du changement des informations.

Dans cette optique d'interopérabilité, un grand nombre de langages de communications pour les agents ont été développés; basés, pour plupart, sur la théorie des actes du discours. Bien que les performatives offertes par ces langages permettent de caractériser les types des messages, ils ne permettent pas encore aux agents de comprendre explicitement les concepts "discutés". Le problème des ontologies reste donc ouvert.

Un autre problème critique est l'allocation de ressources limitées à un bon nombre d'agents. Des mécanismes basés sur les principes économiques ont été proposés pour résoudre ce problème. Dans de telles approches, on suppose que les agents sont égocentrés et ne

cherchent qu'à maximiser leur utilité. Les sous-domaines où on a appliqué ces principes économiques sont l'allocation de ressources, l'allocation de tâches et la négociation. Là aussi la recherche n'a fait qu'entrevoir des techniques rudimentaires basés sur les lois du marché et il reste bien des domaines à couvrir comme par exemple les ventes aux enchères, les comportements acheteur(s) vendeur(s), le partage du marché, etc.

Finalement, il convient de préciser que les chercheurs travaillant sur les SMA d'un point de vue formel se sont presque toujours heurtés à des agents omniscients, c'est à dire des agents ayant des capacités de raisonnement illimitées. Là aussi le problème est très ouvert et il y a présentement pas mal d'équipes qui y travaillent.

Comme on peut encore le constater le domaine des systèmes multi-agents demeure encore aujourd'hui un domaine rempli de défis à surmonter, autrement dit un domaine très ouvert pour la recherche.

Dans le chapitre suivant, nous allons vous présenter les modèles de classification existes tels que la logique floue, les arbres de décision, les réseaux bayésiens, les machines à vecteur de support et les cartes de kohonen pour la classification des arythmies cardiaques.





## III. LES MODELES DE CLASSIFICATION

### III.1. Introduction

Un diagnostic médical est le résultat du raisonnement d'un médecin, décision très souvent prise à partir d'informations incertaines et/ou incomplètes. De nombreuses techniques d'intelligence artificielle (IA) ont été appliquées pour essayer de modéliser ce raisonnement [LAV, 97][LAV, 99]. Ainsi, [SZO, 82] présente l'utilisation détaillée de plusieurs systèmes experts en médecine. Citons, par exemple, des systèmes à base de règles comme MYCIN [SHO 74], [BUC, 84].

En amont de ce raisonnement, il faut aussi être capable de modéliser ces informations incertaines et/ou incomplètes. Les performances des ordinateurs sont de plus en plus remarquables, et avec elles, la nécessité et la possibilité de développer des méthodes de classification automatiques capables et robustes augmentent également. Dans ce contexte, de nombreuses méthodes de classification automatiques sont apparues. Les plus connues sont les méthodes la logique floue [STE, 97], les réseaux de neurones, les réseaux bayésiens, les arbres de décision, les SVM (Support Vector Machine), les cartes de **kohonen**, etc.

En outre, toutes ces méthodes nécessitent un algorithme d'apprentissage qui va permettre de passer d'un espace des exemples  $X$  à un espace dit des hypothèses  $H$  dont l'apprentissage c'est l'acquisition de connaissances et compétences permettant la synthèse d'information [Bes, 07].

Notons que les méthodes -citées ci-avant et que nous allons les voir ci-après une après l'autre- en particulier sont utilisées dans plusieurs systèmes informatiques à savoir les systèmes d'aide au diagnostic médical, les systèmes de télémédecine, les systèmes de télésurveillance, les systèmes de production, les systèmes de reconnaissance de forme, les systèmes de la classification automatique (Learning machine).

### III.2 La classification en IA

La classification est une opération de structuration qui vise à organiser un ensemble d'observation en groupes homogènes et contrastés afin de faciliter l'analyse des informations et d'effectuer des prédictions [Bes, 07], ou, effectuer une classification, c'est mettre en évidence des relations entre des objets et entre ces objets et leurs paramètres. A partir de proximités ou de dissemblances, il s'agit de construire une partition de l'ensemble des objets en un ensemble de classes les plus homogènes possibles [Bel, Che et Ber, 03].

La classification rassemble une famille de méthodes permettant d'automatiser le processus de reconnaissance, elle peut être considérée comme un domaine permettant de définir des algorithmes et techniques susceptibles de classer des objets dont l'aspect est variable par rapport à un objet type. Le but de la classification est donc de classer des observations sur la base d'une série de caractéristiques prédéfinies par apprentissage.

Aujourd'hui on ne peut pas parler d'une théorie de la classification d'objets, mais d'approches différentes dont le but commun est de mettre en évidence, de façon automatique, une structure commune à l'ensemble des objets en présence, par recherche de similarités. Le champ d'application de la classification est particulièrement vaste. Il couvre principalement le domaine du traitement des signaux et d'images. On distingue habituellement deux types de classification:

- ❖ La classification avec ensemble d'apprentissage connu et parfaitement expertisé, ou classification en mode supervisé ;
- ❖ La classification en mode non supervisé : dans ce dernier cas, on ne dispose pas d'un ensemble d'apprentissage préalablement expertisé.

Dans la première partie de ce chapitre, nous allons brièvement rappeler les différentes phases d'un processus de classification d'objets et ensuite les deux types d'apprentissage utilisés, supervisé et non supervisé. La seconde partie de ce chapitre sera consacrée à la présentation des méthodes de classification utilisées en particulier dans le domaine médical [Chi, 05].

### III.2.1 Chaîne de classification d'objets

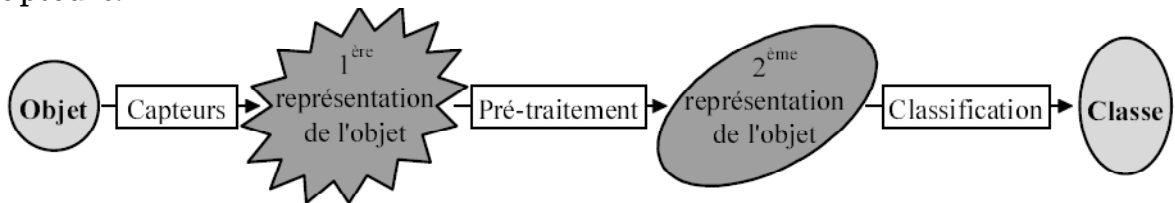
#### III.2.1.1 Les principaux modules de classification d'objets

Un dispositif de classification ou de reconnaissance automatique de formes est généralement conçu comme une chaîne de modules de traitement [Bel, 92]. Ainsi, un système de reconnaissance de formes comporte habituellement les modules suivants :

- Un module d'acquisition : des capteurs mesurent des grandeurs caractéristiques de l'objet à classer (signal, image, ...). Cet ensemble de grandeurs constitue la première représentation de l'objet.
- Un module de prétraitement : il peut être judicieux de modifier des grandeurs brutes issues des capteurs par un algorithme afin de tenir compte des connaissances qui peuvent être disponibles a priori sur le problème. Par exemple à partir de la réponse d'un capteur on peut appliquer un ensemble de filtres destinés à éliminer des bruits indésirables. Ainsi, on obtient une nouvelle représentation de l'objet, plus adéquate pour la classification envisagée. D'autres modules de traitement peuvent élaborer des représentations successives de l'objet ; ces différentes représentations ont généralement pour objectif de réduire la dimension de la représentation, c'est-à-dire de diminuer le nombre de descripteurs de l'objet et d'élaborer des descripteurs de plus en plus pertinents pour la tâche de discrimination à accomplir.
- Un module de classification : l'algorithme de classification considère la dernière représentation de l'objet et décide d'affecter celui-ci à une classe. Cet algorithme peut fournir soit une réponse

binaire à valeurs discrètes (appartenance ou non à une classe) soit une réponse à valeurs continues.

La **FIG 3.1** illustre une chaîne de classification comportant un seul module de prétraitement. On distingue les trois modules et les représentations successives de l'objet. Naturellement, on peut imaginer un dispositif sans module de prétraitement ; dans ce cas l'algorithme de classification travaille directement sur des grandeurs relevées par les capteurs.



**FIG 3.2** : Chaîne de modules de classification

La tâche de l'algorithme de classification est d'autant plus aisée que la représentation de l'objet est pertinente. Par exemple pour la classification des battements ventriculaires prématurés (Bvp), chaque cycle cardiaque est entièrement défini par les paramètres de l'ECG. Si les modules d'acquisition (électrodes) ou de prétraitement (filtrage et détection) ne fournissent pas des descripteurs pertinents à l'algorithme de classification, celui-ci ne pourra pas faire de miracle et distinguer les différentes arythmies.

### III.2.1.2 L'extraction des descripteurs

Habituellement une classification de formes ne se fait pas directement sur des formes brutes (morphologique de l'ECG par exemple), mais plutôt à partir de descripteurs ou paramètres caractérisant les formes. Chacune des  $N$  formes  $F_i$  est représentée par un point  $X_i$  qui est l'espace des paramètres. Il existe deux approches pour caractériser une forme :

- L'approche purement mathématique (analyse en composantes principales, prédiction linéaire, transformée de Fourier, etc.), qui consiste à retenir comme paramètres certains coefficients pertinents,
- L'approche intuitive, qui laisse au spécialiste le soin de définir les descripteurs qui lui semble importants. Cette approche donne souvent de meilleurs résultats, car les paramètres choisis résultent d'une grande expérience et peuvent plus discriminants [Chi, 05].

### III.2.1.3 Procédure de résolution par apprentissage

La classe reçoit une définition qui peut être une définition purement descriptive ou une interprétation, par exemple du type diagnostic médical. Le classifieur qui réalisera le classement des formes doit passer par deux phases [Les, 91], une phase d'apprentissage et une phase de test.

### III.2.1.3.1 Phase d'apprentissage

Le but de l'apprentissage est de découvrir les règles (généralement non déterministes) qui gouvernent et régissent des formes. L'apprentissage est un processus calculatoire qui doit être capable d'amener à une certaine prédiction et à une certaine généralisation. Il existe principalement deux types d'apprentissage, supervisé et non supervisé [Mil, 93]. Dans le premier cas, on doit apprendre des associations (individus, classes); dans le deuxième cas on ne fournit pas d'indications sur les classes.

#### 1- *Apprentissage supervisé*

L'apprentissage est dit supervisé si les différentes familles de formes, ou classes, sont connues *a priori* ainsi que l'affectation de chaque forme à telle ou telle famille. Le processus d'apprentissage en mode supervisé peut se résumer comme suit : on dispose d'un ensemble de formes ou d'objets, qui est un ensemble d'apprentissage A d'objets parfaitement classés, ensemble que l'on se propose de reconnaître ; il peut être présenté sous la forme de paires  $(X_i, C_j), 1 \leq i \leq N, 1 \leq j \leq m, m < N$ , et interprété : « l'objet  $X_i$  appartient à la classe  $C_j$  ».

On construit dans un premier temps une « machine » sachant classer correctement les objets de l'ensemble d'apprentissage A. La description de chaque forme est prise en compte par la tâche d'apprentissage. Le travail d'apprentissage consiste à analyser les ressemblances entre les formes d'une même classe et les dissemblances entre les formes de classes différentes pour en déduire les meilleures séparations possibles entre classes. Ensuite, on fait en sorte que cette « machine » ait une bonne capacité à prédire la classe d'une forme qui n'est pas dans l'ensemble d'apprentissages A. d'où l'exigence de tester cette machine sur un ensemble de validation et d'objets parfaitement classés. Cette phase de validation est nécessaire car elle aussi de savoir si la machine élaborée n'a pas été entraînée à apprendre « par cœur » sans possibilités de généralisation (sur-apprentissage ou over *fitting*) [Bel, 92], [Les, 91].

#### 2- *Apprentissage non supervisé*

On l'appelle aussi, suivant l'approche utilisée, classification automatique, ou apprentissage sans professeur ou encore apprentissage par corrélation [Les, 91]. Ce type d'apprentissage est utilisé dans le cas où on dispose d'une base d'apprentissage dont les classes ne sont pas définies à l'avance : soit constitué d'une part d'une liste d'objets  $X_i$  et d'autre part de la donnée du nombre de catégories, soit enfin uniquement composé d'un ensemble d'objets  $X_i, 1 \leq i \leq N$ .

Il s'agit alors de regrouper les différentes formes en classes en fonction d'un critère de similarité choisi *a priori*. Ce type d'apprentissage permet la construction automatique des classes sans intervention d'un opérateur [Nad, 93]. On distingue essentiellement deux types de méthodes : les processus de coalescence et la

classification hiérarchique [Bel, 92]. Cette approche nécessite cependant d'avoir une bonne estimation du nombre de classes.

### III.2.1.3.2 Phase de test

Cette phase doit permettre l'affectation d'un nouvel objet à l'une des classes, au moyen d'une règle de décision intégrant les résultats de la phase d'apprentissage [Les, 91]. L'objectif est d'obtenir une estimation la plus fidèle possible du comportement du classifieur dans des conditions réelles d'utilisation. Pour cela, des critères classiques comme les taux de classification et les taux d'erreur sont presque systématiquement utilisés. Mais d'autres critères, comme la spécificité et la sensibilité, apportent aussi des informations utiles.

#### 1. Taux de classification et taux d'erreurs

Les taux de classification et d'erreurs permettent d'évaluer la qualité du classifieur  $C$  par rapport au problème pour lequel il a été conçu. Ces taux sont évalués grâce à une base de test qui contient des formes décrites dans le même espace de représentation  $E$  que celles utilisées pour l'apprentissage.

Elles sont aussi étiquetées par leur classe réelle d'appartenance afin de pouvoir vérifier les réponses du classifieur. Pour que l'estimation du taux de reconnaissance soit la plus fiable possible, il est important que le classifieur n'ait jamais utilisé les échantillons de cette base pour faire son apprentissage (la base de test ne doit avoir aucun objet en commun avec la base d'apprentissage et les éventuelles bases de validation). De plus, cette base de test doit être suffisamment représentative du problème de classification [Guy, 98].

En général, quand les échantillons étiquetés à disposition sont suffisamment nombreux, ils sont séparés en deux parties disjointes et en respectant les proportions par classes de la base initiale. Une partie sert pour former la base d'apprentissage et l'autre pour former la base de test. Le découpage le plus courant est de 2/3 pour l'apprentissage et le 1/3 restant pour la base de test. Les performances en terme de taux de classification sont alors déterminées en présentant la classe donnée en résultat  $C(e_j)=s$  à la vraie classe de  $e_j$ . En considérant que la base de test contient  $N$  objets et que sur ceux-ci  $\tau_{corrects}$  sont bien classés par le système, le taux de classification  $\tau_{clas}$  est simplement défini par :

$$\tau_{clas} = \frac{N_{corrects} \cdot 100}{N}$$

Le taux d'erreur  $\tau_{err}$  est défini à partir du nombre d'objets  $N_{err}$  mal classés :

$$\tau_{err} = \frac{N_{err} \cdot 100}{N}$$

## 2. Sensibilité et spécificité

L'évaluation des performances d'un classifieur peut être réalisée par l'appréciation de deux lois statistiques, qui sont la sensibilité  $S_e(i)$  et la spécificité  $S_p(i)$  :

$$S_e(i) = \frac{VP(i)}{VP(i) + FN(i)}$$

$$S_p(i) = \frac{VN(i)}{VN(i) + FP(i)}$$

Où les grandeurs  $VP(i)$ ,  $FN(i)$ ,  $VN(i)$  et  $FP(i)$  sont définies dans le tableau suivant :

	Présence d'événement de classe i	Absence d'événement de classe i
Classification Positive	Vrai Positif $VP(i)$	Faux Positif $FP(i)$
Classification Négative	Faux Négatif $FN(i)$	Vrai Négatif $VN(i)$

**Tab 3.1** : Les définitions des grandeurs  $VP$ ,  $VN$ ,  $FP$  et  $FN$  [Chi, 05] .

La sensibilité  $S_e(i)$  représente la probabilité de bonne classification de la classe  $i$  et la spécificité  $S_p(i)$  est une mesure indirecte de la probabilité de fausse alarme puisque cette dernière n'est rien que  $1 - S_p(i)$ .

### III.2.3 Formalisation mathématique d'un problème de classification

La classification apparait comme une tâche qui consiste à ranger des formes ou objets décrits par un ensemble de variables descriptives en un certain nombre de catégories ou classes définies a priori. Traduit en termes mathématiques, un problème de classification comporte les éléments suivants :

- Une population de  $N$  objets  $O_i$ ,  $1 \leq i \leq N$ ,
- $P$  variables descriptives  $X_i^d$ , qui permettent de décrire les objets ; elles sont aussi appelées plus simplement descripteurs,  $1 \leq d \leq P$ ,
- $C$  classes  $C_k$  dans lesquelles on cherche à ranger les objets ( $k$  variant de 1 à  $C$ ). Résoudre un problème de classification, c'est trouver une application de l'ensemble des objets à classer, décrits par les variables descriptives choisies, dans l'ensemble des classes. L'algorithme ou la procédure qui réalise cette application est appelé classifieur. Les variables descriptives considérées ici sont celles qui sont fournies à l'algorithme de classification. Comme indiqué plus haut, elles peuvent être le résultat d'un prétraitement des variables initiales [Chi, 05].

## III .3. Méthodes de classification

Il existe de nombreuses méthodes de classification en *IA* appliquées dans le domaine médical, parmi celles-ci on trouve la logique floue, les arbres de décision, les réseaux de neurones (notre méthode de classification choisie pour la cardiologie), les réseaux bayésiens, les systèmes expert (en

plus les réseaux de neurones ; on va choisir les systèmes expert pour donner la classe finale), les SVM, etc.

### III.3.1. La logique floue en classification

L'approche par la logique floue permet l'introduction du concept de degré d'appartenance, qui détermine les « forces » avec lesquelles un individu appartient aux différentes classes. Cela repose sur le fait que le concept flou ne cherche pas un point de rupture qui décide de l'appartenance d'un individu à une classe, mais qu'elle raisonne plutôt sur la base d'un intervalle de valeurs. L'idée qui soutient l'approche est la possibilité d'appartenance à la fois à plusieurs classes. Partant dans ce sens, ce sont exclus toutes les méthodes de classification « dure » contraignant les individus à être membre d'une, et une seule, classe. Quoique la probabilité d'appartenance des objets aux classes ne soit pas évidemment une exclusivité des techniques de classification floue, il nous semble que les résultats obtenus des approches floues sont plus pertinents en ce sens qu'ils donnent une matrice des degrés d'appartenance de chaque individu à chaque classe, ce qui n'est pas tout à fait le cas dans les autres analyses « classiques » de classification[MES, 05].

#### III.3.1.1. Conception du classificateur flou

##### III.3.1.1.1. Principe de base

Le problème de diagnostic médical peut être formulé comme suit [Kun, 99] :

Soit  $C = \{C_1 ; C_2 ; \dots ; C_M\}$  un ensemble de  $M$  diagnostics possibles dans le contexte d'un certain problème médical.  $C$  peut être un ensemble d'anomalies du cœur, des types de tissus dans une Image par Résonance Magnétique (IRM) du cerveau ou des types de cellules sanguines, etc. Soit  $x$  une description d'un objet (par exemple, un patient, un morceau de tissu cérébral, une cellule, etc.) sous forme d'un vecteur de réels de  $Q$  éléments :

$$x = [p_1 ; p_2 ; \dots ; p_Q] \in \mathbb{R}^Q.$$

Les composantes du vecteur  $x$  encodent des paramètres tels que des mesures cliniques, des détails physiologiques d'un patient, des résultats de tests, des paramètres d'imagerie telle que l'intensité du niveau de gris, l'aspect circulaire d'une cellule, des électrocardiogrammes, etc.

Un classificateur flou est toute application :  $D : \mathbb{R}^Q \rightarrow [0,1]^M$ , le résultat de la classification est alors donné par :  $D(x) = [\mu_1(x) ; \dots ; \mu_M(x)]$  où  $\mu_i(x)$  représente le degré d'appartenance de  $x$  à la classe  $C_i$  (FIG 3.2).

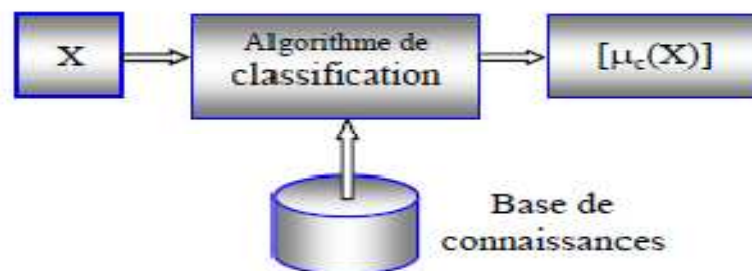


FIG 3.2: Schéma synoptique d'un classifieur supervisé.



### III.3.1.2. Avantages et Inconvénients

Les systèmes à base de logique floue ont des formalismes robustes qui ne nécessitent aucun modèle mathématique, ils permettent l'intégration de connaissances à priori sous forme de règles, ils garantissent le suivi et l'interprétation des processus et des résultats, ils sont facilement implémentables et permettent l'extraction des règles.

Par contre, Ils n'ont aucun apprentissage possible, ils ont une adaptation relativement difficile aux modifications des données, ils nécessitent un processus de défuzzification et l'inconvénient majeur pour nous est qu'ils sont sensibles aux données bruitées, le cas de notre base d'apprentissage.

### III.3.2. Les arbres de décision

Pour certains domaines d'application, il est essentiel de produire des procédures de classification compréhensibles par l'utilisateur. C'est en particulier le cas pour l'aide au diagnostic médical où le médecin doit pouvoir interpréter les raisons du diagnostic. Les arbres de décision répondent à cette contrainte car ils représentent graphiquement un ensemble de règles et sont aisément interprétables. Pour les arbres de grande taille, la procédure globale peut être difficile à appréhender, cependant, la classification d'un élément particulier est toujours compréhensible.

#### III.3.2.1. Exemple

Pour mieux appréhender l'induction des arbres de décision, nous allons reprendre un exemple décrit dans l'ouvrage de **Quinlan (1993)**. Il s'agit de prédire le comportement de sportifs (Jouer ; variable à prédire) en fonction de données météo (Ensoleillement, Température, Humidité, Vent ; variables prédictives) [Qua, 06].

Numéro	Ensoleillement	Température (°F)	Humidité (%)	Vent	Jouer
1	soleil	75	70	oui	oui
2	soleil	80	90	oui	non
3	soleil	85	85	non	non
4	soleil	72	95	non	non
5	soleil	69	70	non	oui
6	couvert	72	90	oui	oui
7	couvert	83	78	non	oui
8	couvert	64	65	oui	oui
9	couvert	81	75	non	oui
10	pluie	71	80	oui	non
11	pluie	65	70	oui	non
12	pluie	75	80	non	oui
13	pluie	68	80	non	oui
14	pluie	70	96	non	oui

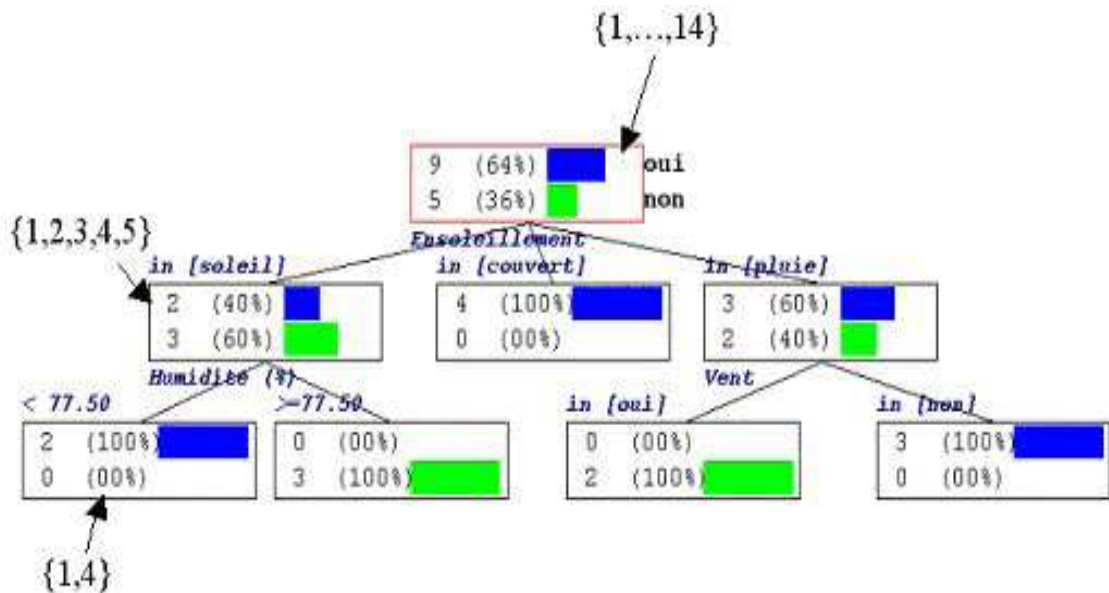
Tab 3.2: Exemple des données

L'algorithme d'apprentissage cherche à produire des groupes d'individus le plus homogène possible du point de vue de la variable à prédire à partir des variables de météo. Le partitionnement est décrit à l'aide d'un arbre de décision.

Sur chaque sommet de l'arbre est décrite la distribution de la variable à prédire. Dans le cas du premier sommet, la racine de l'arbre, nous constatons qu'il y a 14 observations dans notre fichier, 9 d'entre eux ont décidé de jouer (Jouer = oui), 5 ont décidé le contraire (Jouer = non).

Ce premier sommet est segmenté à l'aide de la variable Ensoleillement, 3 sous-groupes ont été produits. Le premier groupe à gauche (Ensoleillement = Soleil) comporte 5 observations, 2 d'entre eux correspondent à Jouer = oui, 3 à Jouer = non.

Chaque sommet est ainsi itérativement traité jusqu'à ce que l'on obtienne des groupes suffisamment homogènes. Elles correspondent aux feuilles de l'arbre, des sommets qui ne sont plus segmentés [Qua, 06].



**FIG 3.3:** Arbre de décision [Qua, 06]

La lecture d'un arbre de décision est très intuitive, c'est ce qui fait son succès. L'arbre peut être traduit en base de règles sans pertes d'informations. Si l'on considère la feuille la plus à gauche, nous pouvons aisément lire la règle d'affectation suivante :

« Si ensoleillement = soleil et humidité < 77.5% alors jouer = oui ».

### III.3.2.2. Avantages et Inconvénients

Les arbres de décisions ont beaucoup d'avantages tels que la manipulation facile des données symboliques, la sollicitation des variables d'amplitudes très différentes et la bonne interprétation des résultats. Par contre ils sont très sensibles au bruit et points aberrants en plus, ils sont valables que pour les échantillons de petite taille et le changement d'une variable changera l'arbre complètement (Problème de mise à jour).

### III.3.3. Réseaux bayésiens

Les réseaux bayésiens sont à la fois :

- Des modèles de représentation des connaissances
- Des "machines à calculer" les probabilités conditionnelles.

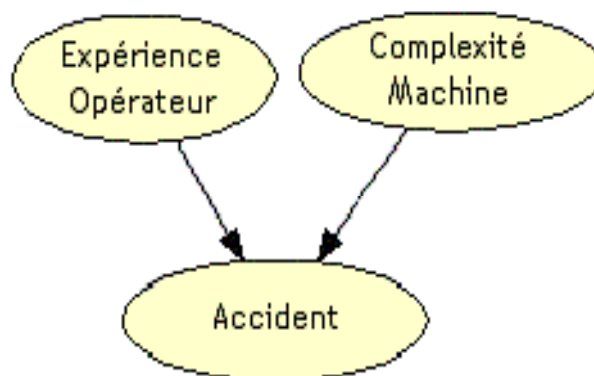
Pour un domaine donné, on décrit les relations causales entre variables d'intérêt par un graphe. Dans ce graphe, les relations de cause à effet entre les variables ne sont pas déterministes, mais probabilisées. Ainsi, l'observation d'une cause ou de plusieurs causes n'entraîne pas systématiquement l'effet ou les effets qui en dépendent, mais modifie seulement la probabilité de les observer.

L'intérêt particulier des réseaux bayésiens est de tenir compte simultanément de connaissances a priori d'experts (dans le graphe) et de l'expérience contenue dans les données. Les réseaux bayésiens sont surtout utilisés pour le diagnostic (médical et industriel), l'analyse de risques, et le datamining [Qua, 06].

#### III.3.3.1. Exemple

Un opérateur travaillant sur une machine risque de se blesser, s'il l'utilise mal. Ce risque dépend de l'expérience de l'opérateur et de la complexité de la machine. «Expérience» et «Complexité» sont deux facteurs déterminants de ce risque (Figure 13: Structure de causalité) Bien sûr, ces facteurs ne permettent pas de créer un modèle déterministe. Si l'opérateur est expérimenté, et la machine simple, cela ne garantit pas qu'il n'y aura pas d'accident. D'autres facteurs peuvent jouer : l'opérateur peut être fatigué, dérangé, etc. La survenance du risque est toujours aléatoire, mais la probabilité de survenance dépend des facteurs identifiés.

Le schéma ci-dessous représente la structure de causalité de ce modèle  
**FIG 3.4:** Structure de causalité.



**FIG 3.4:** Structure de causalité

Le suivant représente la potabilisation de la dépendance : on voit que la probabilité d'accident augmente si l'utilisateur est peu expérimenté ou la machine complexe.

On voit ici comment intégrer des connaissances d'expert (les facteurs déterminants) et des données (par exemple, la table de probabilité d'accident en fonction des déterminants peut venir de statistiques).

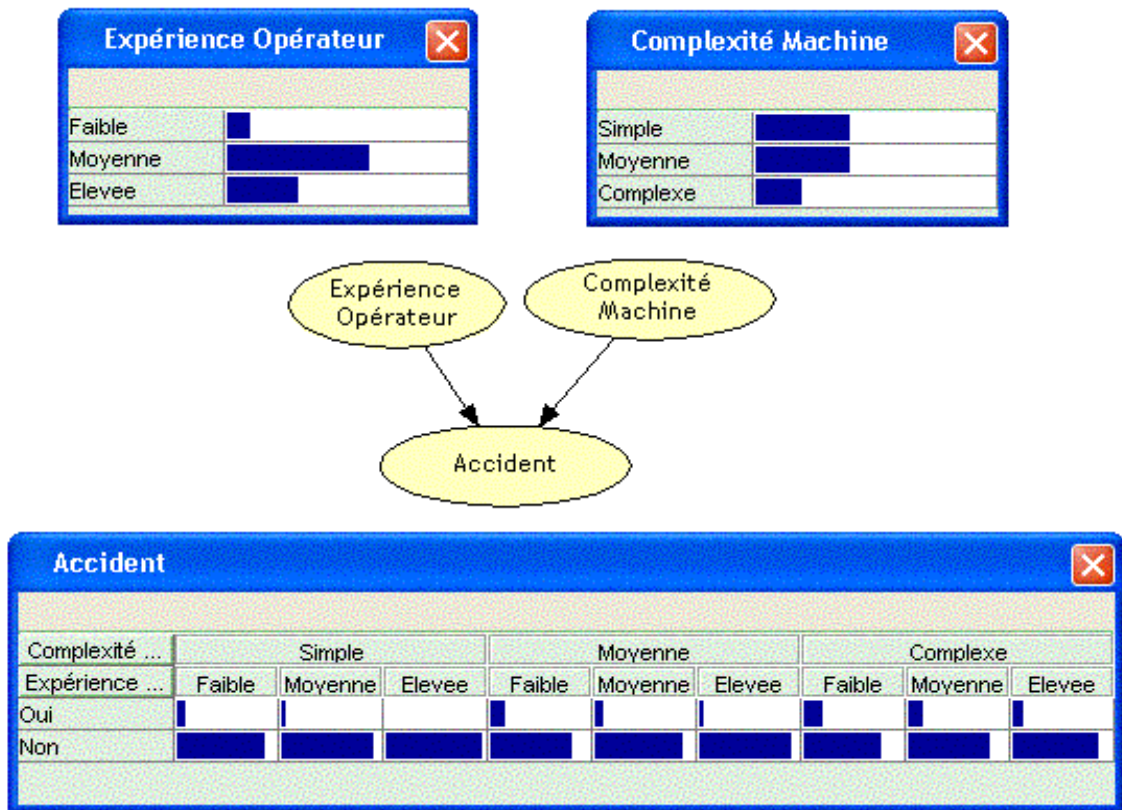


FIG 3.5: Intégration des connaissances d'expert et des données

### III.3.3.2. Avantages et Inconvénients

Les aspects suivants des réseaux bayésiens les rendent, dans de nombreux cas, préférables à d'autres modèles :

*Acquisition des connaissances* : Les réseaux bayésiens donnent la possibilité de rassembler et de fusionner des connaissances de diverses natures dans un même modèle.

*Représentation des connaissances* : La représentation graphique d'un réseau bayésien est explicite, intuitive et compréhensible.

*Utilisation des connaissances* : Un réseau bayésien est polyvalent : on peut se servir du même modèle pour évaluer, prévoir, diagnostiquer, ou optimiser des décisions. Mais les réseaux bayésiens ont pas mal d'inconvénients surtout la complexité des algorithmes :

*Les variables continues* : La plupart des algorithmes développés pour l'inférence et l'apprentissage dans les réseaux bayésiens, aussi bien que les outils disponibles sur le marché pour mettre en œuvre ces algorithmes utilisent des variables discrètes.

*La complexité des algorithmes* : La généralité du formalisme des réseaux bayésiens aussi bien en termes de représentation que d'utilisation les rend difficiles à manipuler à partir d'une certaine taille.

### III.3.4. MACHINE A VECTEURS DE SUPPORT (SVM)

SVM (SUPPORT VECTOR MACHINE) est une méthode de classification qui fut introduite par *Vapnik* en 1995. Le succès de cette méthode est justifié par les solides bases théoriques qui la soutiennent. Il existe en effet un lien direct entre la théorie de l'apprentissage statistique et l'algorithme d'apprentissage de SVM [Vap, 98], [Cal, 02].

Pour deux classes d'exemples donnés, le but de SVM est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Avec SVM, ce classificateur est un classificateur linéaire appelé hyperplan.

Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal [Bes, 07].

Dans [Bes, 07], les auteurs ont utilisé les SVM pour la classification des arythmies cardiaques dont ils ont proposé un système de reconnaissance basé sur les machines à vecteur de support pour la classification de douze rythmes cardiaques. Pour cela, deux types de paramétrages ont été utilisés, à savoir les descripteurs morphologiques et les moments d'ordre supérieur. L'approche proposée a été validée sur la base de données standard MITBIH d'arythmies.

Ils ont fait une étude qui consiste à la classification de 12 classes de battement cardiaques, ils ont cité par exemple le rythme normal N, bloc de branche droit R et gauche L,...etc. La répartition des différents exemples en deux sous bases : base d'apprentissage et base de test.

#### III.3.4.1. Avantages et Inconvénients

L'avantage des SVM est qu'ils s'adaptent facilement aux problèmes non linéairement séparables. Avant de procéder à l'apprentissage de la meilleure séparation linéaire, on transforme les vecteurs d'entrée en vecteurs de caractéristiques de dimension plus élevée. De cette façon, un séparateur linéaire trouvé par un SVM dans ce nouvel espace vectoriel devient un séparateur non linéaire dans l'espace original. Cette transformation des vecteurs se fait à l'aide de noyaux («*kernels*»).

Dans le cas de la classification de textes, les entrées sont des documents et les sorties sont des catégories. En considérant un classificateur binaire, on voudra lui faire apprendre l'hyperplan qui sépare les documents appartenant à la catégorie et ceux qui n'en font pas partie. Selon [Joa, 98a], les SVM conviennent bien pour la classification de textes parce que, premièrement, une dimension élevée ne les affecte pas puisqu'ils se protègent contre le sur-apprentissage. Dans le même sens, il affirme que peu d'attributs sont totalement inutiles à la tâche de classification et que les SVM permettent d'éviter une sélection agressive qui aurait comme

résultat une perte d'information. On peut se permettre de conserver plus d'attributs. Également, une caractéristique des documents textuels est que lorsqu'ils sont représentés par des vecteurs, une majorité des entrées sont nulles. Or, les SVM conviennent bien à des vecteurs dits clairsemés. Un autre aspect positif des SVM est qu'aucun ajustement de paramètres manuel n'est requis, car ils ont l'habileté de trouver automatiquement des paramètres adéquats.

### III.3.5. Les carte de kohonen

Les cartes auto adaptative de **Kohonen** sont des réseaux neuro-mimétiques qui s'inspirent des modèles biologiques de perception. Le réseau simulé pour la classification des battements cardiaque est une SOM bidimensionnelle composée de  $N \times N$  neurones artificiels **FIG 3.6**

- Le vecteur d'entrée  $X = [x_1, x_2, \dots, x_{12}]$  représente les 12 paramètres caractérisant un battement cardiaque, et il est lié à tous les neurones.
- Les connexions extérieures du neurone  $i$  avec le vecteur d'entrée  $X$  se matérialise par un vecteur poids synaptiques  $M_i = [m_{i1}, m_{i2}, \dots, m_{i12}]$ .
- D'autre part, chaque neurone  $i$  de la carte est lié à tous les autres neurones de la carte : interaction entre neurones. Les poids des connexions internes  $w_{ik}$  du neurone  $i$  avec ses voisins sont donnés par la fonction dite du « chapeau mexicain ». Ces connexions dépendent de la distance entre neurones et sont invariants dans le temps.
- La sortie  $S_i$  du neurone  $i$  à l'instant  $t$  s'exprime par :  $S_i(t) = \sigma [\sum_j m_{ij} X_j + \sum_k w_{ik} S_k(t-1)]$ , la sortie  $S_i$  à l'instant  $t$  du neurone  $i$  est donc fonction de deux termes, une entrée totale ( $\sum_j m_{ij} X_j$ ) et une entrée interne résultante d'une somme pondérée (par l'importance des interactions latérales) des signaux émis aux sorties des neurones voisins à l'instant  $(t-1)$  ( $\sum_k w_{ik} S_k(t-1)$ ). La somme globale est évaluée par la fonction d'activation sigmoïde  $\sigma$ .

Pour que la carte atteigne un certain pouvoir de généralisation et de reconnaissance, une phase d'apprentissage est indispensable. Durant cette étape d'entraînement, le réseau s'adapte aux différents types d'arythmies cardiaques et qui sont regroupés dans une base d'apprentissage. Cette phase consiste à modifier les poids des connexions extérieures entre les neurones de la cartes et les entrées.

Pour cela, il existe deux types d'apprentissage : apprentissage supervisé et un autre non supervisé. Dont le premier est une forme d'apprentissage qui se base sur un ensemble de données annotés (Dans notre cas la base de données MIT BIH), c'est à dire où chaque vecteur de l'ensemble a une valeur qui lui est associée.

Dans [Mou, 09], les auteurs ont proposé un système d'aide au diagnostic de quatre (04) arythmies cardiaques très fréquentes qui sont : Extrasystole Ventriculaire (**EV**), Extrasystole Auriculaire (**EA**), Bloc de Branche Droit (**BBD**), et Bloc de Branche Gauche (**BBG**), en plus du battement normal (**N**).

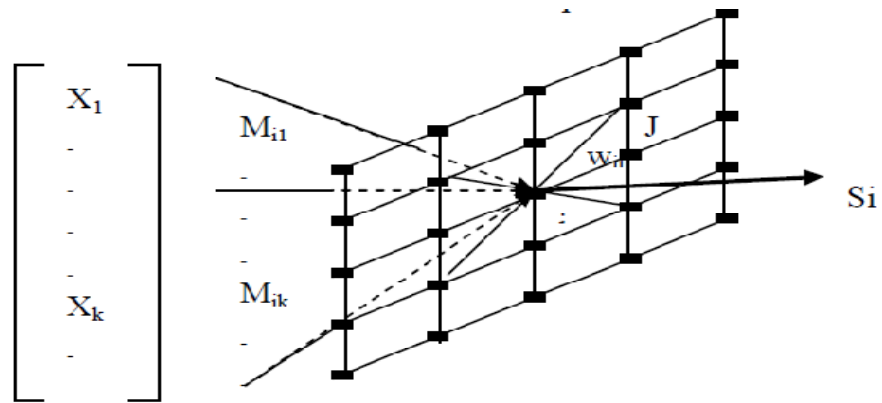


FIG 3.6: La carte auto adaptative de Kohonen 20\*20 [Mou, 09]

Ils ont commencé tout d'abord par localiser les ondes QRS, P, et T et par calculer les paramètres temporelles et morphologiques qui caractérisent un battement en utilisant principalement la technique des ondelettes. Ensuite, ils ont implémenté une carte auto-organisatrice de Kohonen qui est chargée de déterminer le type du battement en fonction de ses caractéristiques. Le système a été validé sur des enregistrements extraits de la base de données MIT-BIH.

### III.3.5.1. Avantages et Inconvénients

Les ancêtres des cartes de Kohonen, les algorithmes comme "k-moyennes", réalisent la discrétisation de l'espace d'entrée en ne modifiant à chaque cycle d'adaptation qu'un seul vecteur référent. Leur processus d'apprentissage est donc très long. L'algorithme de Kohonen profite des relations de voisinage dans la grille pour réaliser une discrétisation dans un temps très court. On suppose que l'espace n'est pas constitué de zones isolées, mais de sous-ensembles compacts. Donc en déplaçant un vecteur référent vers une zone, on peut se dire qu'il y a probablement d'autres zones dans la même direction qui doivent être représentées par des vecteurs référents. Cela justifie le fait de déplacer les neurones proches du vainqueur dans la grille dans cette même direction, avec une amplitude de déplacement moins importante. L'algorithme présente des opérations simples ; il est donc très léger en termes de coût de calculs. Par contre, le voisinage dans les cartes auto adaptatives est malheureusement fixe, et une liaison entre neurones ne peut être cassée même pour mieux représenter des données discontinues.

### III .4. Conclusion

Dans ce chapitre, nous avons introduit la notion de classification, formalisme d'un problème de classification, les différentes méthodes de classification existantes (La logique floue, les arbres de décision, les réseaux bayésiens, les SVM- *Support Vector Machine* et les cartes de *kohonen*) et qui permettent de traiter de manière proche ou similaire de point de vue résultats obtenus avec celles qu'on a choisi pour faire la classification des arythmies cardiaques à savoir le réseau de neurones multicouches et le système expert.

Ainsi, Nous avons donné le strict minimum pour chaque méthode en illustrant chacune d'elles par un exemple.

Finalement, nous avons jugé utile donner pour chaque méthode ces avantages et inconvénients, afin de justifier le choix des réseaux de neurones multicouches et les systèmes expert pour faire la classification des troubles cardiaques.

Dans le chapitre suivant, nous allons vous définir les réseaux de neurones en détaillant le perceptron multicouches, son algorithme d'apprentissage – Rétropropagation du gradient – et de même manière on va détailler les systèmes expert à la fin du chapitre.





## IV. LES RESEAUX DE NEURONES & LES SYSTEMES EXPERTS

### IV.1. Introduction

Apprendre, tout comme rêver, créer ou jouer, fait partie de ses expériences nécessaires qui stimulent et structurent le savoir-faire individuel de tout être humain depuis sa naissance. Qu'il soit biologique (comme marcher, agripper ou regarder avec attention) ou intellectuel (tel que réfléchir, exprimer une opinion ou donner un avis), l'apprentissage est certes un acte fondamental pour l'être humain, mais c'est aussi le processus le plus complexe et le plus subjectif à concevoir. Pas facile finalement de ramener ce mot à un seul synonyme en fonction de ces différents sens puisque « apprendre » une poésie n'est pas ce terme à de l'acquisition de connaissances ou à de l'observation d'événements, il n'en reste pas moins un concept primordial suscitant l'intérêt de nombreuses recherches aussi bien en sciences sociales qu'en informatique.

Dans ce chapitre, plus précisément, dans la première partie, nous nous intéressons à l'application des réseaux de neurones dans le domaine de la cardiologie, plus précisément pour la classification des battements cardiaques.

Dans un premier temps, nous rappellerons les définitions et notations de base relatives aux réseaux de neurones, les différents types des réseaux de neurones, puis nous poursuivrons en exposant les types et la méthodologie d'apprentissage.

Nous présenterons, les étapes de conception d'un réseau de neurones : le choix des entrées et sorties, les couches cachées dans le cas d'un perceptron multicouches et de la structure du réseau etc.....

Et dans la deuxième partie de ce chapitre, nous allons vous présenter une brève introduction sur les systèmes experts, leur architecture, la base de connaissance (base des faits, base des règles), le moteur d'inférence et finalement avantages et inconvénients des systèmes experts.

### IV.2. Historique sur les réseaux de neurones

Depuis fort longtemps, Le rêve de l'humain est de créer une machine dotée d'une forme d'intelligence, pour cela, des recherches menées par des chercheurs pour comprendre comment l'homme fait-il pour penser, raisonner ou même éprouver des sentiments ? Ces recherches ont abouties à l'étude du fonctionnement du cerveau est aussi connue sous le nom de connexionnisme, dont son objectif est de rendre compte de la cognition humaine par des réseaux de neurones.

Brièvement, les premières recherches remontent à la fin du 19e et au début du 20e siècle. Ils consistent en de travaux multidisciplinaires en physique, en psychologie et en neurophysiologie par des scientifiques tels *Hermann Von Helmholtz*, *Ernst Mach* et *Ivan Pavlov*. À cette époque, il s'agissait de théories plutôt générales sans modèle mathématique précis

d'un neurone. On s'entend pour dire que la naissance du domaine des réseaux de neurones artificiels (RNA) remonte aux années 1940 avec les travaux de *Warren Mc Culloch* et *Walter Pitts* qui ont montré qu'avec de tels réseaux, on pouvait, en principe, calculer n'importe quelle fonction arithmétique ou logique. Quelques années plus tard, en 1949, *Donald Hebb* a ensuite proposé une théorie fondamentale pour l'apprentissage.

La première application concrète des réseaux de neurones artificiels est survenue vers la fin des années 1950 avec l'invention du réseau dit « perceptron » par un dénommé *Frank Rosenblatt*.

*Rosenblatt* et ses collègues ont construit un réseau et démontré ses habilités à reconnaître des formes. Malheureusement, il a été démontré par la suite que ce perceptron simple ne pouvait résoudre qu'une classe limitée de problème. Environ au même moment, *Bernard Widrow* et *Ted Hoff* ont proposé un nouvel algorithme d'apprentissage pour entraîner un réseau adaptatif de neurones linéaires, dont la structure et les capacités sont similaires au perceptron.

Malgré tout l'enthousiasme que soulève le travail de *Rosenblatt* dans le début des années 60, la fin de cette décennie sera marquée en 1969, par une critique violente du Perceptron par *Minsky* et *Papert*. Ils montrent dans un livre « *Perceptrons* » toutes les limites de ce modèle, et soulèvent particulièrement l'incapacité du perceptron à résoudre les problèmes *non linéairement séparables*, tels que le célèbre problème du **XOR** (OU exclusif).

Il s'en suivra alors, face à la déception, une période noire d'une quinzaine d'années dans le domaine des réseaux de neurones artificiels.

Il faudra attendre le début des années 80 et le génie de *Hopfield* pour que l'intérêt pour ce domaine soit de nouveau présent. En effet, *Hopfield* démontre en 1982 tout l'intérêt d'utiliser des réseaux récurrents (dits "feed-back") pour la compréhension et la modélisation des processus mnésiques. Les réseaux récurrents constituent alors la deuxième grande classe de réseaux de neurones, avec les réseaux type perceptron (dits "*feed-forward*").

En parallèle des travaux de *Hopfield*, *Werbos* conçoit son algorithme de *rétropropagation* de l'erreur, qui offre un mécanisme d'apprentissage pour les réseaux *multicouches* de type perceptron (appelés **MLP** pour Multi-layer Perceptron), fournissant ainsi un moyen simple d'entraîner les neurones des couches cachées. Cet algorithme de "*back-propagation*" ne sera pourtant popularisé qu'en 1986 par *Rumelhart*.

De nos jours, l'utilisation des réseaux de neurones dans divers domaines ne cesse de croître. Les applications en sont multiples et variées.

## IV.3 Les réseaux de neurones

### IV.3.1. Le neurone biologique

Le neurone biologique est une cellule vivante spécialisée dans le traitement des signaux électriques. Les neurones sont reliés entre eux par des liaisons appelées *axones*. Ces derniers vont eux mêmes jouer un rôle important dans le comportement logique de l'ensemble. Ils conduisent les signaux électriques de la sortie d'un neurone vers l'entrée (*synapse*) d'un autre neurone.

Les neurones font une sommation des signaux reçus en entrée et en fonction du résultat obtenu vont fournir un courant en sortie. (FIG 4.1)

La structure d'un neurone se compose de trois parties :

1. **La somma** : ou cellule d'activité nerveuse, au centre du neurone.
2. **L'axone** : attaché au somma qui est électriquement actif, ce dernier conduit l'impulsion conduite par le neurone.
3. **Dendrites** : électriquement passives, elles reçoivent les impulsions d'autres neurones.

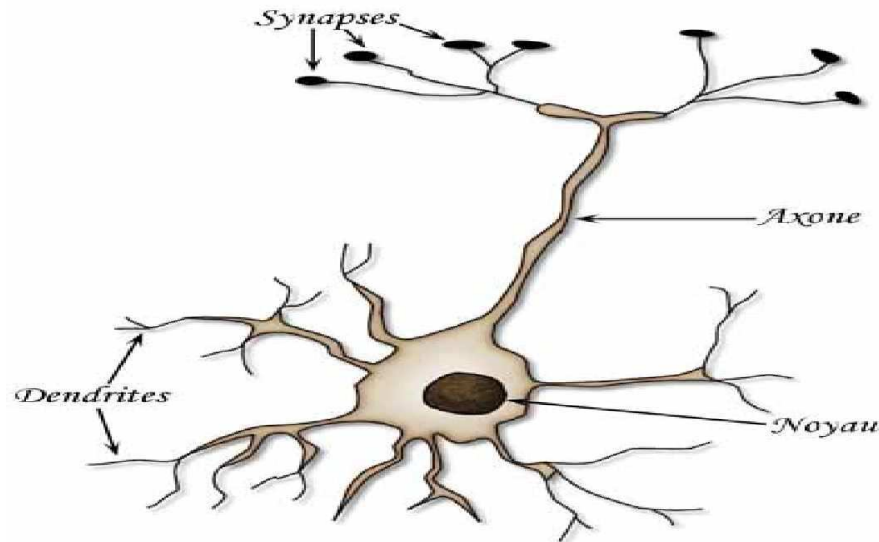


FIG 4.1. Le neurone biologique

### IV .3.2. Le neurone artificiel et principe de fonctionnement

Chaque neurone artificiel est un processeur élémentaire. Il reçoit un nombre variable d'entrées en provenance de neurones en amont ou des capteurs composant la machine dont il fait partie. A chacune de ses entrées est associé un poids représentatif de la force de la connexion. Chaque processeur élémentaire est doté d'une sortie unique, qui se ramifie ensuite pour alimenter un nombre variable de neurones en aval. A chaque connexion est associé un poids. Il est commode de représenter graphiquement un neurone comme indiqué sur la FIG 4.2.

Cette représentation est à l'origine de la première vague d'intérêt pour les neurones formels, dans les années 1940 à 1970 [McC et al. 43], [Min et Pap. 88].

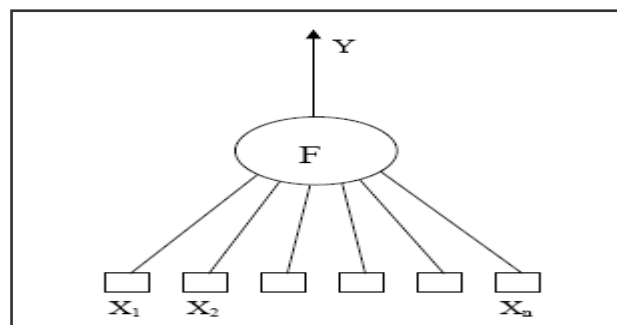


FIG 4.2. Neurone artificiel [Amm, 07].

Le neurone réalise alors trois opérations sur ses entrées :

- *Pondération* : multiplication de chaque entrée par un paramètre appelé poids de connexion,
- *Sommation* : une sommation des entrées pondérées est effectuée,
- *Activation* : passage de cette somme dans une fonction, appelée fonction d'activation.

La valeur calculée est la sortie du neurone qui est transmise aux neurones suivants.

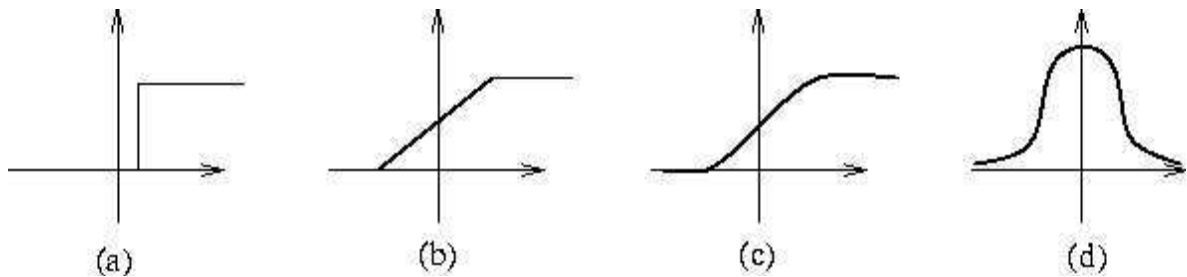


FIG 4.3. Fonctions d'activations :

(a) Seuil, (b) Linéaire par morceaux, (c) Sigmoïde, (d) Gaussienne

La fonction  $f$  est appelée **fonction d'activation** (FIG 4.3). Elle peut être une fonction à seuil, une fonction linéaire ou non linéaire. La fonction sigmoïde se présente comme une approximation continûment dérivable de la fonction d'activation linéaire par morceaux ou de la fonction seuil. Elle présente l'avantage d'être régulière, monotone, continûment dérivable, et bornée entre 0 et 1 :

$$f(x) = \frac{1}{1 + \exp(-x)}$$

La fonction  $f$  peut être paramétrée de manière quelconque. Deux types de paramétrages sont fréquemment utilisés :

- Les paramètres sont attachés aux entrées du neurone : la sortie du neurone est une fonction non linéaire d'une combinaison des entrées  $\{x_i\}$  pondérées par les paramètres  $\{w_i\}$ , qui sont alors souvent désignés sous le nom de poids.

$$y = f \left[ w_0 + \sum_{i=1}^{n-1} w_i x_i \right]$$

- Les paramètres sont attachés à la non linéarité du neurone : ils interviennent directement dans la fonction  $f$  ;
- ➔ On note que la fonction sigmoïde est la plus utilisée dans la pratique.

### IV.3.3. Importance de la fonction d'activation

Chaque neurone formel possède des caractéristiques propres, en particulier un seuil de déclenchement, assimilable à un poids synaptique dont le dépassement implique la décharge du neurone, c'est-à-dire la

transmission d'une information de sortie. La fonction seuil interne à chaque neurone formel assure que la valeur de sommation des potentiels pré-synaptiques ne dépassera pas certaines limites raisonnables (en général l'intervalle 0-1). Elle interdit les évolutions catastrophiques (effets de boucle où les valeurs deviennent de plus en plus grandes). Le perceptron originel, inventé par *Frank Rosenblatt* en 1958, possédait une telle fonction d'activation (sorties toujours ramenées à 0 ou 1). *Minsky* et *Papert* ont démontré en 1969 les limites de cette sorte de perceptron (incapacité de discriminer des classes non linéairement séparables, problème de **XOR**) induisant à attribuer ces limites à tous les réseaux de neurones.

Cependant, les neurones à « sigmoïde » (la fonction douce et progressive) peuvent bien tout modéliser, même avec une seule couche cachée.

#### IV.3.4. Réseau de neurones

Un réseau de neurones peut être considéré comme un modèle mathématique de traitement réparti, composé de plusieurs éléments de calcul non linéaire (neurones), opérant en parallèle et connectés entre eux par des poids.

Les réseaux de neurones artificiels sont des réseaux fortement connectés de processeurs élémentaires fonctionnant en parallèle. Chaque processeur élémentaire calcule une sortie unique sur la base des informations qu'il reçoit.

Les neurones artificiels sont souvent utilisés sous forme de réseaux qui diffèrent selon le type de connections entre les neurones, une cinquantaine de types peut être dénombrée. En guise d'exemples nous citons : le *perceptron* de *Rosenblatt*, le réseau de *Hopfield* etc.

Ces derniers constitués d'un nombre fini de neurones qui sont arrangés sous forme de couches. Les neurones de deux couches adjacentes sont interconnectés par des poids. L'information dans le réseau se propage d'une couche à l'autre, on dit qu'ils sont de type « feed-forward ». Nous distinguons trois types de couches :

**Couche d'entrée** : les neurones de cette couche reçoivent les valeurs d'entrée du réseau et les transmettent aux neurones cachés. Chaque neurone reçoit une valeur, il ne fait pas donc de sommation.

**Couches cachées** : chaque neurone de cette couche reçoit l'information de plusieurs couches précédentes, effectue la sommation pondérée par les poids, puis la transforme selon sa fonction d'activation qui est en général une fonction sigmoïde. Par la suite, il envoie cette réponse aux neurones de la couche suivante.

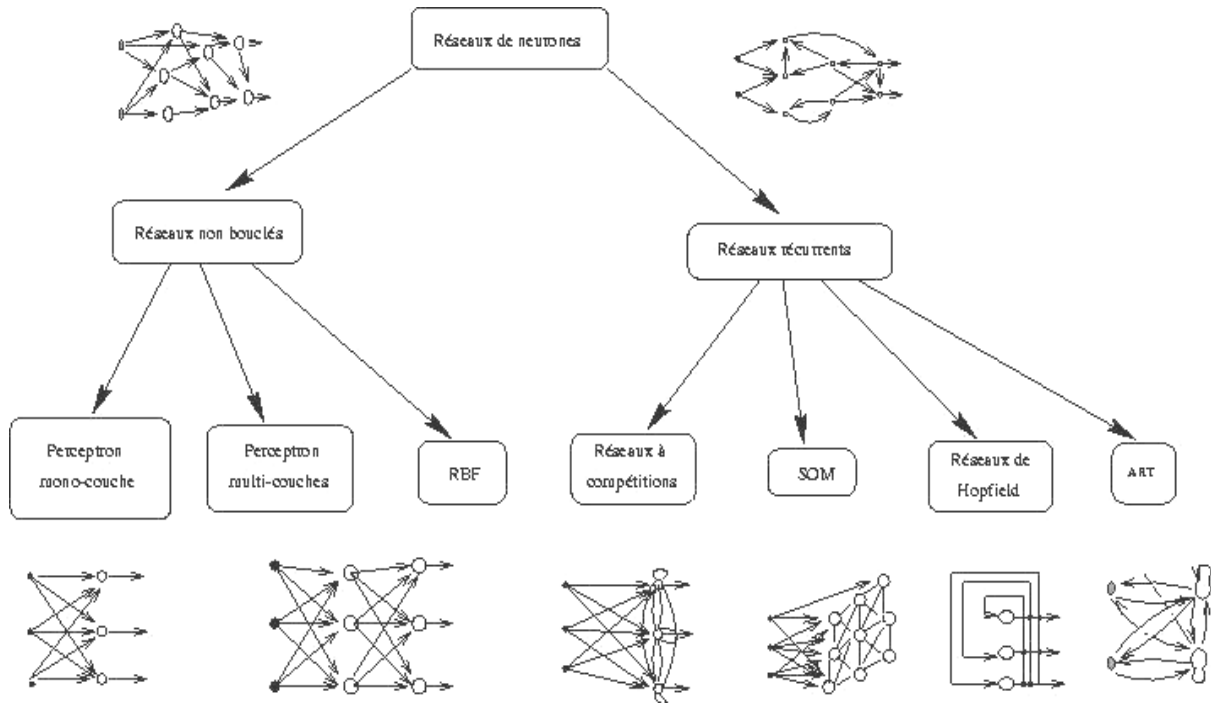
**Couche de sortie** : elle joue le même rôle que les couches cachées, la seule différence entre ces deux types de couches est que la sortie des neurones de la couche de sortie n'est liée à aucun autre neurone.

**IV.3.5. Architecture des réseaux de neurones**

On distingue deux structures de réseau, en fonction du graphe de leurs connexions, c'est-à-dire du graphe dont les nœuds sont les neurones et les arêtes les «connexions» entre ceux-ci :

- ❖ Les réseaux de neurones non bouclés (ou acycliques ou *feed-forward*).
- ❖ Les réseaux de neurones bouclés (ou récurrents, ou Feedback).

Voici une taxonomie des réseaux de neurones artificiels bouclés et non bouclés (FIG 4.4) .



**FIG 4.4:** Taxonomie des réseaux de neurones non bouclés et récurrents [Che, 05]

Un réseau de neurones non bouclé réalise une (ou plusieurs) fonction algébrique de ses entrées par composition des fonctions réalisées par chacun de ses neurones. Dans un tel réseau (FIG 4.5), le flux d'information circule des entrées vers les sorties sans retour en arrière. Si l'on représente le réseau comme un graphe dont les nœuds sont les neurones et les arêtes les « connexions » entre ceux-ci, le graphe d'un réseau non bouclé est acyclique.

Tout neurone dont la sortie est une sortie du réseau est appelé « neurone de sortie ». Les autres, qui effectuent des calculs intermédiaires, sont des « neurones cachés ». Il existe deux types de réseaux de neurones : les réseaux complètement connectés et les réseaux à couche. Le réseau de neurones à une couche cachée et une sortie linéaire est un cas particulier de ce dernier type.

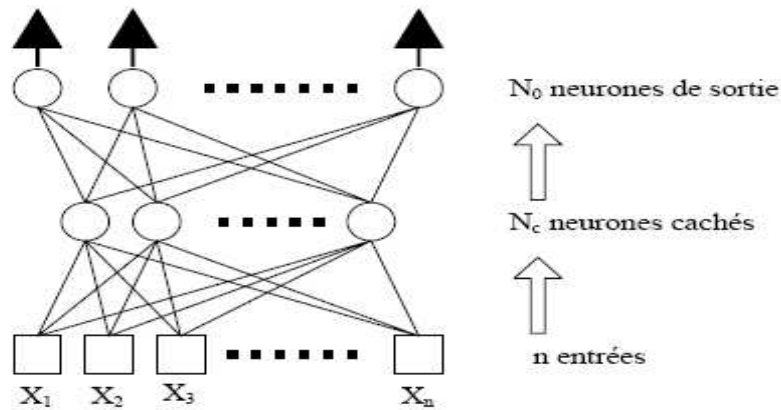


FIG 4.5. Réseau de neurones à  $n$  entrées, une couche de  $N_c$  neurones cachés et  $N_o$  neurones de sortie. [Amm, 07]

#### ✓ Les réseaux de neurones complètement connectés

Dans un réseau complètement connecté, les entrées puis les neurones (cachés et de sortie) sont numérotés, et, pour chaque neurone :

- Ses entrées sont toutes les entrées du réseau ainsi que les sorties des neurones de numéro inférieur.
- Sa sortie est connectée aux entrées de tous les neurones de numéro supérieur.

#### ✓ Les réseaux de neurones à couches

Dans une architecture de réseaux à couches, les neurones cachés sont organisés en couches, les neurones d'une même couche n'étant pas connectés entre eux.

De plus les connexions entre deux couches de neurones non consécutives sont éliminées.

Une telle architecture est historiquement très utilisée, surtout en raison de sa pertinence en classification.

#### IV.3.5.2. Les réseaux de neurones bouclés

L'architecture la plus générale pour un réseau de neurones est le « réseau bouclé », dont le graphe des connexions est cyclique : lorsqu'on se déplace dans le réseau en suivant le sens des connexions, il est possible de trouver au moins un chemin qui revient à son point de départ (un tel chemin est désigné sous le terme de « cycle »). La sortie d'un neurone du réseau peut donc être fonction d'elle-même; cela n'est évidemment concevable que si la notion de temps est explicitement prise en considération.

Ainsi, à chaque connexion d'un réseau de neurones bouclé (ou à chaque arête de son graphe) est attaché, outre un poids comme pour les réseaux non bouclés, un retard, multiple entier (éventuellement nul) de l'unité de temps choisie. Une grandeur, à un instant donné, ne pouvant pas



être fonction de sa propre valeur au même instant, tout cycle du graphe du réseau doit avoir un retard non nul.

Les connexions récurrentes ramènent l'information en arrière par rapport au sens de propagation défini dans un réseau multicouche. Ces connexions sont le plus souvent locales. Pour éliminer le problème de la détermination de l'état du réseau par bouclage, on introduit sur chaque connexion « en retour » un retard qui permet de conserver le mode de fonctionnement séquentiel du réseau (Fig 4.6).

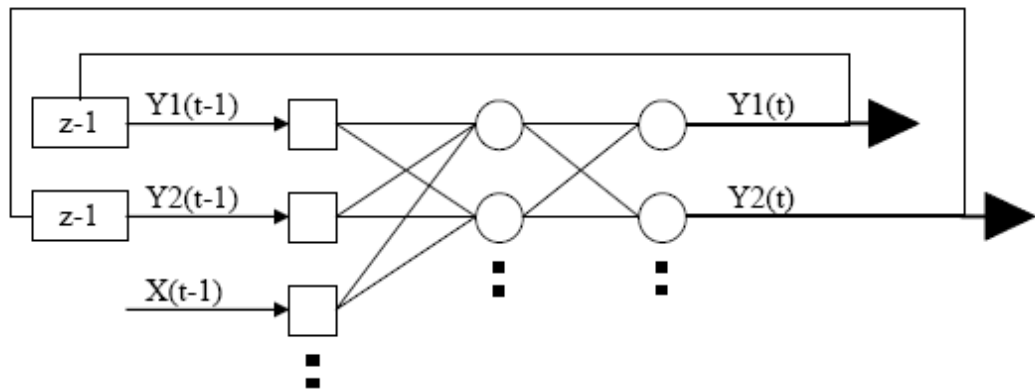


FIG 4.6. Réseau de neurone bouclé. [Amm, 07]

Le graphe des connexions de réseaux récurrents est cyclique. Ces réseaux sont décrits par un système d'équations aux différences.

#### IV.3.5.6. Apprentissage des réseaux de neurones

L'apprentissage d'un réseau de neurones artificiels est induit par une procédure itérative d'ajustement de ses paramètres internes (poids synaptiques et nombres de neurones). Cette procédure d'ajustement est décrite par un algorithme d'apprentissage. Celui-ci détermine alors le comportement du réseau. Ainsi le comportement d'un même réseau diffère selon l'algorithme d'apprentissage utilisé pour modifier ses paramètres. Il existe deux grandes stratégies d'apprentissage [Ben, 06].

Dans la majorité des algorithmes actuels, les variables modifiées pendant l'apprentissage sont les poids des connexions. L'apprentissage est la modification des poids du réseau dans l'optique d'accorder la réponse du réseau aux exemples et à l'expérience. Les poids sont initialisés avec des valeurs aléatoires. Puis des exemples expérimentaux représentatifs du fonctionnement du procédé dans un domaine donné, sont présentés au réseau de neurones. Ces exemples sont constitués de couples expérimentaux de vecteurs d'entrée et de sortie. Une méthode d'optimisation modifie les poids au fur et à mesure des itérations pendant lesquelles on présente la totalité des exemples, afin de minimiser l'écart entre les sorties calculées et les sorties expérimentales. Afin d'éviter les problèmes de surapprentissage, la base d'exemples est divisée en deux parties : la *base d'apprentissage* et la *base de test*. L'optimisation des poids se fait sur la base d'apprentissage, mais les poids retenus sont ceux pour lesquels l'erreur obtenue sur la base de test est la plus faible. En effet, si les poids sont optimisés sur tous les exemples de l'apprentissage, on obtient une précision très satisfaisante sur

ces exemples mais on risque de ne pas pouvoir généraliser le modèle à des données nouvelles. A partir d'un certain nombre d'itérations, le réseau ne cherche plus l'allure générale de la relation entre les entrées et les sorties du système, mais s'approche trop près des points et « apprend » le bruit [Pol et al., 92].

Sur la Fig 4.7 ci-dessous, on peut observer qu'au début de l'apprentissage, pour les premières itérations, l'erreur sur la base d'apprentissage est grande et peut légèrement augmenter étant donné que les poids initiaux sont choisis aléatoirement. Ensuite, cette erreur diminue avec le nombre d'itérations. L'erreur sur la base de test diminue puis augmente à partir d'un certain nombre d'itérations. Les poids retenus sont ceux qui minimisent l'erreur sur la base de test.

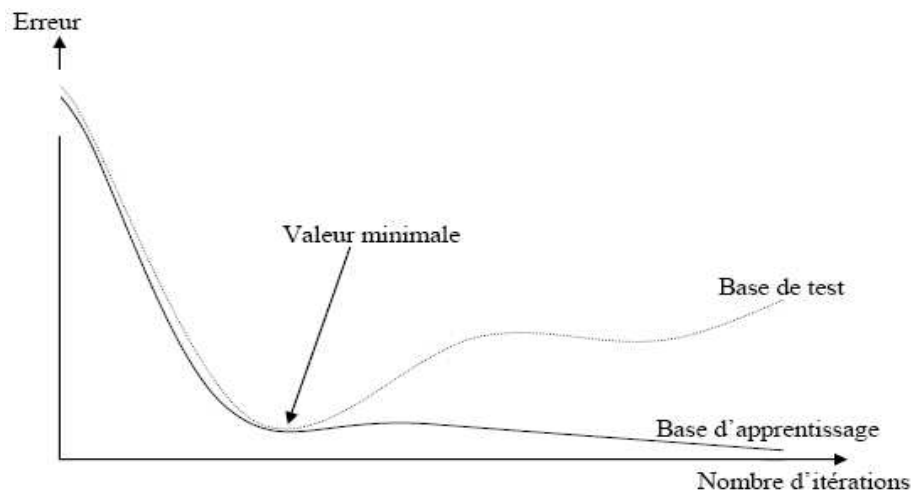


FIG 4.7. Erreur moyenne sur la base d'apprentissage en fonction du nombre d'itérations. [Amm, 07]

### 1. *Sur-apprentissage*

Il arrive qu'à faire apprendre un réseau de neurones toujours sur le même échantillon, celui-ci devient inapte à reconnaître autre chose que les éléments présents dans l'échantillon. Le réseau ne cherche plus l'allure générale de la relation entre les entrées et les sorties du système, mais cherche à reproduire les allures de l'échantillon. On parle alors de surapprentissage : le réseau est devenu trop spécialisé et ne généralise plus correctement.

Ce phénomène apparaît aussi lorsqu'on utilise trop d'unités cachées (de connexions), la phase d'apprentissage devient alors trop longue (trop de paramètres réglables dans le système) et les performances du réseau en généralisation deviennent médiocres.

#### IV.3.6.1. Type d'apprentissage

Il existe de nombreux types de règles d'apprentissage qui peuvent être regroupées en trois catégories [Has, 95] : les règles d'apprentissage *supervisé*, *non supervisé*, et *renforcé*. Mais l'objectif fondamental de l'apprentissage reste le même : soit la classification, l'approximation de fonction ou encore la prévision [Wei et Kul, 91].

### ***1. Apprentissage supervisé***

Un apprentissage est dit supervisé lorsque l'on force le réseau à converger vers un état final précis, en même temps qu'on lui présente un motif. Ce genre d'apprentissage est réalisé à l'aide d'une base d'apprentissage, constituée de plusieurs exemples de type entrées-sorties (les entrées du réseau et les sorties désirées ou encore les solutions souhaitées pour l'ensemble des sorties du réseau).

La procédure usuelle dans le cadre de la classification des battements cardiaques est l'apprentissage supervisé (ou à partir d'exemples) qui consiste à associer une classe spécifique désirée à chaque signal d'entrée. La modification des poids s'effectue progressivement jusqu'à ce que l'erreur (ou l'écart) entre les sorties du réseau (ou résultats calculés) et les résultats désirés soient minimisés.

Cet apprentissage n'est possible que si un large jeu de données est disponible et si les solutions sont connues pour les exemples de la base d'apprentissage.

### ***2. Apprentissage renforcé***

L'apprentissage renforcé est une technique similaire à l'apprentissage supervisé à la différence qu'au lieu de fournir des résultats désirés au réseau, on lui accorde plutôt un grade (ou score) qui est une mesure du degré de performance du réseau après quelques itérations. Les algorithmes utilisant la procédure d'apprentissage renforcé sont surtout utilisés dans le domaine des systèmes de contrôle [Whi et Sof, 92] et [Sut, 92].

### ***3. Apprentissage non supervisé***

L'apprentissage non supervisé consiste à ajuster les poids à partir d'un seul ensemble d'apprentissage formé uniquement de données. Aucun résultat désiré n'est fourni au réseau.

Qu'est-ce que le réseau apprend exactement dans ce cas ? L'apprentissage consiste à détecter les similarités et les différences dans l'ensemble d'apprentissage. Les poids et les sorties du réseau convergent, en théorie, vers les représentations qui capturent les régularités statistiques des données [Fuk, 88]; [Hin, 92]. Ce type d'apprentissage est également dit compétitif et (ou) coopératif [Gro, 88]. L'avantage de ce type d'apprentissage réside dans sa grande capacité d'adaptation reconnue comme une auto-organisation, « *self-organizing* » [Koh, 87]. L'apprentissage non supervisé est surtout utilisé pour le traitement du signal et l'analyse factorielle.

#### **IV.3.6.2. Algorithmes d'apprentissage**

L'algorithme d'apprentissage est la méthode mathématique qui va modifier les poids de connexions afin de converger vers une solution qui permettra au réseau d'accomplir la tâche désirée. L'apprentissage est une méthode d'identification paramétrique qui permet d'optimiser les valeurs des poids du réseau.

Plusieurs algorithmes itératifs peuvent être mis en œuvre, parmi lesquels on note : l'algorithme de rétropropagation (ce que nous allons détailler ci-après), la Méthode Quasi-Newton, La méthode de Newton, etc

#### IV.3.6.2. 1 Algorithme de rétropropagation

L'algorithme de rétropropagation ou de propagation arrière « *backpropagation* » est l'exemple d'apprentissage supervisé le plus utilisé à cause de l'écho médiatique de certaines applications spectaculaires telles que la démonstration de *Sejnowski* et *Rosenberg* (1987) dans laquelle l'ARP est utilisé dans un système qui apprend à lire un texte. Un autre succès fut la prédiction des cours du marché boursier [Ref et al, 94] et [Lee et al, 96] et plus récemment la détection de la fraude dans les opérations par cartes de crédit [Dor et al., 97].

La technique de rétropropagation du gradient (*Backpropagation* en anglais) est une méthode qui permet de calculer le gradient de l'erreur pour chaque neurone du réseau, de la dernière couche vers la première. L'historique des publications montre que l'ARP a été découvert indépendamment par différents auteurs mais sous différentes appellations [Gro, 98]. Le principe de la rétropropagation peut être décrit en trois étapes fondamentales : acheminement de l'information à travers le réseau; rétropropagation des sensibilités et calcul du gradient; ajustement des paramètres par la règle du gradient approximé.

### IV.4. Modélisation à l'aide de réseaux de neurones

Parmi les modèles les plus utilisés dans le domaine des réseaux de neurones est le suivant :

#### IV.4.1. Modèle « boîte noire »

Le terme de « boîte noire » s'oppose aux termes de « modèle de connaissance » ou « modèle de comportement interne » qui désignent un modèle mathématique établi à partir d'une analyse du processus que l'on étudie. Ce modèle peut contenir un nombre limité de paramètres ajustables, qui possèdent une signification physique.

Le modèle « boîte noire » (FIG 4.8) constitue la forme la plus primitive de modèle mathématique : il est réalisé uniquement à partir de données expérimentales ou d'observations ; il peut avoir une valeur prédictive, dans un certain domaine de validité, mais il n'a aucune valeur explicative. Ainsi, le modèle de l'univers selon *Ptolémée* était un modèle « boîte noire » : il ne donnait aucune explication de la marche des astres, mais il permettait de la prédire avec toute la précision souhaitable au regard des instruments de mesure disponibles à l'époque.



FIG 4.8. Diagramme schématique d'un modèle neuronal « boîte noire ». [Amm, 07]

## IV.5. Conception et mise en œuvre des réseaux de neurones

Les réseaux de neurones réalisent des fonctions non linéaires paramétrées. Leurs mises en œuvre nécessitent :

- La détermination des entrées et des sorties pertinentes, c'est à dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser (classification,...).
- La collecte des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones.
- La détermination du nombre de neurones cachés nécessaires pour obtenir une approximation satisfaisante.
- La réalisation de l'apprentissage.
- L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

### IV.5.1. Détermination des entrées/sorties du réseau de neurones

Pour toute conception de modèle, la sélection des entrées doit prendre en compte deux points essentiels :

- Premièrement, la dimension intrinsèque du vecteur des entrées doit être aussi petite que possible, en d'autre terme, la représentation des entrées doit être la plus compacte possible, tout en conservant pour l'essentiel la même quantité d'information, et en gardant à l'esprit que les différentes entrées doivent être indépendantes.
- En second lieu, toutes les informations présentées dans les entrées doivent être pertinentes pour la grandeur que l'on cherche à modéliser : elles doivent donc avoir une influence réelle sur la valeur de la sortie.

### IV.5.2. Choix et préparation des échantillons

Le processus d'élaboration d'un réseau de neurones commence toujours par le choix et la préparation des échantillons de données. La façon dont se présente l'échantillon conditionne le type de réseau, le nombre de cellules d'entrée, le nombre de cellules de sortie et la façon dont il faudra

mener l'apprentissage, les tests et la validation [Bis, 95]. Il faut donc déterminer les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.

Lorsque la grandeur que l'on veut modéliser dépend de nombreux facteurs, c'est-à-dire lorsque le modèle possède de nombreuses entrées, il n'est pas possible de réaliser un « pavage » régulier dans tout le domaine de variation des entrées : il faut donc trouver une méthode permettant de réaliser uniquement des expériences qui apportent une information significative pour l'apprentissage du modèle. Cet objectif peut être obtenu en mettant en œuvre un plan d'expériences. Pour les modèles linéaires, l'élaboration de plans d'expériences est bien maîtrisée, par ailleurs, ce n'est pas le cas pour les modèles non linéaires.

Afin de développer une application à base de réseaux de neurones, il est nécessaire de disposer de deux bases de données, une pour effectuer *l'apprentissage* et l'autre pour *tester* le réseau obtenu et déterminer ses performances.

Notons qu'il n'y a pas de règle pour déterminer ce partage d'une manière quantitative, néanmoins chaque base doit satisfaire aux contraintes de représentativité de chaque classe de données et doit généralement refléter la distribution réelle, c'est à dire la probabilité d'occurrence des diverses classes [Nas, 00].

### IV.5.3. Elaboration de la structure du réseau

La structure du réseau dépend étroitement du type des échantillons. Il faut d'abord choisir le type de réseau : un perceptron standard, un réseau de *Hopfield*, un réseau à décalage temporel, un réseau de *Kohonen*, un Perceptron Multi-Couches, etc.

Par exemple, dans le cas de ce dernier, il faudra aussi bien choisir le nombre de couches cachées que le nombre de neurones dans cette couche.

#### **a. Nombre de couches cachées**

Mis à part les couches d'entrée et de sortie, il faut décider du nombre de couches intermédiaires ou cachées. Sans couche cachée, le réseau n'offre que de faibles possibilités d'adaptation. Néanmoins, il a été démontré qu'un perceptron Multicouches avec une seule couche cachée pourvue d'un nombre suffisant de neurones, peut approximer n'importe quelle fonction avec la précision souhaitée [Hor, 91].

#### **b. Nombre de neurones cachés**

Chaque neurone peut prendre en compte des profils spécifiques de neurones d'entrée. Un nombre plus important permet donc de mieux "coller" aux données présentées mais diminue la capacité de généralisation du réseau. Il faut alors trouver le nombre adéquat de neurones cachés nécessaire pour obtenir une approximation satisfaisante.

Il n'existe pas, à ce jour, de résultat théorique permettant de prévoir le nombre de neurones cachés nécessaires pour obtenir une performance spécifique du modèle, compte tenu des modèles disponibles. Il faut donc nécessairement mettre en œuvre une procédure numérique de conception de modèle.

#### IV.5.4. Apprentissage

L'apprentissage est un problème numérique d'optimisation. Il consiste à calculer les pondérations optimales des différentes liaisons, en utilisant un échantillon.

La méthode la plus utilisée est la rétropropagation, qui est généralement plus économe que les autres en termes de nombres d'opérations arithmétiques à effectuer pour évaluer le gradient.

Pour rendre l'optimisation plus performante, on peut utiliser des méthodes de second ordre. Le calcul est très efficace, mais lourd. Elles ont de nombreuses limitations, quant aux conditions de convergence, sur les dérivées secondes. Des corrections sont proposées pour éviter ce problème, et sont prises en compte par les méthodes dites de Quasi-Newton ou de Newton modifiée.

#### IV.5.5. Validation et Tests

Alors que les tests concernent la vérification des performances d'un réseau de neurones hors échantillon et sa capacité de généralisation, la validation est parfois utilisée lors de l'apprentissage. Une fois le réseau de neurones développé, des tests s'imposent afin de vérifier la qualité des prévisions du modèle neuronal.

Cette dernière étape doit permettre d'estimer la qualité du réseau obtenu en lui présentant des exemples qui ne font pas partie de l'ensemble d'apprentissage. Une validation rigoureuse du modèle développé se traduit par une proportion importante de prédictions exactes sur l'ensemble de la validation.

Si les performances du réseau ne sont pas satisfaisantes, il faudra, soit modifier l'architecture du réseau, soit modifier la base d'apprentissage.

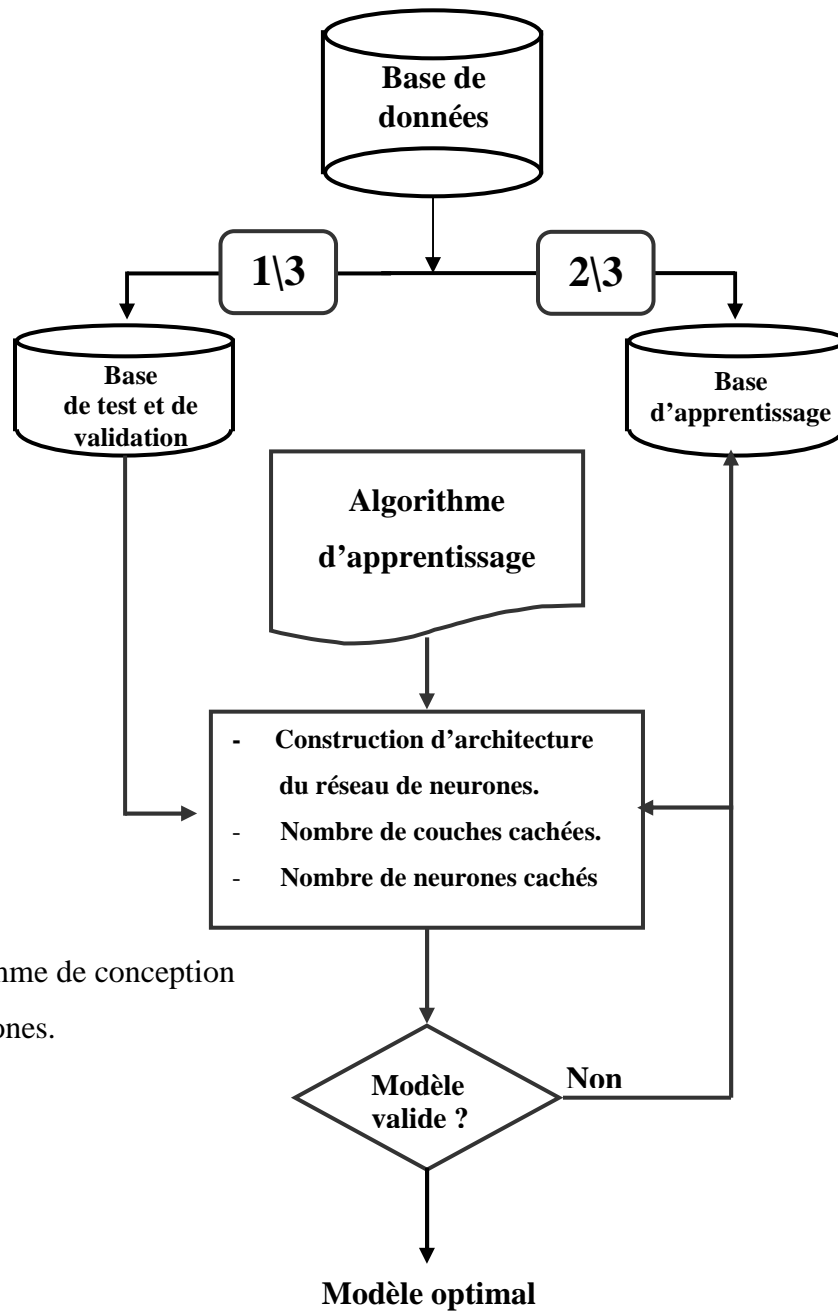


FIG 4.9 : Organigramme de conception d'un réseau de neurones.



## IV.6. Apprentissage par les réseaux multicouches

Un réseau de neurones multicouches (ou Perceptron Multi-Couches « **PMC** ») est composé d'une couche d'entrée, d'une ou plusieurs couches cachées et d'une couche de sortie (Fig 4.5).

### IV.6.1. Algorithme d'apprentissage

L'algorithme d'apprentissage par rétropropagation impose l'utilisation d'une fonction d'activation continue. Cet algorithme consiste à comparer la sortie obtenue et la sortie désirée, puis à propager la différence en sens inverse en modifiant les poids entre chaque couche. C'est un algorithme itératif, qui nécessite la présentation de l'ensemble d'apprentissage un grand nombre de fois. Nous présentons ci-dessous les différentes étapes de cet algorithme :

1. Initialisation des poids du réseau.
2. Présentation du vecteur d'apprentissage à l'entrée du réseau (couche d'entrées).
3. Calcul du vecteur de sortie  $S$  : le système propage les activités neuronales (état excité ou inhibé des neurones) à travers le réseau. Chaque neurone calcule la somme pondérée de ses entrées et transmet le résultat par une fonction de type *sigmoïde* pour produire sa valeur de sortie. Le vecteur de sortie  $S$  est le résultat du calcul de la dernière couche (couche de sortie).
4. Calcul de l'erreur : On compare alors les valeurs de sortie actuelles  $S$  avec les valeurs désirées de  $R$  appelé vecteur de référence (ou encore vecteur cible). On en déduit l'erreur  $J$ , définie comme la somme des carrés des différences observée sur chaque unité de la couche de sortie.
5. Rétropropagation de l'erreur : l'algorithme de rétropropagation du gradient permet récursivement de rétropropager l'erreur de la couche de sortie vers les couches cachées jusqu'à la première couche du réseau.
6. Modification des poids : les poids de chaque neurone sont modifiés, soit à chaque présentation d'un vecteur d'apprentissage (méthode du gradient stochastique), soit après cumul de l'erreur pour un certain nombre de vecteurs d'apprentissage (méthode du gradient standard).

Il y a principalement deux facteurs qui influent sur l'apprentissage. Ce sont la qualité de l'échantillonnage d'apprentissage (les exemples qui constituent la base d'apprentissage) et la diversité des valeurs. En effet, le réseau de neurones généralisera mieux si la qualité de l'échantillonnage est meilleure et si les données des exemples d'apprentissage sont variées. Intuitivement, on est conscient que s'il sait répondre correctement pour un nombre fini de situations les plus diverses, il sera alors plus proche de ce que l'on veut dans une situation nouvelle.

### IV.6.1.1. Algorithme de la rétropropagation du gradient

L'algorithme va être donné dans sa version la plus générale, c'est à dire avec plusieurs couches cachées. On notera  $f$  la fonction d'activation. Une démonstration de l'efficacité de cet algorithme exige de la fonction d'activation qu'elle soit indéfiniment dérivable. On notera  $S_i$  la sortie du neurone  $i$  de la couche de sortie et  $y_i$  la sortie attendue pour ce même neurone. Enfin, pour des neurones d'une couche cachée, on notera  $O_i$  la sortie calculée du neurone.

**Entrée** : un exemple, sous la forme (vecteur\_x,vecteur\_y);  
 Epsilon le taux d'apprentissage  
 Un Perceptron Multicouches avec q-1 couches cachées  $C_1, \dots, C_{q-1}$ , une couche de sortie  $C_q$ .

**Répéter**

Prendre un exemple (vecteur\_x,vecteur\_y) et calculer  $f(\text{vecteur\_x})$

Pour toute cellule de sortie  $i$   $d_i \leftarrow S_i(1 - S_i)(y_i - S_i)$  **finPour**

Pour chaque couche de q-1 à 1

**Pour** chaque cellule  $i$  de la couche courante

$d_i = O_i(1 - O_i) * \text{Somme [pour } k \text{ appartenant aux}$   
 indices des neurones prenant en entrée la  
 sortie du neurone }] de  $d_k * w_{ki}$

**FinPour**

**FinPour**

**Pour** tout poids  $w_{ij} \leftarrow w_{ij} + \text{Epsilon} * d_i * x_{ij}$  **finPour**

**FinRépéter**

La variable " $d_i$ " apparaît deux fois dans le code. Il s'agit de deux variables différentes, car en fait on suppose que les neurones sont numérotés de sorte que l'on puisse associer à un identifiant à un neurone et réciproquement. Par conséquent, le 'i' de " $d_i$ " identifie un neurone et ainsi on peut effectuer la dernière boucle de manière uniforme sans différencier pour la couche de sortie et les couches cachées.

La méthode donne de bons résultats pratiques. Dans la plupart des cas, on rencontre peu de problèmes dus aux minima locaux quand va l'expliquée ci-après, mais il y en a. Toutefois, il est moins performant que d'autres algorithmes de propagation d'erreur : il tend moins rapidement vers des poids plus ou moins optimaux.

### IV.6.2. Problème de l'apprentissage

Un grand nombre de choix s'imposent à celui qui veut entraîner un réseau neuromométique. Chaque choix compte plus d'un écueil potentiel. Voici quelques exemples :

- a) **Présence de minima locaux** : la majorité des apprentissages supervisés procèdent par une minimisation de l'erreur par l'application d'une méthode telle que la descente de gradient. Ce type

de technique peut conduire à une solution sous optimale. En fait, ce problème n'est pas résolu : on ne sait jamais si c'est un minimum local ou le minimum global qui est atteint.

- b) **Mauvais choix de l'architecture** : le comportement des réseaux est gouverné par un ensemble de paramètres architecturaux. On peut citer le nombre de neurones cachés du réseau, le pas d'apprentissage,...etc.
- c) **Surapprentissage** : la courbe C1 de la figure (FIG 4.10) représente l'évolution de l'erreur quadratique sur la base d'apprentissage en fonction du nombre d'itérations. Le but de l'apprentissage du réseau n'est pas simplement d'apprendre une représentation exacte de la base de données d'apprentissage, mais plutôt d'élaborer un modèle statistique capable de régénérer l'ensemble des données potentiellement observables. La généralisation du réseau est donc une fonction très importante et évite que le réseau apprenne parfaitement les exemples de la base d'apprentissage. Afin d'éviter le surapprentissage, plusieurs méthodes ont été proposées. La plus simple utilise trois bases de données distinctes : une base d'apprentissage, une de test et une dite de "validation". Cette dernière base est utilisée pendant l'apprentissage afin d'examiner le comportement du réseau pour des données qui lui sont inconnues. Ainsi, l'apprentissage est arrêté lorsque l'erreur atteint un minimum avec la base de validation (voir courbe C2 (FIG 4.10)).

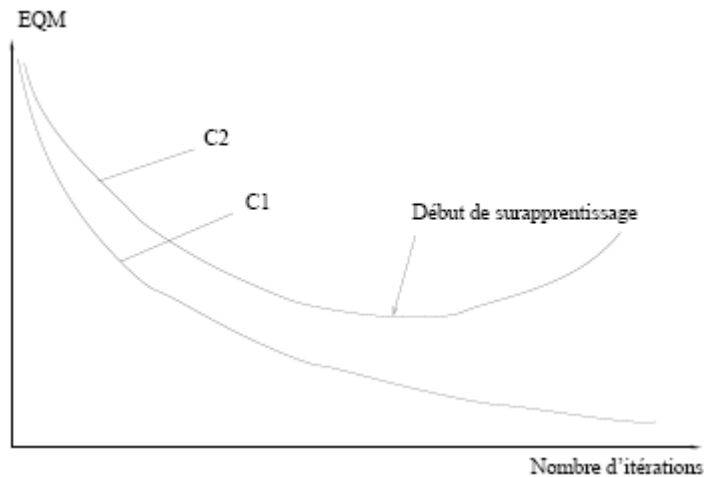


FIG 4.10. Le phénomène de sur-apprentissage d'un RN

- d) **Mauvais codage** : on entend généralement par codage la transformation d'information d'entrée (formes) et de sortie (classes) en des vecteurs de scalaires. Le codage intermédiaire entre l'environnement et le réseau de neurones a une importance capitale : un bon codage doit mettre en évidence, parmi les informations disponibles, celles qui favorisent la résolution du problème. Au contraire, un codage inapproprié peut occulter des informations essentielles et ainsi compromettre les résultats [Tea, 94].

- e) Les détails algorithmiques entourant la réalisation et l'ajustement des réseaux de neurones, ainsi que le perfectionnement de leur capacité d'apprentissage et de généralisation constituent pratiquement un champ d'investigation scientifique en soi. En particulier, pour pallier la difficulté du choix d'une architecture et des paramètres initiaux, l'une des principales méthodes consiste à tester en parallèle plusieurs réseaux de configurations différentes.

Par ailleurs pour accélérer la convergence de l'algorithme de rétropropagation, des méthodes ont été proposées, divers travaux proposent d'introduire du bruit pendant la phase d'apprentissage, soit au niveau de l'ensemble d'apprentissage [Hol, 92], soit dans la règle d'apprentissage [Hay, 96], [Jim, 96], [Ham, 97] et [Hol, 00]. D'autres encore, toujours en vue d'accélérer la convergence du réseau, ont mené des études portant sur le choix de l'initialisation du modèle, [Pat, 95] et [Thi, 97].

Pour le dimensionnement des réseaux, plusieurs travaux concernant le choix optimal de réseaux ont été publiés [Fog, 91], [Hor, 91] et [Hér, 94].

En général, bien que très intéressants, ces travaux restent toujours de portée limitée pour des raisons très variées. En fait, les auteurs ont souvent traité des problèmes différents à l'aide de réseaux définis en fonction du problème, dont la structure reste figée tout au long de l'apprentissage.

Parallèlement à ces études sur les problèmes généraux et sur l'amélioration des capacités d'algorithmes d'apprentissage [Ste, 96], [Thi, 96], [Psi, 99] d'autres recherches ont été menées. L'objectif premier de celles-ci a été de découvrir d'autres modèles de réseaux du point de vue des architectures et des techniques d'apprentissage.

On peut citer par exemple, les travaux de [Fin, 03] qui a utilisé la régression linéaire multiple pour développer un prédicteur pour arrêter l'apprentissage du RN.

## IV.7. Apprentissage : point de vue de l'IAD

Si en psychologie, comprendre l'individu équivaut à étudier les représentations des structures abstraites construites lors de la réalisation d'une tâche cognitive, en Intelligence Artificielle (IA), la modélisation des fonctions de haut niveau de l'esprit humain (raisonnement, langage, apprentissage,...) est mise en avant. En effet, pour l'IA, comprendre l'esprit correspond à en construire un modèle à différents niveaux de granularités, quitte à en oublier parfois toute la complexité humaine.

D'ailleurs, pour tenter de développer au mieux un parallèle entre les états mentaux humains et les représentations symboliques fournies par un ordinateur, certains chercheurs redoublent d'imagination et vont même jusqu'à expliquer la conscience humaine en la modélisant à l'aide d'une multitude d'agents. Cette idée de faire travailler conjointement un certain nombre d'agents autonomes à la résolution de problèmes complexes à

donner naissance à une branche de l'IA, l'Intelligence Artificielle Distribuée (IAD), dont l'objectif est de concevoir des agents artificiels capables de s'organiser efficacement pour accomplir collectivement les tâches qui leur sont demandées. Les paramètres d'imprécisions que pourrait introduire une modélisation mathématique sont davantage pris en compte avec une représentation de type multi-agents, l'apprentissage apparaît même comme un véritable atout pour déterminer un comportement optimal qui répond à la plupart des problématiques et situations qu'un agent peut rencontrer. Dans ce cas précis, l'apprentissage se fonde sur l'observation des autres agents ou de l'environnement et la communication n'est finalement qu'un outil permettant une meilleure coopération entre tous.

#### IV.8. Avantages et inconvénients des réseaux de neurones

On peut donc définir un réseau de neurones artificiels comme un processeur parallèle de traitement d'informations distribuées présentant une propension naturelle à la mémorisation et à l'exploitation de connaissances acquises par l'expérience. Sa structure repose sur une interconnexion massive de cellules élémentaires de traitement d'information, les neurones formels, dont la représentation est un graphe dirigé. En cours d'apprentissage, un réseau peut être amené à modifier sa structure en modifiant les liaisons entre neurones.

La connaissance acquise par un réseau de neurones est encodée par des forces ou des intensités évolutives des connexions établies entre les neurones formels. La force des connexions liant les neurones réfèrent aux valeurs des poids ou coefficients qui définissent les paramètres du modèle interne du réseau. L'intérêt des réseaux de neurones artificiels réside dans le parallélisme de leur structure, leur capacité d'adaptation ainsi que dans leur mémoire distribuée. Ces propriétés sont à la source de la capacité de généralisation d'un réseau et donc de son aptitude à adopter un comportement correct en réponse à des variables d'entrée qui n'ont pas été rencontrées lors de la phase d'apprentissage.

Vus sur un plan mathématique, les réseaux de neurones sont en effet des « approximateurs universels ». (Théorème de Kolmogorov, 1957). Leur utilisation permet de passer directement des données au prédicteur, sans intermédiaire, sans recodage, sans discrétisation, sans simplification ou interprétation sujette à caution « Boîte noire ». Un réseau de neurones possède également une grande résistance au bruit ou au manque de fiabilité des données. L'idée d'apprentissage est plus simple à comprendre que les complexités des statistiques multi-variables. Elle est intuitive. Les réseaux de neurones sont également capables d'analyser des relations spatiales et topologiques. Enfin, le modèle fourni se prête bien à la visualisation, à une recherche de profils-type. Ces propriétés des réseaux de neurones sont intéressantes car elles sont complémentaires des techniques statistiques classiques qui imposent des fortes restrictions. Les régressions se font sur des fonctions de dépendance simples (linéaire, logarithmique) qui ne sont pas toujours très réalistes.

## IV.9. Domaines d'application des RN

Se trouvant dans les domaines où ils nécessitent un calcul vraiment considérable de point de vue complexité, donc ils sont utilisés pour :

- *la régression non linéaire, ou modélisation de données statiques.* Une immense variété de phénomènes statiques peut être caractérisée par une relation déterministe entre des causes et des effets ; les réseaux de neurones sont de bons candidats pour modéliser de telles relations à partir d'observations expérimentales, sous réserve que celles-ci soient suffisamment nombreuses et représentatives.
- *la modélisation de processus dynamiques non linéaires.* Modéliser un processus consiste à déterminer un ensemble d'équations mathématiques qui décrivent le comportement dynamique du processus, i.e. l'évolution de ses sorties en fonction de l'évolution de ses entrées. Ce problème peut être avantageusement résolu par un RN, si le phénomène que l'on désire modéliser est non linéaire.
- *la commande de processus.* Commander un processus signifie imposer un comportement défini à l'avance en fonction de signaux de commande ; l'ensemble commande / processus peut donc être considéré comme un système qui réalise une fonction (non linéaire) que un RN peut approcher.
- *la classification.* En raison de leur propriété d'approximateurs universels, les RN sont capables d'estimer de manière précise la probabilité d'appartenance d'un objet inconnu à une classe parmi plusieurs possibles.

Maintenant, on va vous présenter la partie du système expert, sachant que cette dernière va nous aider, à implémenter le comportement de l'agent d'adaptation en le dotant par un système expert, afin de reproduire le raisonnement, suivant la base des faits et des règles qu'on a.

Le premier système expert fut DENDRAL en 1965. Il permettait d'identifier les constituants chimiques d'un matériau, mais ses règles étaient mélangées au moteur. Le plus connu, peut-être, fut MYCIN en 1972-73, système expert de diagnostic de maladies du sang, avec un vrai moteur et une vraie base de règles. Cependant ses règles étaient affectées de coefficients de vraisemblance qui donnaient à chacune d'entre elles un poids particulier face aux autres. Le moteur produisait un chaînage avant simple tout en calculant les probabilités de chaque déduction, ce qui le rendait incapable d'expliquer la logique de son fonctionnement et de détecter les contradictions. Quant aux experts, ils étaient obligés de trouver des coefficients de vraisemblance pour chacune des conclusions de leurs règles, une démarche compliquée et antinaturelle qui déniait leur capacité de raisonnement.

## IV.10. SYSTEME EXPERT

Un système expert est un logiciel qui reproduit le comportement d'un expert humain accomplissant une tâche intellectuelle dans un domaine précis.

Les systèmes expert sont nés d'une idée simple : rendre les ordinateurs intelligents en leur fournissant la connaissance des experts. Le fondement du concept système expert réside dans la nature explicite de la connaissance. En effet, celle-ci n'est pas absolument pas noyée dans un algorithme, elle est formalisée par un ensemble de règles qui vont être enchaînées pour simuler le raisonnement d'un expert humain.

### IV.10.1. A quoi ça sert ?

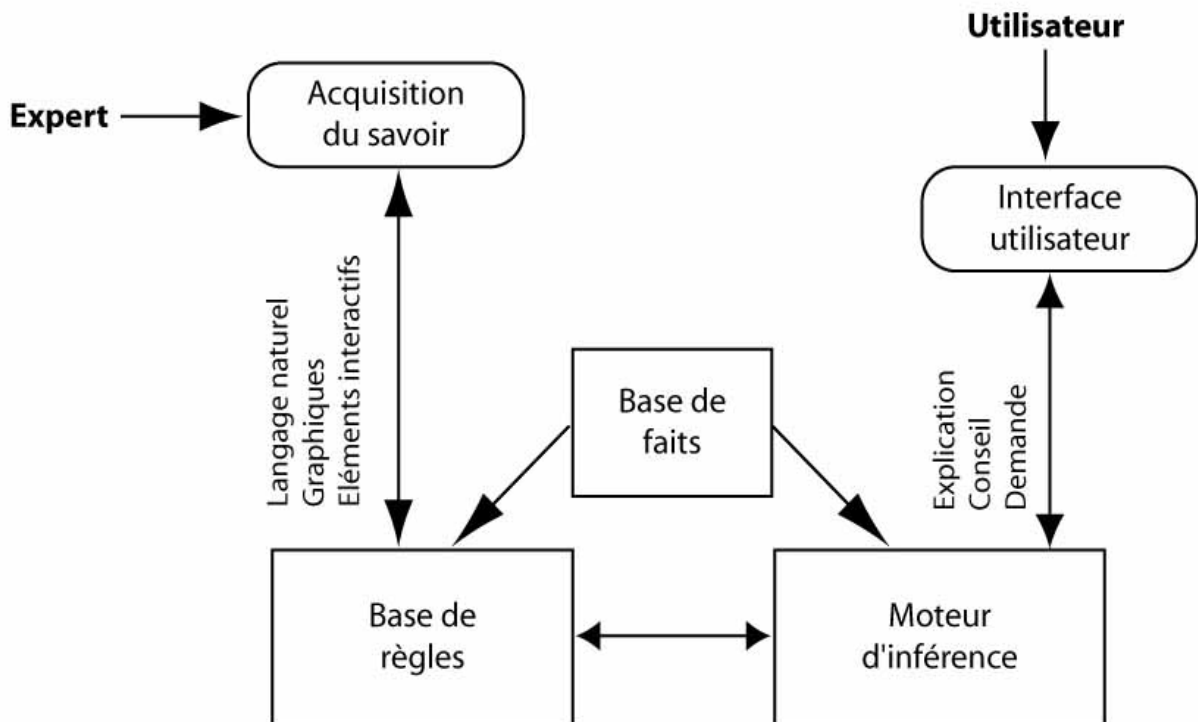
Les systèmes experts ne sont utiles et envisageables que pour des domaines possédant des experts humains. Ces derniers sont des individus connaissant bien le domaine à modéliser et qui sont aussi capable de transmettre leur savoir, ce dernier point est primordial.

Les systèmes experts sont utilisés lorsque les méthodes algorithmiques classiques ne peuvent pas être appliquées avec succès. Ils sont généralement conçus pour résoudre des problèmes de classification ou décision (diagnostic médical, prescription thérapeutique, régulation d'échanges boursiers, etc.).

### IV.10.2. Composantes des systèmes experts

Un système expert peut se composer en deux composantes principales :

1. **La base de connaissance**, elle-même composée d'une **base de règle** qui modélise la connaissance du domaine considéré et d'une **base de fait**.
2. **Le moteur d'inférences** dont le rôle est de raisonner à partir des données contenues dans la base de connaissances.



**FIG 4.11** : Composantes des systèmes experts

L'élément clé d'un système expert réside dans l'indépendance qui existe entre la base de connaissances et le moteur d'inférences. Elle permet une représentation des connaissances sous forme purement déclarative et explicite, c'est-à-dire sans lien avec la manière dont ces connaissances sont utilisées. L'avantage de cette architecture est qu'il est possible de faire évoluer les connaissances du système sans avoir à agir sur le mécanisme de raisonnement.

#### **IV.10.2.1. Base de connaissances**

La base de connaissances d'un système expert se décompose en deux parties. La première partie contient des faits spécifiques du domaine, les connaissances factuelles. On parle alors de base de faits.

La seconde partie, la base de règles, contient les règles qui vont permettre au système de raisonner à partir des faits. Ces connaissances déductives sont représentées par des règles appelées : règles de production.

##### **IV.10.2.1.1. Base de règles**

Comme nous venant de le voir, la base des règles rassemble la connaissance et le savoir-faire de l'expert. Elle n'évolue donc pas au cours d'une session de travail.



Une règle se présente sous la forme : Si X alors Y, X est appelé prémisses et Y conclusion. Alors que X est une conjonction de conditions, c'est-à-dire une suite de comparaisons d'attributs et de valeurs à l'aide d'opérateurs, la conclusion est une affectation, Fait = Valeur.

Grâce à ce type de représentation proche du langage naturel, la connaissance du système sera facilement accessible à l'utilisateur qui pourra ainsi aisément la modifier ou l'agrandir.

De même que dans un système expert, la connaissance est traduite en règles, la méta-connaissance s'exprime par des métarègles (des règles sur la manière d'utiliser les règles).

L'organisation d'une base de connaissances au moyen de métarègles reste essentiellement déclarative, contrairement à toute organisation basée sur une structuration a priori de l'ensemble des règles (écrire les règles dans un ordre donné, etc.).

#### IV.10.2.1.2. Base de faits

La base de faits constitue en quelque sorte la mémoire de travail du système expert. Elle est variable au cours d'exécution et vidée lors de l'exécution se termine. Au début de session, elle contient ce que l'on sait du cas examiné avant toute intervention du moteur d'inférences. Puis elle est complétée par les faits déduits par le moteur ou demandés à l'utilisateur.

Les faits peuvent prendre des formes plus ou moins complexes. Les faits élémentaires peuvent avoir des valeurs :

- Booléennes : Vrai, faux
- Symboliques : Valeurs appartenant à un ensemble fini de symboles.
- Réelles : Valeurs appartenant à un ensemble continu

Un système expert qui n'utilise que des faits booléens est dit d'ordre zéro (0), un système expert qui utilise des faits symboliques ou réels, sans utiliser de variables, est d'ordre 0+. Un système utilisant toute la puissance de la logique du premier ordre est d'ordre 1.

#### IV.10.3. Moteur d'inférences

Pour exploiter cette connaissance, un moteur d'inférences est nécessaire pour relier la description d'un problème aux capacités d'analyse d'une situation donnée. De façon générale, le moteur d'inférences sera capable de répondre à des questions, de raisonner et de tirer les conséquences impliquées par la connaissance incluse dans le système.

Le moteur d'inférences va enchaîner les règles c.-à-d. qu'il va effectuer un chaînage. Il existe plusieurs types de ce dernier, c'est-à-dire plusieurs types de raisonnements :

- Chaînage avant, qui permet de déduire les faits découlant des données initiales.
- Chaînage arrière, qui permet d'atteindre des buts.
- Chaînage mixte, qui est un mélange des deux types de chaînages précédents.

#### IV.10.3.1. Chaînage avant

Le mécanisme de chaînage avant est simple : pour déduire un fait particulier, on déclenche les règles dont les prémisses sont connues jusqu'à ce que le fait à déduire soit également connu ou qu'aucune règle ne puisse être déclenchée.

On va analyser chaque fait ou on va examiner toutes les règles où ce fait apparaît en prémisses. Pour les règles déclenchées, on va affecter les attributs en conclusion des valeurs qui leur correspondent. On dira que les faits ont été propagés, ces attributs affectés feront partie du résultat final de l'expertise ; et, en même temps, ils seront eux-mêmes propagés, on fait cela jusqu'à l'épuisement des faits, et on communique les résultats à l'utilisateur.

#### IV.10.3.2. Chaînage arrière

Le mécanisme de ce type de chaînage consiste à partir du fait que l'on souhaite établir, à rechercher toutes les règles qui concluent sur ce fait, à établir la liste des faits qu'il suffit de prouver pour qu'elles puissent se déclencher puis à appliquer récursivement le même mécanisme aux faits contenus dans cette liste.

Le chaînage arrière est clairement un mécanisme d'induction : on vérifie les hypothèses en remontant depuis les objectifs. On cherche aussi si un fait est possible : celui-ci étant choisi selon des critères extérieurs.

#### IV.10.3.3. Chaînage mixte

Le chaînage mixte utilise les deux chaînages présentés ci-dessus. On peut alors aussi bien raisonner à partir des faits que l'on connaît comme prédicats ou comme objectifs.

Ce sont les caractéristiques qui vont conditionner le chaînage qu'il est judicieux d'utiliser. Ainsi, lorsque les faits sont peu nombreux ou que le but est inconnu, il est préférable d'employer un *chaînage avant*, par contre, dans le cas où les buts sont peu nombreux ou précis, le *chaînage arrière* est préconisé [Gil, 02].

## IV.11. Conclusion

Au début de la cybernétique, de l'automatique et de l'informatique, l'Intelligence Artificielle (IA) classique et connexionniste se sont développés en parallèle.

Après l'article de Minsky et Papert en 1969 démontrant les faiblesses du perceptron, les années 70 ont vu le désintérêt pour le modèle connexionniste et le succès de l'IA classique (cognitive) (manipulant des symboles logiques - comme les systèmes experts).

Mais cette approche a aussi ses limites. D'abord sa lenteur : on est vite confronté à ce que l'on appelle l'explosion combinatoire. La faiblesse est aussi théorique. La connaissance ne se réduit pas au langage. L'intelligence n'est pas qu'un raisonnement mathématique manipulant ces symboles. Une grande partie des experts travaillent de manière intuitive, à partir de la longue expérience qu'ils ont acquise. On a alors beaucoup de mal à faire ressortir les raisonnements qu'ils utilisent.

De plus, d'autres facultés humaines sont mal réalisées par l'IA classique, comme l'adaptation, la généralisation.

Les réseaux neuromimétiques, bien qu'ils ne soient pas non plus une panacée, ont aussi leur mot à dire : en particulier quand il n'existe pas de technologie ou d'algorithme classique mais des bases d'exemples, quand le système étudié est bruité, évolue dans le temps, etc.

En 1982 Hopfield a remis à la mode l'approche connexionniste. Aujourd'hui, les nouvelles technologies du multimédia comme le traitement de l'image et du son demandent une rapidité de plus en plus grande. Les ordinateurs séquentiels s'approchent de leur limite et l'on cherche à paralléliser les microprocesseurs et à trouver les algorithmes pouvant les utiliser à pleine puissance. Or les réseaux de neurones artificiels permettent, par leur conception même, cette parallélisation : ce sont un ensemble d'entités élémentaires travaillant en simultané.

Bien que l'on soit encore très loin des neuro-ordinateurs, d'ores et déjà les ordinateurs actuels pourraient s'équiper de modules spécialisés (cartes son, vidéo) s'inspirant de la technologie des réseaux neuromimétiques. Les liens des réseaux de neurones avec les algorithmes génétiques (pour la modification des poids synaptique, par exemple), la théorie du chaos (modélisation d'un attracteur étrange), le besoin d'algorithmes parallèles, et l'association avec l'IA cognitive permettent d'augurer du développement de la technologie neuromimétique. Bref, il faudra compter avec les réseaux de neurones.

Dans le chapitre suivant, on va vous introduire la partie simulation, qui consiste à étudier les méthodes introduites dans ce mémoire, à l'occurrence du PMC "*Perceptron Multi-Couches*" et les systèmes expert pour la classification des arythmies cardiaque « Autres » et « N » qui sont respectivement (Autres battements anormales tels que : Bloc de branche droit, Bloc de branche gauche, Battement ventriculaire prématuré et dernièrement battement Normal) -pour mieux comprendre ces notions, veuillez s'adresser au premier chapitre, section VIII- ; suivant deux (2) types de descripteurs : le premier, c'est un descripteur basé sur des

caractéristiques temporaires et le deuxième basé sur des caractéristiques morphologiques.

Les deux classes résultantes, seront entrées comme des faits dans le système expert servant à l'opération d'adaptation, qui consiste à donner la classe finale (C'est-à-dire, la fusion des deux résultats en une seule).

Notons que chaque perceptron ou système expert ce n'est qu'un comportement ou – *behaviour* - d'un agent intelligent de classification ou d'adaptation donné.

Dans le chapitre qui suit, nous allons détailler notre contribution, en donnant les étapes que nous avons suivies pour développer les classifieurs neurologiques (temporels et morphologiques) et le système expert pour faire la classification de battements cardiaques, et comment intégrer les agents intelligents dans ce processus.



## V. CONTRIBUTION ET REALISATION

### V.1. Introduction

Ce chapitre a pour objectif de présenter le travail réalisé pour donner une solution de la problématique posée –Une approche multi-agent pour la classification distribuée par descripteurs morphologiques et temporels des arythmies cardiaques en utilisant les réseaux de neurones et les systèmes experts.

Nous allons donc, voir comment développer un réseau de neurones artificiels (ex, PMC perceptron Multi-Couches) de l'étape de sélection de la base d'apprentissage et de test, sélection des descripteurs pertinents temporels ou morphologiques d'un cycle cardiaque étudié et l'architecture des réseaux de neurones (Algorithme d'apprentissage, dimension de ce réseau) et bien sûr discuter les résultats expérimentaux des classifieurs.

Nous allons expliquer le rôle du système expert réalisé, détailler les agents intelligents développés au sein d'une plateforme choisie, rôle de chaque agent, comment attacher les classifieurs réalisés avec ces agents, enfin, nous présenterons un scénario d'exécution de l'application distribuée réalisée.

### V.2 Classification neuronale des arythmies cardiaques

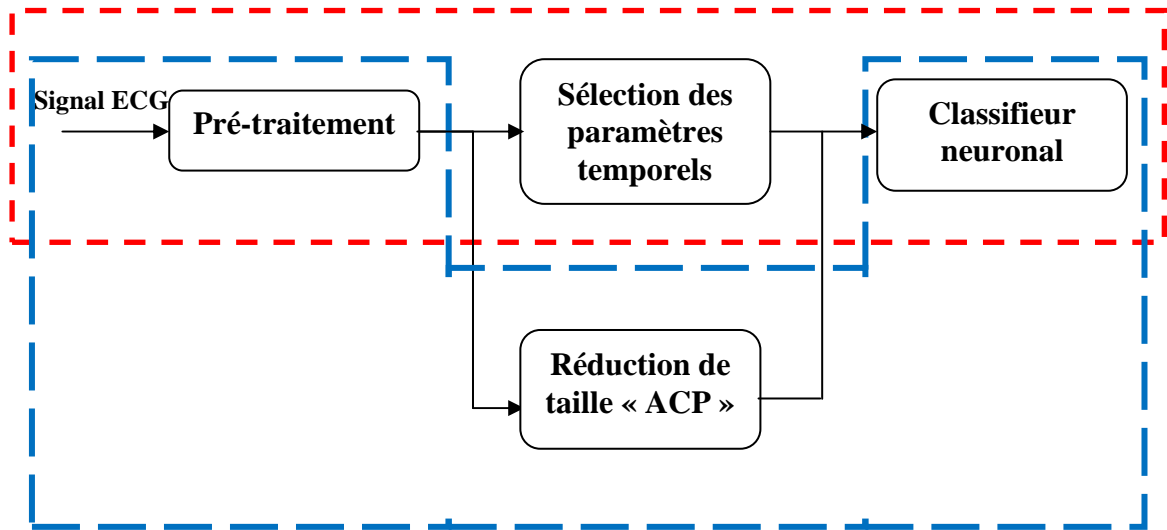
Comme tout modèle de traitement, les classifieurs utilisés en reconnaissance des formes ou d'objets sont amenés à effectuer des tâches complexes (notamment lors de l'apprentissage), à partir d'une base de données réelles.

Pour que le classifieur soit opérationnel en phase d'exploitation, il est primordial que la nature des données et notamment leurs variations soit correctement prises en compte. Le problème se profile à plusieurs niveaux, comportant chacun son lot de difficultés théoriques et pratiques. Il faut disposer premièrement d'un système d'acquisition efficace, qui nous fournira des données de qualité (avec un minimum de bruit). Il faut ensuite pouvoir représenter cette information (possiblement la réduire et la transformer) de manière efficace et compatible avec les étapes de traitement ultérieures. Cette représentation et ses modalités constituent le modèle avec lequel le mécanisme de reconnaissance devra travailler, afin de produire le résultat désiré.

Dans certains domaines tels que le médical, le principal obstacle auquel est confronté le concepteur lors de la mise en œuvre d'un système de classification est la caractérisation des données par des paramètres fiables et pertinents.

L'analyse de la littérature présentée dans le chapitre des modèles de classification montre que, lorsqu'une approche de reconnaissance des arythmies cardiaques est adoptée, les réseaux de neurones artificiels sont une bonne technique de classification.

Pour notre étude, le réseau de neurones utilisé est optimisé par des algorithmes de type supervisé. La stratégie que nous envisageons en œuvre est présentée dans la **FIG 5.0**.



**FIG 5.0** : Schéma global de la classification

Les étapes : « Pré-traitement », « Sélection des paramètres temporels » et « Classifieur neuronal » font partie de la classification neuronale en basant sur des caractéristiques temporelles extraites depuis le signal ECG de l'entrée, puisqu'après le prétraitement (débruitage,...), on passe directement vers l'extraction des paramètres temporelles telles que : les périodes RR0, RRs, PP, le complexe QRS et le rapport RRs/RR0.

Par contre, les étapes : « Pré-traitement », « Réduction de taille « ACP » » et « Classifieur neuronal » font partie de la classification neuronale en basant sur la morphologie du signal ECG en question. Puisque ce dernier contient trop de données (c'est une base de données volumineuse), ce qui nécessite une réduction de taille, donc, il faut avoir une méthode d'analyse numérique qui a les performances nécessaires en termes de minimisation de perte de données et d'optimisation de réduction de taille (compression de données), ACP répond à nos besoins et critères posés.

Notons que vous trouverez les détails nécessaires de la méthode d'Analyse en Composantes Principales « ACP » dans « l'**annexe B** » dans la fin de ce manuscrite.

Et comme nous l'avons précédemment mentionné dans le chapitre IV, dans le paragraphe IV.5, la mise en œuvre de chaque réseau de neurones nécessite :

- 1- La détermination des entrées et des sorties pertinentes, c'est à dire les grandeurs qui ont une influence significative sur le phénomène que l'on cherche à modéliser.
- 2- La collecte des données nécessaires à l'apprentissage et à l'évaluation des performances du réseau de neurones c'est-à-dire la sélection de base d'exemple et base de test ou de validation.
- 3- La détermination du nombre de neurones cachés nécessaires pour obtenir une approximation satisfaisante.

4- La réalisation de l'apprentissage.

4- L'évaluation des performances du réseau de neurones à l'issue de l'apprentissage.

### V.2.1. Sélection de la base d'exemples

Nous avons utilisé, dans le cadre de ce mémoire, la base de données MIT-BIH. Cette dernière se compose de quarante huit enregistrements (deux voies ECG ambulatoire d'environ trente minutes), dont les instants d'apparition et la classe de tous les cycles sont annotés par deux cardiologues différents (pour plus d'information, voir **annexe B**). Les enregistrements comprenant suffisamment de battements normaux ont été choisis pour construire la base d'exemple, voir **tableau 5.1**.

Nous avons rassemblé les battements de chaque enregistrement en deux groupes (classes):

1. Le groupe N est composé des battements différents des « Normal ».
2. Le groupe A est composé des battements autres que normal.

Il est important de remarquer que tous les enregistrements choisis pour faire l'expérimentation sont des battements normaux, mais ils contiennent des cycles cardiaques anormaux.

Toutes ces informations sont récapitulées dans le tableau 5.1 suivant :

Enregistrement	Nombre de battements de type « N »	Nombre de battements de type « A »
100	2239	34
101	1860	05
103	2082	02
105	2526	46
112	2537	02
116	2302	110
203	2529	511
205	2571	85
209	2621	384
210	2423	205
213	2641	610
215	3195	168
230	2255	01

**Tab 5.1 :** Les enregistrements choisis de la base de données MIT-BIH.

La base d'apprentissage affecte la performance des classifieurs neuronaux. Le critère général à respecter est de sélectionner des exemples représentatifs de toutes les classes. Dans notre cas, nous avons préparé deux bases, une d'apprentissage et l'autre de validation ou de test.

Notre expérimentation est composée de 60 cycles de Type N (battement normal) et 36 cycles de Type « A » (mélange entre les BVP, BBDR, BBG), qui vaut 96 cycles pour chaque base choisie parmi les 13 enregistrements sélectionnés.



### V.2.1.1. Sélection des descripteurs temporaires d'un cycle cardiaque

Pour obtenir une meilleure performance d'un classifieur, il vous faut choisir des données pertinentes permettant de décrire de manière assez complète l'échantillon à classer. Ces données sont appelées caractéristiques ou descripteurs.

Dans notre cas, en plus le complexe QRS, des descripteurs permettant de caractériser un battement cardiaque sont retenus. Ils sont comparables aux paramètres sur lesquels sont construites les connaissances médicales.

#### a) Intervalle $RR_p$

Nous appelons  $RR_p$  la distance entre le pic R du présent battement et le pic R du battement précédent. Nous avons divisé la valeur de la distance  $RR_p$  par la moyenne des distances des huit intervalles RR précédents. Ce rapport est au voisinage de la valeur 1 pour un battement Normal et un peu plus petit que 1 pour un battement de type « A ».

#### b) Rapport des intervalles RR

Nous appelons  $RR_s$  la distance entre le pic R du présent battement et le pic R du battement suivant. Dans le cas d'un rythme régulier, le rapport  $RR_s/RR_p$  est un paramètre qui caractérise une classe donnée. Dans le cas d'un rythme régulier « N », ce rapport est voisin de 1, mais il peut largement dépasser cette valeur dans le cas « A ».

#### c) Largeur du complexe QRS

Ce paramètre est important pour l'identification des battements cardiaques, la mesure de ce paramètre a été réalisée en utilisant un algorithme implémenté sous matlab, au laboratoire L.T.S.I à l'université de Rennes 1.

#### d) Distance PP entre deux points minimum et maximum du QRS

Ce paramètre a été aussi retenu comme valeur informationnelle potentiellement pertinente pour chaque cycle cardiaque, sachant que cette distance varie énormément avec la déformation du complexe QRS.

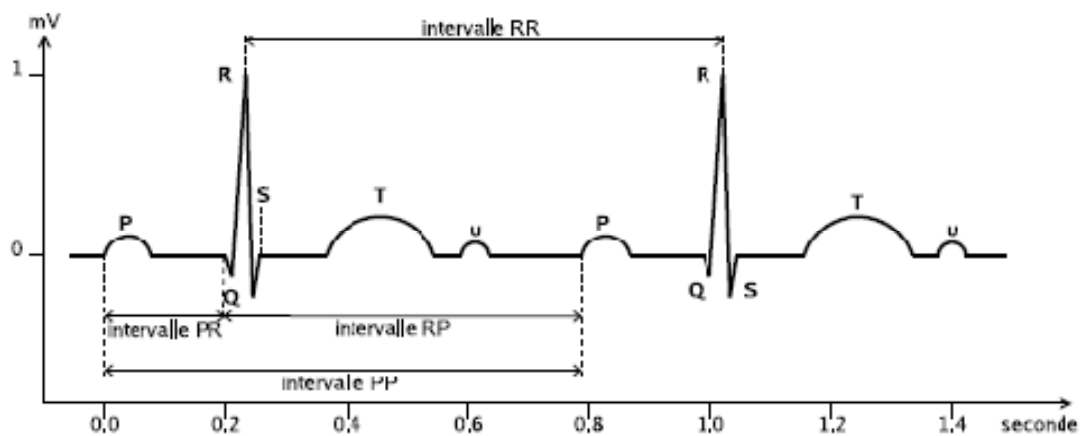


FIG 5.1 : Battements normaux et intervalles d'intérêt.

## V.2.2. Architecture du réseau de neurones

### V.2.2.1. Algorithme d'apprentissage

On a développé nos réseaux de neurones (Par des descripteurs morphologiques et d'autres temporels) sous *Matlab version R2010a*. Au cours de ce travail nous avons utilisé principalement trois algorithmes très classiques en optimisation non linéaire.

1. Quasi-Newton.
2. Levenberg-Marquardt.
3. Méthode du gradient à pas.

En les confrontant sur plusieurs problèmes, il apparait qu'aucune de ces trois méthodes ne prend un avantage considérable sur l'autre. Chaque méthode d'optimisation possède des avantages et des inconvénients. La méthode de Levenberg-Marquardt présente un intérêt pratique car elle peut être utilisée sans avoir à choisir le pas [Che, 05].

Bien entendu, aucune méthode ne conduit à coup sûr au minimum global. Il convient donc de se placer dans des conditions où les minima locaux sont aussi peu nombreux que possible ou correspondent à une erreur suffisamment faible [Che 05].

### V.2.2.2. Dimensionnement du réseau pour chaque classifieur

La structure neuronale est déterminée par le nombre de neurones de la couche d'entrée,  $N_e$  et celui de la couche cachée,  $N_c$ , tous les deux sont variables.

- Couche d'entrée : le nombre de neurones sur cette couche est lié directement au vecteur de caractéristiques :
  - Le classifieur temporaire, le nombre de neurones dans la couche d'entrée est 96 neurones, puisque, nous avons choisi des descripteurs donnant un vecteur de taille est égale à ce nombre (96).
  - Le classifieur morphologique, le nombre de neurones dans la couche d'entrée est lié fortement au nombre de composants principales retenues et qui contribuent à plus de 99% à l'inertie totale du nuage. Le but, est d'avoir une perte de données minimale (moins de 01%).
- Couche cachée : le choix de la taille de la couche cachée est réalisée souvent de façon heuristique et expérimentale, car aucune formule théorique ne peut nous aider à l'optimiser. Donc, nous avons mené une série d'essais pour arrêter le nombre  $N_c$  à mettre sur la couche cachée afin d'avoir un meilleur résultat possible. On a constaté que plus le nombre de neurones cachés est important, plus le nombre d'exemples bien classés est grand, car le réseau, grâce au nombre de paramètres croissant, se spécialise sur la base d'apprentissage. Les fonctions d'activation des neurones cachés sont des sigmoïdes.

- Le nombre d'itérations, fixé dans ces deux essais à 30000, a permis de réaliser un apprentissage satisfaisant. Par ailleurs, il est clair que ce dernier paramètre est important : s'il est trop grand, le réseau risque de réaliser un sur-apprentissage ; au contraire, s'il est trop petit, l'apprentissage ne sera pas conduit à son terme.

### V.2.2.2.1 Réduction par l'ACP

Nous avons appliqué la méthode d'analyse en composantes principales sur la base d'apprentissage et sur des enregistrements de la base de données MIT-BIH., pour voir la variabilité des données par rapports au nombre de composantes principales. En effet, comme nous pouvons le constater dans la figure FIGAA, (La base d'apprentissage), la première composante n'exprime pas plus de 45% de la variabilité totale des données, alors que les 15 premières composantes expriment 99% de la variabilité.

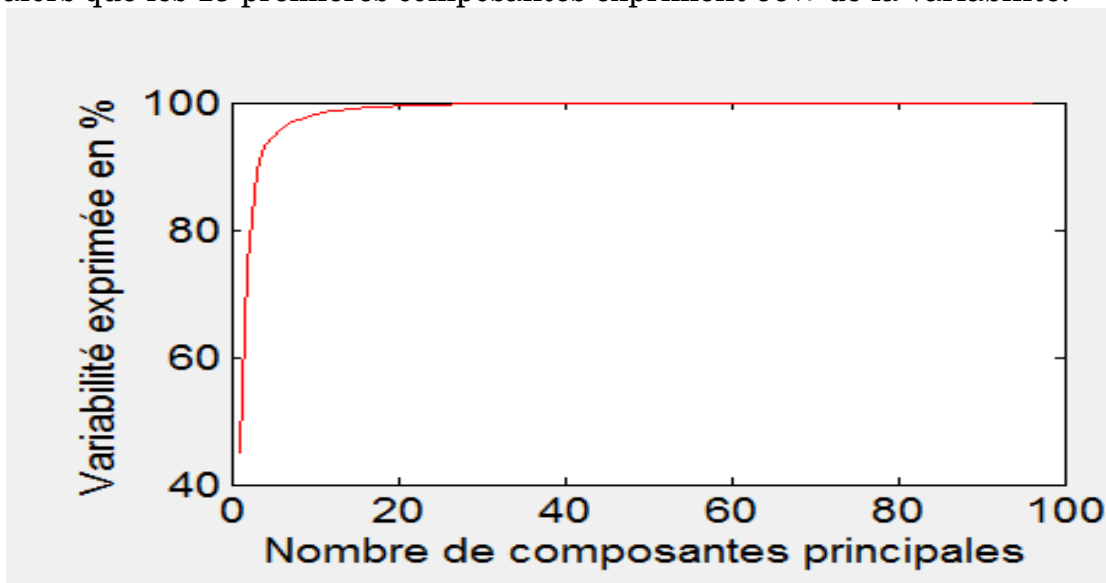


FIG 5.2 :variabilité des données de la base d'apprentissage.

Tandis que pour l'enregistrement 209, la première composante principale n'exprime pas plus de 44% et les 08 premières composantes représentent presque 100% de la variabilité.

Donc, nous remarquons clairement que pour chaque enregistrement, sa morphologie peut être représentée par uniquement une dizaine de paramètres représentatifs. Cette réduction de données permet d'avoir une architecture optimisée des classifieurs d'arythmies.

Le tableau 5.2 récapitule le nombre de composantes principales présentant la variabilité des données, pour chaque enregistrement, toujours en respectant, la contrainte de moins 01% de perte de données.

Échantillon testé	Nbre de composantes principales retenues
Base d'apprentissage	<b>15</b>
ECG 100	<b>16</b>
ECG 101	<b>09</b>
ECG 103	<b>08</b>

ECG 105	13
ECG 112	10
ECG 116	08
ECG 122	16
ECG 205	08
ECG 209	19
ECG 210	18
ECG 213	08
ECG 220	08
ECG 230	09
	<b>Max → 19</b>

**Tab 5.2 :** Nombre de composantes principales retenues pour chaque ECG.

### V.2.3. Résultats expérimentaux des classifications

Dans ce qui suit, on présente de façon résumée les résultats des deux classifieurs obtenus. Les performances sont représentées respectivement par le taux de classification  $\tau_{class}$ , la sensibilité  $S_e$  et la spécificité  $S_p$  déjà définis dans le chapitre des modèles de classification.

Les tableaux suivants (Tab 5.3 et Tab 5.4) récapitulent les résultats des classifications faites sur deux bases d'apprentissages, le premier est une base contenant les descripteurs temporels et l'autre contient les vecteurs propres représentant la base initiale suivant le degré de perte de données.

Enregistrement	$\tau_{class}$ (%)	$S_e$ (%)	$S_p$ (%)
100	92.44	97.12	98.34
101	89.14	96.22	97.33
103	91.87	90.29	99.45
105	93.34	88.45	93.34
112	75.34	87.17	93.40
116	89.77	90.55	96.66
203	95.02	97.33	99.35
205	97,67	79,88	87,77
209	96,77	96,99	90,89
210	92,59	98,77	89,07
213	96,34	90,02	99,9
215	98,56	97,09	100
230	97,55	91,30	95,45
Moyenne	92.40	92.49	96.83

**Tab 5.3 :** Performances du classifieur neuronal à données temporelles.

Enregistrement	$T_{class}$	$S_e$	$S_p$
100	91.24	95.10	92.66
101	93.15	97.21	99.12
103	97.17	93.59	95.25
105	91.20	86.33	90.34
112	85.31	92.30	94.41

116	79.58	91.05	93.76
203	93.41	92.19	95.05
205	99.40	96.22	89.14
209	93.06	96.29	91.90
210	93.42	89.45	93.12
213	96.64	88.17	75.67
215	87,17	94.55	89.33
230	90,81	97.73	95.72
Moy	91.72	92.70	91.64

**Tab 5.4** : Performances du classifieur neuronal à données morphologiques réduites.

### V.3 Développement des agents intelligents

Dans la suite de ce chapitre, nous appellerons : '**Agent1**', l'agent intelligent qui s'occupe de la classification des battements cardiaques doté par un PMC basé sur les caractéristiques temporelles, '**Agent2**', l'agent intelligent qui s'occupe de la classification des battements cardiaques doté par un PMC basé sur les caractéristiques morphologiques compressées à l'aide de la méthode d'analyse en composantes principales (ACP) et dernièrement, '**Agent3**', l'agent qui a comme tâche, le couplage des deux résultats envoyés par l'Agent1 et l'Agent2, par un système expert fonctionnant en chaînage avant et donne à son tour la classe finale de l'ECG en question.

Nous avons choisi la plateforme **Jade** (*Java Agent DEveloppement framework*), car elle supporte la décentralisation et la distribution des tâches, elle est open source, accompagnée par une bonne documentation en plus, elle contient des outils de gestion et de contrôle continu des agents du SMA.

Les comportements des deux agents de classification (**Agent1** et **Agent2**) sont les deux réseaux de neurones qu'on a expliqués précédemment, donc, les entrées de ces réseaux de neurones seront considérées comme entrées de ces agents, et par la suite, les résultats fournis par ces deux derniers, seront considérés comme des faits entrés pour l'Agent3, qui va à son tour donner la classe finale.

Donc, voici l'architecture distribuée de notre application, présentée dans la figure 5.3,

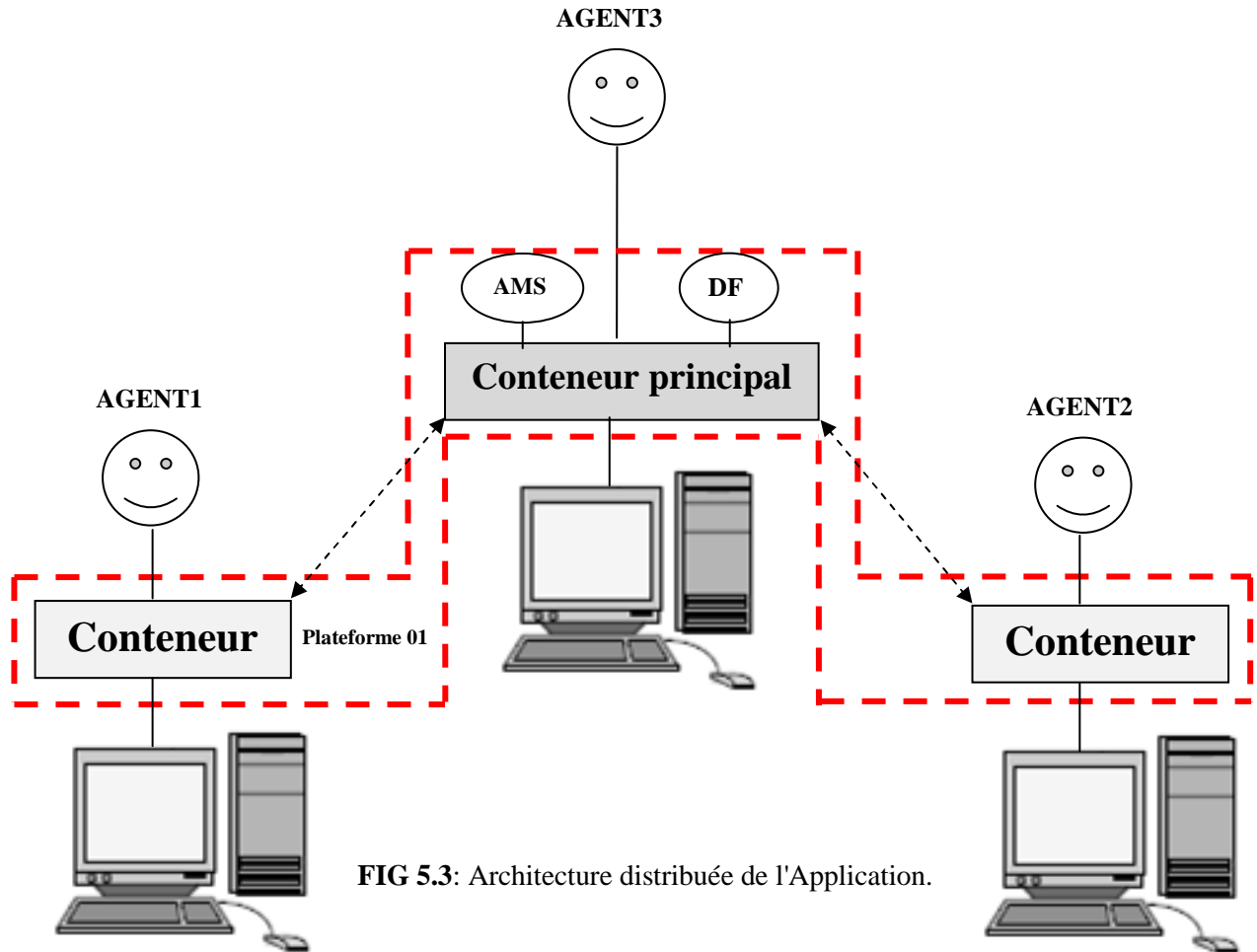


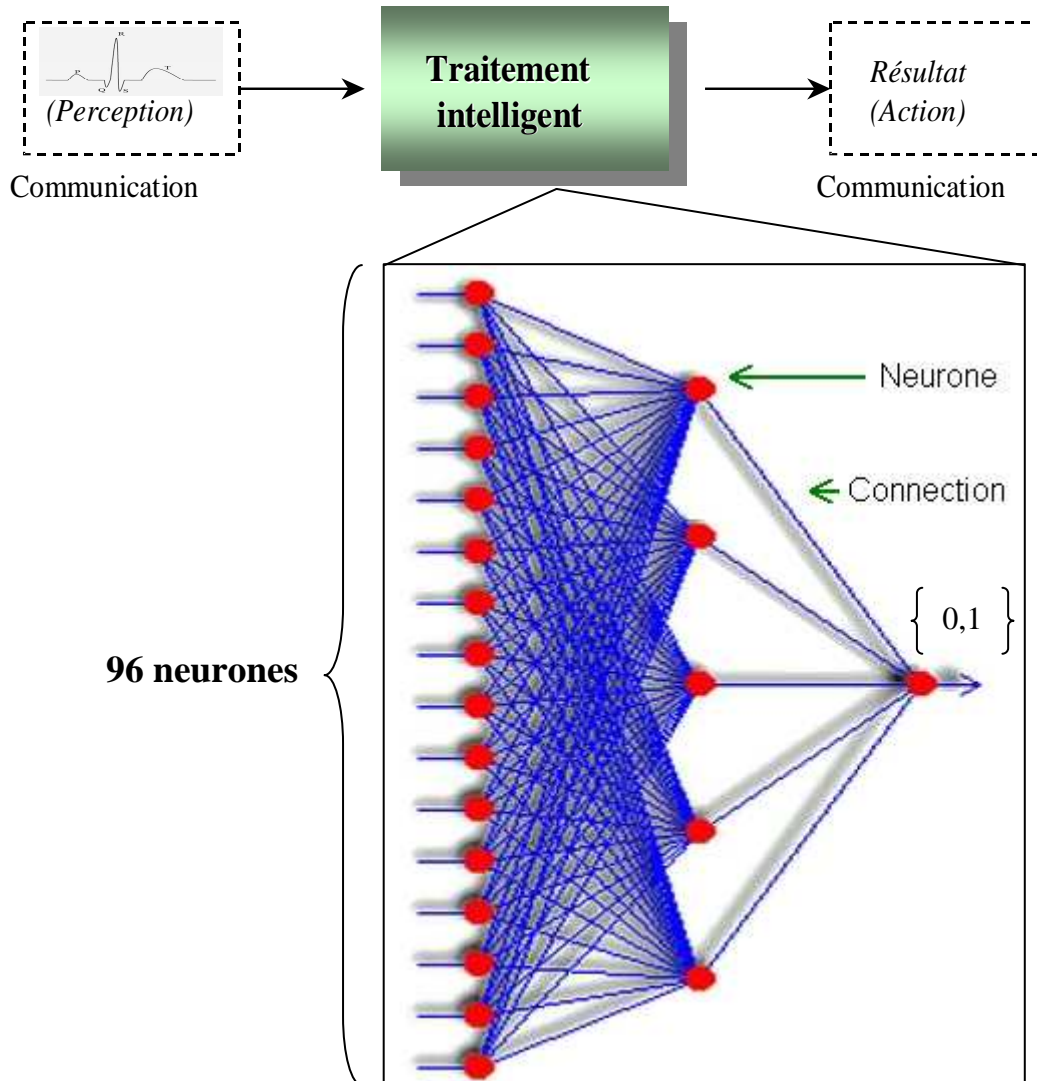
FIG 5.3: Architecture distribuée de l'Application.

Dans la figure précédente, nous avons deux conteneurs 1 et 2 qui sont secondaires et qui sont enregistrés par le conteneur principal de la plateforme Jade, chaque conteneur vit physiquement dans un pc à part, mais logiquement, ils vivent ensemble dans le Framework Jade. Ce dernier fournit des outils tels que l'AMS (Agent Management System) pour la gestion des agents de la plateforme, le DF (Directory Facilitator) qui joue le rôle d'un annuaire, service page jaune.

Jade nous fournit aussi, un langage de communication s'appelle Fipa-Acl, qui est propre aux agents qui vivent dans une plateforme Jade, ce qui est un point très essentiel pour les agents intelligents.

### V.3.1 Architecture des agents de classification

#### a) Agent classifieur à base descripteurs temporels « AGENT1 »



**FIG 5.4:** Architecture de l'agent classifieur AGENT1

La figure 5.4 représente l'architecture minimale de l'agent classifieur à base caractéristiques temporelles, il reçoit comme entrées, le vecteur (96 entrées) contenant les descripteurs temporels tels que : les périodes RR0, RRs, PP, le complexe QRS et le rapport RRs/RR0. Après la réception des données, il fait son apprentissage neuronal supervisé afin de donner la classe (N qui vaut ZERO pour normal ou A qui vaut UN pour autres) de l'ECG en entrée. En fin, il envoie ce résultat vers l'agent d'adaptation (Agent expert) par un message ACL.

## b) Agent classifieur à base descripteurs morphologiques « AGENT2 »

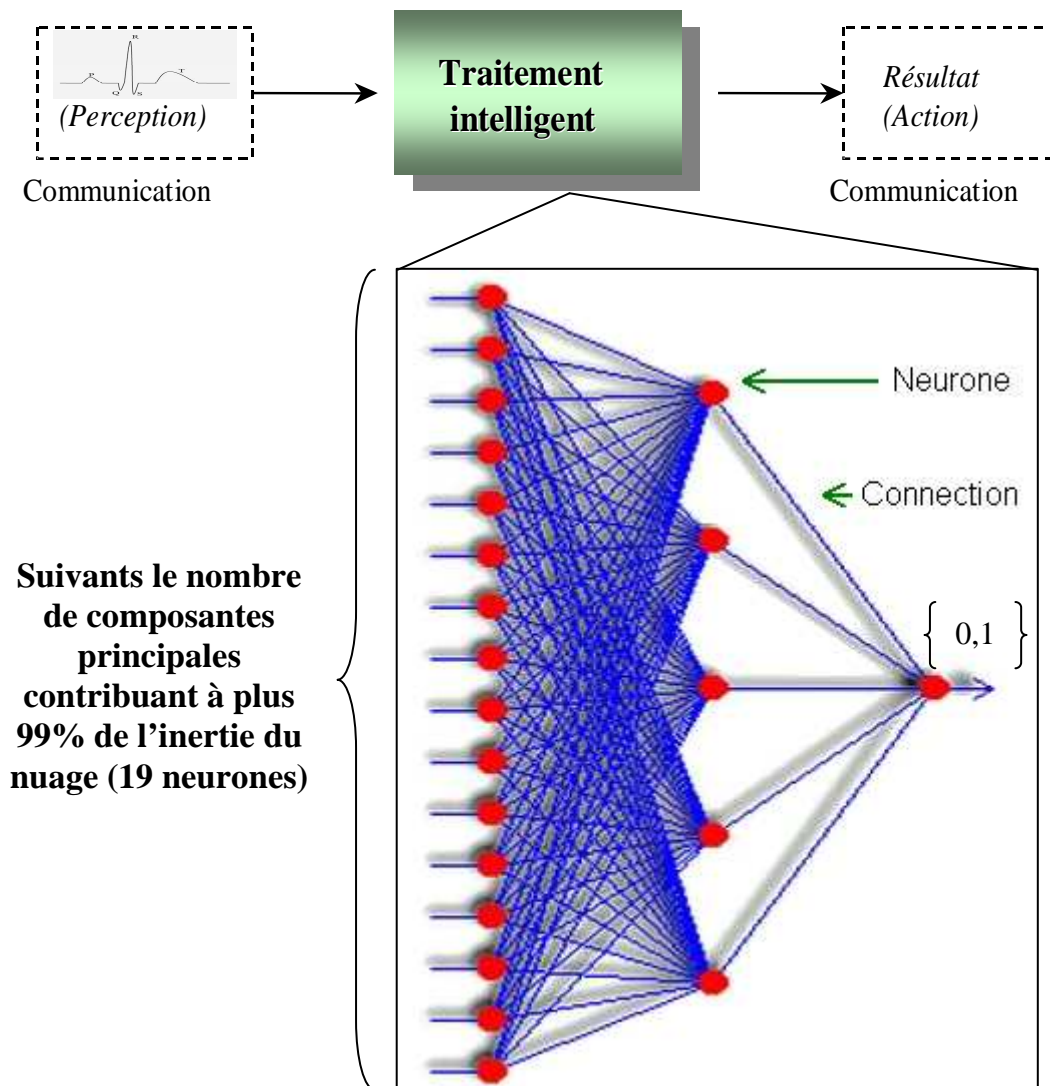
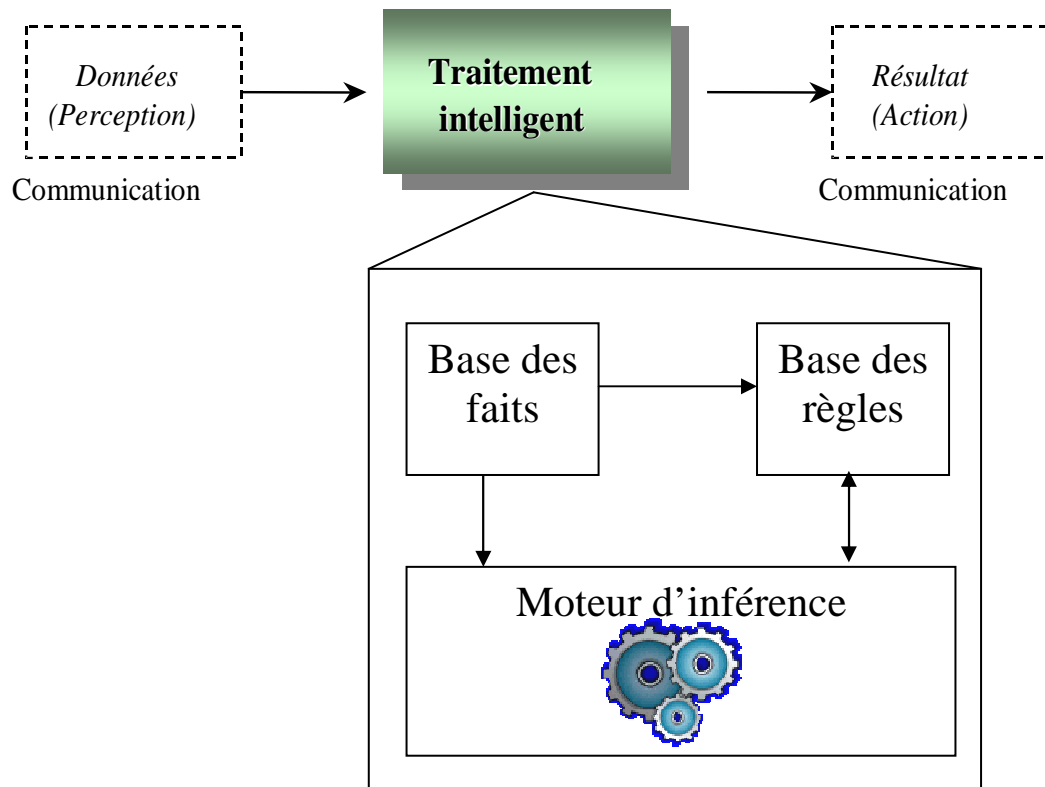


FIG 5.5: Architecture de l'agent classifieur AGENT2

La figure 5.5, illustre l'architecture minimale de l'agent classifieur à base caractéristiques morphologiques, il reçoit comme entrée un vecteur (19 entrées) contenant les composantes principales représentant 99% de l'inertie totale pour chaque enregistrement. Il fait son apprentissage neuronal supervisé, qui donne comme résultat le type de classe (N qui vaut ZERO pour normal ou A qui vaut UN pour autres) à laquelle il appartient cet ECG. Finalement, il envoie ce résultat vers l'agent d'adaptation (Agent expert) par un message ACL.



**Agent d'adaptation « AGENT3 »****FIG 5.6 :** Architecture de l'agent d'adaptation AGENT3

La figure 5.6, représente l'architecture de l'agent d'adaptation qui reçoit les deux messages envoyés auprès des Agent1 et Agent2, contenant les deux résultats d'apprentissage supervisé, ces deux derniers sont les faits qui vont déclencher une inférence par un chaînage avant par le system expert de cet agent, le résultat de l'apprentissage sera la classe finale de l'ECG.

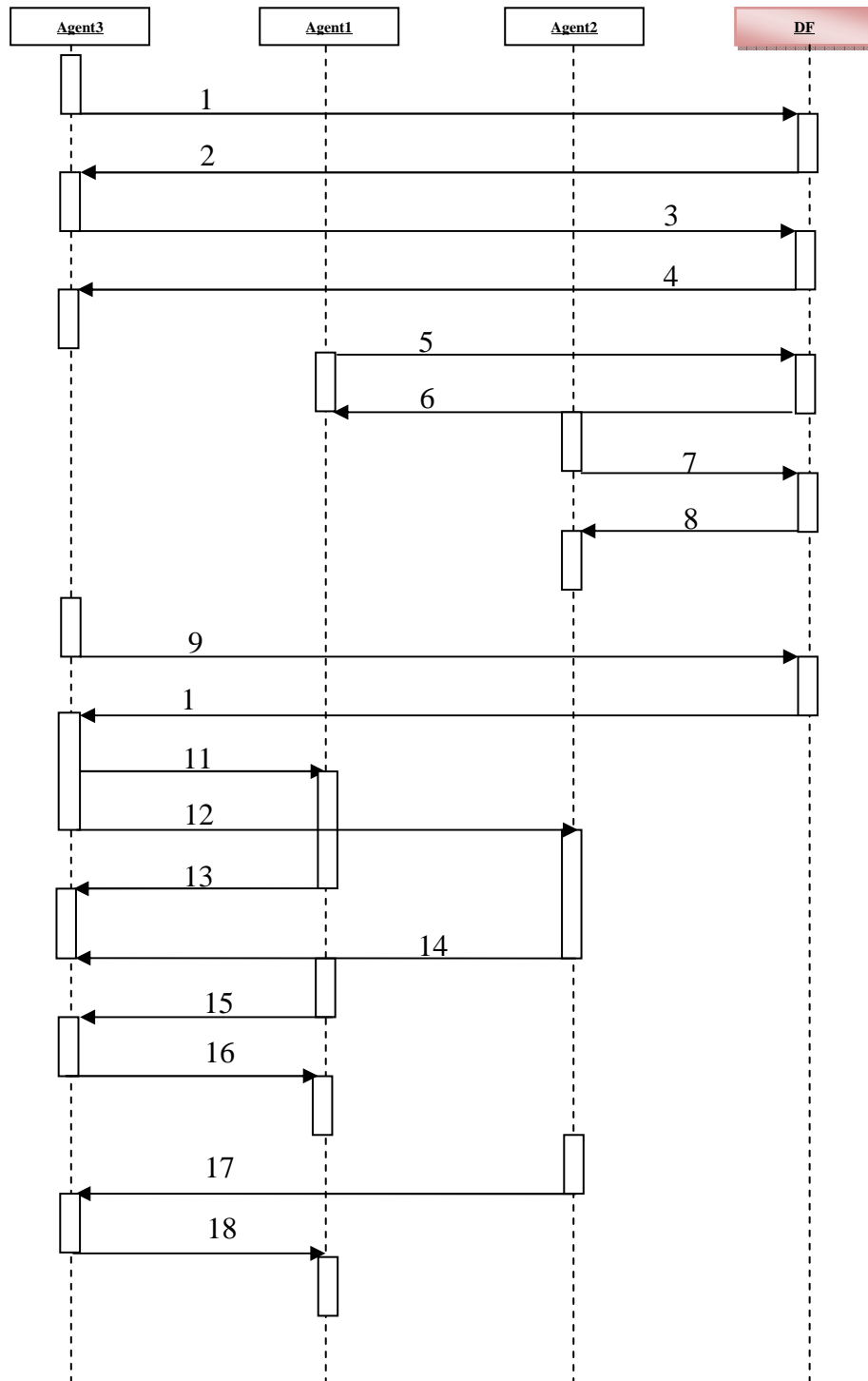
Par ce fait, nous avons fusionné les deux techniques (morphologique et temporel) d'interprétation des ECG, en utilisant les méthodes de l'intelligence artificielle IA, et l'approche multi agents fournie par l'intelligence artificielle distribuée.

### V.3.2 Communication entre les agents du système

Le diagramme suivant présente la communication séquentielle entre les agents du système de monitoring médical, afin de classer les battements cardiaques et détecter l'existence des maladies cardiovasculaires.

La figure 9 représente les interactions entre les agents du système pendant l'opération de l'exécution de l'application.

- Les messages : 1,5 et 7, sont des demandes d'enregistrement des Agents : AGENT1, AGENT2 et AGENT3 dans le DF, ces messages sont de type (REQUEST).
- Les messages : 2,6 et 8, sont des réponses d'acceptation des demandes d'enregistrement des 03 agents du système envoyés par le DF, ces messages sont de type (INFORME).
- Les messages 3 et 9, sont des messages de recherches de type (REQUEST) envoyés de l'AGENT3 d'adaptations vers le DF, pour avoir la liste des agents inscrit dans l'annuaire et fournissant les services de classification Temporelle et Morphologique.
- Les messages 4 et 10, sont les résultats de recherche (Réponses de DF) de type (INFORM) envoyés par le DF vers l'AGENT3 contenant la liste des agents, qui fournissent les services Classification Temporelle et Morphologique inscrits dans l'annuaire.
- Le message 11 (resp, 12) est l'envoi de l'ECG en question vers l'AGENT1 de classification temporelle, (resp, l'AGENT2 de classification morphologique), ce message est de type (CFP).
- Le message 13 (resp, 14) est l'accusé de réception envoyée par l'AGENT1 (resp, l'AGENT2) vers l'AGENT3, ce message est de type (INFORM).
- Le message 15 (resp, 17) est l'envoi du résultat de classification temporelle faite par l'AGENT1 (resp, classification morphologique faite par l'AGENT2) vers l'AGENT3, ce message est de type (INFORM).
- Le message 16 (resp, 18) est l'accusé de réception envoyée par l'AGENT3 vers l'AGENT1 (resp, l'AGENT2), ce message est de type (INFORM) ;
- Les messages 8 et 10, sont des confirmations ou accusés de réception des messages envoyés depuis AGENT1 et AGENT2, envoyés par l'AGENT3. Ces messages sont de type (INFORM).



**FIG 5.7** :Diagramme de communication entre les agents des systèmes.

## V.4. Présentation de l'application

### V.4.1 Agent classifieur neuronal AGENT1

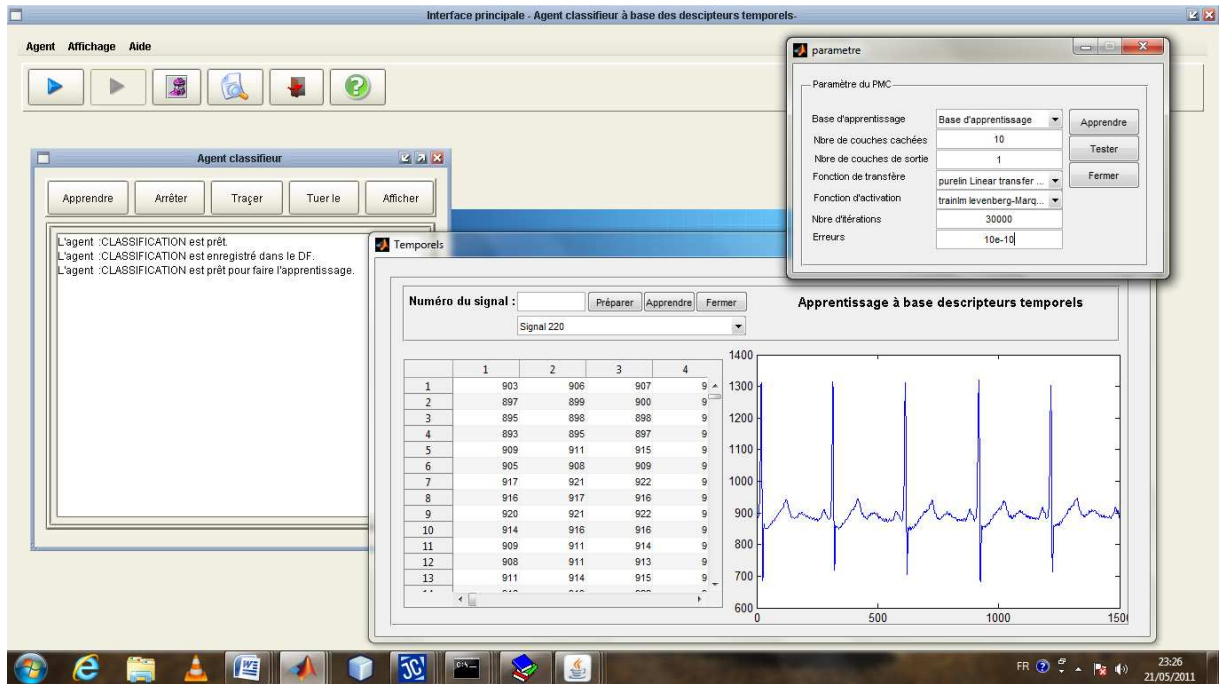


FIG 5.8 : Agent de classification à base des descripteurs temporels.

### V.4.1 Agent classifieur neuronal AGENT2

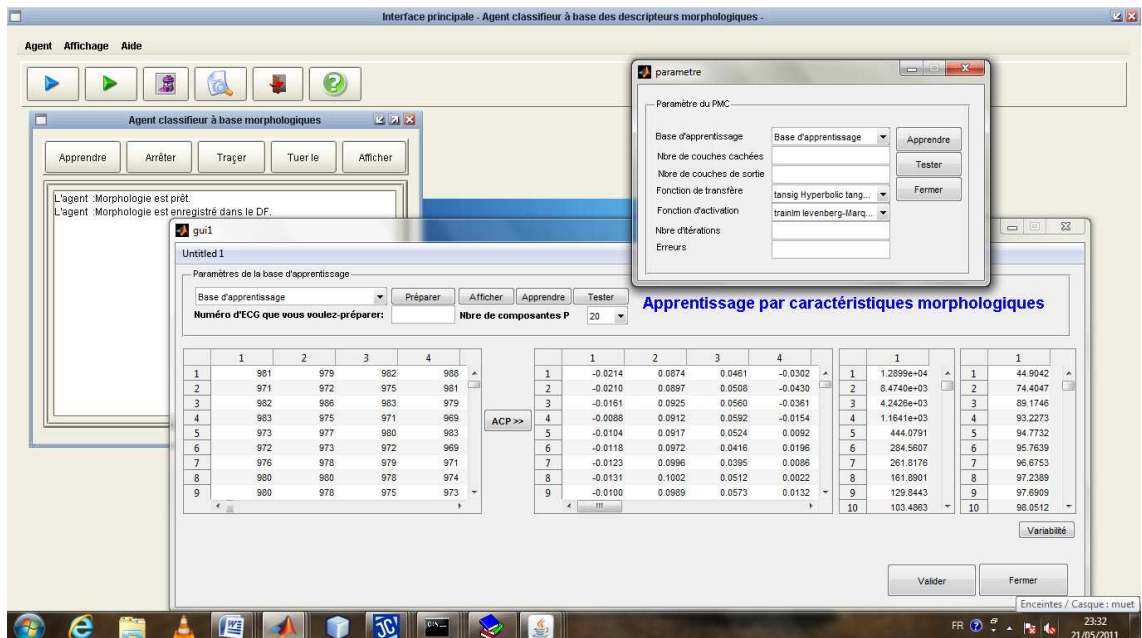


FIG 5.9 : Agent de classification à base des descripteurs morphologiques.

### V.4.1 Agent d'adaptation AGENT3

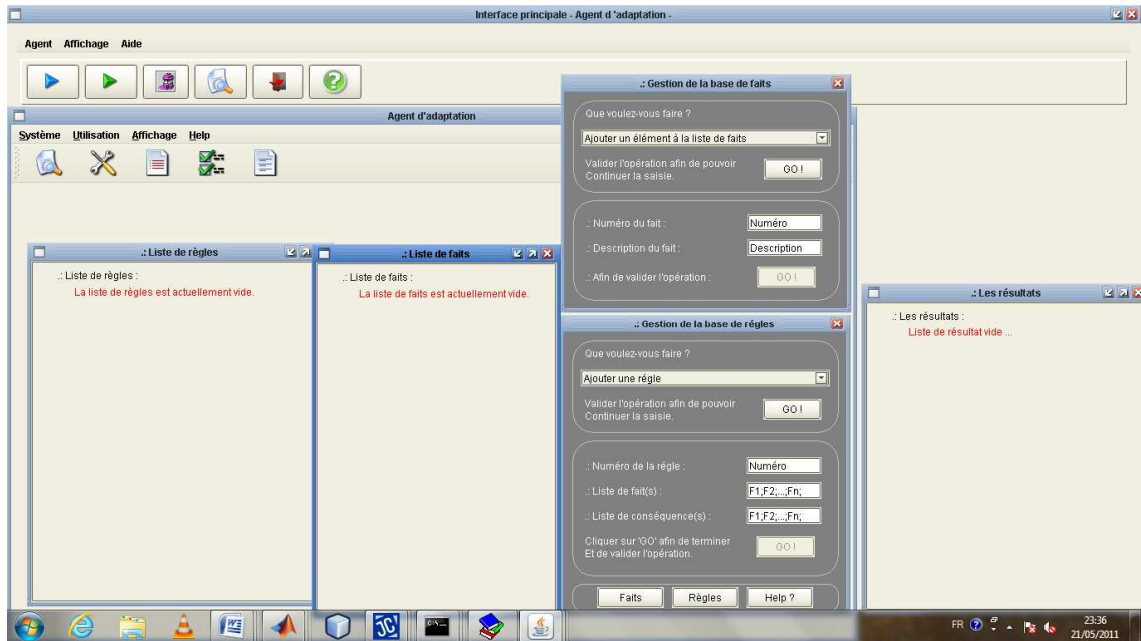


FIG 5.10 : Agent d'adaptation.

La séquence des messages échangés entre les agents du système est illustrée par la figure suivante :

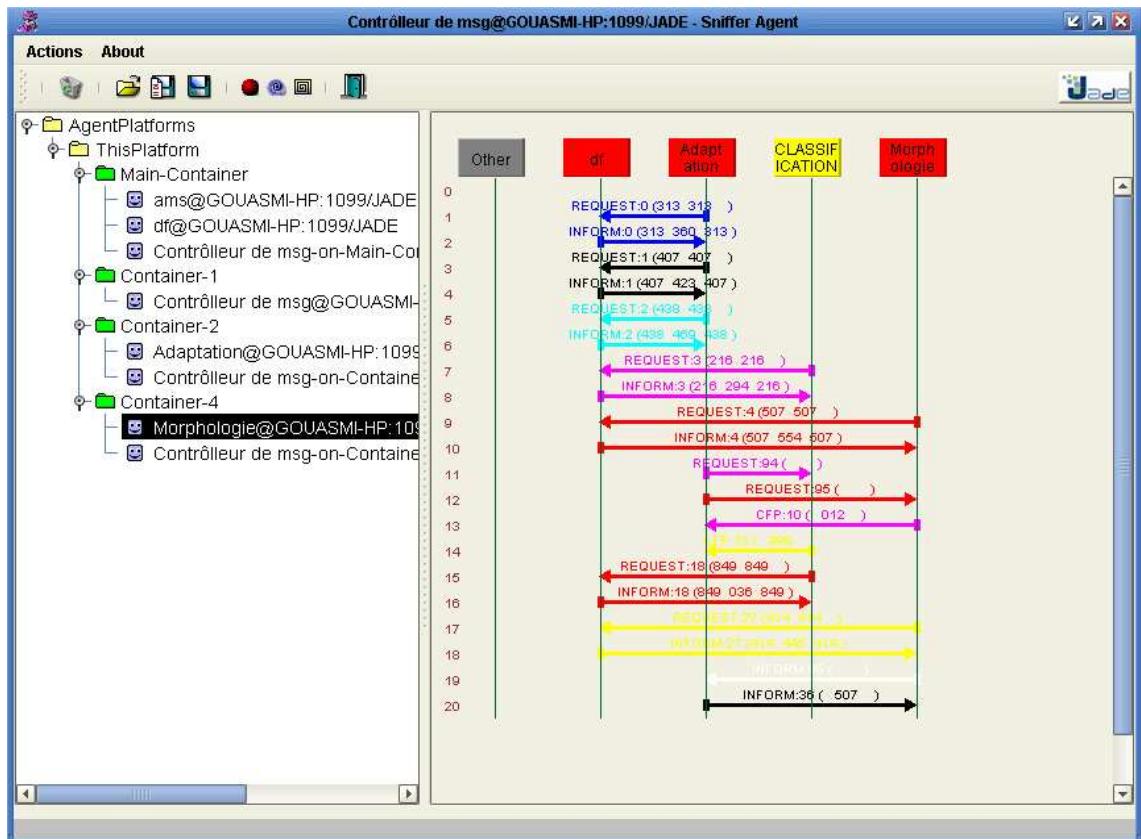


FIG 5.11: SNIFFER, contrôleur des messages échangés entre les agents du système.

Concernant, la base de connaissances de l'agent d'adaptation « AGENT3 », nous avons choisi comme base de fait : {N pour « Normal », A pour « Autre », et I pour « Inconnu »} et comme base des règles :

- $$\left\{ \begin{array}{l} 1) N \text{ et } N \rightarrow N, \\ 2) A \text{ et } A \rightarrow A, \\ 3) N \text{ et } A \rightarrow I \\ 4) A \text{ et } N \rightarrow I \end{array} \right\}.$$

En outre, notre approche oblige l'utilisation du « chaînage avant » pour faire l'inférence, puisque on connaît les prémices en premier temps.

## V.5. Conclusion

Les signaux physiologiques constituent une source d'information qui prend une par de plus en plus importante dans la démarche diagnostique en médecine et qui deviennent de plus en plus facilement accessible du fait du développement de systèmes d'acquisition de signaux. Dans cette thèse nous nous sommes intéressés à un signal physiologique particulier, le signal électrocardiogramme (ECG).

Le signal ECG a un potentiel diagnostique important et son utilisation fait partie intégrante de l'examen clinique d'un patient. Son contenu informationnel est reconnu et exploité essentiellement dans le cadre de l'auscultation cardiaque.

Dans cette thèse, nous nous sommes intéressés à la reconnaissance des battements cardiaques normaux, d'une autre façon de classer les battements cardiaques normaux et les autres qui ne sont pas normaux, pour pouvoir donner un bon diagnostic et dans les meilleurs délais, afin de protéger la santé des humains. Pour cela, nous avons proposé deux méthodes pour la classification, classification à base des descripteurs temporels extraites depuis le signal ECG et l'autre de à base des descripteurs morphologiques. Cette dernière nécessite un traitement de plus, à savoir la réduction de la taille de l'échantillon étudiée par l'une des méthodes d'analyse de données (ACP).

Cette classification est faite par une méthode d'intelligence artificielle, à l'occurrence les réseaux de neurones, qui ont montré leurs performances dans ce domaine, et surtout où se trouvent des données de type numérique et de grande taille (base de données géantes), ils ont la capacité de donner des bons résultats et dans les meilleurs délais, car le temps joue un rôle primordiale dans le diagnostic médical et l'aide de décision dans ce domaine.

En outre, nous avons utilisé les systèmes experts pour la fusion de données, ou pour l'adaptation pour avoir une classe à la fin de l'opération, et bien sûr pour réutiliser les connaissances des experts dans ce domaine.

Nous avons choisi une approche multi-agents pour avoir plus d'autonomie dans les travaux de chaque classifieur, en choisissant la plate forme jade de java pour créer, gérer et contrôler nos agents dans un milieu ouvert.

## CONCLUSION

L'objectif de ce mémoire était l'application d'une approche multi-agents pour le monitoring médical, en utilisant des outils de l'intelligence artificielle pour la classification des battements cardiaques, à l'occurrence, les réseaux de neurones artificiels et les systèmes experts qui ont prouvé leurs performances dans ce domaine (la classification des arythmies cardiaques).

Nous nous sommes intéressés à la classification distribuée et autonome des battements cardiaques, et ceci par l'intermédiaire des agents intelligents, donc nous avons proposé de créer deux agents intelligents de classification, dont le premier reçoit les valeurs descriptives temporelles de l'ECG que nous désirons le classifier, à savoir les distances  $RR_p$ ,  $RR_s$ , le rapport entre ces deux derniers, PP, la durée du complexe QRS, cet agent va faire un apprentissage supervisé en utilisant un réseau de neurones multi-Couches, qui donne une seule sortie (classe de classification, soit « N » pour normal, soit « A », pour autres), le deuxième agent reçoit à son tour des valeurs descriptives morphologiques de l'ECG en question après avoir réduit la taille de l'inertie en entrée par l'intermédiaire de la méthode d'analyse en composantes principales, dont on choisit seulement les composantes principales offrant la moindre perte de données (01 %), et de même cet agent nous donne une seule sortie qui est soit N soit A. Et finalement le troisième agent qui fusionne les deux résultats donnés par les agents de classification neuronaux, cet agent fait réaliser sa tâche en utilisant un système expert afin de faire une inférence pour donner la classe finale de cette classification distribuée.

Les résultats fournis par notre système multi-agent de monitoring ont confirmé le bien fondé du mariage entre les techniques de l'intelligence artificielle classique et celle distribuée.

Cette approche offre des avantages majeurs tels que l'autonomie des tâches réalisées, la distribution dans plusieurs postes de travail permettent de profiter des avantages des réseaux de télécommunication, le temps de traitement est aussi petit puisque le tout se fait en parallèle.

Cette petite recherche sera, dans le proche futur, améliorée –voir les perspectives- afin de donner des solutions efficaces, opérationnelles et moins coûteuses permettant d'identifier les patients à risque de mort subite rythmique et de prendre des mesures préventives selon le cas rencontré et donc des millions de décès pourraient donc être évités.

## PERSPECTIVES

Plusieurs axes de recherche découlent de ce travail.

Poursuivre dans cette voie :

Les résultats de l'abstraction semblent prometteurs. Néanmoins, il reste quelques points à développer :

- Intégrer un module d'extraction de connaissance à partir de données (ECD) au niveau de la chaîne de la classification, plus précisément, au niveau du module d'extraction des descripteurs. Afin d'automatiser cette tâche.
- Intégrer l'induction dans le système expert, pour enrichir la base de connaissances au fur et à mesure (la programmation logique inductive PLI).
- Ajouter le nombre des arythmies étudiées telles que les Extrasystole Ventriculaire ESV, les deux types de bloc de branche (droit et gauche), les battements ventriculaires prématurés Bvp, ... etc.
- intégrer un module d'alarme, émettant des sonores ou bips dans les cas non désirés.
- Ajouter d'autres sources de données telles que la pression artérielle, EEG, facteur d'âge, sexe, pour avoir une classification multi-sources.



# Annexe A

## Analyse en Composantes Principales (ACP)

### a.1 Introduction

L'analyse en composantes principales (ACP) est une technique descriptive permettant d'étudier les relations qui existent entre les variables, sans tenir en compte, a priori, d'une quelconque structure [Jol, 86],[Dia, 96] ; son utilisation pour l'exploitation des données remonte au début du siècle dernier. Elle est principalement issue des travaux de psychologues américains [Pea, 01],[Spe, 04],[Hol, 33]. Le but de l'ACP est d'identifier la structure de dépendance entre des observations multivariées, afin d'obtenir une description ou une représentation compacte de ces dernières. L'idée de base de l'ACP est de réduire la dimension de la matrice des données, en retenant le plus possible les variables présentes dans le jeu des données de départ. Cette réduction ne sera possible que si les variables initiales ne sont pas indépendantes et ont des coefficients de corrélation entre elles non nuls. Ces variables initiales sont transformées en de nouvelles variables, appelées composantes principales. Elles sont obtenues par combinaisons linéaires des précédentes et sont ordonnées et non corrélées entre elles. L'analyse en composantes principales cherche à identifier les vecteurs propres orthonormaux et les valeurs propres de dispersion des variables originelles. Les vecteurs propres orthonormaux sont utilisés pour construire les composantes principales et les valeurs propres sont les variances des composantes principales correspondantes [Jol, 86].

### a.2 Principe de l'analyse en composantes principales

Dans le cadre cette thèse, nous aborderons l'ACP comme une technique de réduction des échantillons de l'ECG dans le cas des descripteurs morphologiques. La théorie sous-jacente à l'analyse en composantes principales est vaste, nous ne passerons donc en revue que les points importants. Du point de vue géométrique l'ACP consiste à effectuer une certaine rotation du repère des variables autour de leurs valeurs moyennes. Cette rotation transforme les  $n$  variables non corrélées. Notons que ce sont justement ces variables transformées que l'on a nommé les composantes principales. Supposons qu'on a un vecteur de données aléatoires  $X_k = [x_1, \dots, x_m]^t \in \mathbb{R}^m$  de moyenne nulle  $\varepsilon\{X(k)\} = 0$  et de matrice de covariance ou d'auto-corrélation  $\Sigma = \varepsilon\{XX^t\} \in \mathbb{R}^{m,m}$ .

En analyse en composantes principales, un vecteur caractéristique  $t \in \mathbb{R}^l$  est associé à chaque vecteur de données dont il optimise la représentation au sens de la minimisation de l'erreur d'estimation de  $\mathbf{x}$  ou la maximisation de la variance de  $\mathbf{t}$ . Les vecteurs  $\mathbf{t}$  et  $\mathbf{x}$  sont liés par une transformation linéaire  $t = P^t \mathbf{x}$ , où la matrice de transformation  $P \in \mathbb{R}^{m,l}$  vérifie la condition d'orthogonalité  $P^t P = I_l$ . Les colonnes de la matrice  $\mathbf{P}$  forment les vecteurs de base orthonormés d'un sous-espace  $\mathbb{R}^l$  de représentation réduite des données. La transformation linéaire s'apparente

ainsi à une projection de l'espace des données de dimension  $m$  vers un sous-espace orthogonal de dimension  $l$ .

Les composantes  $t_j$ , avec  $j=1, \dots, l$ , du vecteur caractéristique  $\mathbf{t}$  représentent les composantes projetées du vecteur de données  $\mathbf{x}$  dans ce sous-espace.

Au sens de l'ACP, la projection  $P$  est optimale si l'erreur quadratique d'estimation des vecteurs de données  $\mathbf{x}$  est minimale. Ce problème d'optimisation s'exprime par :

$$P_{opt} = \operatorname{argmin} J_e(P)$$

Où  $J_e$  représente le critère d'erreur d'estimation de l'ACP. Sous la contrainte d'orthogonalité de la matrice de projection  $P^t P = I_l$ , ce critère aura la forme :

$$\begin{aligned} J_e(P) &= \varepsilon \left\{ \left\| X - \widehat{X} \right\|^2 \right\} \\ &= \varepsilon \left\{ \left\| X - PP^t X \right\|^2 \right\} \\ &= \varepsilon \left\{ (X - Pt)^t (X - Pt) \right\} \\ &= \varepsilon \left\{ (X^t X - 2t^t t + t^t PP^t t) \right\} \\ &= \varepsilon \left\{ \operatorname{trace}(XX^t) - t^t t \right\} \\ &= \operatorname{trace}(\Sigma) - \varepsilon \left\{ t^t t \right\} \end{aligned}$$

Notons que la trace d'une matrice carrée est définie par la somme de ces éléments diagonaux. Du fait que la matrice covariance  $\Sigma$  est indépendante de la matrice des paramètres  $P$ , minimiser  $J_e$  revient à maximiser le second terme  $J_v$  de son expression :

$$J_v(P) = \varepsilon \left\{ t^t t \right\} = \sum_{j=1}^l \varepsilon \left\{ t_j^2 \right\}$$

Ainsi, la minimisation de l'erreur quadratique d'estimation de  $\mathbf{x}$  est équivalente à la maximisation de la variance des projections  $t_j$  des données. En conséquence :

$$P_{opt} = \operatorname{argmin} J_e(P) = \operatorname{argmax} J_v(P)$$

Le problème de l'ACP, considéré sous l'angle de la maximisation de la variance de projection des données, est celui de la détermination des vecteurs propres de la matrice de covariance  $\Sigma$ .

### a.3 Identification du modèle ACP

Soit  $X_k = [X_1, \dots, X_n]^t$  un vecteur d'observations de  $n$  variables aléatoires. On notera  $\bar{X}$  le vecteur correspondant à la moyenne des observations de  $\mathbf{x}$  et  $\Sigma$  la matrice de covariance de ces  $m$  observations. L'estimation des paramètres du modèle ACP se résume en une estimation des valeurs et vecteurs propres de la matrice de covariance  $\Sigma$ . Puisque la matrice de covariance  $\Sigma$  est carrée de dimension  $n \times n$  et symétrique, on peut, selon l'algèbre linéaire, vérifier une relation de type :

$$P^t \sum P = L$$

Où  $L$  est une matrice diagonale et  $P$  est une matrice orthonormale. Les éléments diagonaux de  $L = \{l_1, \dots, l_n\}$  sont les valeurs propres de la matrice de covariance  $\sum$  alors que les vecteurs colonnes de  $P$  sont les vecteurs propres de  $\sum$ .

Les valeurs propres  $l_i$  peuvent être obtenues par la solution de l'équation caractéristique :

$$\det[\sum - l_i I] = 0$$

pour les  $n$  valeurs propres. Connaissant les valeurs propres  $l_i$ , on peut déduire les vecteurs propres de la matrice de covariance par la solution de

$$(\sum - l_i I)A_i = 0$$

et puis,

$$u_i = \frac{A_i}{\sqrt{A_i^t A_i}}$$

pour les  $n$  vecteurs propres. Notons que l'équation précédente est aussi appelée l'équation de normalisation. Il est intéressant de constater qu'ici les éléments des vecteurs propres  $u_i$  sont les cosinus directeurs des axes générés par une rotation dans l'espace des observations. Nous pouvons donc calculer tout un ensemble de valeurs à partir des ces nouveaux axes. En d'autres mots, on peut transformer les variables  $X = [x_1, \dots, x_n]^t$  en un nouvel ensemble de variables  $Y = [y_1, \dots, y_n]^t$  par l'application de

$$Y_i = u_i^t (X - \bar{X})$$

$Y_i$  est la  $i^{\text{ème}}$  composante principale de l'ensemble des observations. Les nouvelles variables  $Y_i$  ont évidemment une moyenne nulle et elles possèdent une variance donnée par  $l_i$  (valeur propre de  $\sum$ ). Les vecteurs de colonnes  $u_i$  de  $P$  sont orthonormaux car ils sont orthogonaux et de modules unitaires. Donc, si l'on utilise ces vecteurs pour obtenir les composantes principales, tel que montré à l'équation précédente, on obtiendra des composantes principales qui ne seront pas corrélés et de variances égales aux valeurs propres de la matrice de covariance  $\sum$ . En appliquant une transformation sur les vecteurs colonnes  $u_i$ , on peut obtenir de nouvelles propriétés intéressantes. L'une de ces transformations consiste à effectuer une mise à l'échelle du  $u_i$  par les valeurs propres de la matrice de covariance  $\sum$  :

$$V_i = u_i \sqrt{l_i}$$

Les composantes obtenues par l'utilisation de  $V_i = [v_1, \dots, v_n]^t$  seront de même unité que les variables originales. Nous retrouvons également une transformation dans laquelle les vecteurs  $u_i$  sont divisés par les valeurs propres de la matrice de covariance  $S$  :

$$W_i = \frac{u_i}{\sqrt{l_i}}$$

Dans ce cas, les composantes obtenues seront de variance unitaire.

Enfin, nous soulignons que l'on peut avoir une mesure de variabilité expliquée par chacune des composantes principales en utilisant les vecteurs transformés de l'équation avant précédente selon

$$\text{var} = V_i V_i^t, \text{ avec } i=1, \dots, n.$$

Les valeurs obtenues par cette dernière équation peuvent être utilisées pour indiquer le pourcentage de la variabilité des caractéristiques représentées par chacune des composantes principales.

#### **a.4 Détermination de la structure du modèle ACP**

L'analyse en composantes principales recherche une approximation de la matrice initiale des données  $X$  par une matrice rang inférieur issue d'une décomposition en valeurs singulières. La question qui se pose alors, et qui a été largement débattue dans la littérature, concerne le choix du nombre de composantes principales qui doit être retenu. De nombreuses règles sont proposées pour déterminer le nombre de composantes à retenir, [Bes, 93],[Him, 94],[Fer, 95],[Val, 99]. La plupart de ces méthodes sont heuristiques et donnent un nombre de composants subjectif. Toutefois, dans le cadre de l'application de l'ACP à la réduction des échantillons de l'ECG, le nombre de composantes à un impact significatif sur la phase finale de classification. Si peu de composantes sont utilisées, on risque de perdre des informations contenues dans les données de départ en projetant certains échantillons dans le sous-espace des résidus et donc avoir des erreurs de modélisation, ce qui provoque une mauvaise caractérisation.

Si par contre beaucoup de composantes sont utilisées, il y a risque d'avoir des composantes retenues (les composantes correspondantes aux valeurs propres les plus faibles parmi celles retenues dans le modèle) qui sont porteuses de redondance, ce qui est indésirable.

Dans leurs travaux de recherche, Qin et al [Qin, 98] ont proposé une technique basée sur la variance de l'erreur de reconstruction des mesures ; ce critère permet de prendre en compte la notion de redondance entre les variables (échantillons de l'ECG dans notre cas).

On trouve dans la littérature plusieurs critères utilisant différentes techniques pour la détermination du nombre de composantes principales, nous allons présenter deux critères heuristiques utilisés pour la réduction de dimension de la taille de l'ECG.

##### **1. Pourcentage cumulé de la variance totale (PVC)**

L'idée de base de cette approche est que chaque composante principale est représentative d'une portion de la variance des mesures de l'information étudiée. Les valeurs propres de la matrice de corrélation sont les mesures de cette variance et peuvent donc être utilisées dans la sélection du nombre de composantes principales. Pour le choix de  $l$ , il faut choisir le pourcentage de la variance totale qu'on veut conserver. Le nombre de composantes est alors le plus petit nombre pris de telle sorte que ce pourcentage soit atteint ou dépassé ; les composantes sont choisies successivement dans l'ordre des variances décroissantes. Le pourcentage de variance expliquée par les  $l$  premières composantes est donné par :

$$PVC(l) = 100 \left( \frac{\sum_{j=1}^l \lambda_j}{\sum_{j=1}^m \lambda_j} \right) \%$$

La variance du bruit étant inconnue a priori, la décision est basée seulement sur le pourcentage de la variance expliquée est un peu arbitraire. Sa capacité à fournir le nombre correct de composantes principales dépendra fortement du rapport signal sur bruit.

## 2. Moyenne des valeurs propres

Cette règle consiste à ne prendre en considération que les composantes pour lesquelles la valeur propre est supérieure à la moyenne arithmétique de toutes les valeurs propres. En particulier, si on travaille sur les données centrées réduites, cela revient à négliger les composantes dont la variance est inférieure à l'unité ( $\frac{1}{m} \text{trace}(\Sigma) = 1$ ). Dans le cas du modèle ACP calculé à partir de la matrice de covariance  $\Sigma$ , la moyenne arithmétique des valeurs propres est donnée par  $\frac{1}{m} \text{trace}(\Sigma)$ .

## Annexe B

### Base de données MIT-BIH

Depuis 1975, les laboratoires de l'hôpital de Beth Israël à Boston et MIT ont réalisé une base de données MIT-BIH qui a été commencée à être distribuée en 1980. Cette base de données contient 48 enregistrements extraits d'une demi-heure des enregistrements ambulatoires à deux voies d'ECG, obtenus à partir de 47 sujets étudiés par le laboratoire d'arythmie de BIH entre 1975 et 1979. 23 enregistrements ont été choisis au hasard d'un ensemble de 4000 enregistrements ambulatoires de 24 heures d'ECG rassemblés d'une population mélangée des patients hospitalisés (60%) et des patients non hospitalisés (40%) à l'hôpital de Beth Israël à Boston (les séries '100'), les 25 enregistrements restants ont été choisis parmi les mêmes enregistrements mais qui en considération des arythmies rarement observés qui ont une signification clinique (les séries '200').

Les enregistrements ont été échantillonnés à une fréquence  $f_e = 360\text{Hz}$  avec une résolution de 11 bits sur une gamme de 10 mV. Deux cardiologues ou plus ont indépendamment annoté chaque enregistrement, environ 110,000 annotations ont été incluses avec la base de données. Le tableau B.1 présente les 48 enregistrements de la base de données MIT-BIH.

N° d'enregistrement Arythmies	N	A	V	F	P	Q
100	2239	33	1	-	-	-
101	1860	3	-	-	-	2
102	99	-	4	-	2028	-
103	2082	2	-	-	-	-
104	163	-	2	-	1380	18
105	2526	-	41	-	-	5
106	1507	-	520	-	-	-
107	-	-	59	-	2078	-
108	1739	4	17	2	-	-
109	-	-	38	2	-	-
111	-	-	1	-	-	-
112	2537	2	-	-	-	-
113	1789	-	-	-	-	-
114	1820	10	43	4	-	-
115	1953	-	-	-	-	-
116	2302	1	109	-	-	-
117	1534	1	-	-	-	-
118	-	96	16	-	-	-
119	1543	-	444	-	-	-

121	1861	1	1	-	-	-
122	2476	-	-	-	-	-
123	1515	-	3	-	-	-
124	-	2	47	5	-	-
200	1743	30	826	2	-	-
201	1625	30	198	2	-	-
202	2061	36	19	1	-	-
203	2529	-	444	1	-	4
205	2571	3	71	11	-	-
207	-	107	105	-	-	-
208	1586	-	992	373	-	2
209	2621	383	1	-	-	-
210	2423	-	194	10	-	-
212	923	-	-	-	-	-
213	2641	25	220	362	-	-
214	-	-	256	1	-	2
215	3195	3	164	1	-	-
217	244	-	162	-	1542	-
219	2082	7	64	1	-	-
220	1954	94	-	-	-	-
221	2031	-	396	-	-	-
222	2062	208	-	-	-	-
223	2029	72	473	14	-	-
228	1688	3	362	-	-	-
230	2255	-	1	-	-	-
231	314	1	2	-	-	-
232	-	1382	-	-	-	-
233	2230	7	831	11	-	-
234	2700	-	3	-	-	-

**Tab B.1** : les 48 enregistrements de la base de données MIT-BIH

N : Battement normal (*Normal beat*).

A : Battement auriculaire prématuré (*Atrial premature beat*).

V : Contraction ventriculaire prématurée (*Premature ventricular contraction*).

F : Mélange entre les battements ventriculaires et les battements normaux (*Fusion of ventricular and normal beat*).

P : Pause.

Q : Battement non classifiable (*Unclassifiable beat*).

## Annexe C

### Les Actes de communication

Les actes de communications proposés par JADE sont : **agree**, **cancel**, **confirm**, **cfp**, **disconfirm**, **failure**, **inform**, **inform\_if**, **query\_if**, **query\_ref**, **reject\_proposal**, **request**, **request\_when**, **request\_whenever**, **subscribe**, **null**, **propose**, **not\_understood**.

L'acte de communication traduit la nature du message envoyé entre agents. Les actes se classent dans des catégories, certaines permettent d'amorcer des protocoles alors que d'autre sont utilisés pour répondre. Ainsi **request**, **cfp**, **request\_when**, **request\_whenever**, **query\_if**, **query\_ref**, **subscirbe** sont utilisés pour amorcer un questionnement ou une demande de proposition et **agree**, **inform**, **failure**, **propose**, **not\_understood**, sont des réponses possibles .Le plus usité est le **not\_understood**, dès qu'un agent ne comprend pas le message reçu, il répond par un message **not\_undertood** pour avoir plus de clarté.

Et voici, le **Tab C.1** : qui explique tous les performatifs :

Actions	Syntaxe	Définition - Sens
<b>Accept Proposal</b>	<b>accept-proposal</b>	Communication de l'accord de l'envoyeur d'effectuer une action qui lui a été préalablement soumise.
<b>Agree</b>	<b>agree</b>	Communication de l'accord de l'envoyeur pour effectuer une action, sans doute dans le futur.
<b>Cancel</b>	<b>cancel</b>	Communication de l'annulation de l'accord donnée préalablement par l'expéditeur pour effectuer une action.
<b>Call for Proposal</b>	<b>Cfp</b>	Communication par l'envoyeur d'une demande d'effectuer une certaine action.
<b>Confirm</b>	<b>confirm</b>	Communication par l'envoyeur de la confirmation de la validité (selon les règles de l'agent) de la proposition préalablement reçue.
<b>Disconfirm</b>	<b>disconfirm</b>	Communication par l'envoyeur de la confirmation du non validité (selon les règles de l'agent) de la proposition préalablement reçue.



<b>Failure</b>	<b>failure</b>	Communication par l'envoyeur de l'échec d'une action essayée.
<b>Inform</b>	<b>inform</b>	Communication par l'envoyeur d'une proposition, pensée vrai par celui-ci.
<b>Inform If</b>	<b>inform-if</b>	Communication par l'envoyeur d'une proposition (pensée vrai par celui-ci), et demande au receveur une confirmation ou une non-confirmation.
<b>Inform Ref</b>	<b>inform-ref</b>	Communication par l'envoyeur d'une demande de l'objet qui correspond à une description envoyée.
<b>Not Understood</b>	<b>not-understood</b>	Communication par l'envoyeur d'un non compréhension d'une action effectuée par le destinataire.
<b>Propagate</b>	<b>Propagate</b>	Communication par l'envoyeur d'un message à propager à des agents dont la description est fournie. Le destinataire du message traite le sous message à propager comme s'il lui était directement destiné et envoie le message " <i>propagate</i> " à l'agent qu'il a identifié.
<b>Propose</b>	<b>propose</b>	Communication par l'envoyeur d'une proposition d'action conditionnée à certains prés conditions données.
<b>Proxy</b>	<b>proxy</b>	Communication par l'expéditeur d'une demande d'une transmission d'un message à des agents dont la description est donnée.
<b>Query Ref</b>	<b>query-ref</b>	Communication par l'envoyeur d'une demande par l'expéditeur de l'objet référencé par une expression.
<b>Refuse</b>	<b>refuse</b>	Communication par l'envoyeur de son refus d'effectuer une action donnée, et en donne les raisons.
<b>Reject Proposal</b>	<b>reject-proposal</b>	Communication, pendant une négociation, par l'envoyeur de son refus d'effectuer des actions.
<b>Request</b>	<b>request</b>	Communication par l'envoyeur d'une demande au destinataire d'effectuer une action.

<b>Request When</b>	<b>request-when</b>	Communication par l'expéditeur d'une demande, au destinataire, d'effectuer une action quand une proposition donnée devient vraie.
<b>Request Whenever</b>	<b>request-whenever</b>	Communication par l'expéditeur d'une demande, au destinataire, d'effectuer une action dès qu'une proposition donnée devient vraie, et à chaque fois que celle-ci redevient vraie.
<b>Subscribe</b>	<b>subscribe</b>	Communication par l'expéditeur d'une demande d'un objet donné par une référence envoyée par l'expéditeur, et de notifier l'agent ayant souscrit dès que l'objet en question change.

**Tab C.1** : Les actes de communication fournis par Jade.

- [Amm, 07] Mohamed Yessin AMMAR, Thèse de doctorat à toulouse, "MISE EN OEUVRE DE RESEAUX DE NEURONES POUR LA MODELISATION DE CINETIQUES REACTIONNELLES EN VUE DE LA TRANSPOSITION BATCH/CONTINU". 2007
- [Bel, 92] Belaid, A, Belaid, Y, « Reconnaissance des formes. Méthodes et application », InterEditions, Paris, 429p, 1992.
- [Bel,Che et Ber, 03] N. Belgacem, M.A Cheikh, F. Berekci Reguig, 2003, «Détection et identification des arythmies cardiaques par application des réseaux de neurones » U. Tlemcen.
- [Ben, 06] Développement d'une méthodologie pour la conduite en sécurité d'un réacteur continu intensifié, Thèse de doctorat, I.N.P. Toulouse, (2006).
- [Bes, 07] *Besrou Rym, Lachiri Zied et Ellouze Noureddine, CRATT 2007* «Classification Des arythmies cardiaques par Machines A Vecteurs De Support » USTO, ORAN.
- [Bes, 93] *Besse, P., and ferre, L, "Sur l'usage de la validation croisée en analyse en composantes principales", Revue de Statistique Appliquée, XLI (1) pp. 71-76, 1993.*
- [Bio, 98] Biot, L. (1998). Evaluation d'un monitoring de la mécanique respiratoire en réanimation adulte. Thèse de médecine, Université Claude Lyon I, Faculté de médecine Lyon Nord.
- [Bis, 95] Neural networks for pattern recognition, Oxford University Press, (1995).
- [Blo et Hil, 80] Blondeau, M. et Hiltgen, M. (1980). Electrocardiographie clinique. Masson.
- [Blo et Hun, 01] Bloch, I. et Hunter, A. (2001). Fusion :General concepts and characteristics. International Journal of Intelligent Systems, 16(10) :1107 -1134.
- [Blo et Maî, 94] Bloch, I. et Maître, H. (1994). Fusion de données en traitement d'images : modèles d'informations et décisions. Traitement du Signal, 11(6) :435-446.
- [BON, 94] Bonabeau Eric, Theraulaz Guy, Intelligence collective, Hermès, Paris, 1994
- [BOU, 92] Bouron T., Structure de communication et d'organisation pour la coopération dans un univers multi-agents, Université Paris VI, (Thèse 3ème cycle, Informatique), Novembre 1992
- [BRA, 95] Brassac Christian, Pesty Sylvie, « Coopération dans les systèmes multi-agents : comportement ou conduite ? », *Decentralized Intelligent Multi-Agent Systems'95*, Krakow, Pologne, 22-24 novembre, 1995.
- [BRE, 98] Brenner Walter, Zarnekow Rüdiger, Wittig Harmut, *Intelligent Software Agents: Foundations and*

- Applications*, Springer-Verlag, Berlin, 1998.
- [Cal, 02] Callut, "Implémentation efficace des Support Vector Machines pour la classification", rapport de mémoire, Faculté des sciences, Bruxelles, 2002.
- [Cau et al, 98] Cauvin, S., Cordier, M.-O., Dousson, C., Laborie, P., Levy, F., Montmain, J., Porcheron, M., Servet, I. et Trav\_e-Massuy\_es, L. (1998).
- [Cha, 01] Chambrin, M.-C. (2001). Alarms in the intensive care unit: how can the number of false alarms be reduced? *Critical Care*, 5(4):184-188.
- [CHA, 02] Imed Jarras et Brahim Chaib-draa, *Aperçu sur les systèmes multi-agents*, 2002
- [CHA, 09] Zid-el-khir CHALABI, Abderhim BOUDJEMAOU, Larbi SAADIA et Nasreddine BERRACHED, SETIT 2009 : « Détection et Classification Automatiques d'Arythmies Cardiaques ».
- [Che, 05] CHEIKH Mohamed Amine, 2005, thèse de doctorat « Analyse du signal ECG par les réseaux neurones et la logique floue : Application à la reconnaissance des battements ventriculaires prématurés ».
- [COH, 95] Cohen P.R., Levesque H.J., "Communicative actions for artificial agents», *Proceedings of the First International Conference on Multi-Agent Systems - ICMAS-95*, AAAI Press/The MIT Press, California, June 1995, 65-72 p.
- [Coi, 93] Coiera, E. (1993). Intelligent monitoring and control of dynamic physiological systems. *Artificial Intelligence in Medicine*, 5:1-8.
- [Côt 98] M. Côté and N. Troudi. Une architecture multi-agent pour la recherche. *L'expertise Informatique*, 3(3), 1998.
- [Côt 99] M. Côté. Une architecture multi-agent et son application aux systèmes financiers. Master's thesis, Département d'Informatique, Université Laval, Janvier 1999.
- [CTM, 98] Collins J., Tsvetovat M., Mobasher B., and Gini M. MAGNET: A Multi-Agent Contracting System for Plan Execution. In *Workshop on Artificial Intelligence and Manufacturing: State of the Art and State of Practice*, pp 63-68, AAAI Press, Albuquerque, NM, August 1998.
- [Deb, 97] M. E. DeBakey and A. M. Gotto, *Le coeur en action*: Institut d'édition Sanofi~Synthelabo, 1997.
- [Dec, 91] Decoste, D. (1991). Dynamic across-time measurements interpretation. *Artificial Intelligence*, 51(1,2,3) :273-341.
- [Dia, 96] Diamantaras K.I., Kung S.Y., "Principal component neural networks, theory and applications", John Wiley and Sons, 1996.
- [Doj et al, 97] Dojat, M., Pachet, F., Guessoum, Z., Touchard, D., Harf, A. et Brochard, L. (1997). N\_eoganesh: a working system for the automated control of assisted ventilation in icus.

- [Doj *et al*, 98] Artificial Intelligence in Medicine, 11:97-117.  
Dojat, M., Ramaux, N. et Fontaine, D. (1998). Scenario recognition for temporal reasoning in medical domains. Artificial Intelligence in Medicine, 14(5).
- [Dor et al., 97] Dorronsoro, J.R., Ginel, F., Sanchez, C., et Cruz, C.S. 1997. Neural fraud detection in credit card operations. IEEE Transactions on Neural Networks, 8(4) : 827–834.
- [Dou, 96] Dousson, C. (1996). Alarm driven supervision for telecommunication networks. ii- on-line chronicle recognition. Annales des Télécommunications, 51(9-10):501-508
- [Ein, 41] W. Einthoven, "The galvanometric registration of the human electrocardiogram, likewise a review of the use of capillary-electrometer in physiology", in *Cardiac classics*: Willius FW-CV Mosby, St Louis, 1941.
- [FER, 95] Ferber Jacques, *Les systèmes multi-agents : vers une intelligence collective*, 1995.
- [FER, 96] Ferber Jacques, « Technologie multi-agent », *Mémento : Les interfaces intelligentes et images de communication et l'images*, France Télécom Recherche, n°8, Octobre 1996.
- [Fer, 95] Ferre L., "Selection of components in principal analysis: A comparaison of methods", Computational Statistics and Data Analysis pp. 669-682, 1995.
- [FIN, 97] Finin Tim, Labrou Yannis, Mayfield James, « KQML as an Agent Communication Language », Software Agents, Bradshaw Jeffrey (Eds.), AAAI/MIT Press, 1997.
- [Fin, 03] Finlay D.D, and al, "Evaluation of a statistical prediction model used in the design of neural network based ECG classifiers:a multiple linear regression approach" proc of the annual IEEE conf on information tecgnology applications in biomedecine, pp.258-260, UK 2003.
- [FIP, 00] Foundation for Intelligent Physical Agents, « FIPA Abstract Architecture Specification », August 2000, <http://www.fipa.org>.
- [Fog, 91] Fogel D.B, "An information criterion for optimal network selection"? IEEE trans. Neural networks, vol 2 nr 5, p. 49067, 1991.
- [FRA, 96] Stan Franklin, Art Graesser, « Is it an agent or just a program ? : a taxonomy for autonomous agents », In : *Intelligent Agents III : Agents Theories, Architectures and Languages, ECAI'96 Workshop (ATAL)*, Springer-Verlag, 1996
- [Fri et al., 02] Fried, R., U., et Imhoff, M. (2002). Pattern recognition in intensive care online monitoring, Chapitre 6. In Computational intelligence techniques in medical diagnosis and prognosis, Springer-Verlag, pages 129-165.
- [Fro, 05] Fromont Elisa, 2005, thèse de doctorat « Apprentissage

- multisource par programmation logique inductive : application \_a la caract\_erisation d'arythmies cardiaques ».
- [Fuk, 88] Fukushima, K. 1988. Neocognitron: a hierarchical neural network capable of visual pattern recognition. *Neural Networks*, 1 :119-130.
- [Gil, 82] Gilbert H.M, « Manuel d'Electrocardiographie », Edition MADSIS 1982.
- [Gro, 98] Grossberg, S. 1998. Birth of a learning law. *International Neural Network Society Newsletter*, 21 : 1-4.
- [Guy, 98] Guyon I, et Al., «What size test set gives good error rate estimates? », *IEEE Trans. Pattern Analysis and Machine Intelligence*. 20(1), pp. 52-64, 1998.
- [Hal, 92] Haldar R., et Rao D.P., Experimental studies on parametric sensitivity of a batch reactor, *Chemical Engineering Technology*, 15, 34-38, (1992).
- [Ham, 97] Hammadi N., Ito H. "Noise injection hidden neurons: A learning technique for enhancing the performance of feedforward networks", *Proceeding Neurap 97 3rd*, Marseille, P. 245-252, 1997.
- [Has, 95] Hassoum, M. H. (1995). *Fundamentals of artificial neural networks*. Cambridge: MIT Press.
- [Hay, 96] Haykin, S., et Li, L. 1996. Nonlinear adaptive prediction of nonstationary signals. *IEEE Transactions on Signal Processing*, 43(2) :526-535.
- [Her, 00] Hernández, A. (2000). Fusion de signaux et de modèle pour la caractérisation d'arythmies cardiaques. Phd, Université de Rennes 1, Rennes, France.
- [Hér, 94] Herault J., Jutten C., " Réseaux neuronaux et traitement du signal", *Hernès*, Paris 1994.
- [Him,94] Himes D.M, and al., "Determination of the number of principal components for disturbance detection and isolation", *Proceedings of ACC*, Baltimore, 1994.
- [Hin, 92] Hinton, G.E. 1992. How neural networks learn from experience. *Scientific American*, 267 : 145-151.
- [Hol, 00] Holger S., Bengio Y., "Boosting Neural Networks", *Neural computation*, 12(8) p. 1869-1887, 2000.
- [Hol, 33] Hotelling H., "Analysis of a complex of statistical variables into principal components", *Journal of Educational Psychology*, Vol. 24, pp. 417-441, 1933.
- [Hor, 91] Hornik K., Approximation capabilities of multilayer feedforward networks, *Neural Networks*, vol.4, pp. 251-257, (1991).
- [Jea, 94] Jean Gay, Michel D, Patrice B, " L'Electrocardiogramme savoir l'interpréter", Edition OPU, Alger, 1994.
- [Jim, 96] Jim K.M, Giles C.L, Horne B.G, " An analysis of noise in recurrent neural networks convergence and generalization", *IEEE Trans.Neural Networks*, Vol. 7, nr6, P. 1424-38, 1996.

- [Jol, 86] Jolliffe I.T., "Principal component analysis", Springer-Verlag, New York, 1986.
- [Kay, 94] Kay J. and Kummerfeld R. J. Customisation and Delivery of Multimedia Information, In Multicomm 94, the Westin Bayshore, Vancouver, British Columbia, CANADA, November 2-3 1994.
- [Koh, 87] Kohonen, T. 1987. Self-organization and associative memory. 2e éd. Springer-Verlag, Berlin.
- [Kua, 98] Kuokka D. and Harada L. Matchmaking for Information Agents, Readings in Agents, M. N. Huhns & M. P. Singh editors, 1998.
- [Kun, 99] Ludmila Kuncheva, Friedrich Steimann: Fuzzy diagnosis. *Artificial Intelligence in Medicine*, 16(2): 121-128, 1999.
- [KUW, 95] Kuwabara Kazuhiro, Ishida Toru, Osato Nobuyasu, "AgenTalk: Coordination Protocol Description for Multiagent System", *Proceedings of the First International Conference on Multi-Agent Systems - ICMAS'95*, AAAI Press/ the MIT Press, 1995, 455 p.
- [LAB, 00] Labrou Yannis, Finin Tim, « History, State of the Art and Challenges for Agent Communication Languages », *Informatik – Informatique*, n°1, Février 2000, 17-24 p.
- [LAB, 93] S. Labidi, W. Lejouad, « De l'Intelligence Artificielle Distribuée aux Systèmes multi-agents », *Rapport de recherche n°2004*, INRIA, Août 1993.
- [LAB, 99] Labrou Yannis, Finin Tim, Peng Yun, « Agent Communication Languages: the Current Landscape », *IEEE Intelligent Systems & Their Applications*, v.14, n°2, 1999, 45-52 p.
- [Lar et Hay, 98] Larsson, J. et Hayes-Roth, B. (1998). Guardian: An intelligent autonomous agent for medical monitoring and diagnosis. *IEEE Intelligent Systems*, 13(1):58-64.
- [LAV, 97] LAVRAC N., KERAVNOU E., ZUPAN B., *Intelligent Data Analysis in Medicine and Pharmacology*, Kluwer, 1997.
- [LAV, 99] LAVRAC N., « Selected techniques for data mining in medicine », *Artificial Intelligence in Medicine*, vol. 16, no 1, 1999, p. 3-23.
- [Lee et al, 96] Lee, K.C., Han, I., et Kwon, Y. 1996. Hybrid neural network models for bankruptcy predictions. *Decision Support Systems*, 18 :63–72.
- [Les, 91] Leschi C. « Aspects fondamentaux et évolutifs de la reconnaissance de formes. Approche pluridisciplinaire, notion d'approximation. », Thèse de doctorat en informatique et automatique appliquée : INSA de Lyon, 407p, 1991.
- [Mar, 76] E. Marey, "Des variations électriques des muscles et

- du cœur en particulier, étudiées au moyen de l'électromètre de M. Lippmann", *CR Acad Sci (Paris)*, vol. 82, pp. 975, 1876.
- [MAY, 96] Mayfield James, Labrou Yannis, Finin Tim, « Evaluating KQML as an Agent Communication Language », In : Wooldridge Michael, Müller Jörg P., Tambe Milind (Eds.), *Intelligent Agents II: Agent Theories, Architectures, Languages – IJCAI'95 Workshop (ATAL)*, Lecture Notes Artificial Intelligence v.1037, Springer-Verlag, 1996, 347-360 p.
- [McC et al. 43] Mc Culloch W.S. et Pitts W., A logical calculus of the ideas immanent in nervous activity, *Bulletin of Math. Biophysics*, vol.5, pp. 115-133, (1943).
- [MES, 05] Ali MESSAOUD, Mohamed BEN MESSAOUD, Abdennaceur KACHOURI et Faïçal SELLAMI, LETI, Laboratoire d'Electronique et de Technologies de l'Information, « Classification des arythmies cardiaques par logique floue à partir de signaux ECG»,
- [Mik et al, 96] Miksch, S., Horn, W., Popow, C. et Paky, F. (1996). Utilizing temporal data abstraction for data validation and therapy planning for artificially ventilated newborn infants. *Artificial Intelligence in Medicine*, 8(6):543-576.
- [MIL, 82] MILLER R., POPLER H., MYERS J., « INTERNIST-1, An Experimental Computerbased Diagnostic Consultant for General Internal Medicine », *N Engl J Med*, vol. 307, 1982, p. 468-476.
- [Mil, 93] Milgram M, « Reconnaissance des formes : Méthodes numériques et connexionnistes », Armand Collin, Paris, 175p, 1993.
- [Min et Pap. 88] Minsky M. et Papert S., *Perceptrons : an introduction to computational geometry*, MIT Press, expanded edition, (1988).
- [Mor, 01] Morris, A. (2001). Rational use of computerized protocols in the intensive care unit. *Critical Care*, 5 :249-254.
- [Mor, 02] Morris, A. (2001). Decision support and safety of clinical environments. *Qual Saf Health Care*, 11 :69-75.
- [Mou, 09] Dhouha Maatar, Zied Lachiri, Unité de Recherche Signal, Image et Reconnaissance de Formes Ecole Nationale d'Ingénieurs de Tunis, « Classification automatique d'arythmies par Hmm utilisant les parametres morphologiques dans L'ECG »
- [MUL, 96] Jörg P. Müller, *The design of intelligent agents. A layer approach*, Lecture Notes of Computer Science v.1177, Springer-Verlag, 1996.
- [Nad, 93] Nadal J.P, « Réseaux de neurones. De la physique à la physiologie », Armand Colin, Paris, 152p, 1993.
- [Nas, 00] Nash J.C, "Compact Numerical Methods for Computers: linear algebra and function minimisation", Ed, Adam



- Hilger, 2000
- [Nod, 98] M. Nodin, W. Bohrer, and A. Hiong Ngu. Semantic brokering over dynamic heterogeneous data sources in infosleuth. Technical Report?, MCC Corporation, 1998.
- [OBR, 98] O'Brien P.D., Nicol R.C., « FIPA – towards a standard for software agents », *BT Technology Journal*, v.16, n°3, July 1998, 51-59 p.
- [OHA, 96] G.M.P. O'Hare, N.R. Jennings (Eds.), *Foundations of Distributed Artificial Intelligence*, Wiley & Sons, New York, 1996.
- [Pat, 95] Patterson D.W. " Artificial neural networks. Theory and applications", neural prentice Hall, Singapore, 460p, 1995.
- [Pea, 01] Pearson K., "On lines and planes of closest fit to systems of points in space", *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science*, Vol, pp. 559-572, 1901.
- [Pie, 98] Pierson, D. J. (1998). Principles and practice of intensive care monitoring, chapitre Goals and indications for monitoring, pages 33-44. McGraw-Hill, Health Professions Division, New York.
- [Pol et al., 92] Pollard S.J.T., Fowler G.D., Sollars C.J. and Perry R. (1992) Low cost adsorbents for waste and wastewater treatment: a review, *The Sci. Total Environ.*, 116, 31-52.
- [Por, 05] Portet François, 2005, thèse de doctorat « Pilotage d'algorithmes pour la reconnaissance en ligne d'arythmies cardiaques
- [Psi, 99] Psychogios D.C, Ungar L.H, "SVD-NET: an algorithm that automatically selects networks, vol.4, nr 3, p.513-15, 1994.
- [Qua, 06] CHU Thanh-Quang, Rapport de stage « Application des méthodes d'intelligence artificielle à l'analyse des données concernant les canalisations de gaz », Institut de la francophonie pour l'informatique, 2006.
- [Qin, 98] Qin S. J. and Dunia R., "Determining the number of principal components for best reconstruction", *Proc. Of the IFAC Symposium on Dynamics and Control of Process Systems*, PP 359-364, Corfu, Greece, 1998.
- [Ref et al, 94] Refenes, A.N., Zapranis A., et Francis G. 1994. Stock performance modeling using neural networks: a comparative study with regression models. *Neural Networks*, 7(2) : 375–388.
- [Rém, 04] Rémi DUBOIS thèse : « Application des nouvelles méthodes d'apprentissage à la détection précoce d'anomalies en électrocardiographie » chapitre 1.
- [Ros, 94] S. J. Rosenschein and G. Zlotkin. *Rules of Encounter: Designing Conventions for Automated Negotiation Among Computer*. MIT Press, Boston, MA, 1994.
- [RUS, 97] *Rationality and intelligence. Artificial Intelligence*, Russell, Vol. 94. 1997.
- [SHO, 93] Yoav Shoham, « Agent-oriented programming »,

- Artificial intelligence*, v.60, n°1, 1993, 51-92 p.
- [Sho, 76] Shortliffe, E. (1976). *Computer-Based Medical Consultations: MYCIN*. Elsevier, New York.
- [SIE, 99] Sierra C. Agent-mediated electronic commerce a European viewpoint, In 7th Journées Francophones pour l'Intelligence Artificielle Distribuée et les Systèmes Multi-agents, La Réunion, November 9 1999.
- [Sil, 04] Silvent ANNE-SOPHIE (2004), Thèse "Extraction de connaissances pour la construction de scénarios médicaux"
- [Sit *et al*, 89] Sittig, D., Pace, N., Gardner, R., Beck, E. et Morris, A. (1989). Implementation of a computerized patient advice system using the help clinical information system. *Computers management and biomedical*
- [Smt, 94] D. C. Smith, A. Cypher, and J. Spohrer. KIDSIM: Programming agents without a programming language. *Communications of the ACM*, 37(7):55-67, July 1994.
- [Spe, 04] Spearman C., "General intelligence objectively determined and measured", *American Journal of Psychology*, Vol. 15, pp. 201-293, 1904.
- [Ste, 96] Steppe J.M, and., "Integrated feature and architecture selection" *IEEE trans. Neural Networks*, Vol. 7 nr 4, p.1007-14, 1996.
- [STE, 97] STEIMANN F., « Fuzzy set theory in medicine », *Artificial Intelligence in Medicine*, vol. 11, no 1, 1997, p. 1-7.
- [Ste, 99] Stein, E. (1999). *Rapid analysis of arrhythmias*. Lippincott Williams and Wilkins.
- [Sut, 92] Sutton, R.S. 1992. Special issue on reinforcement learning. *Machine Learning*, 8 : 1-395.
- [SZO, 82] SZOLOVITS P., *Artificial Intelligence in Medicine*, Westview Press, Inc., Boulder, Colorado (<http://medg.lcs.mit.edu/ftp/psz/AIM82/>), 1982.
- [Szo, 95] Szolovits, P. (1995). Uncertainty and decisions in medical informatics. *Methods of Information in Medicine*, 34 :111-121.
- [Tac, 02] Tâche collective (2002). Evidence-based guidelines for weaning and discontinuing support. *Respiratory Care*, 47(1) :69-90. A collective task force facilitated by the American College of Chest physicians, the American Association for respiratory care, and the American College of Critical Care Medicine.
- [Tea, 94] Teasauro G., Gammo T., "A self-teaching backgrammongram. Achives master-level play", *Neural Computation*, vol 6, pp 215-306; 1994
- [Thi, 96] Thimm G, Moerland P., Fiesler E., "The interchangeability of learning rate and gain in backpropagation neural networks", *Neural Computation*, Vol. 8, p.541-60, 1996.

- [Thi, 97] Thirumalaiah, K., et Deo, M.C. 1997. River stage forecasting using artificial neural networks. *ASCE Journal of Hydrologic Engineering* 3(1) : 26–32.
- [Tro, 99] N. Troudi. Un système multi-agent pour les environnements riches en information. Master's thesis, Département d'Informatique, Université Laval, Avril 1999.
- [Uck, 93] Uckun, S. (1993). Intelligent systems in patient monitoring and therapy management.
- [Val, 99] Valle S., and al., "Selection of the number of principal compenets: The variance of the reconstruction error criterion with a comparaison to other methods", *Industrial & Engineering Chemistry Research*, Vol, 38, pp. 4389-4401, 1999.
- [Vap, 98] V. Vapnik, "Statistical Learning Theory," New York ,1998.
- [Vid, 98] M.J. Vidal, T. Müller, P. Weinstein, and Durfee H. E. The university of michigan digital library service market society. In K. P. Sycara and M. Wooldridge, editors, *Proc. of the 2nd Int. Conf. on Autonomous Agent (Agent-98)*, pages 269-276, 1998.
- [Wal, 93] A. Waller, *An introduction to human physiology*. New-York: Longmans Green, 1893.
- [Wan, 02] Wang, F. (2002). Abstraction temporelle de signal ECG, apprentissage inductif de contraintes temporelles et reconnaissance des arythmies cardiaques. Phd, Université de Rennes 1, Rennes, France.
- [Wei et Kul, 91] Weiss, S.M. & Kulikowski, C.A. (1991), *Computer Systems That Learn*, Morgan Kaufmann.
- [Whi et Sof, 92] D. White, and M. Jordan, *Optimal control: a foundation for intelligent control*, In D. Sofge & D. White (Eds.), *Handbook of Intelligent Control, Neural, Fuzzy, and Adaptive Approaches* , Van Nostrand Reinhold, New York, 1992.
- [WOO, 95a] Michael Wooldridge, Nicholas R. Jennings, « Intelligent Agents: Theory and Practice », *The Knowledge Engineering Review*, v.10, n°2, 1995, 115-152 p.
- [WOO, 98b] Michael Wooldridge, Nicholas R. Jennings, « Pitfalls of Agent-Oriented Development », In: *Proceedings of the 2<sup>nd</sup> Conference on Autonomous Agents (Agents-98)*, Minneapolis, USA, 1998, 385-391 p.
- [WOO, 98b] Michael Wooldridge, Nicholas R. Jennings, « Pitfalls of Agent-Oriented Development », In: *Proceedings of the 2<sup>nd</sup> Conference on Autonomous Agents (Agents-98)*, Minneapolis, USA, 1998, 385-391 p.
- [WWW, 98] Wurman P. R., Wellman M. P. and Walsh W. E. The Michigan Internet AuctionBot: A Configurable and Auction Server for Human and Software Agents, In 2

- [Zar, 01]. International Conferences on Autonomous Agents, Minneapolis, USA, May 1998.  
W. Zareba, P. Maison-Blanche and H. L. Locati, *Noninvasive electrocardiology in clinical practice*. New-York: Futura Publishing Compagny, Inc., 2001.