



HAL
open science

Décision et Cognition en Biomédecine : modèles et Intégration

Vincent Rialle

► **To cite this version:**

Vincent Rialle. Décision et Cognition en Biomédecine : modèles et Intégration. Sciences du Vivant [q-bio]. Université Joseph-Fourier - Grenoble I, 1994. tel-00201144

HAL Id: tel-00201144

<https://theses.hal.science/tel-00201144>

Submitted on 25 Dec 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rapport pour l'obtention
du
Diplôme d'Habilitation à Diriger des Recherches

Titre :

Décision et Cognition en Biomédecine :
Modèles et Intégration

Présenté par :

Vincent Rialle

Maître de conférences en informatique

Soutenu le 1^{er} Juillet 1994

Jury :

Pr J.	DEMONGEOT	<i>Président</i>
Pr M.	FIESCHI	<i>Rapporteur</i>
Pr F.	RECHENMANN	<i>Rapporteur</i>
Pr J.-R.	SCHERRER	<i>Rapporteur</i>
Pr. G.	TIBERGHEN	

Laboratoire TIMC-IMAG

Table des matières

Avant-propos	5
Introduction - Présentation générale de la thématique de recherche.....	7
I. ACTIVITE DE RECHERCHE	31
I.1. Présentation générale de l'activité de recherche	31
I.2. Liste détaillée des thèmes de recherche	33
I.3. Développement des principaux thèmes	39
I.3.1 Premier thème : Modélisation et apprentissage de connaissances dans un univers vaste, incertain et incomplet.....	39
I.3.2 Deuxième thème : Représentation hétérogène de connaissances (Electromyographie) 50	
I.3.3 Troisième thème : Modélisation cognitive : Rapport final de recherche pour le séjour d'une année de Délégation au Québec.....	61
I.3.4 Quatrième thème : Epistémologie de la modélisation cognitive : coordination d'un numéro de revue philosophique	65
I.3.5 Cinquième thème : Modélisation hybride (symbolique et connexionniste) de la compréhension automatique de textes	67
I.3.6 Sixième thème : Modélisation cognitive en psychiatrie : des modèles symboliques aux parallèles et distribués.....	71
II. PUBLICATIONS	89
II.1. Revues nationales	89
II.2. Revues internationales.....	89
II.3. Congrès nationaux et principales communications locales	89

II.4. Congrès internationaux.....	90
II.5. Posters de congrès.....	92
II.6. Mémoires, rapports, thèse.....	92
II.7. Contributions à des ouvrages	93
II.8. Edition d'ouvrages	93
II.9. Soumission en cours	93
III. PERSPECTIVES ET PROJETS DE DEVELOPPEMENT.....	95
III.1. Projet MIX, ESPRIT-III	95
III.1.1 Présentation	95
III.1.2. Contributions spécifiques de l'équipe grenobloise	96
III.2. Projet COGNIMED	97
III.3. Projets franco-québécois	99
III.3.1 Textes et réseaux neuronaux	99
III.3.2 Applications en psychiatrie	100
IV. CONCLUSION	103
ANNEXE - Activité d'encadrement d'étudiants de 3ème cycle.....	105

Avant-propos

Ce rapport est structuré en quatre parties principales et une annexe.

- Une partie introductive présentant de manière synthétique la thématique générale de recherche ainsi qu'un bref état de l'art du domaine dans lequel se placent mes travaux. La présentation des problèmes et des orientations de ce domaine sera appuyée par une bibliographie propre à l'introduction.
- La première partie est consacrée à la description des recherches qui ont été effectuées depuis une dizaine d'années. Outre la description des thèmes de recherche, cette partie inclut la présentation de quelques articles et résumés de travaux.
- La deuxième partie donne une liste complète et structurée des publications, communications, posters, etc.
- La troisième partie, orientée vers le futur, est consacrée aux projets imminents et aux perspectives de recherche et de développement à moyen terme.
- L'annexe présente l'activité d'encadrement d'étudiants de troisième cycle, directement liée aux activités d'enseignement et de recherche.

Introduction - Présentation générale de la thématique de recherche

Cette introduction se donne pour but de présenter de manière synthétique la thématique générale du domaine dans lequel s'inscrit l'axe principal de mes travaux. Cet axe est constitué par la conception de systèmes capables de produire automatiquement une certaine *connaissance* et de la combiner éventuellement avec une connaissance *a priori* donnée par un expert, l'ensemble étant *situé* dans le domaine de la santé. Cet axe constitue mon premier domaine de production scientifique et technologique et d'activité d'enseignement et d'encadrement d'étudiants depuis de nombreuses années, comme le suggéreront cette introduction et les chapitres suivants. Il n'est cependant pas unique étant donné que je me suis intéressé (comme en témoigneront les descriptions de travaux et de projets des parties I à III) à d'autres aspects de l'intelligence artificielle, tels que la place de celle-ci dans la modélisation des processus cognitifs ou dans le domaine de la santé mentale. Il était nécessaire cependant d'unifier cette présentation autour d'un thème productif central. Ce thème sera celui de *l'émergence* et de la *construction automatique de connaissances*.

Par souci de clarté, cette introduction se compose de plusieurs paragraphes portant chacun un titre, et qui permettent une progression graduelle dans la présentation de la thématique. Cette progression débutera par l'exposé du thème et de son contexte (points a et b) et des relations qu'entretient ce thème avec le domaine des sciences de la cognition (point c). Sera ensuite abordée la dichotomie apparente de l'intelligence artificielle entre représentation et émergence de connaissances (points d, e, f et g). Divers aspects de la problématique générale de la construction de classifieurs seront exposés aux points h, i, j et k, le rôle central de l'expert, en l'occurrence du médecin, étant décrit au point j. Cette progression culminera dans le thème de la construction de systèmes hybrides (points l et m) et s'achèvera sur un paragraphe final.

La présentation des problèmes et des orientations du domaine sera appuyée par une bibliographie située à la fin de cette introduction. Cette présentation revêtira à certains endroits la forme d'une argumentation en faveur des orientations retenues, plutôt que celle d'une liste uniforme des divers aspects des questions traitées.

a) Le thème

Le thème central autour duquel s'organisent mes activités de recherche depuis une douzaine d'années est constitué par la modélisation informatique des processus de raisonnement et de décision, avec application majeure à l'expertise médicale. Ce thème s'inscrit dans le champ de l'intelligence artificielle (IA) au sens large du terme, c'est-à-dire l'ensemble des techniques "computationnelles" qui concourent à la conception de systèmes artificiels capables de simuler des phénomènes de perception, de raisonnement et d'action et qui gagnerait à être re-nommé sous le terme d'*autonomie artificielle* (terme de plus en plus employé et qui nous paraît plus conforme à la réalité que celui d'IA). Il s'inscrit donc également dans le champ plus large des *sciences de la cognition*.

Un tel domaine étant très vaste, cette introduction mettra l'accent sur l'orientation vers des systèmes dotés de capacités d'apprentissage et d'unification entre connaissances déclaratives et connaissances apprises à partir d'une expérience concrète. Orientation qui reflète plus particulièrement mes travaux de ces dernières années.

Le domaine médical est très présent dans l'ensemble des recherches, sans toutefois en constituer l'unique champ d'application (*e.g.* (Rialle, 1994)). Il représente plutôt une source quasi inépuisable de problématiques décisionnelles plus ou moins complexes et tient souvent lieu de banc d'essai ou d'épreuve de vérité des modèles étudiés. Il constitue également le domaine dans lequel j'évolue journallement depuis 14 ans et que je connais donc particulièrement. Il faut également ajouter que le champ médical opère ici une sorte de *réduction* de la problématique générale des systèmes à base de connaissances et, qui sans cela, serait trop vaste à appréhender. Par exemple, la *représentation* concerne ici des formes verbales (symptômes, maladies...) et non pas toute forme de représentation envisagée par les sciences de la cognition ou la philosophie de l'esprit. De même, l'apprentissage et l'émergence de connaissances se fera à partir de fichiers très structurés comme l'on en rencontre constamment en médecine, et non pas à partir de récits libres ou encore d'expériences sensori-motrices comme en robotique.

Cependant, l'une des difficultés majeures que rencontre l'entreprise de conception et de réalisation de systèmes d'aide à la décision capables d'apprentissage est la multiplicité, voire la profusion, des approches théoriques et d'algorithmes que l'on rencontre aujourd'hui (cf. points *h* et *i* notamment). Cette entreprise ne saurait cependant faire l'économie d'une étude de ces diverses approches, sous peine d'ignorer entièrement certaines voies de recherches productives et d'en surestimer d'autres. La contrepartie de cette attitude est l'apparent éclatement des travaux dans des directions différentes et la difficulté d'approfondir l'une de ces directions sans délaisser les autres. Mais c'est aussi le défi de l'intégration de ces approches dans des architectures logicielles complexes (cf. points *i*, *k* et *l*), défi qui, s'il n'est pas relevé, risque de laisser la recherche en matière d'algorithmes d'apprentissage et d'adaptation dans une certaine *schizophrénie*. L'honneur d'opérer cette synthèse revient principalement à l'informatique et au génie logiciel qui ont ici pour première mission de concevoir des outils capables d'assumer correctement cette intégration. Il n'est pas inutile de souligner que ces problèmes d'intégration ont des aspects très complexes, comme en témoignent les projets COGNIMED et MIX présentés dans ce mémoire.

b) L'arrière-plan du thème

Le développement de l'utilisation de l'ordinateur en ce point sensible de l'activité médicale que constitue la décision (diagnostique, thérapeutique, pronostique) est pratiquement continue depuis l'aube de l'informatique. Quoique toujours difficile à mettre réellement en œuvre dans une activité clinique régulière (Rialle, 1989), le système informatique d'aide à la décision n'en reste pas moins le lieu d'une intense activité de recherche (en ce qui nous concerne, voir les thèmes I.2.1, I.2.2, I.2.4, I.2.5, I.2.7, I.2.8, I.2.9, I.2.14, I.2.18 et I.2.19). L'entreprise consistant à programmer une machine afin qu'elle produise des raisonnements habituellement attendus d'un spécialiste met au premier plan quelques difficiles questions relatives au *pourquoi* et au *comment* de tels systèmes. Rappelons

ici qu'une littérature scientifique très abondante a été produite depuis plus de trente ans dans le domaine de l'IA en général, et depuis plus de quinze ans dans les applications de l'IA à la Médecine (IAM). Le caractère abondant de cette production scientifique tient essentiellement au fait que le domaine de la santé offre un champ d'application *a priori* pertinent aux recherches, et souvent aux "essais et erreurs", de l'IA.

c) L'orientation dans le sens des sciences de la cognition

On constate cependant une orientation d'évolution marquée dans le sens des *sciences de la cognition* ou *sciences cognitives* (e.g. : (Tiberghien, 1986; Andler, 1987; Patel et Evans, 1989; Varela, 1989; Andler, 1992; Beuscart et Fieschi, 1992; Tiberghien, 1993)). Si un phénomène de mode n'est sans doute pas totalement absent de cette évolution, celle-ci découle d'une manière naturelle du rôle de plus en plus prépondérant que prennent les systèmes artificiels d'apprentissage de connaissances en IAM. Le constat sinon d'échec du moins d'immenses difficultés de mise en œuvre des systèmes à bases de connaissances déclaratives dans le domaine spécifique de l'aide au diagnostic médical¹, ouvre la porte à un retour en force de *l'expérience* et de la *mémoire cumulée de l'activité décisionnelle* du praticien au cours des mois et des années de pratique. Dans cette optique, l'expérience, emmagasinée sous forme de *bases de cas* (BCa) dans la mémoire de l'ordinateur, prend le pas, dans la construction d'un système d'IAM, sur la connaissance figée et laborieusement élaborée dans une base de connaissances (BCo) .

Cette introduction tentera de préciser ce passage de la représentation à l'émergence de connaissances qui consacre en quelque sorte l'immersion de l'IAM dans les sciences de la cognition.

d) Deux axes fondamentaux de l'IA

Favorisée par la dissémination des micro-ordinateurs dans les lieux de recherche et d'activité (un C.H.U. possède tout particulièrement cette double vocation d'activité d'expertise et de recherche), une floraison d'approches théoriques et algorithmiques du raisonnement automatique — visant à reproduire divers domaines d'activité de l'intelligence humaine et recherchant une forme de consécration à travers des applications pertinentes — a lieu depuis une trentaine d'années. Deux axes centraux se dessinent avec une précision désormais manifeste :

- d'une part, l'approche *représentationniste* qui se propose de modéliser le raisonnement à partir de connaissances directement énoncées — stipulées — par un expert et explicitement représentées dans l'ordinateur.

Rappelons que le problème de la *représentation* en intelligence artificielle est des plus complexes (Newell, 1980)(Meunier, 1992)(Winograd, 1993). Nous ne saurions le traiter ici dans son ensemble,

¹ Ce domaine n'inclut pas ceux de la recherche sur le génome ou de la cartographie cérébrale, dans lesquels de grandes bases de données et connaissances sont actuellement construites, mais qui ne se situent pas directement dans le champ de l'aide à la décision et aux soins médicaux.

c'est pourquoi nous entendrons simplement par *représentation*, aussi bien l'*activité* consistant pour un expert à expliciter son savoir — en utilisant la langue naturelle et diverses aides telles que schémas, dessins, etc. — que l'aboutissement de cette activité consistant en une *re-présentation* de ce savoir dans un langage formel d'expression de connaissances. Il reste bien entendu que la représentation, au sens strict du terme, est omniprésente dans les productions de l'informatique : tout y est *re-présenté* sous forme binaire au niveau du logiciel. Mais la définition que nous en avons retenue se place uniquement sur le plan de la *modélisation* des processus de décision.

Cette représentation emprunte les formalismes préétablis — classiques en intelligence artificielle — d'accueil de connaissances (tels que règles de productions, schémas, scripts, réseaux sémantiques...) (Nilsson, 1988)(e.g. dans le domaine médical : (Fieschi, 1984; Rialle *et al.*, 1991)). Les connaissances qu'ils permettent d'accueillir peuvent être de type livresque, encyclopédique, académique, ou représenter un savoir-faire pratique, une technique, des habitudes de raisonnement propres à un expert, et peuvent conduire à une *classification heuristique* (Clancey, 1985) (à ne pas confondre avec la classification par apprentissage développée ci-après). Elles sont dans tous les cas le résultat d'une *stipulation*, c'est-à-dire d'un acte d'autorité de la part d'une personne compétente. Cette orientation résulte directement du courant cognitiviste qui a longtemps prévalu dans les sciences de la cognition. Ce courant postule (Varela *et al.*, 1991) que la cognition peut être définie comme une *computation*, c'est-à-dire comme une série d'opérations définies ou accomplies sur des symboles ou des valeurs numériques. L'intelligence artificielle, en tant que discipline de la modélisation des connaissances et des raisonnements sous forme de symboles régis par des systèmes formels, en est son principal fleuron. Elle a produit notamment un nombre considérable de systèmes dits "experts". Production qui a conduit cependant à un nombre non moins considérable de désillusions quant à la "faisabilité" réelle, la maintenance, l'évaluation, et l'utilisation concrète de ces systèmes (en particulier ceux dits de *première génération*) ainsi qu'à leur rentabilité socio-économique dans le domaine biomédical.

- d'autre part, l'approche quelquefois qualifiée d'*émergentiste* et dans laquelle la connaissance, le comportement et l'adaptation sont le produit d'une expérience et se produisent au cours d'un apprentissage.

Le terme *émergence* est en général réservé aux systèmes composites (réseaux d'automates (Demongeot *et al.*, 1985)), dans lesquels un certain nombre d'unités de calcul (automates, *neurones formels*) travaillent en parallèle. Chacune d'elles possède son propre comportement consistant à calculer à tout instant l'état (actif ou passif par exemple) qui doit être le sien, en fonction des données qui lui sont fournies par ses unités voisines. Le résultat de ce calcul, l'état de l'unité, devient à son tour une donnée disponible à son voisinage. L'activité simultanée de tous ces processus locaux donne lieu à un comportement global qui est dit *émerger* de ce concert d'activités locales. Dans les systèmes dits *cognitifs*, ces comportements de *haut* niveau — niveau symbolique — possèdent leur propre "syntaxe", leurs propres règles d'interrelations, sans référence apparente au niveau local (sub-symbolique). Ici se pose le problème crucial en sciences de la cognition des liens de continuité entre l'activité symbolique de la pensée et une activité sub-symbolique qui plonge ses

racines dans le substrat neuronal (Changeux, 1982; Smolensky, 1986; Varela *et al.*, 1991; Dennett, 1993).

Par émergence nous entendons la faculté d'obtenir un dispositif (en l'occurrence un programme d'ordinateur (Forrest, 1991)) doté de capacités de classification à partir d'un ensemble représentatif de données d'observation, sans que cette faculté *ne soit à aucun moment explicitement pré-définie*. D'un point de vue pratique et algorithmique, les théories et les systèmes proposés dans cette voie opèrent à partir d'un ensemble de cas soigneusement analysés et colligés dans la base de cas. Les renseignements permettant une organisation globale des données selon des classes ne reposent nulle part ailleurs que dans les données elles-mêmes (voir la structure d'une base de cas définie au point *h*). La capacité décisionnelle émerge de celles-ci grâce à des mécanismes spécifiques d'apprentissage, d'auto-organisation, de construction d'arbres de décision, etc., issus de diverses approches théoriques (abordées au point *i*).

Précisons de plus que la notion même d'émergence est, jusqu'à un certain point, indépendante de celle de représentation : des systèmes cognitifs naturels et artificiels peuvent en effet former une "connaissance" de leur environnement au cours de leur évolution sans nécessairement construire des *représentations* internes (Brooks, 1991; Varela *et al.*, 1991).

e) Problèmes de la représentation

La contrainte de la représentation dans les systèmes à base de connaissances (SBCo) est une contrainte très forte qui apparaît peu satisfaisante à plusieurs égards :

- d'une part, elle oblige l'expert à *reproduire* et à rendre explicite son propre modèle cognitif, ce qui va de toute évidence bien au-delà de la simple énonciation d'une série de règles de production fournies à un moteur d'inférence général (qu'il soit, selon les taxonomies de l'IA, d'ordre zéro, zéro-plus ou un, opérant en chaînage avant ou arrière, en logique monotone ou non-monotone, etc.). L'élaboration d'une base de connaissances dans un domaine tant soit peu complexe (c'est-à-dire digne d'un tel investissement) est un processus de longue haleine, nécessitant une expertise non seulement du domaine, mais aussi de l'acquisition et de la représentation des connaissances (travail de *l'ingénieur cognitif*, cf. I.2.8), et enfin des technologies de la connaissances (*knowledge engineering tools*), pour un résultat souvent incertain et coûteux.
- d'autre part, elle rend le modèle créé *dépendant*, le plus souvent, de son créateur (elle a en quelque sorte constamment *besoin* de lui pour sa mise à jour et sa maintenance).

Sans les supprimer entièrement, les approches de l'émergence permettent de *desserrer* considérablement ces contraintes, notamment :

- en diminuant notablement le long processus d'élaboration "manuelle" des représentations explicites *a priori* (une représentation explicite *a posteriori* peut néanmoins être créée, *e.g.* algorithmes de production d'arbres et règles de décision).
- en ajoutant une *mobilité* considérable dans le processus de construction du classifieur. Il est en effet souvent plus aisé de rajouter des cas dans une BCa que des connaissances explicites dans un SBCo, ces cas pouvant émaner de divers spécialistes (alors qu'il est difficile de construire une

représentation explicite des connaissances pratiques de plusieurs experts à la fois, les *représentations* issues de l'expérience variant considérablement d'une personne à l'autre). Après mise à jour de la BCa, il suffit d'appliquer à nouveau l'algorithme d'apprentissage (non sans avoir vérifié que les conditions d'applications de l'algorithme restaient remplies) pour obtenir un nouveau classifieur (qu'il reste ensuite à évaluer).

f) Problèmes de l'émergence

Cependant la voie de l'émergence ne va pas sans rencontrer d'importantes difficultés qu'il est nécessaire de souligner ici. Sans entrer dans le débat qui oppose la représentation et l'émergence dans les sciences de la cognition, nous soulignerons simplement que les classifieurs dont nous disposons à l'heure actuelle ne produisent pas encore une connaissance structurellement et fonctionnellement complexe :

- les méthodes affiliées à l'IA permettent tout au plus d'obtenir des listes de règles d'inférences ou des arbres de décision (aisément réductibles à des ensembles de règles) qui restent des structures de représentation élémentaires;
- les méthodes connexionnistes conduisent le plus souvent à des classifieurs de type *boîte noire* (dans l'état actuel de l'art) dans lesquels aucune structure explicite de connaissance ne transparait;
- l'analyse des données multivariée (analyse discriminante, analyse de régression,...) conduit à des résultats interprétables par un spécialiste de l'analyse multivariée et en général aisément traduisibles sous forme de règles de production (le système MÉNINGE en est un exemple (Robert, 1989; François *et al.*, 1990))

Il faut donc souligner que ces "connaissances artificielles", même si les techniques qui concourent à les produire se situent dans le vaste domaine de l'apprentissage automatique (*machine learning*), ne constituent pas un modèle cognitif de l'activité d'un expert. En effet, pour qu'un ensemble de connaissances et de mécanismes d'inférence puisse prétendre au statut de modèle cognitif, il doit pouvoir rendre compte des multiples mécanismes cognitifs mis en œuvre par l'expert lorsqu'il étudie un cas, à commencer par celui de la coexistence presque simultanée des opérations fondées sur la familiarité ("le cas présent me rappelle le cas de M. X") avec celles fondées sur l'analyse logique ("j'observe ceci, j'en déduis cela").

La difficulté de prendre en compte la temporalité d'un raisonnement illustre bien le caractère limité des structures de connaissances obtenues de manière automatique. La quasi totalité des applications médicales effectives, c'est-à-dire implémentées sur ordinateur et validées, ne prend en compte que des *informations immédiatement disponibles*. Par informations immédiatement disponibles, nous entendons des informations préenregistrées dans un fichier accessible par le système décisionnel. Le problème de la modélisation d'une *démarche inscrite dans le temps*, telle que la demande puis l'attente de résultats d'examens complémentaires, la prise en compte de la *variation* de dosages au cours du temps, la génération de plans d'action, etc., relève autant du domaine de la robotique médicale ou du monitoring que de celui de l'aide à la décision proprement dite (un système de

monitoring par exemple pourra être doté de modules de décision sur informations immédiatement disponibles).

Ici se pose la question de l'interactivité du système : faut-il persister à vouloir construire des systèmes de type anthropomorphique, capables de dialoguer avec des utilisateurs (en posant des questions lorsque nécessaire au raisonnement et en répondant à des demandes d'explications de la part de l'utilisateur), apanage des systèmes experts classiques ? On trouve de toute évidence, dans la capacité de dialogue et d'explication en langage clair des raisonnements, une supériorité de l'intelligence artificielle symbolique sur les méthodes connexionnistes (Kodratoff, 1989). Néanmoins, les problèmes méthodologiques et déontologiques posés par la réalisation de tels systèmes restent encore largement ouverts (Patel et Evans, 1989).

Enfin, une autre limitation des méthodes d'apprentissage automatique réside dans la difficulté à prendre en compte les *valeurs manquantes*. Cette prise en compte est très difficile à systématiser comme le voudraient la plupart des méthodes d'apprentissage automatique, alors qu'une connaissance explicite édictant un raisonnement ou une réponse diagnostique particulière en l'absence de tel ou tel renseignement peut être aisément rendue dans une base de connaissances.

g) De la nécessaire complémentarité des approches

On aura compris à travers les paragraphes précédents que les deux approches, représentationniste et émergentiste, ne sont pas fondamentalement opposées. Elles possèdent chacune leurs intérêts et leurs limites et correspondent plutôt à des attitudes complémentaires au plan de la phylogénèse du savoir : les connaissances académiques, consensuelles, dans un domaine quelconque d'investigation sont le plus souvent le fruit d'un long processus de production au sein d'une communauté scientifique. Une connaissance académique émerge de l'interaction de plusieurs individus, même si celle-ci est formulée, dans son expression finale, par un seul d'entre eux. Toute connaissance doit ainsi *d'abord* émerger avant d'être formulée dans un système de symboles.

La coexistence de la représentation et de l'émergence qui se manifeste au niveau collectif prend sans doute ses racines dans la structure même du système cérébral. Il est connu que l'hémisphère gauche du cerveau commande les activités rationnelles fondées sur des représentations logiques, et que l'hémisphère droit commande celles qui sont non raisonnées (sentiments, pressentiments, intuition, sens de l'esthétique, etc.). Il est également reconnu que les connaissances académiques sont souvent d'ordre général et que, particulièrement en médecine, il faut ensuite une pratique assidue de cette connaissance sur des cas concrets pour qu'elle devienne effective. Il y a donc chez l'expert coexistence entre d'une part ses connaissances générales, encyclopédiques, acquises de ses maîtres, et d'autre part la mémoire cumulée des cas qu'il a traités et qui deviennent source de progrès dans sa pratique.

Il n'est donc pas étonnant que l'IAM se soit penchée sur la coexistence de ces deux dimensions de la cognition — représentation et émergence —, et ait tenté de produire des systèmes dit *hybrides* (cf. points *l* et *m*).

h) Présentation des notions de *classifieur* et de *base de cas*

La capacité d'élaborer des catégories puis de classer des objets ou des phénomènes perçus suivant ces catégories constitue l'un des fondements de la cognition humaine (Rosch, 1978; Meunier, 1992). Il n'est par conséquent pas étonnant de constater que la plupart des systèmes de raisonnement artificiel reviennent à pratiquer une forme de classification. C'est donc autour de la notion de *classifieur*, en tant que capacité cognitive plus ou moins reproduites dans des outils logiciels, que se situent la majeure partie des travaux d'analyse décisionnelle à partir de données observées dans le domaine médical.

Quelle que soit l'orientation théorique adoptée, si le système résultant a pour but de donner une réponse de type classificatoire, nous considérerons à la suite de nombreux chercheurs, tels Weiss et Kulikowski (Weiss et Kulikowski, 1991) et Thornton (Thornton, 1992), qu'un tel système peut être rattaché à cette catégorie générale de processus dénommée "classifieur".

D'une manière pratique, construire un modèle classifieur consiste à élaborer un processus capable d'associer par lui-même, à toute description de cas, sa classe d'appartenance, cette classe pouvant aussi bien représenter une entité taxinomique (un syndrome, une maladie) qu'une décision, un comportement, etc.

La catégorie des classifieurs inclut notamment les systèmes experts à visée classificatoire et le raisonnement fondé sur des cas dans sa dimension classificatoire. Elle présente de plus des relations importantes avec les notions d'*autonomie artificielle* développées dans les sciences cognitives (Varela et Bourguin, 1992). Selon les théories qui les sous-tendent, certains classifieurs pourront être exprimés sous forme de connaissances explicites (cas de l'induction d'arbres de décision par exemple), d'autres, tels que la plupart des classifieurs neuronaux, ne donneront lieu à aucune représentation aisément repérable et interprétable (il en va ainsi notamment des *réseaux de Hopfield* (Hopfield, 1982)).

Quelques éléments sur une structure formelle simplifiée de base de cas

Un élément commun aux méthodes automatiques de construction de classifieurs est constitué par la base de cas.

On peut structurer formellement une base de cas autour de trois ensembles : O, E et K, définis de la façon suivante :

- Soit O un ensemble de m variables (souvent appelées propriétés ou attributs selon le contexte) servant à décrire les observations.

$$\mathbf{O} = \{ \mathbf{O}_j \mid j \in [1, m] \}$$

Dans le domaine médical par exemple, les variables correspondent aux entités porteuses d'informations et jugées utiles par les spécialistes du domaine pour l'évaluation de l'état d'un patient. Ces entités peuvent être l'âge, la glycémie ou la teinte de la peau, etc.

Chaque variable sert à décrire un aspect particulier de la réalité observée et possède son propre ensemble de valeurs.

Cet ensemble peut être

- lexico-sémantique (valeurs qualitatives). Exemple : pour la teinte de la peau, l'ensemble de valeurs sera par exemple {normale, bleutée, pâle, rosée, rouge},
- ou numérique (discret ou continu), telle que la température qui pourra appartenir par exemple à l'intervalle [35, 42], ou le nombre de grossesses (variable numérique discrète).

• Soit un ensemble K de p classes (appelés également, selon le domaine d'application, concepts, diagnostics, réponses, etc.) :

$$K = \{ K_s \mid s \in [1, p] \}$$

• Soit E un ensemble de n cas, chaque cas étant représenté par un vecteur de m valeurs : les valeurs respectivement prises par le cas sur les éléments de O .

$$E = \{ E_i \mid i \in [1, n] \}$$

$E_i = (O_1(E_i), \dots, O_m(E_i))$, soit encore :

$$E_i = (a_{i1}, \dots, a_{im})$$

On appelle *cas pré-classé* un couple (E_i, k_i) dans lequel E_i représente un cas et k_i représente la classe qui lui a été attribuée (en général par un expert).

Une base d'apprentissage est une base de cas composée de cas pré-classés.

N.B. Les cas sont également appelés, suivant le domaine : stimuli, entrées, intrants, données, etc.

On peut schématiquement représenter (Fig. 1) un ensemble de cas dotés de leur classe par un tableau de n lignes représentant les cas, et de $m + 1$ colonnes représentant respectivement les m variables du domaine et la classe.

L'élément a_{ij} désigne la valeur sur la variable O_j du cas E_i , et l'élément k_i désigne la classe du même cas.

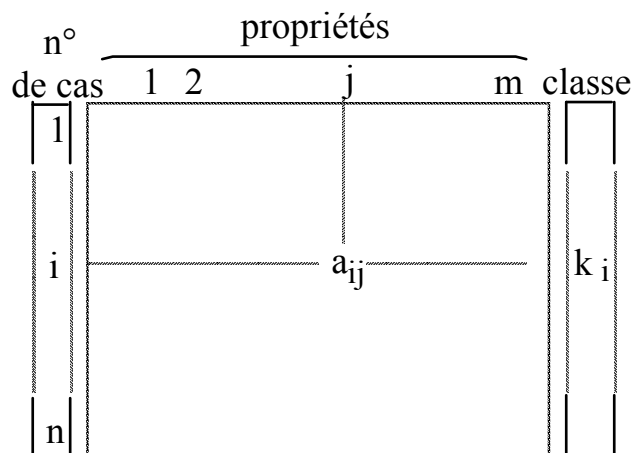


Fig. 1 : Représentation schématique d'une base de cas pré-classés.

Remarques :

Cette formalisation correspond à une simplification d'une réalité phénoménologique qui peut être beaucoup plus complexe. Par exemple, des variables continues peuvent être "vécues" comme discrète sur certains sous-domaines ou dans certaines conditions (e.g. l'âge, la température corporelle,...). Il peut y avoir plusieurs diagnostics possibles pour un cas donné, sans possibilité de trancher en faveur de l'un d'entre eux. Etc.

De plus le formalisme de base de cas qui viens d'être exposé ne couvre pas les problèmes de **représentation de raisonnements** ou de **représentation d'histoires** (comme le voudrait notamment le raisonnement fondé sur des cas (Riesbeck et Schank, 1989)). Enfin, ce formalisme répond mal aux bases de cas dont le nombre de variables prises en compte varie d'un cas à l'autre (e.g. le domaine de l'électromyographie).

La construction automatique d'un classifieur suppose donc deux composantes fondamentales (Fig. 2):

1°) une base de cas pré-classés qui soit le plus possible représentative du domaine,

2°) un algorithme (issu d'une théorie) de construction du classifieur.

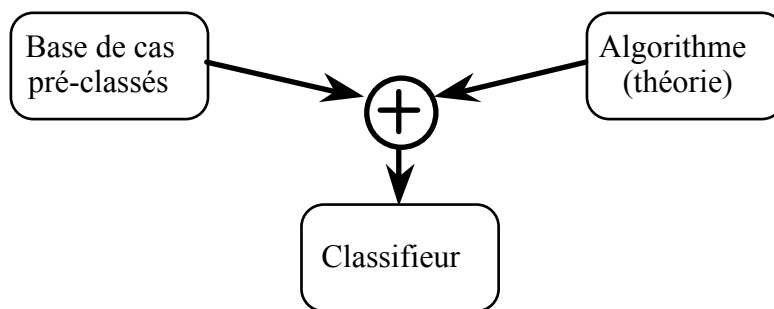


Fig. 2: Composantes fondamentales intervenant dans la construction d'un classifieur.

L'algorithme est appliqué sur la base de cas pour produire le classifieur. On peut également considérer, à la suite de C.J. Thornton (Thornton, 1992) que l'algorithme recherche, dans l'espace des hypothèses de classifieurs, celui qui satisfait le mieux les contraintes constituées par la base de cas (*i.e.* celui qui classera correctement le plus grand nombre de cas).

Phases d'élaboration :

La construction automatique du classifieur se fait alors selon deux phases (Fig. 3) :

1°) la **phase d'apprentissage** : il s'agit dans cette phase de construire le classifieur (*i.e.* le processus) capable de classer **par lui-même** les exemples de la base de cas (processus de classification). Cette construction utilise un algorithme issue d'une théorie particulière (analyse des données, induction automatique, algorithmique génétique, etc.). Le classifieur résultant peut revêtir

des formes aussi différentes qu'un ensemble de règles de production (muni d'un moteur d'inférences), un réseau de neurones, un arbre de décision exprimé dans un langage algorithmique, etc.

2°) la **phase d'utilisation du modèle**, dans laquelle le classifieur est mis à contribution pour déterminer la classe d'un cas qui lui est présenté (processus de prédiction). Le préalable à toute utilisation courante d'un classifieur est **l'évaluation** de ses performances selon une méthodologie rigoureuse (évoquée au point *i*).

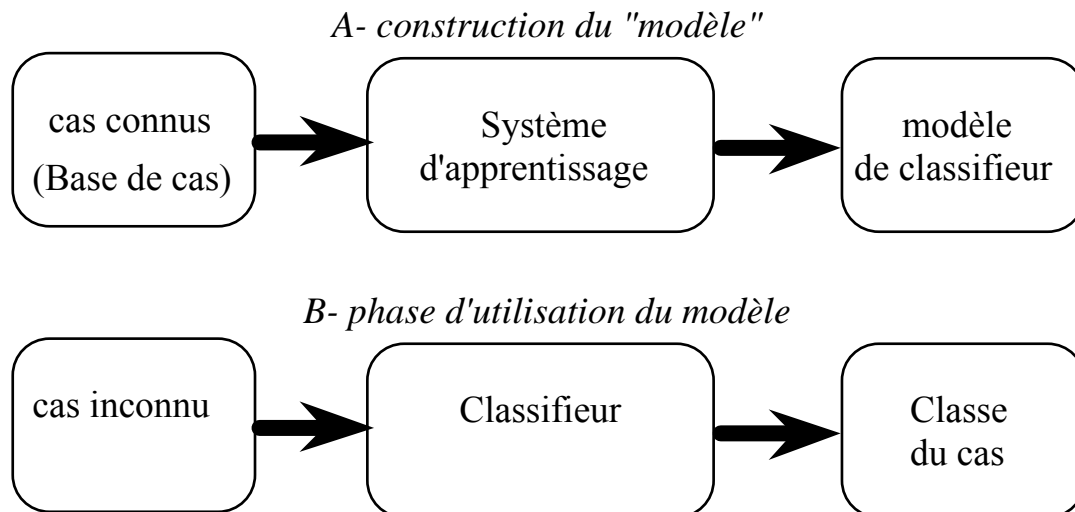


Fig. 3 : Phases de mise en œuvre d'un classifieur.

i) **Richesse et diversité des origines théoriques des classifieurs**

La notion de classifieur ainsi que les problèmes d'implémentation logicielle des divers types de classifieurs constituent un important pôle de recherche en informatique et intelligence artificielle. En témoignent les récents projets ESPRIT qui ont été consacrés à ce thème :

- le projet STATLOG (*Comparative testing and evaluation of statistical and Logical Learning Algorithms on large-scale applications for classification, prediction and control*) (Réf. 5170) (Henery *et al.*, 1991);
 - le projet ILP (*Inductive Logic Programming*) (Réf. 6020) (CCE, 1992);
 - le projet MLT (*Machine Learning Toolbox*) (Réf. 2154) (MLT-Consortium, 1993)
- et le projet MIX (*Modular Integration of Connectionist and Symbolic Processing in Knowledge-Based Systems*) (Réf. 9119) qui vient tout juste de débiter (février 1994) et auquel nous participons (cf. point *m*).

Les méthodes de construction automatique de classifieurs sont d'origines diverses. Certaines d'entre elles s'inscrivent dans le courant de l'intelligence artificielle *classique* (Kodratoff, 1988; Kodratoff et Diday, 1991) et proposent un apprentissage symbolique automatique par des méthodes de généralisation et de spécialisation (Michalski et Chilausky, 1980; Michalski, 1983; Ganascia, 1992)

ou par construction d'arbres de décision à partir de critères tels que l'entropie de l'information de Shannon (Quinlan, 1986). D'autres sont fondées sur des réseaux de neurones formels et s'inscrivent donc dans le courant connexionniste (Rumelhart *et al.*, 1986), complémentaire ou concurrent de l'IA classique (Memmi, 1989; Bournet *et al.*, 1991). Le courant probabiliste fournit principalement des classifieurs bayésiens (Salamon *et al.*, 1976; Cheeseman *et al.*, 1988). D'autres méthodes s'inscrivent dans le courant de l'analyse statistique multivariée (Robert, 1989; Robert, 1991) ou de la régression qualitative (Breiman *et al.*, 1984). Enfin, le courant de l'algorithmique génétique (Holland, 1975; Goldberg, 1989) apporte depuis quelques années une contribution substantielle à la construction de classifieurs et de bases de connaissances (Bonelli et Parodi, 1991) (Venturini, 1992).

Quelle que soit la filière théorique de construction du classifieur se posent deux problèmes essentiels, que nous n'aborderons pas en détail ici mais qui sont inhérents à nos projets :

- a) la prise en compte de l'incertitude liées aux données et aux décisions, particulièrement dans le domaine médical (Robert, 1991);
- b) l'évaluation et la validation des systèmes décisionnels. De nombreux travaux ont été conduits dans ce domaine. On lira notamment : (Fieschi et Joubert, 1985; François, 1990; Weiss et Kulikowski, 1991).

Voici un bref rappel de trois grandes sources théoriques de production de classifieurs.

1.1 L'induction automatique d'arbres et règles de décision

L'objectif de l'induction automatique est la production d'arbres et règles de décisions à partir d'exemples de situations (Quinlan, 1986; Crémilleux, 1991; Kodratoff et Michalski, 1993; Michalski *et al.*, 1993). Etant donné un ensemble disjoint de classes, un arbre de décisions est une structure récursive permettant d'exprimer le processus de classification séquentielle par lequel un "objet", décrit par un ensemble de propriétés, est assigné à l'une de ces classes. Les algorithmes d'induction automatique sont relativement nombreux et peuvent être classés en plusieurs "familles". Parmi celles-ci, la plus utilisée actuellement est la famille TDIDT (Top-Down Induction of Decision Trees). Un algorithme de cette famille permet de construire un arbre de décision optimal dont les nœuds sont les propriétés, les arcs représentant les valeurs prises par les nœuds, et les feuilles les diagnostics. Le choix de chacune des variables constituant les divers nœuds de l'arbre repose sur un critère de coût ou de bénéfice telle que la diminution d'entropie de l'information ou encore la maximisation du test du Chi2 avec la variable "classe".

L'arbre une fois construit est utilisé pour classer des cas décrits selon le même ensemble de propriétés. Il suffit alors, pour un cas donné, de rechercher dans l'arbre le chemin qui le décrit et d'en déduire sa classe.

1.2 Le connexionnisme

Le paradigme connexionniste est fondé sur :

- une modélisation mathématique du neurone (*neurone formel*) sous la forme d'une unité élémentaire de traitement. Cette unité possède plusieurs *entrées* d'information, un *état* qui la caractérise à tout instant, et une fonction interne, appelée *fonction de transfert*, *fonction d'activation*

ou *fonction neurone*, qui calcule l'état dans lequel doit être mis le neurone en fonction des entrées. Les états prennent leur valeur la plupart du temps dans un ensemble binaire (état actif, état passif) mais peuvent également être calculés sur une échelle continue;

- la construction de *réseaux de neurones formels* (RNF) composés d'un ensemble d'unités de traitement, sur l'étude des propriétés émergentes de ces réseaux ainsi que sur l'étude des divers types de réseaux et de leur adéquation à divers types de comportements cognitifs ou de problèmes computationnels (classification, catégorisation, reconnaissance de forme, apprentissage, etc.). La plupart de ces réseaux sont *auto-adaptatifs*, modifiables par *entraînement*, et sont naturellement *massivement parallèles*.

Plusieurs types de réseaux neuronaux donnent des résultats intéressants en matière de classification (Weiss et Kapouleas, 1989; Shavlik *et al.*, 1991). Nous nous intéresserons plus particulièrement au réseau neuro-mimétique de Kohonen fondé sur un principe d'*auto-organisation*.

i.3 L'auto-organisation

L'auto-organisation constitue l'un des thèmes majeurs du connexionnisme (Carpenter et Grossberg, 1987; Kohonen, 1988). L'approche de T. Kohonen (Kohonen, 1982) par exemple consiste en une modélisation dans des termes mathématiques, algorithmiques et graphiques d'un phénomène neurologique de mieux en mieux connu : celui de l'*engrammation* d'informations sensorielles dans différentes zones du cerveau. Le modèle qu'il propose est une généralisation de la notion de projections topographiques directes entre deux structures, connue sous le nom de *retinotectal mapping*. La justification biologique de ce modèle repose sur l'existence d'aires neuronales spécifiques à chaque type de stimulation sensorielle. La surface de ces aires est plus ou moins proportionnelle à la fréquence ou à l'intensité des stimulations: c'est le cas par exemple: a) des projections rétino-topiques, somat-topiques et tonotopiques sur les premières aires sensorielles ; b) des aires somat-topiques du cortex moteur.

Une aire corticale correspondant à une stimulation sensorielle donnée, par exemple le pouce de la main droite, sera d'autant plus étendue que la stimulation sensorielle sera fréquente, et sa surface aura tendance à s'accroître avec l'intensification du stimulus. A l'inverse, une aire corticale correspondant à une stimulation sensorielle de plus en plus rare deviendra de plus en plus restreinte. Si nous pouvions observer les aires corticales somat-topiques, nous serions alors en mesure de dire si tel ou tel stimulus est fortement ou faiblement représenté.

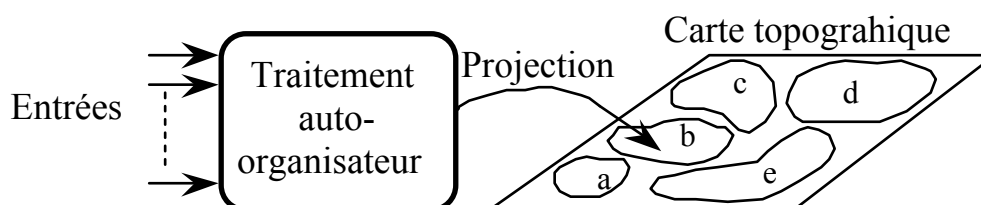


Fig. 4 : Processus d'auto-organisation : les stimuli d'entrée sont "classés" dans des zones spécifiques au terme du processus d'apprentissage auto-organisateur.

Le modèle d'auto-organisation de Teuvo Kohonen est fondé sur ce principe : plus un stimulus est fréquent, plus sa représentation sera développée sur une *carte* dont chaque point est un neurone formel. Cette carte est généralement un tableau à une ou deux dimensions d'unités de calcul (*i.e.* neurones, *processing units*) (Fig. 4). Nous décrivons ce modèle d'une manière succincte et renvoyons le lecteur intéressé par les aspects algorithmiques aux nombreuses publications qui le décrivent en détail.

Chaque cellule de la carte est reliée à une série de cellules d'entrées. Ces cellules transmettent les informations correspondant, dans une perspective neuro-mimétique, aux stimulations sensorielles (ou stimuli) et, dans la perspective d'un classifieur, aux variables de la BCa. Pour rester dans cette seconde perspective et pour la clarté de l'exposé, on peut dire que chaque neurone de la carte possède une *mémoire* pouvant mémoriser un cas (cette mémoire est inscrite dans ce que l'on nomme les *liaisons synaptiques* entre le neurone et les cellules d'entrée). Une fonction de ressemblance permettra, au cours de l'algorithme, de dire si tel neurone a dans sa mémoire une ressemblance avec tel ou tel cas (cette fonction n'est autre, la plupart du temps, qu'une simple distance euclidienne entre le vecteur d'entrée et le vecteur *mémoire* du neurone).

Avant le processus d'auto-organisation, on introduit dans la mémoire de chaque neurone, des valeurs aléatoires. Il n'a donc au début du processus aucun lien de ressemblance avec ses voisins et ne représente donc aucune catégorie particulière de cas (*i.e.* c'est l'état d'inorganisation originel).

Le processus d'auto-organisation va consister à *rendre* ces neurones de plus en plus représentatifs de cas précis de la base d'apprentissage et à faire en sorte que des cellules d'une même zone de la carte aient dans leur mémoire des valeurs peu éloignées les unes des autres. On dit également, de manière imagée, que ce processus *apprend* à chaque neurone à réagir de manière sélective aux stimuli auxquels il est le plus sensible, et à *influencer* dans le même sens ses voisins. De sorte qu'à l'issue du processus d'apprentissage auto-organisateur, les neurones d'un même voisinage auront tendance à répondre aux mêmes stimuli.

Ce modèle présente un intérêt évident en matière de classification automatique, si l'on considère que l'ensemble des informations à classer représente l'ensemble des stimuli que doit apprendre le réseau. Celui-ci va *projeter* dans des aires spécifiques de la carte les stimuli qui appartiendront à une même classe. Il restera ensuite à effectuer *l'interprétation sémantique* de la carte dans les termes du domaine d'application (attribution d'une "étiquette" à chaque neurone).

Une question actuelle consiste à savoir si l'analyse des représentations topographiques apportées des informations supplémentaires, sur des liaisons entre les stimuli notamment, par rapport à d'autres méthodes telles que l'analyse en composantes principales ou l'analyse factorielle discriminante.

Des applications médicales de la classification selon la méthode de Kohonen sont désormais nombreuses (en ce qui nous concerne: (Iordanova *et al.*, 1992; Beltrami et Rialle, 1993)).

i.4 Algorithmique génétique et générateur de diversité

A un niveau très général d'analyse, on observe que la nature offre deux grands paradigmes naturels de comportements intelligents : le cerveau et l'évolution. Le premier paradigme est à l'œuvre au niveau de l'individu, le second au niveau de l'espèce. Tous deux font l'objet de modèles d'apprentissage et de décision en sciences cognitives. Si le premier paradigme semble actuellement

le plus employé dans la modélisation cognitive, à travers notamment la recherche en pleine expansion sur les réseaux neuronaux, le second est de plus en plus étudié et appliqué dans des *algorithmes génétiques* (AG). La notion d'algorithme génétique, présentée pour la première fois en 1975 par John H. Holland (Holland, 1975), a été considérablement développée au cours des années 80 (Goldberg, 1989; Davis, 1991).

Le principe de l'algorithme est simple et repose sur un codage des problèmes et de leurs solutions sous la forme de chaînes d'éléments de base, insécables (*building blocks*), toute chaîne pouvant être rompue *entre* chaque élément de base, à l'image des chromosomes qui constituent de véritables listes de caractéristiques d'un individu. Le codage prend habituellement la forme d'une chaîne binaire très structurée, de longueur fixe ou variable selon le type de problème. On génère tout d'abord une *population* de solutions potentielles à un problème donné, sous la forme de telles chaînes, puis on sélectionne, au moyen d'une mesure d'ajustement, les éléments de la population qui satisfont le plus les contraintes (ou caractéristiques) de la solution recherchée. Des *opérateurs génétiques* sont ensuite appliqués à cette population d'*élite* de manière à obtenir une nouvelle population possédant, dans son ensemble, plus de caractéristiques de la solution que la précédente génération. On réitère le processus sur plusieurs générations, jusqu'à l'obtention d'une génération de solutions optimales parmi lesquelles il suffira de retenir la meilleure.

Plusieurs applications de l'algorithmique génétique à la construction de classifieurs ont été déjà expérimentées avec succès, à commencer par les travaux de celui qui en est le premier promoteur, John Holland. L'algorithmique génétique est maintenant utilisée dans une grande variété de domaines de recherche utilisant la simulation sur ordinateur (Davis, 1991).

Sa contribution à la modélisation de processus cognitifs est des plus remarquables : Bersini et Varela (Bersini et Varela, 1991) font remarquer que les classifieurs de Holland tout comme le système immunitaire appartiennent à une classe de processus biologiques, ou plus exactement *bio-mimétiques*, dont l'adaptabilité repose sur une *génération* ininterrompue de nouveaux *opérateurs* capables de prendre en compte une série d'événements totalement imprévisibles dans l'environnement du système. La survie et l'adaptation se paient au prix d'une *invention* perpétuelle de comportements. Invention qui nécessite en quelque sorte de se *libérer*, en présence d'un événement totalement inconnu, des comportements pré-définis et inscrits dans la mémoire du système (cette *inscription* ce fait après découverte d'une solution viable, *i.e.* d'une adaptation. L'événement et le comportement adéquat sont mémorisés et passent ainsi dans le domaine du connu). Une fonction cognitive similaire, nommée *générateur de diversité* a été introduite par Dehaene et Changeux (Dehaene et Changeux, 1989) dans un modèle de fonctionnement du cortex pré-frontal. Ce générateur est chargé de créer des pré-représentations de l'environnement à partir d'informations élémentaires, ainsi que des règles de conduites dans ces pré-représentations. Le système doit ensuite évaluer la pertinence de chacune de ses pré-représentations afin d'en sélectionner la plus adéquate pour assurer sa pérennité ou atteindre ses objectifs.

j) Place de l'expert dans le processus d'élaboration des classifieurs

La démarche de construction d'un classifieur par voie d'émergence repose sur des connaissances expertes à divers niveaux du processus (nous prendrons ici encore le cas d'un expert médical) :

- le premier niveau est celui du découpage de la réalité perçue par le praticien en éléments simples (les signes, symptômes, syndromes, maladies...). L'acte même de sélection des traits pertinents (les *variables*) qui serviront à décrire systématiquement les cas dans la BCa (âge, sexe, température, ...), est celui d'un expert, en ce qu'il requiert des connaissances pratiques et académiques approfondies du domaine d'application (la BCa ne saurait à l'évidence conduire à un modèle de décision valable si la nosographie sur laquelle elle est établie était approximative ou erronée). Cependant, de même que dans un système expert certaines règles de déduction peuvent être sujettes à controverses, des choix nosographiques pratiqués par un spécialiste peuvent ne pas être partagés par tous ses confrères. Cette sensibilité de la base de cas à son créateur peut se révéler importante dans certains domaines. Elle sera encore augmentée par l'attribution éventuelle de *coefficients de pondération* des variables, nécessaires lorsque l'on veut modéliser les différents niveaux d'influence de ces variables sur le raisonnement.

La responsabilité du spécialiste se mesure également à l'importance des erreurs de conception : un trait peut avoir été négligé lors de la création de la BCa et se révéler ultérieurement suffisamment porteur d'information (*e.g.* suffisamment discriminant) pour devoir figurer dans la liste des variables. L'ajout d'un tel trait à une base comportant déjà un certain nombre de malades est généralement impossible, étant donné qu'il est très difficile de retrouver ces malades dans l'état où ils étaient lors de l'examen afin d'en déduire sur chacun d'eux la valeur du trait ajouté.

- le second niveau auquel intervient l'expertise humaine est bien sûr celui du diagnostic. Comme indiqué précédemment, l'attribution du diagnostic à partir du tableau symptomatologique est l'acte par excellence de l'expert, acte que le classifieur devra chercher à reproduire.

Reste ensuite à aborder la question de la *place* occupée par la méthode théorique (intervenant sous la forme du programme informatique de *calcul* du classifieur à partir de la BCa) par rapport à celle du médecin dans le processus complet de réalisation du classifieur. Cette question générale consiste en particulier à savoir si cette méthode — fondée sur des principes de logique ou de mathématique rigoureux et, en cela, entièrement hors de l'influence de l'expert (hormis d'éventuels "réglages" manuels de paramètres) — constitue une contrainte, un biais méthodologique ou une gêne quelconque pour le médecin. Cette question, qui ne saurait être réglée en quelques phrases dans le cadre de ces lignes, met en jeu plusieurs niveaux de réponse. On peut se poser en premier lieu la question de la place des principes théoriques de classification par rapport à celui du sujet humain. Sans pénétrer dans ce débat, il convient de rappeler ici que la seule classification de référence prise en compte est celle du médecin (concrétisée par la BCa) et que le modèle théorique n'a d'autre objet que de tenter de la reproduire et de reproduire son incertitude (aucun système artificiel ne saurait en effet construire par lui-même une base de cas diagnostiqués, hypothèse qui rendrait d'ailleurs caduque la nécessité d'un classifieur). Cependant, ce premier niveau de la question fait immédiatement surgir un second niveau, de nature éthique et déontologique celui-là : comment la mise à disposition d'outils d'analyse décisionnelle de précision va-t-elle influencer la pratique médicale même, entre un prolongement des capacités cognitives du médecin et la menace, encore toute spéculative, d'un remplacement dans certains cas du médecin par l'outil décisionnel? Dans le

domaine de l'IAM, comme dans tout domaine développant des technologies avancées d'une impressionnante efficacité, un développement de la responsabilité et de la vigilance du médecin quant aux possibles effets pervers de ces technologies est de toute évidence d'une impérieuse nécessité. Ce n'est qu'à l'intérieur de ce cadre éthique et déontologique que peuvent être appréciés à leur juste valeur les outils cognitifs développés ici.

k) Mise en œuvre de l'émergence dans le domaine médical : le projet COGNIMED

Le projet COGNIMED (cf. III.2) a pour but d'intégrer au sein d'une même structure logicielle un certain nombre de modèles de classification automatique et d'apprentissage de connaissances ou comportements décisionnels. Il est né du double constat de l'affinement considérable de ces modèles depuis une quinzaine d'années et de l'intérêt croissant qu'ils présentent en matière d'aide à la décision médicale. Il inclut donc :

- une étude approfondie de ces modèles,
- la réalisation d'une "plate-forme" logicielle permettant leur implémentation aisée, leur expérimentation et leur évaluation comparée dans divers domaines de la médecine.

Cette plate-forme permettra notamment le choix entre divers principes de similarité entre cas, proposera divers algorithmes d'intelligence artificielle pour l'analyse décisionnelle et l'apprentissage automatique de comportements diagnostiques à partir d'exemples, et offrira des facilités pour l'évaluation des méthodes et leur comparaison. Conçu selon une approche modulaire, à la manière d'une *boîte à outils* évoluée avec des facilités d'ajouts de modules (*i.e.* nouveaux algorithmes classifieurs, nouveaux principes de similarité,...), il offrira une présentation graphique des résultats et un interface-utilisateur convivial.

Une telle plate-forme logicielle devrait s'insérer, afin de s'ajouter aux outils d'analyse mis au service du corps médical et dès lors que les conditions seront réunies, dans un système d'information hospitalier tel que celui qui est actuellement en cours de développement au CHU de Grenoble.

l) Systèmes hybrides symboliques et connexionnistes

L'un des principaux obstacles rencontrés par les modèles exclusivement fondés sur l'un des paradigmes de modélisation en cours à l'heure actuelle (traitement symbolique, connexionnisme, algorithmes génétiques, etc.) est leur manque de *robustesse*. Chacun de ces paradigmes peut s'avérer excellent dans un champ délimité d'application, mais ne constitue en quelque sorte qu'une simple pièce du puzzle des qualités nécessaires à un modèle pour approcher le domaine du *cognitif*. Plusieurs auteurs se sont mesurés à la question de la fragilité des modèles et des moyens d'y échapper (*e.g.* (Holland, 1986; Sun, 1991)).

L'approche des *systèmes hybrides symboliques et connexionnistes* (Gallant, 1988; Feldman, 1989; Hendler, 1989; Memmi, 1989; Amy *et al.*, 1990; Rialle *et al.*, 1991; Sun, 1991; Giacometti, 1992; Giacometti *et al.*, 1992) constitue l'une des réponses à ce problème du manque de robustesse des modèles. Cette approche hybride cherche à associer les qualités des réseaux neuronaux formels et celles de l'intelligence artificielle symbolique dans le but de construire des modèles dits *cognitifs*,

censés être plus performants que ne le seraient des modèles fondés sur une approche unique. Elle consiste à l'origine à faire cohabiter, au sein d'un même environnement logiciel, une représentation symbolique des phénomènes traités, avec une représentation parallèle et distribuée généralement associée au courant du connexionnisme. Diverses sortes *d'hybridation* ont été étudiées, depuis des architectures fondées sur une simple coopération entre modules connexionnistes et symboliques, jusqu'à une intégration plus ou moins complète des deux approches. Ces divers modes d'intégration de l'intelligence artificielle symbolique et des réseaux neuronaux ont été classifiés par divers auteurs (Amy *et al.*, 1990; Sun et Bookman, 1992). Sun et Bookman (1992) proposent par exemple quatre catégories : (a) le simple couplage de deux modules — l'un symbolique, l'autre connexionniste — entièrement indépendants et s'apportant mutuellement des informations; (b) les systèmes essentiellement symboliques de traitement de l'informations mais possédant des composants neuronaux (*e.g.* le système BoltzCONS proposé par (Touretzky, 1989)); (c) les systèmes entièrement connexionnistes proposant une émergence des symboles à partir des interactions neurales (*e.g.* le système RAAM de (Pollack, 1990)); (d) les réseaux connexionnistes localistes, dans lesquels certaines unités de traitement sont volontairement associées à des symboles par le concepteur du système (*e.g.* le système proposé par (Pinkas, 1992)). Citons également, dans le domaine médical, l'approche *macro-connexionniste* de Pham et Degoulet (Pham et Degoulet, 1991). Actuellement (cf. point *m*), les systèmes hybrides tendent à intégrer dans l'architecture logicielle toutes formes d'apprentissage et d'aide au raisonnement (incluant l'algorithmique génétique et le raisonnement fondé sur des cas). Dans tous les cas, ce type de logiciel est d'une mise au point délicate et relève encore largement du domaine de la recherche (cf. projet MIX, III.1). Il semble néanmoins que les systèmes hybrides constituent l'une des voies les plus prometteuses pour la modélisation de phénomènes complexes, mettant en jeu à la fois des processus d'apprentissage et d'adaptation et des représentations symboliques complexes. D'une manière générale, la recherche sur les systèmes hybrides se développe de plus en plus et tend vers la conception de véritables *machines cognitives* (Amy, 1989).

m) Projet MIX

Le projet MIX (pour *Modular Integration of Connectionist and Symbolic Processing in Knowledge-Based Systems*) auquel nous participons est un projet ESPRIT-III de type *Basic Research Action* (Réf. 9119) récemment approuvé par la Commission des Communautés Européennes. Géré par un consortium de 6 équipes européennes dont celle de l'IMAG (Bernard Amy, LIFIA, et Vincent Rialle, TIMC), ce projet a pour objectif de clarifier l'impact théorique de l'intégration des méthodes symboliques et des méthodes connexionnistes d'apprentissage automatique. Il se propose à cet effet de produire un environnement logiciel d'intégration de ces méthodes et d'effectuer quelques applications dans le domaine de l'industrie et de la médecine. Ce projet sera détaillé au point III.1 de ce mémoire.

n) Conclusion

Cette introduction avait pour but de présenter le contexte et la thématique générale des travaux que nous conduisons en matière de systèmes d'aide au diagnostic dans le domaine médical. Elle a conduit à situer les travaux qui se déroulent actuellement en IAM dans le contexte plus vaste des sciences cognitives, et de développer le point de vue de *l'émergence* comme approche tant théorique que pragmatique de la construction d'outils d'aide à la décision, le domaine médical constituant en ce qui nous concerne un creuset d'expérimentation particulièrement riche pour ces recherches.

Ces systèmes reposent sur des concepts d'apprentissage automatique de décisions plus que de connaissances pré-définies et figées, d'adaptation à de nouvelles situations plus que de comportements répétitifs et stéréotypés. Au célèbre "goulot d'étranglement" que constitue la construction d'une base de connaissances déclaratives, ces systèmes opposent une utilisation intensive des bases de cas, véritables prolongements sur support informatique de la mémoire à long terme du médecin, et donnent le pas à *l'émergence* de capacités classificatoires automatisées.

Selon l'approche émergentiste, la connaissance propre à conduire à une décision doit en effet être "apprise" de la BCa au moyen d'une méthode, d'un algorithme approprié issu d'un éventail assez large de théories plus ou moins concurrentes.

Une abondance de sources théoriques et méthodologiques n'a cessé d'inspirer et de sous-tendre au cours de ces années les recherches visant à reproduire dans une machine des comportements d'experts.

Face aux difficultés diverses et parfois insurmontables rencontrées par le courant des systèmes experts, particulièrement en médecine, une tendance à la réintégration de l'expérience du spécialiste — sous la forme de bases de cas — dans le processus de construction de systèmes décisionnels s'est fortement développée ces dernières années.

Cette tendance s'intègre parfaitement dans les orientations qui se développent dans le contexte plus vaste des sciences de la cognition. A la capacité cognitive de catégorisation / classification, essentielle dans la construction de l'intelligence humaine, répond en intelligence artificielle un large éventail de méthodes théoriques et algorithmiques de construction automatique de *classifieurs*. Quelques-unes de ces méthodes, parmi les plus représentatives, ont été esquissées afin d'illustrer ce propos.

Cependant, quelle que soit l'orientation théorique empruntée, l'expert, détenteur du savoir et de l'expérience, reste la pierre angulaire de la construction du système à apprentissage automatique. C'est pourquoi nous nous sommes particulièrement attachés à mettre en lumière les places que devraient occuper, dans les productions d'IAM, d'une part le médecin en tant que *sujet connaissant*, d'autre part l'*expérience médicale* produite par lui et recueillie dans les *Bases de cas* pour devenir la source d'apprentissage du système artificiel. Sujet, connaissance et expérience cumulée sont primordiales dans le contexte des sciences cognitives dès lors qu'il s'agit de concevoir des systèmes cognitifs artificiels qui soient de véritables prolongements des capacités cognitives du spécialiste. A cet égard, on observe que l'expertise humaine se trouve présente aussi bien dans la démarche déclarative des SBCo que dans celle de l'émergence. Seul le *degré d'intervention* de l'expert dans la construction du système diffère considérablement. Dans le cas des systèmes à bases de connaissances déclaratives, c'est *toute* la connaissance qui est requise *sous une forme représentée*

pour parvenir au diagnostic, alors que dans le cas des systèmes fondés sur telle ou telle approche de l'émergence, seules les connaissances a) *nosographiques* (la grille médicale d'observation qui préside à la structuration de la base de cas), b) *structurelles* (les définitions concernant le type des variables, leur domaine de variation, etc.) et c) *décisionnelles cas par cas* (l'attribution cas par cas d'une décision diagnostique, thérapeutique...) sont investies dans le système.

La "griffe" de l'expert reste néanmoins fortement présente dans le cas des systèmes d'apprentissage. Il faut en effet une connaissance d'expert pour parvenir à une liste pertinente de traits définissant la structure de la base de cas; vient ensuite l'alimentation de cette base par des cas de maladies observés dans une pratique régulière. Cette activité de recueil de cas assortis pour chacun d'eux du résultat diagnostique ou thérapeutique est également le produit d'une connaissance experte. La démarche de l'émergence est ainsi étroitement associée aux connaissances du spécialiste.

Une contrepartie cependant importante des systèmes d'apprentissage automatique est qu'ils demeurent, en l'état actuel des connaissances en IA, essentiellement limités à des tâches de classification/prédiction automatiques. Bien que la classification se situe au cœur de l'activité cognitive intelligente, l'ambition générale de l'IA dépasse cependant la seule modélisation de cette fonction.

Les prochaines années devraient voir se clarifier les relations entre architectures à bases de connaissances complexes (architectures "objets", IA distribuée), pratiquant une classification symbolique et les systèmes de construction automatique de connaissances tels que nous les avons présentés ici.

Enfin, terminons sur une préoccupation d'ordre éthique, particulière au domaine médical, en rappelant qu'avec le développement de ses performances et des facilités d'utilisation offertes par le génie logiciel dans son ensemble (un classifieur performant peut être construit en quelques minutes à partir d'une base de cas), le *système artificiel de décision* peut se voir conféré une part de plus en plus importante de *l'autorité de décision*. Ce qui, de toute évidence n'apparaît pas souhaitable sur un plan déontologique ni réaliste sur un plan médical. Cependant, l'apparente facilité de construction et d'utilisation vers lesquelles on s'achemine avec les systèmes d'apprentissage automatique et les systèmes hybrides n'est pas sans risque. Les avancées théoriques qui ont été esquissées, relayées par le génie logiciel qui les mettent à la portée de nombreux utilisateurs, peuvent jouer le rôle d'un moteur surpuissant dans un véhicule circulant sur les routes souvent sinueuses du domaine de la santé: des dérapages sont à prévenir. Une information et le développement d'une formation appropriées devront dans tous les cas accompagner la diffusion de ces méthodes et outils diagnostiques.

Bibliographie

- Amy B. (1989), Contextual cognitive machines. In: G. Tiberghien (Ed.), *Advances in cognitive sciences*. Vol 2 : Theory and applications, New York: Ellis Horwood.
- Amy B., Giacometti A. & Gut A. (1990), Modèles connexionnistes de l'expertise. *Proc. NEURO NIMES '90*, Nanterre, France: EC2, 99-119.
- Andler D. (1987), Une nouvelle science de l'esprit : intelligence artificielle, sciences cognitives, nature du cerveau. In: (Ed.), *Le Débat*. 47, Paris: Gallimard.

- Andler D. (Ed.), (1992), *Introduction aux sciences cognitives*. Paris: Gallimard.
- Beltrami E. & Rialle V. (1993), La problématique du profil type des abuseurs sexuels. *Revue Sexologique*, 1(2), 55-78.
- Bersini H. & Varela F.J. (1991), Hints for adaptive problem solving gleaned from immune networks. In: P. Schwefel & R. Männer (Ed.), *Parallel Problem Solving From Nature*. Lecture Notes in Computer Science n° 496, Berlin: Springer-Verlag. 343-354.
- Beuscart R. & Fieschi M. (1992), Cognition et systèmes experts médicaux. *Le courrier du CNRS* 79, 101.
- Bonelli P. & Parodi A. (1991), An Efficient Classifier and its Experimental Comparison with two Representative Learning Methods on three medical domains. In: R. K. Belew & L. B. Booker (Ed.), *Proc. 4th Int Conf on Genetic Algorithms*. San Mateo, California: Morgan Kaufmann. 288-295.
- Bournet P., Reggia J. & Samuélidès M. (1991), *Une approche connexionniste de l'IA*. Toulouse, France: Tecknea.
- Breiman L., Friedman J., Olshen R. & Stone C. (1984), *Classification And Regression Trees*. Monterrey, California: Wadsworth.
- Brooks R.A. (1991), Intelligence without representation. *Artificial Intelligence* 47, 139-159.
- Carpenter G. & Grossberg S. (1987), A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics and Image Processing* (37), 54-115.
- CCE (1992), *ESPRIT: Specific Search and technological development programme in the field of information technology- Results and Progress 1991-1992*. Luxembourg: Commission of the European Communities.
- Changeux J.P. (1982), *L'Homme neuronal*. Paris: Fayard.
- Cheeseman P., Self M., Kelly J., Stutz J., Taylor W. & Freeman D. (1988), Bayesian classification. *Proc. AAAI '88*, 607-611.
- Clancey W.J. (1985), Heuristic Classification. *Artificial Intelligence* 27, 289-350.
- Crémilleux B. (1991), *Induction automatique : aspects théoriques, le système ARBRE, applications en médecine*. Thèse de Doctorat d'Informatique. Université J. Fourier, Grenoble.
- Davis L. (1991), *The handbook of Genetic Algorithms*. Van Nostrand Reinhold.
- Dehaene S. & Changeux J.P. (1989), A simple Model of Prefrontal Cortex Function in Delayed-Response Tasks. *Journal of Cognitive Neuroscience* 1(3), 244-261.
- Demongeot J., Goles E. & Tchuente M. (Ed.), (1985), *Dynamical Systems and Cellular Automata*. New York: Academic Press.
- Dennett D.C. (1993), *La conscience expliquée*. Paris: Odile Jacob ; tr. par P. Engel, de: *Consciousness explained*, 1991, Boston: Little-Brown-and-Co.
- Feldman J.A. (1989), Connectionist Representation of Concepts. In: R. Pfeifer, Z. Schreter, F. Fogelman-Soulié et al. (Ed.), *Connectionism in perspective*. Amsterdam, The Netherland: North-Holland Publ. 25-45.
- Fieschi M. (1984), *Intelligence Artificielle en Médecine*. Paris: Masson.
- Fieschi M. & Joubert M. (1985), Some Reflections on the Evaluation of Expert Systems in Medicine. *Methods of Information in Medicine* 25, 15-21.
- Fodor J.A. & Pylyshyn Z.W. (1988), Connectionism and cognitive architecture: A critical analysis. *Cognition* 28(1-2), 3-71.
- François P. (1990), *Evaluation des systèmes informatiques d'aide à la décision médicale : étude du système MENINGE*. Thèse de Génie Biologique et Médical. Université J. Fourier, Grenoble.
- François P., Robert C., Crémilleux B., Bucharles C. & Demongeot J. (1990), Variables processing in expert system building: Application to the aetiological diagnosis of infantile meningitis. *Medical Informatics* 15(2), 115-124.
- Gallant S.I. (1988), Connectionist expert systems. *Communications of the ACM* 2(31), 152-169.
- Ganascia J.-G. (1992), L'apprentissage par Généralisation. *Le Courrier du CNRS* 79, 71.
- Giacometti A. (1992), *Modèles hybrides de l'expertise*. Thèse de doctorat. Ecole Nationale Supérieure des Télécommunications, Paris.
- Giacometti A., Iordanova I., Amy B., Vila A., Reymond F., Abaoub L., Dahou M. & Rialle V. (1992), A Hybrid Approach to Computer Aided Diagnosis in Electromyography. *14th Int Conf IEEE Engin Med Biol Soc* 14, 1012-1013.
- Goldberg D.E. (1989), *Genetic algorithms in search, optimisation and machine learning*. Reading, Massachusetts: Addison-Wesley.
- Hendler J.A. (1989), On the need of hybrid systems. *Connection Science (special issue on Hybrid Connectionist / Symbolic Systems)* 1(3),
- Henery R.J., King R., Sutherland A., Mitchell J.M.O. & Brazdil P. (1991), StatLog : Comparative Testing of Statistical and Machine Learning Algorithms. In: (Ed.), *ESPRIT - Information Processing and Software: Results and Progress of Selected Projects*. DG XIII, Bruxelles: C.E.E. 398-411.
- Holland J.H. (1975), *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan: University of Michigan Press.

- Holland J.H. (1986), Escaping brittleness: The possibilities of General Purpose Learning Algorithms Applied to Parallel Rule-Based Systems. In: R. Michalski, J. Carbonel & T. Mitchel (Ed.), *Machine Learning*, 2, San Mateo, CA: Morgan-Kaufmann. 594-623.
- Hopfield J.J. (1982), Neural Networks and physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Science* **79**, 2554-2558.
- Iordanova I., Rialle V. & Vila A. (1992), Use of Unsupervised Neural Networks for Classification Tasks in Electromyography. *Proc. 14th Int Conf IEEE Engin Med Biol Soc* **14**, 1014-1015.
- Kodratoff Y. (1988), *Leçons d'Apprentissage Symbolique Automatique*. Toulouse, France: CEPADUES édition.
- Kodratoff Y. (1989), IA et approche connexionniste. *Bulletin de liaison de la recherche en informatique et en automatique, INRIA* (124), 16-21.
- Kodratoff Y. & Diday E. (1991), *Induction symbolique et numérique à partir de données*. Toulouse, France: CEPADUES Editions.
- Kodratoff Y.T. France & Michalski R.S. (1993), *Apprentissage symbolique: une approche de l'intelligence artificielle, tome 2*. Toulouse, France: CEPADUES-Edition.
- Kohonen T. (1982), Self-organized formation of topologically correct feature maps. *Biological Cybernetics* **43**, 59-69.
- Kohonen T. (1988), *Self-organization and associative memory*. Berlin: Springer-Verlag.
- Memmi D. (1989), Connectionism and Artificial Intelligence. *Proc. NEURO-NIMES '89*, Nanterre, France: EC2, 17-34.
- Meunier J.G. (1992), Le Problème de la catégorisation dans la représentation des connaissances. *Intellectica* (13-14), 353-356.
- Michalski R.S. (1983), A Theory and Methodology of Inductive Learning. *Artificial Intelligence* **20**, 111-161.
- Michalski R.S., Carbonell J.G. & Mitchell T.M. (1993), *Apprentissage symbolique: une approche de l'intelligence artificielle, tome 1*. Toulouse, France: CEPADUES-Edition.
- Michalski R.S. & Chilauskys R.L. (1980), Knowledge acquisition by encoding expert rules versus computer induction from examples - a case study involving soybean pathology. *Int. J. of Man-Machine Studies* **12**, 63-87.
- MLT-Consortium (1993), *MLT Project: Final Public Report*. Bruxelles: Commission of the European Community.
- Newell A. (1980), Physical Symbol Systems. *Cognitive Science* **4**, 135-183.
- Nilsson N. (1988), *Principes d'Intelligence Artificielle*. Toulouse, France: CEPADUES-Editions ; tr. par de: *Principles of Artificial Intelligence*, 1980, Los Altos, CA: Morgan-Kaufmann.
- Patel V.L. & Evans D.A. (Ed.), (1989), *Cognitive Science in Medicine : Biomedical Modelling*. Cambridge, Massachusetts: MIT Press.
- Pham K.M. & Degoulet P. (1991), MOSAIC : Medical Knowledge Processing Based on a Macro-Connectionist Approach to Neural Networks. In: (Ed.), *MEDINFO : Proc. 6th Congress on Medical Informatics*. Vol I, 82-86.
- Pinkas G. (1992), Represented Unrestricted First-Order Logic Formulas in Connectionist Networks. In: R. Sun & L. A. Bookman (Ed.), *Proc. of the AAAI-92 Workshop "Integrating Neural and Symbolic Processes - The Cognitive Dimension"*. San Jose, California: 23-30.
- Pollack J.B. (1990), Recursive Distributed Representations. *Artificial Intelligence* **46**, 77-105.
- Quinlan J.R. (1986), Induction of Decision Trees. *Machine Learning* **1**(1), 81-106.
- Rialle V. (1989), Pour une intégration des systèmes experts en milieu clinique : quelques réflexions sur la problématique de l'utilisation des systèmes à bases de connaissances en pratique clinique courante. In: (Ed.), *Actes des Journées Francophones d'Informatique Médicale de Montpellier*. Rennes: Editions de l'Ecole Nationale de la Santé Publique. 279-285.
- Rialle V. (1994), De la Représentation à l'Emergence: Décision et Cognition en Intelligence Artificielle Biomédicale. In: D. H. Joly (Ed.), *Biomédecine 2000*. Nanterre, France: EC2.
- Rialle V., Ohayon M., Amy B. & Bessière P. (1991), Medical Knowledge Modeling in a Symbolic-Connectionist Perspective. *Proc. 3th An. Int. Conf. IEEE Engin. in Med. and Biol. Society*, 13(3), IEEE, 1109-1110.
- Rialle V., Vila A. & Besnard Y. (1991), Heterogeneous knowledge representation using a finite automaton and first order logic: A case study in Electromyography. *Artificial Intelligence in Medicine* **3**(2), 65-74.
- Riesbeck C.K. & Schank R.C. (1989), *Inside Case-Based Reasoning*. Lawrence Erlbaum Associates.
- Robert C. (1989), *Analyse descriptive multivariée - Application à l'intelligence Artificielle*. Paris: Flammarion.
- Robert C. (1991), *Modèles statistiques pour l'I.A.: L'exemple du diagnostic médical*. Paris: Masson.
- Rosch E. (1978), Principles of categorization. In: E. Rosch & B. B. Lloyd (Ed.), *Cognition and Categorization*. Hillsdale, New Jersey: Lawrence Erlbaum.

- Rumelhart D.E., McClelland J.L. & PDP-Research-Group (Ed.), (1986), *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*. Vol 1: Fondations. Cambridge, Massachusetts: MIT Press/Bradford Books.
- Salamon R., Bernardet M., Samson M., Derouesne C. & Grémy F. (1976), Bayesian methods applied to decision making in neurology – methodological considerations. *Methods of Information in Medicine* **15**, 174-179.
- Shavlik J.W., Mooney R. & Towell G. (1991), Symbolic and Neural Learning Algorithms: An Experimental Comparison. *Machine Learning Journal* **6**(2), 111-143.
- Smolensky P. (1986), Information Processing in Dynamical Systems: Foundation of Harmony Theory. In: D. E. Rumelhart, J. L. McClelland & PDP-Research-Group (Ed.), *Parallel Distributed Processing: Exploration in the Microstructure of Cognition*. Cambridge, MA: MIT Press. 194-281.
- Sun R. (1991), *Integrating Rules and Connectionism for Robust Processing*. PhD. Thesis. Brandeis University, Computer Science Dept, Waltham, MA.
- Sun R. & Bookman L.A. (Ed.), (1992), *Proc. of the AAAI-92 Workshop "Integrating Neural and Symbolic Processes - The Cognitive Dimension"*. San Jose, California:
- Thornton C.J. (1992), *Techniques in Computational Learning: An Introduction*. London: Chapman & Hall.
- Tiberghien G. (1986), Psychologie cognitive, sciences cognitives et technologie de la connaissance. In: J. L. Le Moigne (Ed.), *Intelligence des mécanismes, Mécanismes de l'intelligence*. Paris: Fayard.
- Tiberghien G. (1993), Questions de modélisation et de simulation cognitives. In: J. F. Le Ny (Ed.), *Intelligence Artificielle et Intelligence Naturelle*. Paris: Presses Universitaires de France. 43-69.
- Touretzky D.S. (1989), *BoltzCONS : Dynamic Symbol Structures in a Connectionist Network*. Tech. Rep. No. CMU-CS-89-172, Carnegie Mellon University.
- Varela F., Thompson E. & Rosch E. (1991), *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press ; tr. fr. 1993. *L'inscription corporelle de l'esprit*, tr. par V. Havelange, Paris: Le Seuil.
- Varela F.J. (1989), *Connaître les sciences cognitives: tendances et perspectives*. Paris: Le Seuil.
- Varela F.J. & Bourguine P. (Ed.), (1992), *Toward a Practice of Autonomous Systems*. Proc. 1st Europ Conf Artificial Life. Cambridge, Massachusetts: MIT Press.
- Venturini G. (1992), Characterizing the Adaptation Abilities of a Class of Genetic based Machine Learning Algorithms. In: F. J. Varela & P. Bourguine (Ed.), *Toward a Practice of Autonomous Systems*. Cambridge, Massachusetts: MIT Press. 302-309.
- Weiss S.M. & Kapouleas I. (1989), An empirical comparison of pattern recognition, neural nets, and machine learning classification methods. In: (Ed.), *Proc. IJCAI 1989*. San Mateo, CA: Morgan Kaufman. 781-787.
- Weiss S.M. & Kulikowski C.A. (1991), *Computer Systems that Learn : Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning and Expert Systems*. San Mateo, California: Morgan Kaufmann.
- Winograd T. (1993), Heidegger et la conception des systèmes informatiques. *Intellectica* **17**(2), 51-78.

I. ACTIVITE DE RECHERCHE

I.1. Présentation générale de l'activité de recherche

A) Les thèmes centraux de l'activité de recherche depuis 1984 :

Comme il a été indiqué dans l'introduction, le thème central autour duquel s'organisent mes activités de recherche depuis une dizaine d'années est constitué par le modélisation informatique des processus de raisonnement et de décision, avec application majeure à l'expertise médicale. Rappelons que ce thème s'inscrit dans le champ de l'intelligence artificielle au sens large du terme, c'est-à-dire l'ensemble des techniques "computationnelles" qui concourent à la conception de systèmes artificiels capables de simuler des phénomènes de perception, de raisonnement et d'action et qui gagnerait à être re-nommé, selon le terme de F. Varela, *autonomie artificielle*. Il s'inscrit donc également dans le champ plus large des *sciences de la cognition*.

Rappelons également que le domaine médical est très présent dans l'ensemble des recherches, sans toutefois en constituer l'unique champ d'application. Il représente plutôt une source quasi inépuisable de problématiques décisionnelles plus ou moins complexes et tient souvent lieu de banc d'essai ou d'épreuve de vérité des modèles étudiés.

B) Les recherches comprennent :

- une modélisation des connaissances mises en œuvre dans les processus décisionnels.
- une modélisation des raisonnements d'experts intégrant :
 - les connaissances, académiques et de savoir-faire pratiques,
 - l'expérience cumulée au cours d'une pratique de l'expertise.
- une étude des systèmes de classification et de production de connaissances inférentielles à partir de données d'observation . Le terme anglais *Machine learning* est le plus à même de représenter cette classe de systèmes dont le but commun est l'aide automatisée à décision mais dont les algorithmes ressortissent à différents domaines de recherche (logique, théorie de l'information, analyse des données, connexionnisme, algorithmique génétique,...).
- une approche hybride (symbolique et connexionniste) de l'analyse de textes en langage naturel. Il s'agit ici d'une orientation récente mais néanmoins prometteuse,

issue de mon année d'étude post-doctorale à l'université du Québec à Montréal (voir E, I.3.5 et III.3).

C) Les domaines de test et de validation sont :

- la génétique médicale : système GENESYS (de 1984 à 1987, thèse nouveau régime)
- l'analyse des gaz du sang et les soins intensifs : système GDS (de 1987 à 1990)
- le diagnostic électrophysiologique des neuropathies
 - système à base de connaissances NEUROOP
(étudié dès 1986, intensivement depuis 1990)
 - système hybride SHADE : base de connaissances et base de données
(depuis 1991, en cours)
- la psychiatrie et psychologie :
 - recherche de profils types de sujets auteurs de violences sexuelles
(1992, en cours)
- l'analyse cognitive de bases de données textuelles et médicales (1992, en cours)

D) Une activité de réflexion sur :

- la pratique de l'ingénieur cognitif;
- la place des systèmes experts au sein de l'institution hospitalière;
- l'activité de modélisation des processus de décision chez l'homme et l'informatisation de ces modèles.

E) Une coopération scientifique de type internationale comprenant :

- un séjour de recherche d'une année au Canada, dans deux laboratoires de recherche: l'un à l'Université de Montréal (Centre de Recherche Fernand Seguin - 4 mois), l'autre à l'Université du Québec à Montréal (Centre de Recherche en Cognition et Information ATO-CI - 8 mois), durant l'année universitaire 1991-92.
- plusieurs projets de coopération franco-québécois, avec des liens étroits de recherche avec des chercheurs québécois, comme en témoignent les parties I.3.3, I.3.4, I.3.5, I.3.6, III.3.
- un projet européen (ESTEEM) consacré à l'aide à la décision en Electromyographie (en cours depuis 1991).

F) Tous les travaux de recherche mentionnés :

- ont donné lieu à publications (voir II),
- ont été ou sont soutenus par des subventions de recherche (voir Annexe 3),
- ont donné matière à une importante activité d'enseignement et d'encadrement d'étudiants de troisième cycle de l'université Joseph Fourier (voir Annexe 1).

La suite du rapport développe les thèmes centraux de recherche, à travers notamment les publications auxquelles ils ont donné lieu et les perspectives de développement.

I.2. Liste détaillée des thèmes de recherche

Cette liste reprend de manière exhaustive les thèmes de recherche qui ont été abordés depuis le début de mes activités de chercheur à la Faculté de Médecine de Grenoble (avril 1980). Certains de ces thèmes n'ont pas été poursuivis (e.g. la docimologie), tandis que d'autres ont connu un fort développement. Cette remarque explique la longueur variable de la présentation de chaque thème.

I.2.1. Extraction de connaissances dans un domaine vaste, évolutif et incertain

- application à la Génétique Médicale - utilisation conjointe de concepts d'intelligence artificielle, de bases de données et d'outils mathématiques tels que l'Analyse de Données, l'Analyse Discriminante, l'Analyse en Composantes Principales.

Travaux qui ont conduit à ma thèse de Doctorat de Génie Biologique et Médical : *Aide au diagnostic et à l'apprentissage dans un domaine médical incertain, incomplet et évolutif - Etude des méthodes existantes et proposition d'une méthodologie nouvelle*. Soutenue le 20 février 1987.

Travaux ayant obtenu un soutien financier de la Fondation pour la Recherche Médicale.

Collaboration avec la Clinique Médicale Infantile (Dr. A. JOANNARD) du C.H.U. de Grenoble; collaboration scientifique: Pr. J. DEMONGEOT.

I.2.2. Ergonomie cognitive et communication avec un système expert en langage naturel spécialisé.

Ce thème donné lieu à une communication au congrès COGNITIVA 1985 et un court article. Les recherches en analyse du langage naturel ont ensuite été laissées de côté jusqu'en 1992, date à laquelle j'ai repris de manière intensive et selon une approche différente ce thème de recherche (voir I.3.5 et III.3).

I.2.3. Aide à l'Enseignement et Docimologie

Test de connaissances par Q.C.M. sur micro-ordinateur et analyse docimologique des résultats d'examens ou concours.

I.2.4. Système expert d'aide au diagnostic électromyographique

En collaboration avec le Service d'Electromyographie du CHU A. MICHALLON (Dr. Mme A. Vila, du Service EFSN, Dr. Mme Garrel, Hôpital A. Michallon, Grenoble).

Ce travail est effectué dans le prolongement historique des travaux de D. ZIEBELIN (système expert MYOSYS débuté vers 1982), et fait intervenir une équipe pluri-disciplinaire : domaine

médical : Dr. A. VILA, Dr. L. ABAOUB, Dr. F. REYMOND; informatique et intelligence artificielle : V. RIALLE et étudiants de DEA de Génie Biologique et Médical et de DESS-IDC.

Nous poursuivons depuis plusieurs années des recherches sur la faisabilité d'un système informatisé d'aide au diagnostic dans ce domaine. C'est ainsi qu'une maquette initiale de système expert, proposée par D. Ziebelin en 1984, fut la première présentée dans le domaine de l'Electromyographie. Nous avons ensuite utilisé les techniques de l'Analyse des Données comme moyen de production de connaissances nouvelles et proposé l'utilisation de scores comme moyen de discrimination d'une maladie dans un système expert. Nous avons également effectué une recherche comparative destinée à trouver le meilleur outil de réalisation d'un système expert en électromyographie. C'est à la suite de ces travaux, nous avons réalisé un premier prototype opérationnel, NEUROP, de système à base de connaissances dans le domaine du diagnostic électrophysiologique des neuropathies.

NEUROP a été réalisé sur micro-ordinateur Macintosh à l'aide du langage Prolog. Il est actuellement en cours de transfert sur station de travail. Ce système propose des hypothèses diagnostiques à partir des résultats de tests électro-physiologiques (potentiel d'action sensitif, vitesses de conduction nerveuse sensitive et motrice, etc.) réalisés sur le malade. Il fournit une interprétation des résultats à trois niveaux : a) le segment nerveux (atteinte axonale, démyélinisante, mixte, sévère ou modérée, etc.); b) le nerf (atteinte focale, focale multiple, diffuse, etc.); c) l'ensemble des nerfs (mononeuropathie, polyneuropathie symétrique ou asymétrique, etc.). L'aide fournie est importante étant donné le nombre de segments nerveux étudiés (de 5 à 50) et le nombre de combinaisons diagnostiques possibles. Très interactif, il a pour originalité, sur le plan théorique, d'inclure dans la base de connaissances un automate fini destiné à conduire un raisonnement spatial à partir de la répartition des lésions le long des fibres nerveuses. NEUROP est actuellement en cours d'évaluation au sein du Laboratoire d'Electromyographie du Centre Hospitalier et Universitaire de Grenoble.

La réalisation du système d'aide au diagnostic des neuropathies poursuit un triple objectif :

- d'une part faciliter l'exploration nerveuse et musculaire d'un patient. Deux moyens de contribuer à cette exploration sont offerts : premièrement, la proposition d'un guide pertinent et optimisé d'exploration neuro-musculaire ; deuxièmement, la génération d'hypothèses diagnostiques attachées à chaque nerf ou muscle et à chaque segment neuromusculaire étudié, ainsi que l'établissement d'un diagnostic global final.

- d'autre part donner au jeune médecin, ou au futur spécialiste en électromyographie, un outil d'auto-enseignement, de test et d'auto-évaluation de ses connaissances. Ce deuxième aspect occupe une place importante dans notre recherche. L'électromyographie est en effet une tâche complexe et délicate, nécessitant un haut niveau de spécialisation, et requérant la présence de l'expert. Dans ce contexte un outil interactif d'apprentissage et d'entraînement à la pratique de l'examen et du diagnostic permet, d'une part à l'enseignant, de démultiplier ses capacités d'enseignement, et d'autre part à l'étudiant d'approfondir, vérifier ou tester ses connaissances à volonté et sans mobiliser l'expert. Ceci, grâce à la consultation de Bases de Données anatomiques et bibliographiques incluses dans le système expert.

- enfin, améliorer les connaissances du domaine en préparant le terrain pour un certain nombre d'études ultérieures, notamment : études statistiques, classification automatique de malades par analyse de données et réseau connexionniste.

Les médecins notent pour leur part une amélioration significative de leur démarche d'examen électrophysiologique des nerfs par suite de la rigueur et de la systématisation imposées par l'entreprise de modélisation informatique.

Ces travaux ont été soutenus par la REGION RHONE-ALPES en 1990. Ils s'inscrivent maintenant dans le cadre du projet européen ESTEEM.

Le projet ESTEEM (European Standardized Telematic tool to Evaluate EMG knowledge based systems and Methods) est un projet européen en cours dans le cadre du programme "Advanced Informatics in Medicine" (A.I.M.) de la Commission des Communautés Européennes). Ce projet international regroupe neuf partenaires de la Communauté Européenne, dont le laboratoire TIMC-IMAG en collaboration avec le Département d'Electromyographie. Il devrait durer vraisemblablement jusqu'à fin 1994.

I.2.5. Système expert d'Analyse des Gaz du Sang et de Conseils d'Urgence en Réanimation

En collaboration avec les Services de Réanimation du C.H.U. de Grenoble (Dr. J.C. PEYRIN, Dr. B. FAUVAGE, Dr. P. GIRARDET) et le Service de Physiologie de la Faculté de Médecine (Dr. F. GRIMBERT, Dr. P.BACONNIER); collaboration scientifique: Pr. J. DEMONGEOT.

Dans ce cadre:

- étude d'une représentation mixte des connaissances par "objets" et par règles de production,
- utilisation du système de développement de systèmes experts NEXPERT-OBJECT et langage PROLOG,
- coordination générale des travaux de construction de la base de connaissances. Equipe pluridisciplinaire composée de : a) 2 médecins anesthésistes-réanimateurs, b) 2 physiologistes, c) 3 informaticiens,
- saisie directe des résultats d'analyse du sang sur les appareils de mesure et réalisation d'un logiciel serveur et d'un réseau reliant plusieurs sites hospitaliers (Laboratoire d'Analyse et Services d'Urgence).

Ces travaux, achevés en 1990, ont obtenu un soutien financier de la Fondation pour la Recherche Médicale, du Pôle Régional Rhône-Alpes de Génie Biologique et Médical, de la REGION RHONE-ALPES, ainsi que de l'Université Joseph Fourier de Grenoble, en 1987, 1988, 1989 et 1990.

I.2.6. Co-organisation, avec le Pr. J. DEMONGEOT, T. HERVE et Ch. ROCHE, de la conférence : "Intelligence Artificielle et Sciences Cognitives"

Tenue à Grenoble dans les locaux de l'Institut National Polytechnique, les 18, 19 et 20 février 1987. Cette conférence a réuni pendant trois jours une centaine de spécialistes dans trois axes de recherches constitués par : les langages orientés objets, les sciences cognitives et les systèmes experts en biomédecine.

Co-édition des actes du congrès ("Artificial Intelligence and Cognitive Sciences", Manchester University Press, New-York, 1988).

I.2.7. Responsabilité, organisation et animation d'un groupe de recherche en Intelligence Artificielle et Médecine :

séminaires bi-mensuels d'une durée d'une heure et demie, d'octobre à juin, en 1986-1987, 1987-1988 et 1988-1989, soit une trentaine d'exposés sur des sujets tels que :

- Analyse des systèmes experts médicaux existants et de travaux en cours
- Traitement du langage naturel
- Modélisation du raisonnement incertain

I.2.8. Pratique intensive, depuis 1983, et essai de théorisation de la fonction de cogniticien

Champ d'application: intelligence artificielle et médecine.

I.2.9. Analyses statistiques et Analyses des Données (analyses discriminantes et en composantes principales) en électromyographie

Pratique des logiciels d'analyse suivants: Statgraphics, Statview, Systat, sur Macintosh et compatible PC.

I.2.10. Participation au groupe de recherche NORIA sur les réseaux de neurones et l'intelligence artificielle

Groupe local à la région de Grenoble, d'abord dans le cadre du LASCO (Laboratoire de Sciences Cognitives, direction: Pr G. Tiberghien) en 1989-1990; puis dans le cadre du Pôle Rhône-Alpes de Sciences Cognitives (CNRS, direction: Pr G. Tiberghien) depuis 1991. Thèmes : informatique, neurosciences et psychologie cognitive.

I.2.11. Collaboration scientifique avec le laboratoire NEURINFO

(de NEURones et INFOrmatique, directeur E. Sanchez) de l'Institut Méditerranéen de Technologie (Technopôle de Château-Gombert, Marseille) dans le cadre d'un projet de représentation du raisonnement incertain dans le système expert NEUROP (présenté ci-dessus), au moyen de la théorie des possibilités (projet soutenu par le Pôle de Génie Biologique et Médical de la Région Provence-Côte d'Azur). Maquette terminée en août 1991. Evaluation sur 30 dossiers de malades du Département d'Electromyographie (Hôpital A. Michallon, Grenoble) : dernier trimestre 1991.

I.2.12. Collaboration scientifique avec le Centre de Recherche En Psychologie COgnitive de l'Université d'Aix-Marseille

(CREPCO, Directeur : J.P. Caverni, URA CNRS n°182)

Dans le cadre d'un projet PIRTTEM - CNRS intitulé: *Formalisation des Processus mentaux dans une activité d'expertise: évaluation de connaissances*. Travaux sur deux ans, à partir d'octobre 1990. Mon rôle était de définir une architecture capable de simuler l'activité cognitive d'un expert (en

l'occurrence un correcteur de textes en langue naturelle d'élèves du secondaire) au moyen d'unités symboliques et d'unités connexionnistes coopérantes.

I.2.13. Collaboration scientifique avec le Centre Fernand Seguin de Recherche en Psychiatrie (Université de Montréal)

Dans le cadre du projet de recherche *Intelligence Artificielle en Psychiatrie*, soutenu par la REGION RHONE-ALPES (voir Subventions de recherche). J'ai en outre obtenu une *Bourse d'Excellence* de la REGION RHONE-ALPES pour l'année universitaire 1991-1992 pour favoriser cette recherche. J'ai travaillé à temps plein sur ce projet de septembre à décembre 1991.

Les principaux résultats sont :

1 • Voyage à Orlando (Floride), du 30 octobre au 3 novembre 1991, pour présenter deux communications :

a) RIALLE V., M. OHAYON M.: "Computers and Artificial Intelligence in Psychiatry: Brief History and State of the Art". In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1280-1281.

b) RIALLE V., M. OHAYON M., AMY B., BESSIÈRE P.: "Medical Knowledge Modeling in a Symbolic-Connectionist Perspective" In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1109-1110.

2 • préparation, avec le Dr. E. Stip du Centre F. Seguin, d'un workshop sur la Modélisation Cognitive dans le cadre du Congrès canadien sur l'étude du cerveau, du comportement et des sciences cognitives, deuxième réunion annuelle, Université Laval, Québec, 13 - 14 juin 1992.

3 • Voyage dans la ville de Québec pour la présentation d'une communication au Congrès canadien sur l'étude du cerveau, du comportement et des sciences cognitives sus-mentionné.

Titre de la communication : Modélisation Cognitive: Présentation, éléments d'analyse et discussion.

4 • réalisation d'un logiciel graphique interactif en Prolog pour l'expérimentation de théories sur l'inférence analogique, sur Macintosh, avec interface utilisateur graphique convivial.

5 • co-rédaction et soumission de l'article suivant :

RIALLE V., STIP E. & O'CONNOR K. (1994), Discussion of the issues and problems raised by computer-mediated psychotherapy. *Humane Medicine*. sous presse.

6 • co-rédaction et soumission de l'article suivant :

RIALLE V. & STIP E. (1994), Modélisation cognitive en psychiatrie : des modèles symboliques aux parallèles et distribués. *Journal of Psychiatry and Neurosciences*.19(3), sous presse.

I.2.14. Réalisation d'un logiciel d'analyse de fichiers de malades par réseau de neurones formels de type Kohonen

En collaboration avec Melle Irèna Iordanova (en année post-doctorale 1990-1991) et le Centre ATO.CI de Montréal. Programmation en C sur Macintosh. Deux applications ont donné lieu à publication, une en électromyographie et l'autre en psychiatrie :

- Une communication sur une application en électromyographie a été réalisée au 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Paris, octobre 1992. Titre : Use of topological feature map for EMG classification.
- Un article a été écrit avec le Docteur E. Beltrami de Montréal (cf. point 17).

I.2.15. Réalisation d'un logiciel graphique interactif en Prolog pour l'expérimentation de théories sur l'inférence analogique à partir de figures géométriques composites

Dans le cadre du projet de recherche *Intelligence Artificielle en Psychiatrie*, pré-cité.

I.2.16. Collaboration scientifique avec le Centre de Recherche en Cognition et Information ATO-CI (Université du Québec à Montréal - UQAM -)

Direction scientifique : Pr. Jean Guy Meunier

Cette collaboration est effectuée dans le cadre d'une délégation pour l'année 1991-92 dont voici le thème tel qu'il a été transmis au Ministère de l'Éducation Nationale :

La Délégation de M. V. Rialle à l'Université du Québec à Montréal (UQAM), au sein du Centre de Recherche en Cognition et Information ATO-CI, spécialisé dans le traitement cognitif de l'information, a pour objectif de contribuer au développement des sciences cognitives à l'UQAM et favoriser un courant d'échanges scientifiques entre cette université et l'Université Joseph Fourier de Grenoble. Monsieur Rialle exercera son activité de recherche dans une équipe formée notamment de linguistes, psychologues, informaticiens, philosophes et mathématiciens, dont le caractère éminemment interdisciplinaire semble particulièrement favorable au développement de la recherche en sciences cognitives. Les travaux porteront sur l'intégration de méthodes issues du courant connexionniste (réseaux de neurones, parallélisme de traitement) et du courant des algorithmes génétiques, aux méthodes fondées sur l'intelligence artificielle symbolique et l'analyse statistique, développées depuis une vingtaine d'années au Centre ATO-CI. Ils se situeront d'une part dans le prolongement de recherches déjà commencées, telles que la validation d'un modèle de réseau de neurones propre au Centre ATO-CI (modèle EIDOS), d'autre part dans l'expérimentation de modèles non encore utilisés dans ce Centre. Ces travaux devraient contribuer au débat actuel sur l'opposition ou la complémentarité entre les approches symbolique et connexionniste pour la modélisation de processus cognitifs.

Cette collaboration se prolonge actuellement à travers le projet COGNIMED (point 19 et III.2).

I.2.17. Recherche en psychiatrie pour la détermination du profil de l'abuseur sexuel (coopération scientifique Franco-Québécoise)

L'Université du Québec à Montréal possède son propre secteur de recherche en psychiatrie, dirigé par le Dr. E. Beltrami. Ce secteur est membre du Centre ATO-CI cité au paragraphe précédent. Le Dr. Beltrami s'intéresse particulièrement, depuis plusieurs années, à l'aide au diagnostic par système

expert. J'ai travaillé avec lui sur une méthode connexionniste d'aide à la classification automatique et au diagnostic, dans le domaine des troubles du comportement et des abus sexuels. Cette méthode est fondée sur la carte topographique du réseau auto-organisant de Kohonen.

Année 1992-1993.

I.2.18. Intégration d'algorithmes classifieurs (projet COGNIMED)

Le projet COGNIMED a pour objectif principal d'intégrer au sein d'une même structure logicielle la plupart des modèles de classification automatique et d'apprentissage de règles et d'arbre de décision. Il est présenté d'une manière plus détaillée dans la partie III.2.

I.2.19. Recherche et développement des systèmes d'aide à la décision hybrides (système SYNHESYS-SHADE et projet ESPRIT MIX)

Une première réalisation concrète de système hybride symbolique et connexionniste a été réalisée en électromyographie. Il s'agit de l'intégration, au sein d'un même système d'aide à la décision, de la modélisation symbolique des connaissances et des raisonnements (approche spécifique de l'intelligence artificielle) et de la modélisation par réseaux de neurones formels (approche spécifique du connexionnisme). Dans cette optique, un noyau de système a été développé au LIFIA : le système hybride SYNHESYS. Ce noyau est utilisé dans le domaine médical pour la réalisation du système SHADE (Système Hybride d'Aide au Diagnostic en Electromyographie). Les recherches sur l'intégration entre niveau symbolique et niveau connexionniste des connaissances pour la modélisation de l'expertise ont donné lieu à de nombreux développements. La recherche a été conduite dans une coopération scientifique tri-partite entre le laboratoire TIMC-IMAG (équipe de Vincent Rialle), le laboratoire LIFIA-IMAG (équipe de Bernard Amy) et l'équipe du Département d'Electromyographie (Service d'Exploration Fonctionnelle du Système Nerveux, Dr Garrel) de l'Hôpital A. Michallon, dirigée par le Dr. Annick Vila.

La suite de travaux sur les systèmes hybrides s'inscrit maintenant dans le cadre du projet MIX, ESPRIT-III n° 9119, décrit au point III.1.

I.3. Développement des principaux thèmes

Cette partie présente de façon plus développée quelques thèmes de recherche abordés durant les 7 dernières années. S'ils ont chacun leur spécificité, ces thèmes restent néanmoins inter-reliés par le thème central de la décision et de la cognition en médecine.

I.3.1 Premier thème : Modélisation et apprentissage de connaissances dans un univers vaste, incertain et incomplet

Reproduction de l'article :

RIALLE V., Data Analysis as an aid to Learning and Knowledge Base Making in a medical field. In *Artificial Intelligence and Cognitive Sciences*, J. Demongeot, T. Hervé, V. Rialle et C. Roche (Eds), Manchester University Press, New York, 1988, 375-386.

DATA ANALYSIS AS AN AID TO LEARNING AND KNOWLEDGE BASE MAKING IN A MEDICAL FIELD

Résumé :

Nous proposons un système d'aide à la construction de connaissances expertes en génétique médicale dans lequel la modélisation du domaine est fondée sur la représentation des symptômes sous forme de couples (entité symptomatique, valeur). Les entités symptomatiques (e.s) sont hiérarchisées en un arborescence représentant le corps humain médical ; un certain nombre d'entre elles sont regroupées au sein de scores constituant de super-entités-symptomatiques. A chaque e.s au score est associé un acte vectoriel, l'ensemble des axes constituant l'espace vectoriel dans lequel se situent les syndromes de la base.

Le projet consiste à associer en une démarche commune l'aide au diagnostic et la production de connaissances nouvelles obtenues par Analyse en Composantes Principales et Analyse Discriminante effectuées sur la base de données dans sa représentation vectorielle. Deux étapes sont à distinguer : 1°) L'analyse statistique de la Base de Donnée conduisant à une méta-connaissance sur le domaine et visant à réduire rapidement l'éventail des syndromes candidats à l'explication d'un tableau clinique. Il est réalisé périodiquement et en préalable au processus dynamique d'apprentissage. 2°) Le processus dynamique d'apprentissage et de reproduction de connaissances nouvelles. Ce processus est déclenché lors d'une démarche d'aide au diagnostic pour un tableau symptomatologique particulier (base de fait) et en l'absence de connaissances suffisantes contenues dans la base de connaissances. Il consiste en une exploration, par Analyse des Données, du sous-domaine impliqué par la base de faits. Le système propose alors soit une réponse diagnostique, soit l'orientation de l'examen vers telle ou telle entité symptomatique. Chaque couple (situation, action) nouveau est ajouté à la Base de Connaissances. Le système réalise l'intégration, au sein d'un même environnement des éléments suivants :

- Base de données syndromique et symptomatique
- module d'analyse multivariée
- base de connaissances
- moteur d'inférence
- interface utilisateur

Il regroupe actuellement une centaine de syndromes et près de 700 symptômes.

Keywords: Artificial intelligence, Expert systems, Diagnosis aiding, Automatic knowledge learning, Data analysis, Medical genetics.

1 General presentation

At the moment, expert systems theory expects at least one human expert of a concerned domain. Nevertheless, there exist a few domains which are too vast, uncertain, incomplete and evolutive to be totally mastered by any one specialist, and for which there do not exist experts able to construct a knowledge base usable as a knowledge engineering tool. However, it seems that assistance to

reasoning and decision making is of paramount importance in such domains (particularly in the medical field).

We propose a knowledge extraction helping system with an application in medical genetics (Edwards, C. N., Buyse, M. L., 1986), (Veloso, M., Feijoo, M.J., 1986), (Rialle, V., 1985). This system aims at production of inferential knowledges issued of a pool of observed datas such as syndroms, symptoms, frequencies, etc. On one hand, it offers the physician an assistance in his own approach to scientific research and, in the other hand, it constitutes a real automatic knowledge learning system. With this end in view, we propose a method and some tools for the exploration of the concerned domain descended from the Data Analysis (Benzecri, J.P. & all., 1981), (Macgibbon, B., Preus, M., 1979), (Gouvernet, J., 1979), (Preus, M., Ayme, S., 1983). In fact, the present work is situated at the cross-road of a number of scientific fields such as Medicine, Artificial Intelligence, Automatic Learning, Uncertain Reasoning, Mathematical modelling of decision problems, and Data Base theory.

One guiding idea of this work is that there is not a qualitative difference between knowledge which is elaborated and matured by any human expert and one which is stemmed from a mathematical analysis of experimental datas. Numerical methods of analysis such as Principal Coordinate Analysis, Roc-Analysis, etc., may give a point of view upon a question as relevant and useful as the expert one, and may safely and efficiently replace the latter when the domain is too vast. This point of view is reliable because of the explicit theories and experimental observations and verifications it is based on.

However, numerical analysis should not dissimulate the presence, and consequently the error source, of the expert who defines the variables to observe according to his proper experience.

2 Domain modelling

In the expert system point of view, Medical genetics is caraterized by: 1°) its vastness (several thousands of symptoms and syndroms), 2°) its uncertainty, 3°) its uncomplete state, 4°) the rapid increasing of knowledge. The domain modelling tries to keep in account these features and propose the following notions:

A - Symptoms, symptomatical entities and scores:

a - The symptom constitutes the basic element of the system from which the others will be elaborated. In our application, symptoms are subdivided into four categories according to increasing levels of observation cost or observation difficulties. For this reason, every symptom is marked by a number comprised between 1 and 4 with the following meanings:

1: immediat clinic: symptoms which are visible or identifiable without mediatory instruments (visual observation, by palpating, etc.),

2: extensive clinic: identifiable symptoms by means of a mediatory instrument (mesure equipement, auscultation, etc.), in the place where the patient is observed,

3: standard paraclinic: symptoms which are revealed by a commonly used exam (like radiography) elsewhere than the place where the patient is observed,

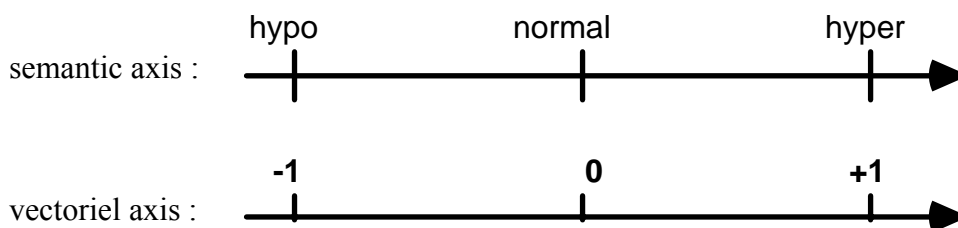
4: extensive paraclinic: symptoms which are revealed by a long or expensive paraclinical exam (biopsy, scanner, NMR, etc.).

A symptom s corresponds to the occurrence of the event: "the symptomatical entitie se takes the value v ": $s = \{ se \blacklozenge v \}$.

b - a symptomatic entitie (s.e.) is a spatio-temporal localization of symptoms. It often corresponds to an anatomic or functional localization where one expects to find symptoms or abnormallities corresponding to one or more illnesses. Symptomatic entities are arranged as a tree representing the human body . This s.e. tree induces a symptoms tree. In such a tree, the hierarchy corresponds to the imprecision level of the symptom descriptions (in the litteracy, syndroms are often discribed with general or imprecise symptoms). The accurate symptoms are the leaves of the symptoms tree.

A symptomatical entity owns an oriented semantic axis (as shown on picture 1) on which are distributed the various possibles values (ordinal values) of the symptoms. Thus, for each of the symptomatic entities, we may define a bijection between its semantical axis and the vectorial space **R**: to each ordinal value, we attribute an [a priori] chosen interval. The word **axis** points out a symptomatical entity's vectorial subspace.

Example: telorism



Each symptom will be represented by a vector:

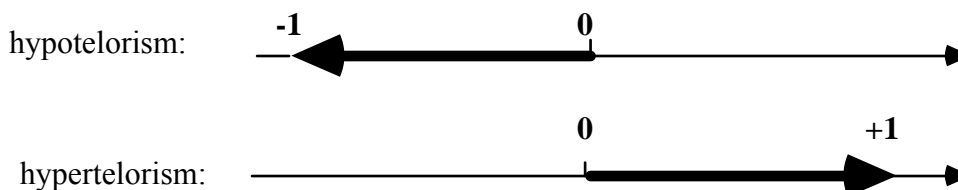


Figure 1.

c - a score (Robert, C., Zarski, M., Demongeot, J., 1986) is a grouping of a certain number of symptomatical entities corresponding to one of the following possibilities: a physiologic function, an anatomic localization, a similar abnormality in the body, etc.

The major interest of a score is that it allows to strongly reduce the number of symptomatical entities and, consequently, it reduces the number of axes of the working space. A second interest is the possibility it offers to establish transversal links inside the s.e. tree. Indeed, in the pure s.e. tree, the links are only ascending or descending ones. Nevertheless, a number of s.e. may have a similar finality such as motionness, development, shape and so on.

B - Syndroms:

a - Semiological representation: this representation reflects the physician's vision of the patient, in which a syndrom is expressed as a more or less precise list of symptoms. If sd represents a syndrom and s_1 through s_n represent symptoms, we notice: $sd = \{ s_1, \dots, s_n \}$.

b - vectorial representation: this representation is adapted to a numerical analysis of sets of syndroms. It express the coordinates of a syndrom (called a syndrom-point) in a vectorial space whoses axes are constituted with symptomatical entities or scores.

Thus, if m represents the number of axes of the vectorial space, a syndrom will be represented as a m components vector in this space:

$$sd = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{pmatrix} = x_1 \mathbf{u}_1 + \dots + x_m \mathbf{u}_m$$

The Descriptive Precision Coefficient (DPC) of a syndrom is calculated with the number t of terminal symptoms and the syndrom total number of symptoms:

$$DPC = \frac{t}{n} \quad DPC \in [0, 1]$$

The DPC will take place in the choice of the syndroms which are possible causes in the patient symptomatic panel. This choice will have a consequential effect on the system's choice of the s.e. to propose for observation.

Finally, let us precise that a syndrom may have several DPC depending on which observation level or vectorial subspace the analysis is made. For exemple, a syndrom may be well defined at the first level (immediat clinic) of symptom observation (in this case the DPC of level 1 will be close to 1), at the very time when the same syndrom will be badly defined at level 3 (standard paraclinic) (in this case, the DPC of level 3 will be close to 0). This last precision reflects the fact that a syndrom may be better or les well described according to the level on which one observes it.

C - working spaces : complete space and conjectural spaces:

A working space is a multidimensional space (let us put m dimensions) constituted with the set of m s.e. and scores which are involved in the semiologic description of the syndroms. We distinguish two types of working spaces: the complete space (unic) and the conjectural spaces.

The complete space comprises the wholeness of the symptomatical entities and scores which are in the data base. This space will be essentially used by a static preparatory study of the complete domain (all syndroms in the space of all axes).

The conjectural spaces are the various subspaces involved in a particular diagnosis research. Each of them corresponds to the following couple of sets at a given time: (set of observations, candidate syndroms).

3 Preparatory static study

The preparatory static study consists of an exploration of the domain which is situated outside the dynamical learning process (that will be exposed in 4). This study intends to construct a part of the knowledge base which is directly deductible from the data base. It leads to a five step analysis, related to the five following spaces:

- first step: complete space,

- second through fifth steps: respectively the first through the fourth observation level subspaces.

For each of these steps, the process is:

1° Construction of the bidimensional matrix (syndroms x axes). This matrix comprises:

. in rows: every syndroms of the data base which has some not null components in the current level of observation,

. in columns: the coordinates of the syndroms in the current subspace.

2° Construction of the corresponding variance-covariance matrix and diagonalisation of this matrix.

3° Collection and analysis of the eigenvalues and eigenvectors of the matrix.

An eigenvector is expressed with the list of its components on the considered observation level. let us define the word "contribution" as a synonym of "component value".

The analysis gives a list of the main normalized eigenvectors whose inertia is superior to 80 per cent of the whole cloud inertia and, for each of these eigenvectors, the sorted list of the axes which have the best contributions.

We extract from these lists the axes which appear the most frequently and with the strongest contributions. These axes are named pertinent axes.

Then, we propose to make an exploration of the syndrom cloud projected on plans determined by couples of pertinent axes. These various projections may be interpreted in a same way as a Principal Component Analysis, and may lead to various knowledges in the following examples:

if score 60 > 9 then multiple pterygium syndrom

if score 63 > 5 then Charlie syndrom

if score 60 and score 63 are nul then palatal split (certainty 0.7)

if score 62 = 0 then Apert syndrom or pterygium poplite

More generally, we may say that: 1°) a score allows us to pinpoint one or several isolated syndroms upon which it presents a maximum value; 2°) when two scores are nul, this allows us to discriminate in favour of a particular set of syndroms. Thus, such a couple induces a dichotomic approach which right away excludes half part of the syndroms set. The couples of selected scores are arranged in a decreasing order based on the discriminant power of the scores and s.e.. The repetition of the dichotomic procedure on these sorted plans leads to a rapid decreasing of the remaining candidate syndroms set (picture 2).

This preparatory static study has been applied to our data base and has induced the following results: the analysis of the 12 first eigenvectors given by the Jacobi algorithm revealed the 14 pertinent axes of prime importance from the 67 axes contained in the first observation level. Then we have successively analysed the projection on the entire syndrom set of the data base onto the 7 plans issued from the eigenvectors analyse. Each plan has permitted the separation of the syndroms into two significant groups: on one hand, those which had a projection on the origin of the axes, on the other hand, those whose projection were outside the origin. This analysis has resulted in the production of 7 dichotomic rules (notated: DR₁,...,DR₇) whose interest is to reinforce the patient observation process by prompting the physician to confirm or infirm the symptoms that belong to

the pertinent axes. To know the absence or the occurrence of these symptoms will permit to rapidly select the subset which eventually contains the researched syndrom.

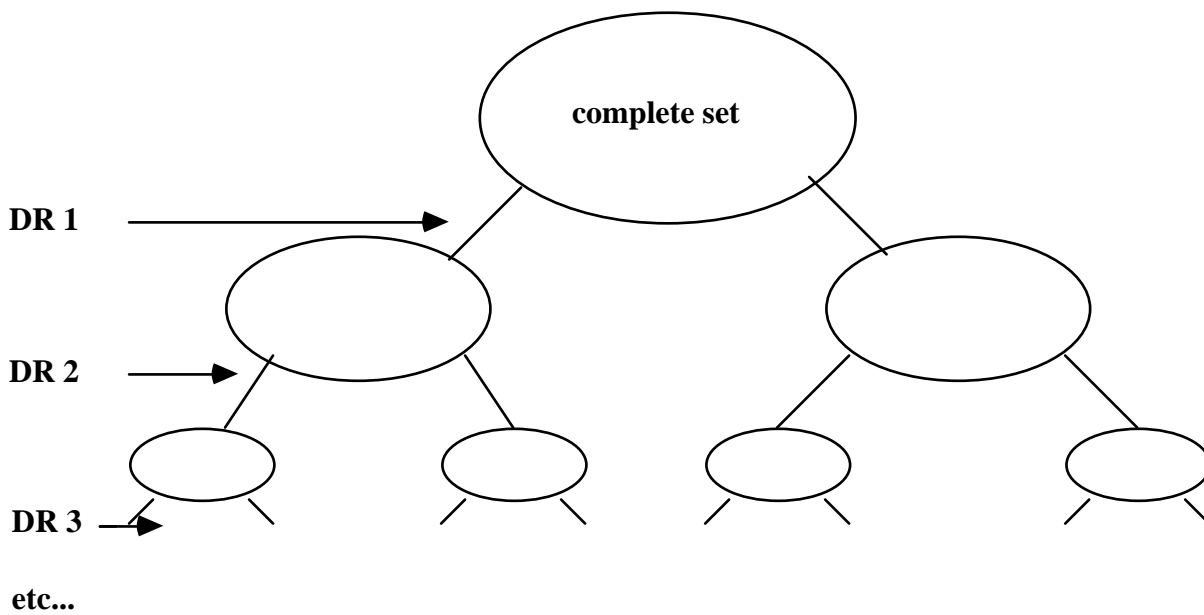


Figure 2.

4 Dynamic production of knowledge and learning

Dynamic production of knowledge (Walker, M.G., Blum, R.L., 1986) and learning (Michalsky, R.S., 1983), (Davis R., 1979), (Politakis, P.G., Weiss S.L., 1984) take place during a patient observation and diagnosis research. The system questioning and the answer analyse are made by phases squared with the various observation levels. At each phase, the system has to answer to this question : "is the Factual Base state (FBs) totally new, or has it been observed in the past ?". So, the meta-rule used as the basic postulate is:

IF The actual state of the the Factual Base is known from the system

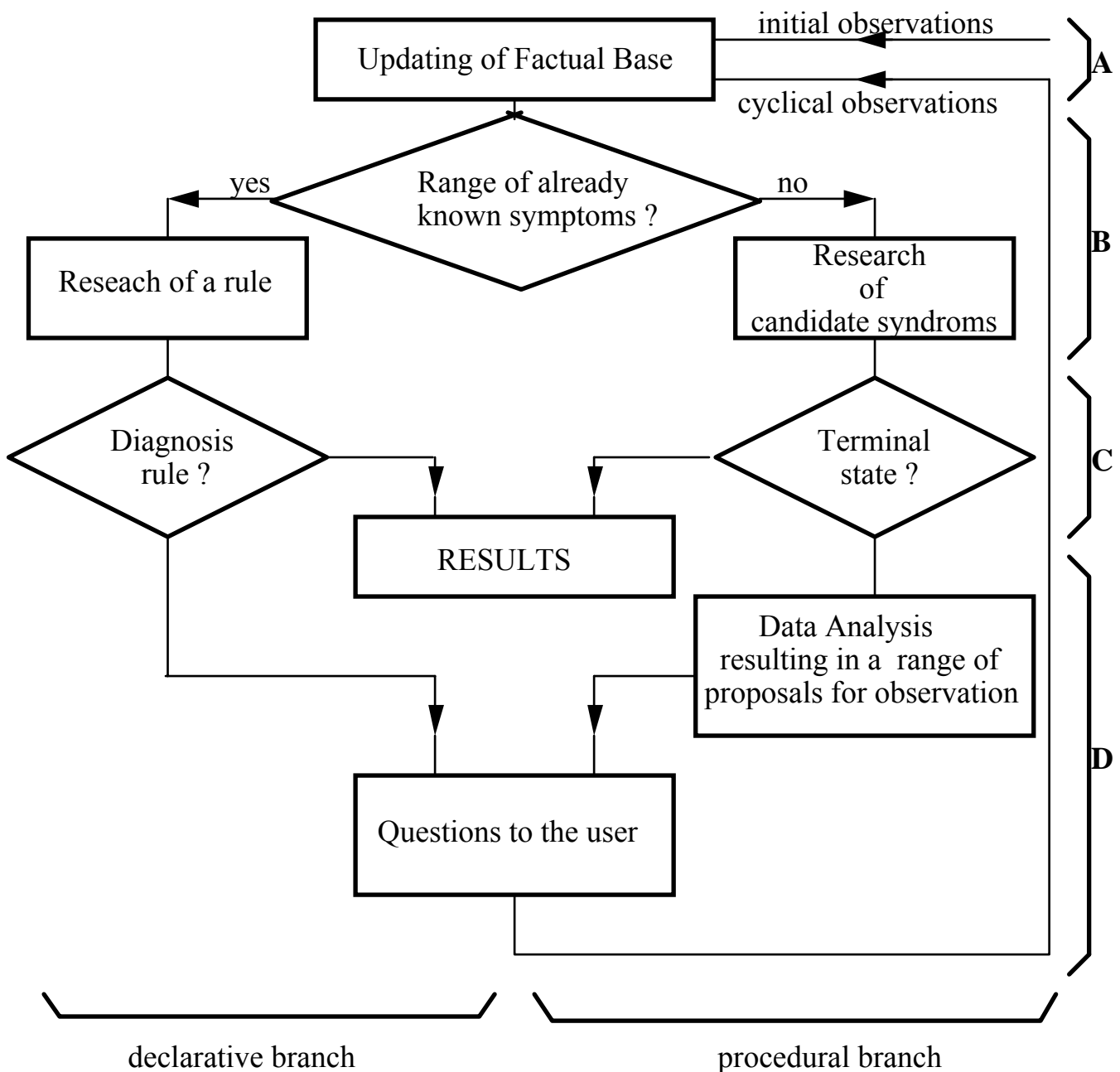
THEN 1° trigger a Data Analysis,
 2° memorise the couple (FBs, conclusion or question) deduced from
the Data Analysis

ELSE trigger the corresponding knowledge

Every time the system has produced an answer to the Factual Base state, by means of Data Analysis, this answer is stored in a Learning Base as a couple (FBs, consequence). The answer of the system is, more precisely, the second element of the couple (FBs, conclusion or question) and may be either a syndrom or syndrom set proposal, or an s.e. to propose the physician to observe.

The Learning Base is embedded in the Knowledge Base and exclusively gathers the Data Analysis obtained knowledges.

The figure 3. shows the complete and ideal cycle covering the thorough study of a patient:



- A : situation**
- B : action choice**
- C : test**
- D : actions**

Figure 3.

5 Prospective methodology

We are going to expose a methodology in a learning situation made of one to possibly four phases. This methodology is applicable under the hypothesis of a vast data base, made of the maximum number of symptoms and syndroms. The examples shown here are merely an illustration of the method. The latter include a procedural and a declarative aspect.

The analysis of the syndroms projection on the plans generated by the couples of pertinent axes of the first observation level (see the preparatory static study) has revealed the existence of two

remarquable groups of syndroms on each of the 7 plans. These two groups are respectively composed of:

- the syndroms that are situated at the origin of the axes (coordinate: 0,0),
- the syndroms that have at least one non null coordinate.

Starting from this fact, we propose, as we will see in the first phase, a dichotomic approach of reduction of the candidate syndroms set. Here is the description of the first two phases.

First phase:

It comprises the generating process of the diagnosis hypothesis at observation level 1, and includes 3 steps:

- **step one**, data base initialising: The physician provides the system with the list of observed symptoms. the system complete this list with the symtoms that are situated over the observed symptoms in the symptoms hierarchy. This adding of less precise symptoms is necessary since many syndroms, in the litteracy, are imprecisely defined with non terminal symtptoms (such as abnormal hands). these inexplicit symptoms are rarely those which are transmitted by the physician to the system. The observed symptoms such as precise congenital abnormalities may be seen as details, but nevertheless, they have to capture the syndroms referred to by less descriptive ones.

This step one of the first phase comprises a reinforcing process of the patient observation by mean of the level 1 pertinent axes. This process consists in the proposal of a series of complementary observations corresponding to the symptoms belonging to the pertinent axes and not belonging to the actual Factual Base. The confirmation or invalidation of these symptoms by the physician will speed up the the reduction of the candidate syndroms range since the corresponding axes are the most discriminant with regard to this range.

- **step two**, reduction of the candidate syndroms range by means of the dichotomic rules built up thanks to the preparatory static study. Theoretically, there may be an explosion of possible intersections (2^n intersections for n rules). In reality, this number has to be strongly decreased because of a number of empty intersections.

There are three types of ways from the root of the dichotomic tree to its leaves:

- the way leading to an intersection endowed with a lot of syndroms (more frequent type). This way causes the continuation of the process through the second phase.

- the way leading to a single syndrom or a restricted number of syndroms. This way causes the end of the analysing process and the proposal of this syndrom or range of syndroms for the physician's valuation.

- the way leading to an empty intersection. This case may square with the occurrence of a new syndrom.

The last two types result in the ending of the learning process, whereas the first one leads to a decreased range of syndroms which is offered to the second phase of the analysis.

- **step three**, fuller reseach of candidate syndroms by means of nearness analysis: this opeation is an analysis of nearness between the syndrom-points and the patient-point in the vectorial space. As a result of this analysis, we retain as as candidates the syndrom-points whose distance to the patient-point is lower than an experimentaly fixed threshold. This threshold is the radius of a sphere centered on the patient-point.

The set A1 of the candidate syndroms constructed by nearness analysis may be completed by the set A2 of the syndroms that are directly selected by coincidence with the patient symptoms panel. If the set A2 - A1 is not empty, the system will fire a procedure to examine the axes which are very charged by the syndroms of this set.

The nearness research of candidate syndroms and the research by Factual Base coincidences are two methods that reinforce each other. Seemingly, it is possible to envisage the use of other methods. Equally, it is not out of the question that a complementary declarative process be used to strengthen the phase of generating hypothetical diagnosis.

At the end of this first phase, we obtain a complete selection of candidate syndroms. This selection may be empty or have only a few syndroms. In this case, the system proposes to the physician the results of its investigation. There is at this point the production of a new knowledge in so far as the observation panel has led to a terminal state of the enquiring process which is the failure of the system or the diagnosis proposal. After that, the physician has to decide if he retains or rejects this response in the Learning Base.

Second phase:

The second phase of this study intends to analyse the syndroms selection descended from the first phase. It consists of a Principal Components Analysis (PCA) of the projection of these candidate syndroms on the vectorial subspace constituted from the second observation level axes.

The first two eigenvectors of the canonical PCA reveal the more discriminant axes. The research method of the pertinent axes lies in retaining the axes (number inferior to 4 or 5) that have the strongest contributions in the first eigenvector, then in the second one. This pertinent axes range allows the systems to make a request of complementary informations to the physician (as a confirmation or invalidation of the symptoms brought by these axes).

This action results in a new state of the Factual Base by the adding of new symptoms in their validated or invalidated aspect. As a consequence, this new state induces a reduction of the candidate syndroms set. Two ways are possible depending on the state of this revised syndrom set:

- either this set is empty, or has a single element, or is sufficiently restricted to be proposed as a result of the investigation. In the same manner as in the first phase, this possibility leads to the concluding of the process.

- or this set still contains a lot of syndroms. This case results in the activating of the third phase of the investigation which consists in the analysis of this set in the third observation level subspace, in exactly the same manner as in the second phase. The same actions may be repeated until the fourth level.

6 Conclusion

In this paper, we have dealt with a particular kind of scientific research field characterized by the vastness of its knowledge object set, the rapid increasing of it, its uncertainty, and therefore its perpetual incompleteness. The Medical Genetics field is representative of such fields that in spite of their features need the production of a decision making knowledge.

As a response to this decision making knowledge production request, the use of the Data Analysis is justified because of the very nature and aims of this discipline. As a matter of fact Data Analysis concentrates a number of features able to bring solutions to the foreseen problem. Namely:

- It is adapted to the processing of large data sets,

- it reveals empirical relations between variables,
- it furnishes a rigorous mathematical background to the data modelling,
- it may offer a spacial visualisation for complex situations or phenomenas,
- therefore it produces valuable decision elements and counsels, capable of being admitted into a knowledge based system.

As a consequence, we may consider either the Data Analysis as an "expert entity" to whom we may propose observation datas to, and who gives back his conclusions, or as a rigorous and efficient tool, usable by a specialist to increase, reinforce or verify his personal knowledges.

Decision making as a goal is equally shared both by Data Analysis and by Expert Systems. So, it is not vain to associate the two in the domains where they are both efficient. This intention is all the more justified as if the expert were endowed of the capacity to mentally realise the data modelling, and construct the inertia matrix, he would obviously use such a capacity in his efforts to take a decision.

The integration of the Data Analysis in the expert system environment as a learning and knowledge base making tool does not go without some difficulties since this discipline is devoted to the treatment of numbers whereas expert systems are rather consecrated to the treatment of symbolic knowledges and variables. To this aspect of the problem, we have proposed a solution consisting to establish a bijection between the semantic and vactorial axes. It is obvious that the behaviour of the system depends on the art and the maner this bijection is realised. Nervertheless, the method has yet shown a certain number of results that augurs well of its pertinency and efficiency.

References

- Benzecri, J.P. & all. (1981). *Pratique de l'Analyse des Données*. ed. Dunod, Paris, 1, 2, 3.
- Davis R. (1979). Interactive Transfert of Expertise: Acquisition of New Inference Rules. *Arificial Intelligence*, **12**, 121-57.
- Edwards, C. N. & Buyse, M. L. (1986). Building a computerized international network for medical expertise: the birth defects information system. *Proceedings of IA-BIOMED 86 conference*, Centre de Recherche en Informatique Médicale, Montpellier (France), 115-9.
- Gouvernet, J. (1979). Apport de Méthodes de Classification en Génétique Médicale, application aux maladies osseuses constitutionnelles. *Medical thesis*, Faculté de Médecine de Marseille (France).
- Macgibbon, B. & Preus, M. (1979).The Distorted Shell Method of Clustering for Syndrom Classification. *American Journal of Human Genetics*, **31**, 498-507.
- Michalsky, R.S. (1983). A theory and Methodology of Inductive Learning. *Arificial Intelligence*, **22**, 111-61.
- Politakis, P.G. & Weiss S.L. (1984). Using Empirical Analysis to Refine Expert System Knowledge Bases. *Arificial Intelligence*, **22**, 23-48.
- Preus, M., Ayme, S. (1983). Formal Analysis of Dismorphism: objective methods of syndrom definition. *Clinical Genetics*, **23**, 1-16.
- Rialle, V. (1985). Génération de Connaissances Expertes - Application à la Génétique Médicale. *Proceedings of the Ffth International Workshop on Expert Systems*, Avignon (France), **5**, 901-13.
- Robert, C., Zarski, M. & Demongeot, J. (1986). Utilisation de la notion de score comme outil décisionnel en médecine, proposition pour la fabrication automatique de règles de production. *Proceedings of IA-BIOMED 86 conference*, Centre de Recherche en Informatique Médicale, Montpeliler (France), 234-44.
- Veloso, M. & Feijoo, M.J. (1986). DYSMOR: Computer Based Differencial Diagnosis of Dymorphic Syndromes.MEDINFO 86, *Elsevier SCientific Publishing*, North-Holland, 216-8.
- Walker, M.G., Blum, R.L. (1986). Towrad Automated Discovery from Clinical Data Bases: the RADIX Projet. MEDINFO 86, *Elsevier SCientific Publishing*, North-Holland, 32-6.

I.3.2 Deuxième thème : Représentation hétérogène de connaissances (Electromyographie)

Présentation de l'article :

RIALLE V., VILA A. & BESNARD Y., Heterogeneous knowledge representation using a finite automaton and first order logic: a case study in electromyography. *Artificial Intelligence in Medicine*, 3 (2), 1991, 65-74.

HETEROGENEOUS KNOWLEDGE REPRESENTATION USING A FINITE AUTOMATON AND FIRST ORDER LOGIC: A CASE STUDY IN ELECTROMYOGRAPHY

Abstract :

In a certain number of situations, human cognitive functioning is difficult to represent with classical artificial intelligence structures. Such a difficulty arises in the polyneuropathy diagnosis which is based on the spatial distribution, along the nerve fibres, of lesions, together with the synthesis of several partial diagnoses. Faced with this problem while building up an expert system (NEUROP), we developed a heterogeneous knowledge representation associating a finite automaton with first order logic. A number of knowledge representation problems raised by the electromyography test features are examined in this study and the expert system architecture allowing such a knowledge modeling are laid out.

Keywords: Medical expert systems, Heterogeneous knowledge representation, Finite automata, Electromyography.

1. Introduction

The various kinds of knowledge and reasoning used in expert systems (ES) have been carefully analyzed and classified over several years [6,11,17]. Nevertheless some types of knowledge remain difficult to represent by means of classical structures (production rules, frames, semantic nets, etc.) commonly used in expert systems. We were faced with this kind of problem while building up an expert system for the electrophysiological diagnosis of neuropathies (NP). The first version of knowledge about NP included in our system was described in 1987 [20].

Several projects have already been completed in the field of artificial intelligence applied to electromyography (EMG). The MUNIN system [2,3] is one of the major realizations. This project is based on a neo-bayesian approach of the inferential process. The reasoning framework is constituted with a probabilistic causal network in which each node has a concurrent conditional probability table. The main task of the system is to calculate probabilities of terminal nodes (illnesses) when some test result values are fed into the network.

The PC-KANDID system was developed in Prolog by A.Fuglsang-Frederiksen's team [7,13]. The system is founded on a logical rule based inference, without uncertainty management, and on an interactive cycle: planning - test - diagnosis. It covers most EMG diagnoses. NEUREX [12], an application of the Parsimonious Covering Theory, is intended to be a general framework for neurological localization and diagnosis. A first outline of an expert system designed for electromyography: MYOSYS, in Prolog language, was outlined in our laboratory by D. Ziebelin [19] in 1984. In contrast to the former systems, which deal with wide fields, the ADELE system [4] is an expert system devoted to the diagnosis of a single disease: the carpal tunnel syndrome. The MYOLOG system [8] realized in Prolog is intended for the diagnosis of cervical radiculopathies and brachial plexus neuropathies, from spontaneous and voluntary muscular activities and clinical data. A system resulting from a neural network and an augmented transition network has been developed in Lisp by P.P. Jamieson [10]. Finally, we should mention the increasing use of artificial neural nets to classify patients into predetermined groups of pathologies [15].

We do not intend to present a new EMG expert system. Our goal is to expose an artificial intelligence modeling problem in a particular field of medical knowledge and practice. The choice of concepts described in this study is based on 2 major reasoning features that lead to NP diagnosis: this diagnosis is based, on one hand, on spatial distribution of lesions along the nerve fibres and, on the other hand, on synthesis of the different local diagnoses (at the nerve segment level, then at the whole nerve level).

Analysis of the lesions distribution along the nerves involves a spatial reasoning that may be either systematic or guided by heuristics. The highlighting of such heuristics, which is part of cognitive psychology [1], seems obviously more complex than the use of systematic reasoning. Nevertheless, the latter, even if described in propositional logic, would still be cumbersome with regards to the considered reasoning features. Therefore, the knowledge modeling solution we propose is heterogeneous: it is based on a finite automaton [9,14] used jointly with three other knowledge sources. First order logic is used in almost every part of the system : inference engine, finite automaton and production rules. The overall system is written in Prolog [5,16,18].

First we will briefly describe the main characteristics of an EMG examination. Then, we will lay out the general structure of the EMG knowledge leading to the diagnosis of neuropathies. A number of knowledge representation problems raised by the EMG test characteristics will be tackled and the choice of a heterogeneous knowledge representation using a finite automaton, as a solution to these problems, will be explained. Due to the number of knowledge sources, we will not describe them extensively but we will focus in greater detail on the description of the automaton and on the overall expert system architecture which make its use possible.

2. EMG principles

EMG consists of different suitable electrophysiological techniques aimed at studying the peripheral nerve system (PNS), muscles and neuromuscular junction. Neuropathies are diseases of the PNS and especially of the nerves innervating the face and limb muscles. Lesions detected by the EMG examination affect either axons of the fibres (*axonal lesions*) or myelin sheath (*demyelinating lesions*); if both structures are concerned, we call them *mixed lesions*. The neuropathy diagnosis is

provided by the EMG examination of several nerves, mainly median and ulnar nerves for the upper limbs, peroneal and tibial nerves for the lower limbs.

The diagnosis activity involves 2 main steps, with a specific type of knowledge for each step. These two steps are as follows:

- *Step A*: successive tests are performed on each chosen nerve. The choice of nerves is made according to the patient's symptoms and signs. These tests are included into charts that have been standardized in our laboratory. They are chosen according to the clinical findings — i.e. the diagnosis hypotheses —, and to the nerve under study. The number of segments that we usually explore is variable: generally between 1 and 2 for sensitive fibres, and between 2 and 5 for motor fibres. For example, in the median nerve study, there are 2 stimulation points — palm and wrist — for the sensory nerve conduction test, and 5 segments for the motor nerve conduction test.

Electrophysiological parameters to be analyzed on the sensory and muscle action potentials are:

a) sensory fibres: sensory action potential (SAP) and sensory nerve conduction velocity (SCV) on both segments; amplitude ratio of the SAP for the second segment.

b) motor fibres:

- for the first segment: amplitude and distal latency of the compound muscle action potential (CMAP),

- for the next four segments: amplitude of the CMAP, amplitude ratio of the successive CMAP, and motor nerve conduction velocity (MCV).

Each such examined nerve is given a diagnosis, whose process will be described below.

- *Step B*: the overall diagnosis is made up from the synthesis of the diagnoses for each nerve (*local synthesis*) and for the various nerves (*overall synthesis*).

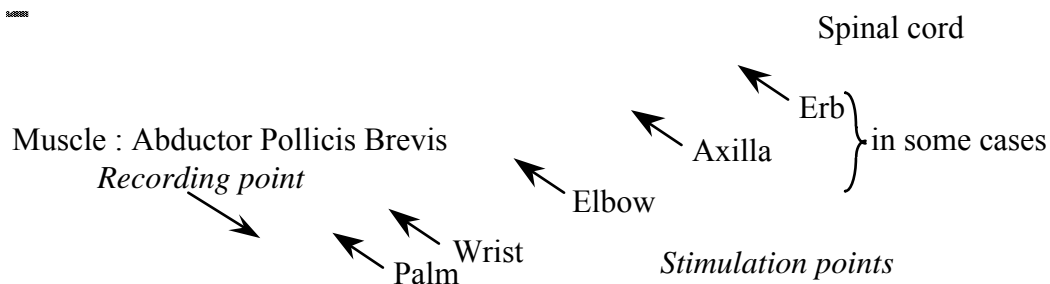


Fig. 1. Motor median nerve segmentation

3. Knowledge description and representation

3.1. knowledge levels

These two stages of the examination lead to the three following levels of knowledge (Fig. 2):

Level 1 : analysis of one nerve segment (step A,1),

Level 2 : analysis of a total nerve (step A,2): synthesis of the segment analyses,

Level 3 : analysis of the whole nerves (step B): synthesis of the nerve analyses.

Thus, both level 2 and level 3 correspond to the synthesis of the preceding level.

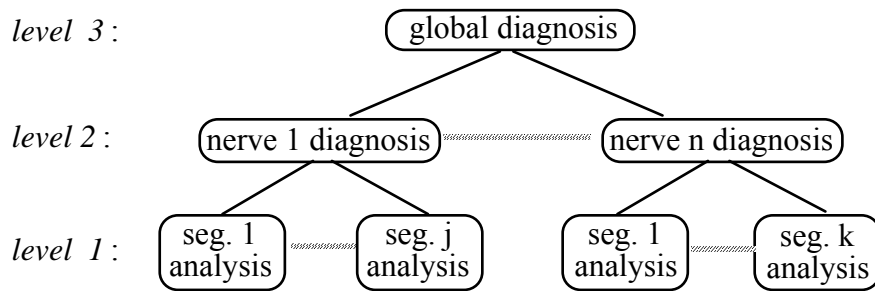


Fig. 2. Diagnosis analysis tree

3.2. Knowledge representation at different levels

Level 1: there are three types of segments depending, on the one hand, on the kind of fibre under consideration (sensory or motor), on the other hand, on the position of the segment with regards to the detection point. For a given type of segment, the diagnosis is given by about ten production rules. All electrophysiological variables, such as amplitudes, velocities, amplitude ratios, etc., are continuous. Their values are interpreted by an interpretation module. This module creates the corresponding semantic variables getting their values in sets such as {normal, decreased, very decreased} or {normal, increased, very increased}. These semantic variables are used as premises to the segment diagnosis rules (Fig. 3).

	<i>if</i>	amplitude at the wrist	is	very decreased
		<u>amplitude at the wrist</u>		
<i>and</i>	<i>if</i>	ratio: amplitude at the palm	is	normal
<i>and</i>	<i>if</i>	velocity at the wrist	is	normal or decreased
	<i>then</i>	lesion of the segment	is	severe axonal
—————				
	<i>if</i>	amplitude at the wrist	is	normal
		<u>amplitude at the wrist</u>		
<i>and</i>	<i>if</i>	ratio: amplitude at the palm	is	normal or decreased
<i>and</i>	<i>if</i>	velocity at the wrist	is	decreased
	<i>then</i>	lesion of the segment	is	mild demyelinating

Fig. 3. Sample of diagnosis rules applied to a motor segment of the median nerve

Level 2: Each segment whose diagnosis is different from the *normal* label is *pathological* (we do not take into consideration, at this stage, the kind of lesion: axonal, demyelinating or mixed). The knowledge leading to the diagnosis of the nerve is based on the spatial distribution of the pathological segments along the nerve (Fig. 4).

In level 2 there are three possible kinds of pathologies, determined according to the following rules:

- *focal neuropathy:* only one segment is affected
- *multiple focal neuropathy:* several non-contiguous segments are affected
- *diffuse neuropathy:* several contiguous segments are affected

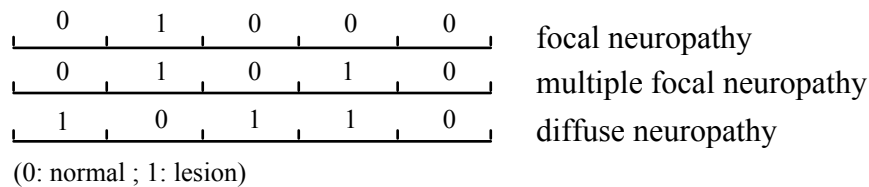


Fig. 4. Examples of spatial distribution of lesions

Level 3: As in level 2, knowledge leading to the diagnosis of all the nerves is based on a spatial distribution of the nerve lesions as well as on the number of lesions, with moreover a notion of symmetry (Fig. 5) (symmetrical nerves are called *homologous*).

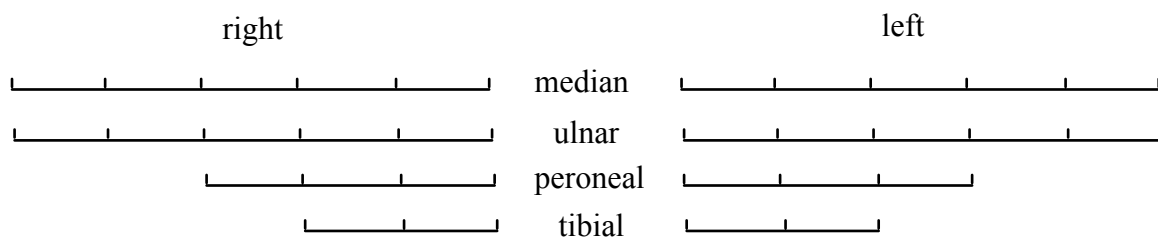


Fig. 5. Diagram of spatial distribution of the segments

The possible overall diagnoses are the following: { focal mono-neuropathy, multiple focal neuropathy, diffuse mono-neuropathy, symmetrical poly-neuropathy, asymmetrical poly-neuropathy, uncertain diagnosis, normal examination }. The processing rules of association between the affected nerves and the diagnosis are as follows:

- focal mono-neuropathy: one effected segment (a.s.) on only one nerve
- multiple focal mono-neuropathy: several non-contiguous a.s.
(distributed or not on the whole nerves)
- diffuse mono-neuropathy: at least two contiguous a.s. on the same nerve
- symmetrical poly-neuropathy: at least two contiguous a.s. on homologous nerves
- asymmetrical poly-neuropathy: at least two contiguous a.s. on non-homologous nerves
- normal examination: no a.s.
- uncertain diagnosis: any other case

4. Problems of knowledge representation at level 2

Each of the three knowledge levels that have just been described needs its own model of representation. Level 1 corresponds to a simple classification knowledge that directly associates, for any segment, the electrophysiological values to the diagnosis. This knowledge does not pose any problem of representation.

Level 2 is where a representation problem starts. Indeed, as previously stated, the diagnosis of the whole nerve needs a particular element of reasoning: the *relative location of lesions*.. The practitioner's appreciation of this location stems from an empirical knowledge. The distribution of the segments affected by a lesion could be represented by a categorical variable, called *relative location of the lesions*, the value of which is in the set : {unique isolated lesion, non-contiguous isolated lesions, contiguous lesions}. The associated diagnosis is obtained by application of rules described in point 3.

The most commonplace solution would consist in creating such a variable and proposing the following choice to the user (by means of a dialogue screen and whenever the reasoning needs it): a) no segment affected, b) a unique segment affected, c) several isolated segments affected, d) several contiguous segments affected. Obviously, the consequence of this solution would be to considerably diminish the degree of *intelligence* of the system, since it is possible to deduce the answer from the state of the factual base.

A second solution would consist in representing all the possible figures in production rules such as: *if* seg₁ is V₁ and seg₂ is V₂ and ... and seg_n is V_n *then* nerve_k is W_k; with V_i ∈ { normal, pathological } and W_k ∈ { normal examination, focal neuropathy, multiple focal neuropathy, diffuse neuropathy }.

Nevertheless, this solution has two major drawbacks:

- *First*: the number of possibilities for one nerve with n segments is 2ⁿ. Since, first, n is comprised between 1 and 5 and, second, there are at least 20 couples of rachidian nerves likely to be analyzed, we obtain an explosion of potential combinations, even in first order logic.
- *Second*: from a scientific or medical point of view these production rules have no interest since they are simple basic operation rules.

5. Finite automaton

Taking into account these knowledge representation problems, we made the choice of inserting in the knowledge base and in the inference engine an extremely simple and safe mechanism with regards to its principle as well as its recursive representation in first order logic: the *finite automaton*. This solution allows an almost immediate association between the list of segment statements and the nerve diagnosis.

5.1. Definition

A finite automaton consists of a finite set of states and a set of transitions which make it possible to move from one state to another when a symbol, from a predetermined set of input symbols, is fed into the automaton. The finite automaton is formally defined by a 5-tuplet (Q, Σ, δ, q₀, F) in which

Q is the finite set of states, Σ is a finite input alphabet, q_0 belongs to Q and designates the automaton initial state, $F \subseteq Q$ is the set of final states, and δ is the transition function associating each couple (q,a) of $Q \times \Sigma$ with an element p' of Q : $\delta(q,a) = p'$.

With regards to our application, the finite set of states is $Q = \{ \text{start}, n, f_a, f_b, m_f_a, m_f_b, d \}$. The symbols n, f_a, m_f_a and d represent normal, focal, multiple focal and diffuse states respectively. These states are represented by large triangles in the transition diagram (Fig. 6). The symbols f_b and m_f_b (small triangles in the diagram) correspond to states which repeat a principal state and yet direct the analysis towards a new principal state.

The set of final states is: $F = \{n, f_a, f_b, m_f_a, m_f_b, d\}$.

As indicated in 3.2, the normal segments are symbolized by 0 and the affected ones by 1, hence: $\Sigma = \{ 0, 1 \}$.

For example, the following input chain : $[0,1,0,1,0]$, called *chain of segment states*, reflects the state of a nerve whose second and fourth segments are affected, the three others being normal. Therefore the definition of an automaton for the analysis of a nerve is : $(Q, \Sigma, \delta, \text{start}, F)$.

The symbolic functioning is the following: the chain of segment states is read from left to right or from right to left (the direction is not important). Each reading of a new value from the input chain introduces a *state transition* from the previous state. For example, let n be the automaton state at a given time of the analyzing process, and let 1 be the new input value, then the automaton moves to the state f_a (i.e. $\delta(n,1) = f_a$). The initial state, before the first input, is the *start* state. The final state (after the last input) corresponds to the nerve diagnosis. Since there are several transition possibilities (exactly 2) for every state, the automaton is called *nondeterministic*.

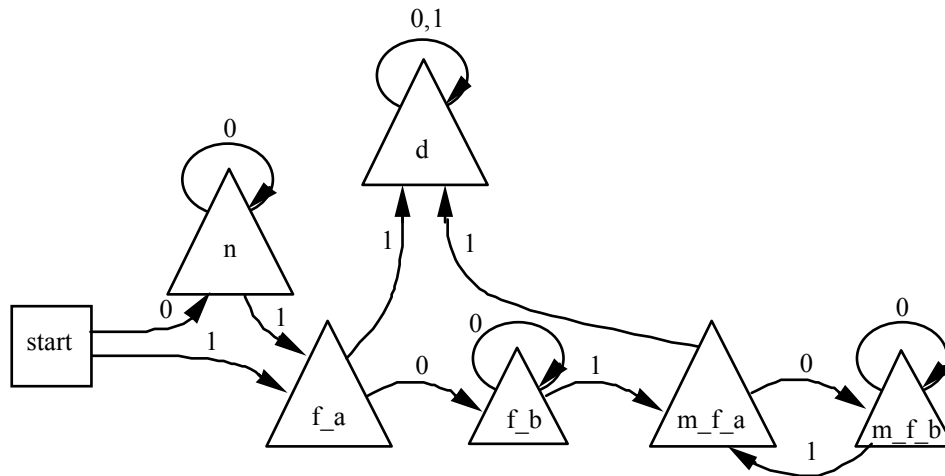


Fig. 6. Transition diagram of the finite automaton

n : normal state	f_a, f_b : focal neuropathy
d : diffuse neuropathy	m_f_a, m_f_b : multiple focal neuropathy

The transitions defining the automaton and expressed in relational form, as in Prolog language, are shown in Figure 7. The automaton mechanism is expressed in the recursive predicate shown in Figure 8.

```

transition ( start , 0 , n ).
transition ( start , 1 , f_a ).
transition ( n , 0 , n ).
transition ( n , 1 , f_a ).
transition ( f_a , 0 , f_b ).
transition ( f_a , 1 , d ).
transition ( f_b , 0 , f_b ).
transition ( f_b , 1 , m_f_a ).
transition ( m_f_a , 0 , m_f_b ).
transition ( m_f_a , 1 , d ).
transition ( m_f_b , m_f_b , 0 ).
transition ( m_f_b , 1 , m_f_a ).
transition ( d , 0 , d ).
transition ( d , 1 , d ).

```

Generical form : transition (<former state>,<current input>,<resulting state>)

Fig. 7. Automaton transition relations

```

automaton ([], Final_state, Final_state).
automaton ([_], d, d).
automaton ([T|Q], Previous_state, Final_state):-
    transition ( Previous_state, T, New_state),
    automaton (Q, New_state,Final_state),!.

```

Fig. 8. Finite automaton predicate

5.2. Utilization

The finite automaton described above simulates the knowledge of level 2. One can consider that it is in itself a small-scale system (inference engine + knowledge base) in which the motor is constituted by the recursive predicate *automaton*, and the knowledge base is represented by the transition set. Thus, the automaton insertion in the expert system takes place at the inference engine level as well as at the knowledge base level.

The system architecture includes 4 cooperating knowledge sources successively used through the 4 phases of the diagnosis process (Fig. 9). A special inference engine supervises the successive use of the knowledge bases. These knowledge bases communicate via a central working memory, as in the case of a *blackboard*.

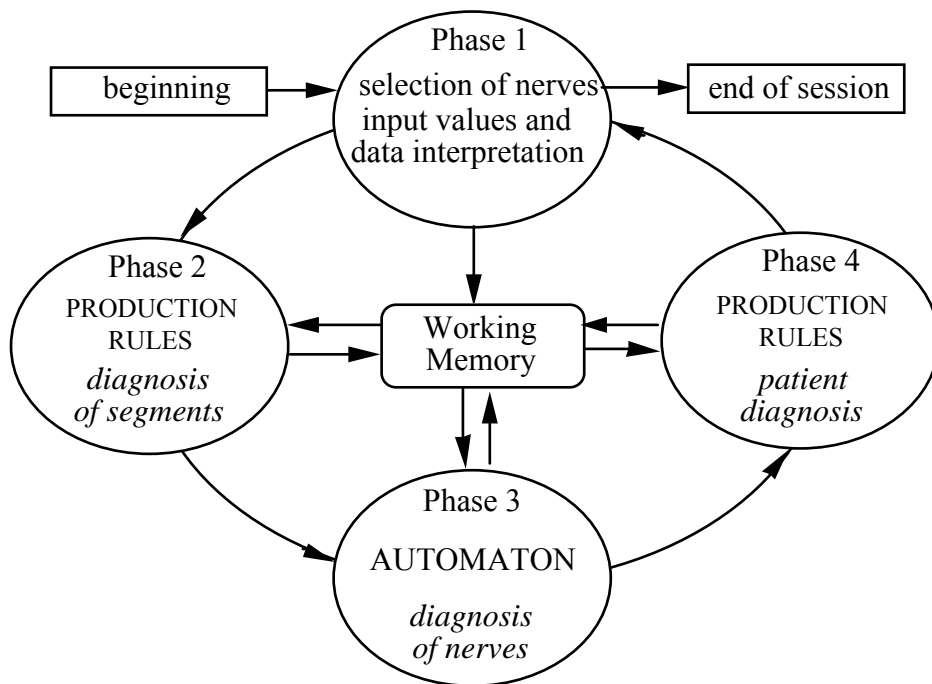


Fig. 9. Location of the automaton in the diagnosis research process

The 4 temporal phases of the process are the following:

- phase 1: nerve selection by the physician, input and interpretation of electrophysiological parameters,
- phase 2: inference cycles on each segment (segment diagnosis),
- phase 3: automaton activation on each nerve (nerve diagnosis),
- phase 4: inference cycles on all nerves (patient diagnosis),

The *working memory* allows result sharing between the different inference phases, i.e.:

- from phase 1 to phase 2: sharing of the electrophysiological parameters,
- from phase 2 to phase 3: sharing of the diagnoses of segments,
- from phase 3 to phase 4: sharing of the diagnoses of nerves.

6. Conclusion

We showed that there are some scientific knowledge fields for which classical structures such as production rules, frames or semantic nets are not suitable. EMG is a typical example of such fields: the physician performs a spatial reasoning from the distribution and number of abnormal segments. Such a reasoning is without doubt the result of different procedures depending on the physician (competence, experience, etc.). We leave it to cognitive psychology to study human functioning in establishing such a diagnosis. Thus, the model we described *is not a priori* a human function model. The synthesis diagnosis is set up by a finite automaton which repeats a sequential, spatial and numeric reasoning. This reasoning is rather similar to that of a beginner's or a very well organized but nervous person's.

The automaton stands among three other knowledge sources, whose functioning is supervised by an appropriate inference engine. The insertion of this automaton in the core of the expert system has

notably influenced the architecture of the whole system. First order logic is used in the construction of the automaton as well as in the representation of knowledge of levels 1 and 3.

This solution allows a significant decrease of the number of production rules and of the system response time. The current version of the system does not use any probability or uncertainty weighting. A new version using an approximate reasoning is under study.

This heterogeneous knowledge representation is operational in the NEUROOP expert system, written in Prolog and currently under evaluation in the EMG Laboratory of the teaching hospital of Grenoble.

7. Discussion

The paradigm of expert systems includes the idea that knowledge accumulated in a knowledge base should reflect the expert's knowledge as exactly as possible. Classical knowledge representation structures such as production rules, frames or semantic nets must allow an easy *reading* of this knowledge for maintenance and updating purposes. This legitimate point of view has contributed to the wide development of knowledge based systems. However, such an opinion is limited to knowledge that easily lends itself to a symbolic representation (such a feature does not in any way preclude the intrinsic knowledge complexity).

It is not so when the conceptual model cannot be easily translated into classical structures such as production rules, frames, semantic nets or mixed models combining two or three of these structures. Even though it is not closely copied from these usual artificial intelligence structures, the automaton model is yet connected to the ES paradigm at least through its declarative form in first order logic. Moreover, as indicated in point 5, this model could be considered as a small-scale production system. Particularly, there is a strong analogy between the transition relation set as shown in fig. 7 and a small knowledge base. Especially, this transition set could be easily altered by addition, modification or withdrawal of transitions, according to a possible knowledge improvement. This will in fact be the case in the next version of the automaton, when specialists include knowledge related to *conduction blocks* (null amplitude of a sensory or compound muscle action potential) in the expert system.

Nevertheless the similitude to the ES paradigm remains limited. Indeed, there is no reason for affirming that the model constituted by the finite automaton reflects an expert's way of reasoning (it is difficult to imagine that an expert reasons as mechanically as the automaton does).

However, one major question raised from the adoption of a procedural knowledge representation in an expert system is: is it necessary to represent, with an absolute accuracy, the expert's way of reasoning? Is this accuracy not achieved in some cases at the cost of a considerable complication of the knowledge base?

The weakness that one could reproach the automaton model with is that it does not use any heuristic knowledge — a major attribute of human intelligence —: the way to achieve the synthesis diagnosis of a nerve is an absolutely systematic and exhaustive reasoning. However, this weakness is compensated by the relatively low number of nerve segments (5 maximum) which minimizes the advantage of heuristic rules.

It should also be observed that, to be understood, the automaton we have devised needs very few concepts about the theory of finite automata, and constitutes an elegant solution to the spatial reasoning knowledge representation. As far as they are concerned, the medical experts who participated in the knowledge base building consider the adoption of the notion of the finite

automaton as a practical substitute to a human knowledge model whose conception is beyond their reach. This way, we can state that such a solution constitutes to some extent a contribution of artificial intelligence to the medical expertise modeling.

Acknowledgements. This work has been supported by the REGION RHONE-ALPES via the Pôle Rhône-Alpes de Génie Biologique et Médical under grant 509004/12 GBM.

References

- [1] J.R. Anderson, Cognitive Psychology, *Artificial Intelligence* 23 (1984) 1-11.
- [2] S. Andreassen, S.K. Andersen, F.V. Jensen, M. Woldbye, A. Rosenfalck, B. Falck, U. Kjærulff and A.R. Sorensen, MUNIN - An expert system for EMG, *Electroencephalography and Clinical Neurophysiology* 66 (1987) S4.
- [3] S. Andreassen and M. Wellman, MUNIN - On the case for probabilities in medical expert systems - a practical exercise, in: J. Fox, M. Fieschi and R. Engelbrecht (Eds.), *Lecture Notes in Medical Informatics* 33 (Springer Verlag, New York, 1987) 149-160.
- [4] A. Blinowska and J. Verroust, Building an expert system in electrodiagnosis of neuromuscular diseases: prototype system, *Electroencephalography and Clinical Neurophysiology* 66 (1987) S10.
- [5] I. Brakto, *Prolog Programming for Artificial Intelligence* (Addison-Wesley Publishing, Massachusetts, 1986).
- [6] W.J. Clancey, Heuristic Classification, *Artificial Intelligence* 27 (1985) 289-350.
- [7] A. Fuglsang-Frederiksen, J. Ronager and S. Vingtof, PC-KANDID: An expert system for electromyography, *Artificial Intelligence in Medicine* 1 (1989) 117-124.
- [8] R. Gallardo, M. Gallardo, A. Nodarse, S. Luis, R. Estrada, L. Garcia and O. Padron, Artificial Intelligence in the electromyographic diagnosis of cervical roots and brachial plexus lesions, *Electroencephalography and Clinical Neurophysiology* 66 (1987) S37.
- [9] J.E. Hopcroft and J.D. Ullman, *Introduction to automata theory, languages and computation* (Addison-Wesley Publishing, Massachusetts, 1979).
- [10] P.W. Jamieson, Symbolic analysis of electromyographic data, *Electroencephalography and Clinical Neurophysiology* 66 (1987) S50.
- [11] E.T. Keravnou and J. Washbrook, Deep and Shallow models in medical expert systems, *Artificial Intelligence in Medicine* 1 (1989) 11-28.
- [12] J.A. Reggia, S. Tuhim, S.B. Ahuja, T. Pula, B. Chu, V. Dasigi and J. Lubel, Plausible Reasoning During Neurological Problem Solving: The Maryland NEUREX Project, in: R. Salamon, B. Blum and M. Jorgensen (Eds.), *Proc. MEDINFO 86* (Elsevier Science Publishers, Amsterdam, 1986) 17-21.
- [13] J. Ronager, S. Vingtof, J.O. Kjærum and A. Fuglsang-Frederiksen, An EMG expert assistant in clinical neurophysiology (PC-KANDID), *Theoretical Surgery* 3 (1988) 48.
- [14] A. Salomaa, *Theory of automata* (Pergamon Press, Oxford, 1969).
- [15] C.N. Schizas, C.S. Pattichis, I.S. Schofield, P.R. Fawcett and L.T. Middleton, Artificial neural net algorithms in classifying electromyographic signals, in: *Proc. First IEE International Conference on Artificial Neural Networks* (Institution of Electrical Engineers, London, 1989) 134-138.
- [16] E. Shapiro and L. Sterling, *The art of Prolog* (M.I.T. Press, Massachusetts, 1987).
- [17] J.F. Sowa, *Conceptual structures: information processing in mind and machine* (Addison-Wesley Publishing, Massachusetts, 1984).
- [18] P.A. Subrahmanyam, The software engineering of expert systems: Is Prolog appropriate? *IEEE transactions on software engineering* SE-11 (1985) 11.
- [19] A. Vila, D. Ziebelin and F. Reymond, Experimental EMG expert system as an aid in diagnosis, *Electroencephalography and Clinical Neurophysiology* 61 (1985) S240.

- [20] A. Vila, L. Kress, V. Rialle, C. Robert and D. Ziebelin, Are expert-systems an aid for diagnosing neuropathies ?, *Electroencephalography and Clinical Neurophysiology* 66 (1987) S109.

I.3.3 Troisième thème : Modélisation cognitive : Rapport final de recherche pour le séjour d'une année de Délégation au Québec

Présentation du rapport :

RIALLE V., *Rapport final de recherche pour le séjour d'une année de Délégation au Québec*. Présenté à la Région Rhône-Alpes comme compte rendu de *Bourse d'Excellence* allouée pour ce séjour. Octobre 1992.

En raison de la longueur du rapport, seuls l'introduction générale, les objectifs, les résultats et la table des matières sont présentés afin de donner une vue d'ensemble des recherches effectuées.

RAPPORT FINAL DE RECHERCHE POUR LE SÉJOUR D'UNE ANNÉE DE DÉLÉGATION AU QUÉBEC

soutenu par une Bourse d'Excellence de la REGION RHONE-ALPES

INTRODUCTION GENERALE

Ce rapport expose les travaux de recherche qui ont été effectués à l'Université de Montréal et à l'Université du Québec à Montréal, durant l'année universitaire 1991-92, et pour lesquels une bourse d'Excellence m'avait été accordée par la Région Rhône-Alpes.

Etant donné que le principal axe de recherche se situait dans le projet "Intelligence Artificielle en Psychiatrie", les parties I à IV sont les reproductions intégrales du rapport final de ce projet, remis à la Région Rhône-Alpes.

Un addendum a été placé après la conclusion pour rendre compte d'importants travaux en intelligence artificielle et sémantique textuelle, réalisés durant le premier semestre de cette année. Ces travaux ont conduit à une publication de plus de 50 pages qui n'a pu, en raison de sa taille, être ajoutée aux annexes, déjà volumineuses. Les futures collaborations scientifiques envisagées incluront largement ce dernier domaine de recherche.

RAPPEL DES OBJECTIFS INITIAUX ET DU CONTEXTE DE LA RECHERCHE

Objectifs initiaux

La recherche d'une rationalisation de la démarche diagnostique est au centre des préoccupations de la Psychiatrie moderne. Cette rationalisation passe notamment par la définition d'un langage médical commun à l'ensemble des psychiatres. L'objectif de ce projet est de mettre au service de la recherche nosologique et de la modélisation des phénomènes psychopathologiques en psychiatrie les ressources de l'intelligence artificielle.

Pour remplir cet objectif, deux moyens opérationnels sont proposés :

1- l'évaluation d'un système expert : ADINFER dans une étude comparative de nosographies;

2 - la construction de modèles de connaissances, de raisonnements et de phénomènes pathologiques en psychiatrie. Ces modèles seront fondés d'une part sur l'intelligence artificielle symbolique, d'autre part sur les réseaux de neurones artificiels (systèmes hybrides symboliques-connexionnistes). Ils seront axés principalement sur le raisonnement analogique et ses perturbations psychopathologiques et constituent une recherche appliquée innovante en matière d'IA et connexionnisme.

Contexte de la recherche

Le projet de recherche informatique en psychiatrie vient concrétiser une coopération scientifique entre d'une part le Centre de Recherche en Psychiatrie de Montréal (récemment re-nommé Centre de Recherche Fernand Seguin — C.R.F.S.) sous la direction Scientifique du Pr. M. Ohayon, et d'autre part le Département de Mathématiques, Statistiques et Informatique Médicale (Laboratoire TIMB-IMAG, Pr. Jacques Demongeot) de la Faculté de Médecine de Grenoble. Le Centre de Recherche Fernand Seguin, affilié à l'Université de Montréal, et est installé dans l'hôpital psychiatrique Louis-Hippolyte Lafontaine. Un séjour de deux semaines dans ce centre, en février 1991, avait permis de jeter les bases de cette coopération.

Ce projet de recherche a été proposé à la REGION RHONE-ALPES dans le cadre son programme pluri-annuel INTELLIGENCE ARTIFICIELLE, SYSTEMES EXPERTS ET APPLICATIONS en 1991, et a obtenu une subvention de 80 000 F. Il répond notamment à la volonté de la REGION RHONE-ALPES de développer des relations scientifiques avec le Québec.

Une procédure de mise en *délégation* a été ouverte, afin que je puisse me consacrer, durant une année à Montréal, à ce projet. La procédure est ouverte à partir de mon invitation, à titre de *professeur invité* dans le département de Psychiatrie, de la part du Pr. Ohayon et du Docteur F. Borgeat, Directeur du Département de Psychiatrie de l'Université de Montréal.

La collaboration scientifique s'est déroulée en deux étapes :

- une première étape (étape A) durant laquelle je me suis consacré à temps plein au projet, dans les locaux du Centre Fernand Seguin, à Montréal. Cette étape a duré du 23 août 1991, date de mon arrivée à Montréal, au 31 décembre 1991. Elle a été principalement consacrée à la modélisation du raisonnement analogique en psychiatrie, et s'est conclue sur un pré-rapport partiel de recherche, de 22 pages, remis au Professeur Ohayon (présenté en annexe 1).
- une seconde étape (étape B), du 1er janvier 1992 au 2 août 1992 (date de mon retour à Grenoble), durant laquelle je me suis consacré à temps partiel au projet. Cette étape s'est également déroulée sur place à Montréal, mais à partir d'un nouveau laboratoire d'accueil¹. Elle a été consacrée à l'étude de l'impact de l'intelligence artificielle et des sciences cognitives en psychiatrie, et a conduit à plusieurs publications internationales avec des chercheurs du Centre F. Seguin.

RESUME DES RESULTATS

Les résultats obtenus au cours de ce séjour à l'Université de Montréal et à l'Université du Québec ont presque tous été soumis pour publication dans des revues et des congrès internationaux. Tous

¹Le changement de laboratoire est survenu pour des raisons administratives et a influencé d'une manière positive la collaboration scientifique entre mon laboratoire et l'université du Québec, via le REGION RHONE-ALPES.

les écrits concernant la psychiatrie et l'intelligence artificielle figurent en annexe de ce rapport. D'autres travaux dans le domaine des sciences cognitives mais non directement liés à la psychiatrie n'ont pas été ajoutés aux annexes, déjà considérablement volumineuses.

Résumé des étapes A et B confondues

1 • Voyage à Orlando (Floride), du 30 octobre au 3 novembre 1991, pour présenter deux communications :

a) RIALLE V., M. OHAYON M.: "Computers and Artificial Intelligence in Psychiatry: Brief History and State of the Art". In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1280-1281. (Annexe 2 du rapport de la Région)

b) RIALLE V., M. OHAYON M., AMY B., BESSIÈRE P.: "Medical Knowledge Modeling in a Symbolic-Connectionist Perspective" In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1109-1110. (Annexe 3 du rapport de la Région).

2 • préparation, avec le Dr. E. Stip du Centre F. Seguin, d'un workshop sur la Modélisation Cognitive dans le cadre du Congrès canadien sur l'étude du cerveau, du comportement et des sciences cognitives, deuxième réunion annuelle, Université Laval, Québec, 13 - 14 juin 1992.

3 • Voyage à la ville de Québec pour la présentation d'une communication au Congrès canadien sur l'étude du cerveau, du comportement et des sciences cognitives sus-mentionné.

Titre de la communication : Modélisation Cognitive: Présentation, éléments d'analyse et discussion. (Annexe 4 du rapport de la Région).

4 • réalisation d'un logiciel graphique interactif en Prolog pour l'expérimentation de théories sur l'inférence analogique, sur Macintosh, avec interface utilisateur graphique convivial.

5 • co-rédaction et soumission de l'article suivant :

RIALLE V., STIP E. & O'CONNOR K., Discussion of the issues and problems raised by computer as an aid to psychotherapy, *Humane Medicine*. (annexe 5 du rapport de la Région).

6 • co-rédaction et soumission de l'article suivant :

RIALLE V. & STIP E. : La Modélisation Cognitive: Principes et Exemples en Psychiatrie. *Journal of Psychiatry and Neurosciences*. (annexe 6 du rapport de la Région).

Il faut mentionner, en plus de cette liste, les remerciements qui m'ont été adressés par le docteur Alain Lesage du C. R. F.S. dans son article *La formation d'interviewers à l'utilisation de questionnaires standardisés en psychiatrie: travaux réalisés avec l'Examen Psychiatrique Standardisé*, publié dans le numéro 91 de la revue *Acta Psychiat. Belg.* Je lui avais gracieusement consacré le temps nécessaire à la mise au point de son manuscrit d'une trentaine de pages, au mois de novembre dernier.

Quelques mots complémentaires sur l'étape B

Le changement de laboratoire, survenu en janvier 1992, n'a que peu influencé le développement des recherches commencées à l'automne précédent: celles-ci se sont poursuivies (les points 2,3,5 et 6 ont été conduits durant cette période).

Le second laboratoire d'accueil est le Centre de Recherche en Cognition et Information, ATO-CI (dirigé par le Pr. Jean-Guy Meunier), de l'Université du Québec à Montréal (UQAM). Ce laboratoire, très bien placé parmi les laboratoires d'intelligence artificielle du Québec (80 personnes, 5 ordinateurs parallèles Volvox en réseau, 5 stations de travail, 10 stations micro-ordinateur, 15 millions de dollars canadiens de dotation reçus l'année dernière pour un vaste projet dans le domaine du parallélisme, de l'intelligence artificielle et du connexionnisme) s'est vivement intéressé à mes travaux. Plusieurs axes de travail ont donc été ouverts, parmi lesquels un axe spécifiquement psychiatrique répondant à mes préoccupations dans ce domaine.

L'UQAM possède son propre secteur de recherche en psychiatrie, dirigé par le Dr. E. Beltrami et membre du laboratoire d'accueil ATO-CI. Le Dr. Beltrami s'intéresse particulièrement, depuis plusieurs années, à l'aide au diagnostic par système expert. J'ai donc travaillé avec lui sur un système connexionniste d'aide à la classification automatique et au diagnostic, dans le domaine des troubles du comportement et des abus sexuels. L'annexe 7 donne un résumé des travaux qui ont été conduits. Ce résumé est la copie d'un rapport officiel remis à la direction du Centre ATO-CI. Un projet de coopération, clairement évoqué dans ce dernier document, est en cours de mise au point.

Table des matières

INTRODUCTION GENERALE	2
I. RAPPEL DES OBJECTIFS INITIAUX ET DU CONTEXTE DE LA RECHERCHE	2
I.1 Objectifs initiaux	2
I.2 Contexte de la recherche	2
II. RESUME DES RESULTATS	3
Résumé des étapes A et B confondues	4
Quelques mots complémentaires sur l'étape B	5
III. LISTE DES PUBLICATIONS	5
IV. CONCLUSION	6
ADDENDUM	7
ANNEXE 1:	
Pré-rapport de décembre 1991	
ANNEXE 2 :	
Communication : Computers and Artificial Intelligence in Psychiatry	
ANNEXE 3 :	
Communication : Medical Knowledge Modeling in a Symbolic-Connectionist Perspective	
ANNEXE 4 :	
Communication : Modélisation Cognitive: Présentation, éléments d'analyse et discussion	
ANNEXE 5 :	
Version de travail de l'article : Discussion of the issues and problems raised by computer as an aid to psychotherapy	
ANNEXE 6 :	
Article : La Modélisation Cognitive : Principes et Exemples en Psychiatrie	

ANNEXE 7 :

Rapport de recherche en psychiatrie au Centre ATO-CI

I.3.4 Quatrième thème : Epistémologie de la modélisation cognitive : coordination d'un numéro de revue philosophique

J'ai écrit le texte suivant — sur la problématique de la modélisation — en juin dernier à Montréal, alors que j'étais en année de Délégation, sur l'invitation de Madame Josiane Boulad-Ayoub, Directrice du Département de Philosophie de l'Université du Québec à Montréal (UQAM). Ce texte a été accepté par la direction de la revue LEKTON, Publication de ce Département de Philosophie (voir annexe 5), pour constituer le thème central d'un prochain numéro de cette revue (volume IV n°1).

Cette jeune revue — qui se définit comme interdisciplinaire et destinée à la discussion de questions théoriques fondamentales et de problématiques nouvelles — propose chaque année deux numéros ayant un thème propre. Son objectif est de favoriser échanges et débats dans le domaine de la philosophie et des sciences humaines au sein de la communauté scientifique. Elle est de plus en plus connue des milieux francophones et francophiles nord-américains intéressés par les sciences de la cognition.

La revue est en cours de réalisation avec une série d'articles de fond signés de quelques personnalités scientifiques faisant autorité dans le domaine.

Reproduction du sujet en préparation, tel qu'il a été accepté par la direction de la revue :

Projet de numéro de LEKTON sur le thème de la modélisation en sciences de la cognition

V. Rialle

Montréal, le 23 juin 1992

Présentation du thème

La modélisation des processus mentaux à l'œuvre dans les diverses activités cognitives humaines suscite un intérêt croissant dans le monde scientifique d'aujourd'hui. Modèle cognitif, simulation cognitive, auto-organisation, adaptation, émergence de comportements globaux à partir de comportements locaux, sélectionnisme génétique, mentalisme darwinien, énaction, etc., constituent autant de notions émergentes et présentes dans la modélisation neuro-psycho-cognitive. LEKTON se propose d'aborder toute une série de questions, des plus naïves aux plus embarrassantes, relatives à ces notions qui modèlent le paysage de la recherche cognitive en cette veille du troisième millénaire, et se place à l'avant-scène des questions épistémologiques posées par les sciences cognitives.

La psychologie cognitive et l'intelligence artificielle se sont retrouvées, par des voies différentes, sur le même terrain de la représentation de la mémoire, de l'intelligence et de la simulation cognitive, lui-même fertilisé par les progrès des neurosciences et de l'immunologie. Des modèles neuro-psycho-mimétiques fleurissent maintenant dans une grande variété de domaines: sciences du langage, sciences de la vie et de la santé, sciences humaines et sociales, technologies de la communication, robotique, intelligence artificielle,... L'ordinateur, s'il ne règne en maître à penser, règne en tout cas en maître d'ouvrage, étant donné qu'aucune des simulations de modèles parallèles et distribués ne saurait se concevoir sans faire appel à lui. La référence au connexionnisme pour l'explication des phénomènes mentaux et des comportements humains a reçu jusqu'à maintenant quelques sévères critiques. L'intérêt pour les modèles parallèles et distribués ne cesse néanmoins de se développer et d'exciter la curiosité de nombreux chercheurs dans les diverses branches des sciences cognitives. Le secret de cet engouement réside certainement dans les étonnantes capacités des réseaux neuronaux formels (RNF) à épouser des comportements neuro-biologiquement et psychologiquement vraisemblables. Leur apparente supériorité par rapport aux techniques de modélisation symbolique, proposées par l'intelligence artificielle classique, est due à plusieurs facteurs dont le principal est peut-être la possibilité qu'ils offrent d'expérimenter aisément des modifications de paramètres, des perturbations, voire des "lésions" structurelles et fonctionnelles, et d'en mesurer les conséquences dans le comportement général du modèle. Les effets induits par ces manipulations paramétriques sont ensuite rapprochés des fonctionnements cognitifs observés dans le règne animal et chez l'homme. Pour l'heure, les recherches se situent essentiellement au niveau de la mesure de l'adéquation entre le modèle et la réalité : jusqu'à quel point le RNF donne-t-il les réponses attendues et se comporte-t-il comme l'original vivant ? L'enthousiasme actuellement suscité par les RNF se situe essentiellement dans la recherche de cette adéquation et des mécanismes d'apprentissage et d'adaptation qui la rendent possible. L'induction de connaissances concernant le sujet humain à partir de l'observation du réseau et de la manipulation paramétrique reste néanmoins une question centrale. Que peut-on observer sur un modèle formel qui puisse informer ou répondre à une question concernant l'Homme ? Dans quelles conditions peut-on effectuer le mouvement inverse d'induction de phénomènes biologiques ou neuro-psychologiques, non directement observables chez le sujet humain, à partir des réponses et des comportements observables sur le RNF ? En bref, pourra-t-on élaborer des modèles mathématiques de la *psyché* et les utiliser pour inférer des comportements humains ? Va-t-on vers un modèle mathématique de l'esprit ? peut-on recréer technologiquement la pensée, et quel est l'intérêt de cet objectif prométhéen ?

La référence informatique, devenue de plus en plus présente depuis deux décennies, agit-elle comme un puissant outil conceptuel ou comme un biais méthodologique réducteur ? Le connexionnisme, la programmation parallèle (devenue possible par la disponibilité d'ordinateurs du même type), offrent-ils une alternative à la modélisation linéaire et aux modèles computationnels strictement séquentiels ?

I.3.5 Cinquième thème : Modélisation hybride (symbolique et connexionniste) de la compréhension automatique de textes

Il s'agit, comme indiqué dans la partie I.1, d'un domaine de recherche novateur bien que la recherche sur le langage naturel ait constitué pour moi-même, dès 1984, un centre d'intérêt important (voir I.2/[2] et article : RIALLE V., Du Langage Naturel Spécialisé au Système Expert. *Proceedings COGNITIVA 1985*, Paris, 3-7 juin 1985, 371-375).

Le séjour de délégation post-doctorale au Centre ATO-CI (présenté en I.1/E et repris en III.3) a été l'occasion d'un réinvestissement important de ce thème de recherche, selon une approche toutefois entièrement différente de celle de 1984. Le séjour de recherche à Montréal a notamment donné lieu au texte présenté ici, texte qui constitue en fait un rapport de recherche. En raison de la longueur de ce texte, seuls l'avant propos, l'introduction générale et la table des matières seront présentés afin de donner une vue d'ensemble des recherches abordées.

Présentation de l'article :

RIALLE V., Notes sur le connexionnisme, l'analyse automatique de textes et le parallélisme. *Cahiers du Centre de Recherche en Cognition et Information ATO-CI*. Université du Québec à Montréal, 1(1), 1992.

NOTES SUR LE CONNEXIONNISME, L'ANALYSE AUTOMATIQUE DE TEXTES ET LE PARALLÉLISME

Avant propos :

Le présent travail a été réalisé dans le cadre d'un stage post-doctoral (sans constituer pour autant un rapport de stage), lui même inscrit dans le cadre d'une coopération scientifique entre le Centre de Recherche en Cognition et Information ATO-CI (Pr. Jean-Guy Meunier), Université du Québec à Montréal, et le Laboratoire TIMB (Traitement de l'Information et Modélisation Biomédicale) (Pr. Jacques Demongeot), l'Université Joseph Fourier, Faculté de Médecine de Grenoble.

Il a pour but d'apporter une contribution à l'analyse automatique de textes par des modèles connexionnistes, dans le cadre des recherches en analyse textuelle et cognitive du Centre ATO-CI. Il tente donc une sorte d'introduction à la problématique du traitement parallèle et distribué du texte (TPDT), tel qu'il a été perçu durant ce séjour.

Il est également le fruit de nombreuses et fructueuses interactions avec les chercheurs du Centre ATO-CI investis dans ce tournant, aussi bien technologique que méthodologique, constitué par l'arrivée des ordinateurs parallèles et le rapprochement entre sciences cognitives et traitement de bases de données textuelles. Si, parmi les idées exprimées ici, certaines devaient inspirer la réalisation de logiciels d'analyse textuelle connexionniste au Centre ATO-CI, le but de ce travail serait atteint.

Introduction générale

L'analyse automatique du texte autant que le connexionnisme constituent, chacun à leur manière, des domaines très actifs de recherche. Leur rapprochement dans le but de construire des systèmes évolués d'analyse linguistique et textuelle est donc chargé de promesses. Ce rapprochement ne va pas, néanmoins, sans poser un certain nombre de problèmes, tant sur le plan théorique et conceptuel, que sur le plan purement technique et pratique.

Une précision s'impose en préambule au présent travail quant à l'expression *analyse de texte* qui sera souvent employée dans l'exposé (e.g. analyse automatique de texte...). L'expression *analyse de texte* est polyvalente, sinon ambiguë (MEUNIER, 1990), du fait de son utilisation dans diverses disciplines, dans des contextes et pour des objectifs différents. Le sens dans lequel elle est communément utilisée par un informaticien mérite donc d'être précisé.

Les types d'analyse auxquelles on peut prétendre au moyen de l'ordinateur débutent inexorablement au niveau des *chaînes de caractères* pour s'élever peu à peu et au pris d'efforts de modélisation, puis de programmation, considérables vers des structures de données et des raisonnements plus ou moins complexes sur ces structures. Elle n'atteint pas, loin s'en faut, une analyse argumentative du discours contenu dans un texte, telle que cette expression pourrait porter à le penser chez un lecteur d'obédience non informaticienne. Il s'agit simplement ici de traitements automatiques de l'information textuelle et de modélisation de certains processus cognitifs à l'œuvre chez un lecteur ou nécessaires pour un objectif particulier (indexation, résumé, recherche de thèmes, "fouille", levée d'ambiguïté, etc.). Malgré quelques velléités de l'intelligence artificielle, la structure discursive du texte (discours politique, roman poème...) reste largement hors de portée du traitement automatique.

Le travail présenté ici essayera notamment d'apporter quelques éléments de réponse aux nombreuses questions posées par un projet de modélisation parallèle et distribuée (MPD), dans une perspective connexionniste, du traitement textuel. Par exemple : en quoi consiste un système connexionniste? quels en sont les modalités, les différents types, les caractéristiques, les limites, etc.? quelles fonctions d'analyse textuelle peuvent leur être assignées? quels sont les travaux significatifs réalisés dans le domaine du traitement connexionniste du texte? quelle peut être la place et l'apport du parallélisme dans ce domaine? quelles sont les voies d'accès au niveau *sémantique* de l'analyse automatique du texte? etc.

Ce rapport est structuré en trois parties et une annexe :

- L'objet de la première partie est une présentation succincte d'un certain nombre d'approches connexionnistes de l'analyse du langage pratiquées jusqu'à aujourd'hui. Les travaux rapportés se réclament, plus ou moins explicitement, de courants très variés : linguistique, psycholinguistique, sémantique, philosophie du langage, psychologie cognitive, etc.

Etant donné leur nombre, un choix a dû être fait, qui s'est porté sur ceux qui traitent des aspects sémantiques, et plus particulièrement de la désambiguïsation sémantique.

- La deuxième partie présente une étude de modélisation conduite à travers la construction progressive de *modèles*. Cette partie possède un avant-propos : la présentation succincte de quelques notions de base du connexionnisme, ainsi qu'une courte présentation de la problématique des intrants textuels de réseaux connexionnistes.

Les essais de modèles qui sont ensuite exposés s'inscrivent dans une progression du plus simple au plus complexe. Le découpage en modèles simples s'accompagne d'une visée pédagogique dont le

but est d'"apprivoiser" au mieux les possibilités de modélisation proposées par le courant des réseaux de neurones formels (RNF).

L'orientation des modèles proposés reprend les idées générales esquissées dans le projet ALEX du Centre ATO-CI, à savoir : "Dans une perspective connexionniste, un document textuel est défini comme un ensemble d'unités d'information (UNIF) couvrant plusieurs centaines de pages. Ces unités peuvent être de plusieurs types (mots, formes morphologiques, lexies complexes, formes lemmatisées, etc.). (Ces UNIFS sont identifiées par des analyseurs produits par le Centre d'ATO, lesquels seront utilisés dans le présent projet)."

A l'horizon plus ou moins lointain des essais de modélisation présentés ici se trouve l'objectif de réalisation d'un système de *sémiotique cognitive* (MEUNIER, 1988, RASTIER, 1991).

- La troisième partie est un ensemble de notes au sujet du parallélisme de traitement. Celui-ci intervient de deux façons : d'une manière implicite (parfois explicite) dans la majorité des modèles connexionnistes; d'une manière explicite dans un certain nombre de tâches qui ne relèvent pas du connexionnisme (recherches parallélisées de chaînes de caractères, langages de programmation parallèle, tris parallèles, etc.).

- l'annexe présente quelques développements sur la notion de réseau de neurones formels et ne constitue en aucun cas un document de référence sur ce sujet. Elle vise seulement à fixer quelques idées, répondre éventuellement à quelques questions et rappeler quelques notions de base du connexionnisme.

Table des matières :

AVANT PROPOS	4
INTRODUCTION GÉNÉRALE	4
PREMIÈRE PARTIE	6
Quelques travaux et modèles en analyse du langage et connexionnisme	6
Introduction	6
Modèle général de COTRELL et SMALL	7
Modèle général de WALTZ et POLLACK	8
Modèle de désambiguïsation sémantique de VICTORRI	11
Modèle de désambiguïsation sémantique de GALLANT	12
Représentations structurées et modèles connexionnistes de ELMAN	14
Modèles stochastiques	15
DEUXIÈME PARTIE	19
Essais de modélisation connexionniste de l'analyse automatique de textes	19
II.1. Quelques mots d'introduction au paradigme connexionniste	19
Pourquoi le terme engrammation ?	21
II.2. Quelques mots d'introduction à l'analyse connexionniste du texte	21
Le concept d'UNIF	23
II.3. Problématique générale de construction des intrants	23
II.4. Apprentissage de co-présences et réseaux lexical du texte	24
II.4.1. Elaboration du modèle	24
II.4.2. Mode de lecture du texte	24
II.4.3. Le réseau	25

II.4.4. Algorithme de construction du RLT	26
II.4.5. Vers une désambiguïsation sémantique sans artéfacts	27
II.5. Notes pour un modèle de désambiguïsation grammaticale	29
II.6. Modèles de type mémoire associative pour la catégorisation et la classification	31
II.7. Désambiguïsation catégorielle et processus stochastique	32
TROISIÈME PARTIE	34
Notes a propos du parallélisme	34
Le modèle de AKINGBEHIN et CONRAD	34
Limitations des modèles séquentiels et potentialités du parallélisme en modélisation cognitive	35
CONCLUSIONS ET PERSPECTIVES	38
BIBLIOGRAPHIE	41
ANNEXE	48
A. Modèle discret du neurone formel	48
B. Réseaux de neurones formels	49
C. Réseaux de neurones à attracteurs	50
D. Espace des états et distance de Hamming	51
E. Réseau de Hopfield	52
F. Réseaux neuronaux et états énergétiques	52
G. Minima énergétiques et mémorisation	54
H. Minima énergétiques et recuit-simulé	54
RÉSUMÉ DES ABRÉVIATIONS :	55

I.3.6 Sixième thème : Modélisation cognitive en psychiatrie : des modèles symboliques aux parallèles et distribués.

Présentation de l'article :

RIALLE V. & STIP E., Modélisation cognitive en psychiatrie : des modèles symboliques aux parallèles et distribués. *Journal of Psychiatry and Neurosciences*. 19(3), 1994, sous presse.

Modélisation cognitive en psychiatrie : des modèles symboliques aux parallèles et distribués

Vincent Rialle*, Ph D, Emmanuel Stip**, MD, Msc.

*Université J. Fourier, *Laboratoire TIMB*, Faculté de Médecine de Grenoble
Université du Québec à Montréal, *Centre de Recherche en Cognition et Information ATO-CI*.

**Centre de recherche Fernand Séguin, Hôpital L.H. Lafontaine, Département de psychiatrie, Université de Montréal

Travail présenté au : Second Annual Meeting, Canadian Society for Brain, Behaviour, and Cognitive Science. Université Laval, Québec, 13 - 14 Juin 1992

Address reprint: V. Rialle, Université Joseph Fourier, Faculté de Médecine de Grenoble, France 38706 La Tronche Cedex. Tél. : (33) 76 63 71 11. Fax : (33) 76 51 86 67. E.mail : rialle@timb.imag.fr

Résumé. Les techniques computationnelles de la représentation de connaissances et de la modélisation, tant de l'expertise médicale que de phénomènes neuro et psychocognitifs, captent l'intérêt des chercheurs du domaine de la santé mentale depuis les origines de l'informatique. Ce travail a pour but de présenter à la fois les diverses techniques de modélisations et quelques réflexions et applications marquantes, essentiellement dans le champ de la psychiatrie, mais aussi dans celui de la psychanalyse. L'intelligence artificielle et le courant connexionniste y ont tour à tour exercé une certaine fascination. Un accent plus particulier sera porté sur les modèles parallèles et distribués qui ne cessent actuellement de se développer et d'exciter la curiosité des chercheurs en santé mentale. Le secret de cet engouement réside certainement dans les étonnantes capacités des réseaux neuronaux formels à épouser des comportements neuro-biologiquement et psychologiquement vraisemblables. Leur apparente supériorité par rapport aux techniques de modélisation symbolique, proposées par l'intelligence artificielle classique, est due à plusieurs facteurs dont le principal est peut-être la possibilité qu'ils offrent d'expérimenter aisément des modifications de paramètres, des perturbations, des "lésions" structurelles et fonctionnelles et d'en mesurer les conséquences dans le comportement général du modèle. Les perturbations induites par ces manipulations paramétriques sont ensuite rapprochées des dysfonctionnements cognitifs observés sur des malades psychiatriques.

Mots clés: Psychiatrie et modélisation, modélisation cognitive, sciences cognitives, intelligence artificielle, réseaux de neurones, connexionnisme.

Summary. Since the advent of computer technology, mental health researchers have been captivated by the computational approach to the representation of knowledge and to modelling, not only in medical expertise but also in neuro and psychocognitive phenomena. The object of this article is to

present the various modelling techniques, as well as to expand on the attributes and remarkable applications, mainly in the field of psychiatry, but also in psychoanalysis. Artificial intelligence and Connectionism have each in turn fascinated researchers. The emphasis will be on Parallel Distributed Processing models, which continue to develop and arouse the curiosity of mental health researchers. The driving force for development in this field most certainly stems from the remarkable capacity of Formal Neural Networks to approximate neuro-biological and psychological behaviour. Their obvious superiority, compared to the symbolic modelling techniques inherent to classical Artificial Intelligence, arises from several factors, principally the potential for experimentation with parameter modification, perturbations, structural and functional lesions, and to assess the consequential model behaviour. The perturbations created through parameter modification can be matched with cognitive disorders observed in psychiatric patients.

Key words: Neural networks, Parallel Distributed Processing, Cognitive modelling, A.I., models in psychiatry, connectionism.

Introduction

La relation entre les deux univers, mental et neuronal, demeure de type biunivoque, se complexifie. Cette complexité qui augmente encore plus en psychiatrie nécessite la mise en place de modèles. On assiste, depuis que l'ordinateur est devenu d'un accès aisé et courant, à une croissance considérable du nombre de modèles et de leurs simulations. Ceux-ci sont rapidement diffusés dans la littérature scientifique et alimentent l'imaginaire des chercheurs dans leurs projets de conception de nouveaux modèles. Ces modèles sont ensuite investis dans une impressionnante diversité d'applications. Étant donné les interrelations perpétuelles qu'ils entretiennent constamment, les ranger selon une taxonomie précise est une entreprise quasiment vouée à l'échec. Beaucoup de modèles empruntent en effets à plusieurs domaines théoriques : calcul probabiliste, théorie combinatoire, logique formelle... La manière de présenter ces modèles varie donc selon les hypothèses et les champs d'où sont issues ces hypothèses: neurochimiques, psychophysiologiques, cognitivistes.

Ce travail a pour but de présenter les principales techniques de modélisations ainsi que quelques réflexions et applications marquantes, essentiellement dans le champ de la psychiatrie, mais aussi dans celui de la psychanalyse. L'intelligence artificielle, le courant connexionniste et la modélisation génétique y ont tour à tour exercé une certaine fascination. Un accent plus particulier sera néanmoins porté sur les modèles parallèles et distribués qui ne cessent actuellement de se développer et d'exciter la curiosité des chercheurs en santé mentale. Nous donc allons voir en quoi la modélisation cognitive (Imbert, 1987, Varela, 1989, Demongeot *et al.*, 1989, Walker, 1990, Kaplan *et al.* 1990, Rialle, 1992, Tiberghien, 1992) est une source de progrès pour la modélisation en psychiatrie.

Pour la clarté de la présentation, nous ordonnerons ces modèles autour de trois axes principaux : l'intelligence artificielle symbolique (IAS), le connexionnisme (C) et la modélisation génétique. L'IAS et le connexionnisme se partagent la majorité des modélisations. D. Memmi (1989) en a rappelé les origines concomitantes. Ces deux domaines de recherche sont nés en effet à la même époque que leur inséparable outil d'expression qu'est l'ordinateur. Elles furent engendrées dans les années 40 par les pères fondateurs des sciences et techniques informatiques, de la cybernétique et de la simulation cognitive. Alors qu'Alan Turing, en 1936, énonce les principes d'une "machine universelle" et que John Von Neumann décrit, en 1945, l'architecture de l'ordinateur séquentiel (*machine de Von Neumann*), des neurobiologistes, W. McCulloch et W. Pitts (1943), proposent le premier modèle mathématique de fonctionnement du neurone. Ils montrent notamment la capacité d'un réseau de neurones formels à reproduire certaines fonctions logiques. Quelques années plus tard, Hebb (1949) introduit la notion d'apprentissage par ajustement des poids synaptiques. C'est à

partir de 1958 que le premier réseau de neurones formels opérationnel — le Perceptron — est décrit par Rosenblatt (1961).

La modélisation stochastique plonge ses racines dans l'histoire des sciences : on peut par exemple la faire remonter au Révérend Thomas Bayes (1763) qui, voici deux siècles, proposa la célèbre formule qui porte aujourd'hui son nom, et qui connaît un usage intensif depuis une trentaine d'année dans les modèles de décision probabiliste sur ordinateur. A l'autre extrémité de l'histoire, c'est-à-dire depuis moins de vingt ans, le paradigme du sélectionnisme génétique a fait son apparition dans la résolution de problèmes et, plus récemment encore, dans le *mentalisme darwinien* (Dehaene et Changeux, 1989).

I. Les modèles computationnels et leurs paradigmes

L'intelligence artificielle

L'intelligence artificielle Symbolique a vu le jour à la suite des travaux de Newell (1980), McCarthy (1977) et C. Shannon (1964) principalement, et fut dès ses origines l'objet d'un fort engouement. Elle promettait cependant beaucoup plus qu'elle ne tenait (Dreyfus, 1984). L'attribut *symbolique* dans l'expression *intelligence artificielle symbolique* a pour but de préciser de quelle intelligence artificielle il s'agit ici. L'expression plus simple d'*intelligence artificielle* est de plus en plus souvent employée pour désigner l'ensemble des techniques utilisées en informatique pour programmer des processus complexes. Historiquement néanmoins, l'intelligence artificielle se définit comme la science ou la technique du raisonnement symbolique. Elle est fondé sur l'utilisation de symboles plus ou moins élaborés et propres à l'intelligence naturelle, en opposition avec les autres techniques informatiques, fondées par nature sur la manipulation de chiffres et le calcul numérique. Les spécificités de l'IAS (Kodratoff, 1989), peuvent être exprimées en quelques points essentiels:

- l'IAS s'applique à développer la *transparence des processus* qu'elle modélise: la notion de *boîte noire* réalisant une tâche de manière occulte à l'utilisateur lui est, en tout état de cause, totalement étrangère.
- elle privilégie la *forme déclarative* à toute autre forme d'expression des connaissances. Ce principe permet à la fois de *mettre à jour* les connaissances (ajouts, modification, suppression) et de les *représenter* aussi bien graphiquement qu'en langage clair.
- un système d'IAS suppose qu'un nombre considérable de connaissances doivent être déclarées; Aux connaissances essentielles, qui peuvent être très nombreuses, il faut en effet ajouter des connaissances secondaires et souvent évidentes pour l'utilisateur (connaissances pragmatiques ou de type *outil de travail*).
- la capacité d'*explication en langage clair* (langage de l'utilisateur) du raisonnement est une caractéristique fondamentale de l'IAS.
- les thèmes principaux sur lesquels s'étendent les recherches sont nombreux. Parmi eux, on peut citer : la représentation des connaissances et la modélisation des raisonnements, l'apprentissage (induction automatique, généralisation, spécialisation, processus d'analogie, etc.), la collaboration entre agents intelligents (IA distribuée), etc.

Fer de lance de l'intelligence artificielle, le modèle du système expert s'est répandu dans les années 70 et semble avoir atteint son apogée durant la décennie 80. Il se compose :

- (a) d'une *base de connaissances* regroupant des connaissances plus ou moins granulaires ou élémentaires, représentées sous forme *règles d'inférences*, d'*objets*, de *scripts*, etc.
- (b) d'une *base de faits* destinée à recevoir les descriptions des problèmes soumis au système.
- (c) d'un *moteur* chargé de mettre en œuvre ces connaissances lors de la soumission du problème au système. Le résultat en est un *raisonnement symbolique*, plus ou moins *calculatoire* s'il est dirigé par des valeurs d'incertitude sur les faits ou les connaissances élémentaires.

(d) d'une interface-utilisateur la plus conviviale possible, de façon à rendre parfaitement aisée la description des problèmes par l'utilisateur et l'explication des raisonnements par le système. S'il a montré maintes fois ses limites au cours des deux précédentes décennies, ce modèle a néanmoins ouvert la voie à une prodigieuse quantité de réalisations. Il a maintenant tendance à basculer du domaine de la recherche à celui des techniques maîtrisées et couramment appliquées. Il suppose néanmoins un art et une science de l'expression et de la représentation des connaissances par des experts et des spécialistes de la cognition (*cogniticiens*).

Le paradigme connexionniste (C)

Le paradigme connexionniste (Hopfield, 1982, Amit, 1982, McClelland *et al.*, 1986, Rumelhart *et al.*, 1986, Hinton & Sejnowski, 1986, Fodor & Pylyshyn, 1988, Kohonen, 1988) est fondé sur :

- Une modélisation mathématique du neurone (*neurone formel*) sous la forme d'une unité élémentaire de traitement. Cette unité possède plusieurs *entrées* d'information, un *état* qui la caractérise à tout instant, et une fonction interne, appelée *fonction de transfert*, *fonction d'activation* ou *fonction neurone*, qui calcule l'état dans lequel doit être mis le neurone en fonction des entrées. Les états prennent leur valeur la plupart du temps dans un ensemble binaire (état actif, état passif) mais peuvent également être calculés sur une échelle continue.

- la construction de *réseaux de neurones formels* (RNF) composés d'un ensemble d'unités de traitement, l'étude des propriétés émergentes de ces réseaux ainsi que l'étude des divers types de réseaux et de leur adéquation à divers types de comportements ou de problèmes (classification, catégorisation, reconnaissance de forme, apprentissage d'aptitudes cognitives, etc.). La plupart de ces réseaux sont *auto-adaptatifs*, modifiables par *entraînement*, et sont naturellement *massivement parallèles*. A un niveau plus imagé que réellement fondé, ces réseaux peuvent être comparés à des réseaux de neurones biologiques: les entrées simulent les synapses, l'état du neurone formel simule l'axone et les dendrites, et l'unité de traitement simule le soma du neurone naturel. Ce qui est compris sous le terme de réseaux de neurones formels (ou plus simplement *réseaux de neurones*) est avant tout un ensemble de *modèles mathématiques*. Le transfert de ces modèles dans des circuits intégrés ne constitue actuellement qu'une faible part du volume des publications sur les RNF, les modèles transférés étant presque toujours en retard sur les modèles théoriques, en perpétuelle et rapide évolution.

On distingue deux grandes familles de RNF déterminées par le type d'apprentissage : supervisé ou non-supervisé (chaque famille ayant d'importantes subdivisions).

L'apprentissage supervisé suppose une phase d'entraînement du réseau dans laquelle un *tuteur* indique en permanence à celui-ci ses erreurs par rapport au comportement correct. Une procédure dite de *rétropropagation du gradient de l'erreur* permet de corriger ces erreurs au fur et à mesure que sont présentés les exemples à apprendre.

Dans l'apprentissage non supervisé, ce tuteur n'existe pas : le réseau est livré à lui-même pour *découvrir* son environnement et en mémoriser les caractéristiques.

On distingue d'autre part plusieurs *générations* de RNF. En ce qui concerne les réseaux à apprentissage supervisé, la première génération a produit des réseaux entièrement connectés; la seconde correspond aux efforts d'optimisation des réseaux par la spécialisation de certains *neurones cachés* et par une diminution parfois importante du nombre de connexions. Pour les réseaux à apprentissage non supervisé, la recherche de règles d'apprentissage efficaces, biologiquement et psychologiquement vraisemblables, est au centre des préoccupations. On constate actuellement une importante recherche d'agencements judicieux de plusieurs types de réseaux au sein d'une même architecture (*architectures connexionnistes hybrides*). Une question fondamentale à laquelle devra tenter de répondre la recherche actuelle en réseaux de neurones formels est : comment définir le *meilleur* réseau pour une application donnée? Il n'existe actuellement aucune réponse théorique, mais seulement quelques principes de plus en plus précis et quelques équipes de chercheurs qui travaillent sur ce sujet. Il est toutefois possible de classer les modèles connexionnistes en fonction des types de problèmes qu'ils peuvent résoudre. On trouve deux types principaux de problèmes : la

mémorisation de formes dans un sens très large (reconnaissance, discrimination, etc.), et la résolution de problèmes.

- La mémorisation de formes inclut la classification et la catégorisation automatiques, et la recherche de caractéristiques dominantes (Kohonen, 1988, Proulx et Bégin, 1990).

- La résolution de problèmes consiste en la recherche d'états du réseau qui correspondent à des solutions plus ou moins approchées du problème (Hopfield et Tank, 1985, Hinton & Sejnowski, 1986). Il s'agit en l'occurrence essentiellement de problèmes d'optimisation combinatoire. On conçoit alors le réseau de telle sorte qu'il *engramme*, dans ses connexions synaptiques, les *contraintes* qui définissent le problème. L'approche de la solution est conduite au moyen d'une fonction de l'ensemble des neurones et des connexions synaptiques appelée *énergie* du réseau. Le processus inférentiel conduisant à la solution plus ou moins approchée du problème est obtenu par un mécanisme de *relaxation* du réseau vers un état énergétique minimal. La solution ainsi obtenue correspond à un état dans lequel un maximum de contraintes sont satisfaites (*i.e.* un minimum de tensions persiste).

La modélisation génétique

Principe général

Un autre courant de modélisation se développe actuellement, fondé sur le paradigme du sélectionnisme génétique, avec le modèle computationnel de l'*algorithmique génétique* (Goldberg, 1989, Davis, 1991). Le concept d'algorithme génétique a été présenté pour la première fois par J. H. Holland (1975). Son principe en est le suivant : on génère de manière aléatoire une *population* de solutions potentielles à un problème donné; on sélectionne ensuite, au moyen d'une mesure appropriée, les éléments de la population qui regroupent le maximum de caractéristiques (ou contraintes) de la solution recherchée. Des *opérateurs génétiques* sont ensuite appliqués à cette population d'*élite* de manière à obtenir une nouvelle population possédant, dans son ensemble, plus de caractéristiques de la solution que la génération précédente. Ce processus est réitéré sur plusieurs générations jusqu'à l'obtention d'une génération de solutions optimales, au sein de laquelle il suffit ensuite de retenir la meilleure.

Les solutions potentielles, éléments de la population, sont des codages de caractéristiques du problème, à l'image des gènes et des chromosomes qui codent de véritables listes de caractères d'un individu. Le codage prend habituellement la forme d'une chaîne binaire très structurée, de longueur fixe ou variable selon le type de problème. L'hypothèse génétique est utilisée pour modéliser la recherche d'une adaptation, c'est-à-dire d'une solution adaptée à un ensemble de conditions (survie d'une espèce, optimisation en recherche opérationnelle...). Ici encore l'ordinateur est en quelque sorte le promoteur de ce modèle. Il en rend possible des simulations et en atteste la validité à travers des exemples concrets de résolutions de problèmes (Davis, 1991).

Par une amusante ironie de la recherche scientifique, l'algorithme génétique répond parfaitement à un type particulier d'objectif, historiquement assigné au modèle du système expert qui, lui, est incapable d'y répondre: il s'agit de l'objectif de recherche d'une solution à un problème donné sachant qu'il n'existe aucun algorithme déterministe pour y parvenir. L'algorithme génétique utilise un mécanisme aléatoire convergent de recombinaison de fragments de solution. Dans ce cas, le détail des opérations de recombinaison est impossible, non seulement à énoncer mais aussi à prévoir par un humain en raison de l'aspect intrinsèquement probabiliste de l'algorithme. Le système expert décrit au contraire, dans sa base de connaissances, un cheminement logique de recherche de solutions. Ce cheminement est parfaitement prévisible et dirigé par le moteur d'inférence du système. Néanmoins, la comparaison entre algorithme génétique et système expert s'arrête là, étant donné que les deux approches comblent des besoins qui ne sont généralement pas comparables.

Le mentalisme Darwinien

La notion de mentalisme darwinien a été introduite par J.P. Changeux *et al.* pour rendre compte de la capacité d'un modèle à *créer* des solutions potentielles pour son adaptation et concourir à la résolution de ses divers problèmes. Ils introduisent à cet effet un élément essentiel au modèle : le

générateur de diversité. Ce générateur est chargé de créer des pré-représentations de l'environnement et des règles de conduites dans ces pré-représentations. Ces règles sont construites à partir d'informations élémentaires. Le système doit ensuite évaluer la pertinence de chacune de ses pré-représentations afin d'en sélectionner la plus adéquate pour assurer sa pérennité ou atteindre ses objectifs. Dehaene et Changeux (1989) ont proposé un modèle de ce type, correspondant à des fonctions du cortex pré-frontal. Alors que l'algorithme génétique présenté précédemment est conçu dans un champs totalement extérieur au connexionnisme, le modèle de Dehaene et Changeux correspond à la simulation d'un réseau neuronal, et s'inscrit entièrement dans une perspective connexionniste.

Différences et complémentarités entre approches symbolique et connexionniste

Deux approches computationnelles différentes

L'intelligence artificielle symbolique et le connexionnisme correspondent tous deux à des paradigmes fondamentalement différents de représentation des processus cognitifs:

a) l'intelligence artificielle postule une représentation des connaissances et des raisonnements au moyen de symboles (Newell, 1980), c'est-à-dire d'éléments d'un langage. Le mot symbole est pris ici dans son sens le plus élémentaire: celui de représentation d'un concept, d'une idée ou d'un objet (par exemple : les symboles mathématiques, l'alphabet, les mots d'un langage, etc.). Elle privilégie une représentation propositionnelle de la connaissance et, en cela, pratique un *a priori* fondamental : la connaissance peut être "discrétisée" (découpée en éléments), et représentée sous la forme d'unités symboliques (représentation + sens) qu'il s'agit ensuite d'interpréter de manière automatique. Ce postulat est lourd de conséquences et renvoie notamment aux oppositions entre écoles *mentaliste* et *béavioriste* en psychologie expérimentale (Tiberghien, 1986).

b) le paradigme connexionniste est fondé sur la métaphore neuronale: systèmes et programmation "massivement" parallèles, au moyen de réseaux d'unités de traitement (neurones formels) mimant, jusqu'à un certain point et de manière très limitée, le fonctionnement de neurones biologiques. Le nombre de ces unités est très variable et dépend du type de réseau: il peut aller de quelques unités à plusieurs milliers. Chaque unité est connectée avec l'ensemble ou une partie de l'ensemble des unités. Elle réagit individuellement — selon une fonction qui lui est propre — aux messages qui lui parviennent des autres unités via ses *liaisons synaptiques*. Cette réaction donne lieu à un message qui est à son tour propagé vers d'autres unités ou vers la sortie du réseau.

Macro-structures et micro-structures cognitives

Un préalable à toute comparaison entre IAS et C consiste à situer ces deux approches dans l'ensemble des niveaux de modélisation cognitive (Amy *et al.*, 1990). Le sujet humain possède des chaînes de traitement des informations qui vont des récepteurs sensoriels jusqu'au système nerveux central. Elles sont capables d'opérations les plus élémentaires et automatiques, telles que la reconnaissance de formes ou de sons, jusqu'aux processus les plus complexes (raisonnements, langage, états intentionnels, affects, intuitions, croyances...). Ainsi, dans l'échelle de complexité des traitements opérés par le cerveau, on associe généralement l'IAS au niveaux "élevés" de complexité, alors que le connexionnisme est plus généralement associé aux opérations de perception.

Historiquement et jusqu'à une période récente, le champ d'application du connexionnisme est plutôt la modélisation de fonctions perceptives et de reconnaissance de formes, alors que l'IAS, à travers les systèmes experts principalement, s'est attachée à représenter des connaissances ou des situations complexes, à modéliser et simuler des raisonnements sur ces représentations, puis à expliquer dans le langage de l'utilisateur tous ses développements. Ce n'est que récemment que divers auteurs, comme abordé plus loin, ont vu dans le connexionnisme et quelques techniques algorithmiques associées, de remarquables analogies avec notre fonctionnement neuropsychologique.

Apprentissage et flexibilité

Apprentissage, entraînement, flexibilité, adaptativité sont des attributs plus spécifiques du connexionnisme que de l'IAS. On retrouve constamment, dans les diverses tendances du

connexionnisme ainsi que dans quelques études comparatives entre IAS et C (Memmi, 1989, Amy *et al*, 1990), la constatation que la première se heurte à l'aspect figé des concepts enfermés dans ses symboles : la connaissance modélisée dans une base de connaissances revêt un caractère rigide et peu propice à des modifications significatives du modèle. Lorsqu'on ajoute ou supprime une règle de production par exemple, les effets sur l'ensemble du processus de raisonnement ou sur la cohérence générale de la base de connaissances peuvent être considérables ou difficiles à maîtriser. D'autre part, bien que l'apprentissage symbolique automatique et l'induction de connaissances à partir de données représentent un aspect important de l'IAS (Kodratoff, 1991), cet apprentissage ne possède pas un caractère aussi fondamental que celui pratiqué dans les RNF. Dans la plupart de ceux-ci en effet, l'apprentissage est la condition *sine qua non* de leur fonctionnement. Cependant, là encore la comparaison doit être largement modulée par la différence des niveaux structurels propres aux deux approches : un apprentissage de raisonnements complexes n'entretient que peu de rapports avec un apprentissage de reconnaissance de formes.

Adressage versus association

L'une des différences majeures qui distinguent IAS et C concerne les moyens de mémorisation et d'accès aux informations. La notion de *mémoire* en IAS est fondamentalement différente de celle que propose le connexionnisme. Dans le premier cas, la mémoire est conçue comme une structure "adressable" : une *information* est mémorisée dans un *lieu* particulier de cette structure. Le seul moyen d'accès à ce lieu est son identificateur, qui n'est autre qu'une *adresse*. La non-connaissance de cette adresse interdit l'accès à l'information qui y est contenue. Dans le cas du connexionnisme, la notion de *lieu* où trouver une information s'estompe au profit de celle d'*accès associatif* à l'information mémorisée: on accède à cette information en présentant au réseau un fragment de celle-ci (*e.g.* récupération d'une figure complète à partir d'une figure incomplète ou *bruitée*). La notion d'adresse disparaît en tant que localisation au profit de celle d'*information délocalisée*: l'information contenue dans le réseau est répartie dans les poids synaptiques, les fonctions neurones, le type de connectivité ainsi que tout autre paramètre du réseau. Cette conception *distribuée* de la mémoire est, semble-t-il, plus proche d'un modèle neuro-cognitif ou psycho-cognitif que ne l'est la conception localiste. La mémoire humaine semble en effet plutôt fonctionner par des processus associatifs très complexes, sur lesquels il reste encore beaucoup à apprendre, et qui ont vraisemblablement peu de rapports avec des adressages de zones physiques parfaitement délimitées, comme dans le cas d'un ordinateur. Ce point de vue réplique, une fois encore, le débat entre le courant localisationniste et celui d'associationniste ou globaliste qui a eu lieu en neuropsychologie, depuis l'époque des premiers constats sur les relations neuro-anatomiques et pathologiques. En psychiatrie, l'évolution des idées pour ceux qui remarquent les arguments biologiques des maladies mentales, est également sensible aux débats localisation-distribution. "Can schizophrenia be localized in the brain?" en est un exemple (Andreasen, 1986). Le PETscan, la Résonance magnétique nucléaire objectivent des modifications morphologiques et fonctionnelles dans la schizophrénie (Volkow & Wolf, 1991) mais nous savons aussi que les associations des réseaux de neurotransmetteurs dopaminergiques, sérotoninergiques sont nécessaires pour mieux comprendre les effets des antipsychotiques. Il en est de même des associations de réseaux dans la dépression et l'autisme. Les troubles neuropsychologiques dans la schizophrénie, même s'il existe une prédominance pré-frontale, ne sont interprétables que dans une distribution des traitements de l'information.

Capacités de dialogue et d'explication

Étant très largement dévolue à la représentation et au traitement d'informations symboliques, l'IAS possède un avantage incontestable sur le connexionnisme en ce qui concerne l'*explication* de ses processus en langage clair, ainsi que la fonction de *dialogue* avec un utilisateur de logiciel. Des travaux considérables ont été conduits en IAS en ce qui concerne l'interaction entre utilisateur et ordinateur. Ces travaux empruntent à un nombre non moins considérable de travaux théoriques (raisonnement symbolique, linguistique informatique, ergonomie cognitive, etc.). Le connexionnisme ne s'est pas fixé de tels buts, et ne peut être comparé, dans ce domaine, à l'IAS.

Bien que des recherches se développent dans ce sens, la connaissance acquise par entraînement dans un RNF est encore difficilement représentable d'une manière symbolique (sauf pour des réseaux parfaitement localistes, dans lesquels on a fixé *a priori* la signification symbolique de certaines cellules).

Modèles hybrides symboliques et connexionnistes

L'un des principaux obstacles rencontrés par les modèles exclusivement fondés sur l'un des paradigmes de modélisation en cours à l'heure actuelle (traitement symbolique, connexionnisme, algorithmes génétiques, etc.) est leur manque de *robustesse*. Chacun de ces paradigmes s'avère souvent excellent dans un champ délimité de performances, mais ne constitue en quelque sorte qu'une simple pièce d'un puzzle particulier, celui des qualités nécessaires à un modèle pour mériter le qualificatif de cognitif. Plusieurs auteurs se sont mesurés à la question de la fragilité des modèles et des moyens d'y échapper (*e.g.* Holland, 1986, Sun, 1991).

L'approche des *systèmes hybrides symboliques et connexionnistes* (Hendler, 1989, Memmi, 1989, Amy, 1989, Feldman, 1989, Touretzky, 1989, Amy *et al.*, 1990, Rialle *et al.*, 1991, Sun, 1991, Giacometti *et al.*, 1992) constitue l'une des réponses à ce problème du manque de robustesse des modèles. L'approche hybride recherche l'association des qualités des réseaux neuronaux formels et celles de l'IAS dans le but de construire des modèles cognitifs (principalement des systèmes d'aide au raisonnement et d'apprentissage de connaissances ou de comportements) plus performants que ne le ferait chacune des composantes de l'association prise individuellement. Elle consiste à faire cohabiter, au sein d'un même environnement logiciel, la représentation symbolique des phénomènes traités, avec une représentation parallèle et distribuée. Diverses sortes d'*hybridation* sont étudiées, depuis des architectures fondées sur une simple coopération entre modules connexionniste et symbolique, jusqu'à une intégration plus ou moins complète des deux approches. Ces divers modes d'intégration de l'IAS et des réseaux neuronaux ont été classifiés par plusieurs auteurs (*e.g.* Amy, 1989, Sun et Bookman, 1992). Sun et Bookman (1992) proposent notamment quatre catégories : (a) le simple couplage de deux modules — l'un symbolique, l'autre connexionniste — entièrement indépendants et s'apportant mutuellement des informations; (b) les systèmes essentiellement symboliques de traitement de l'informations mais possédant des composants neuronaux (*e.g.* le système BoltzCONS proposé par Touretzky, 1989); (c) les systèmes entièrement connexionnistes proposant une émergence des symboles à partir des interactions neurales (*e.g.* le système RAAM de Pollack, 1990); (d) les réseaux connexionnistes localistes, dans lesquels certaines unités de traitement sont volontairement associées à des symboles par le concepteur du système (*e.g.* le système proposé par Pinkas, 1992).

Dans tous les cas, ce type de logiciel est d'une mise au point délicate et relève encore largement du domaine de la recherche. Il semble néanmoins que les systèmes hybrides constituent l'une des voies les plus prometteuses pour la modélisation de phénomènes complexes, mettant en jeu à la fois des processus d'apprentissage et d'adaptation et des représentations symboliques complexes.

Notions d'état énergétique

Les essais de modélisation qui seront esquissés au point II nécessitent pour leur compréhension un préambule sur la notion d'état énergétique d'un RNF et sur les algorithmes qui permettent de prendre en compte cet état.

Réseaux neuronaux et états énergétiques

Un réseau de neurones peut être considéré comme un système de particules analogue à ceux qui sont étudiés par la physique des particules et la physique stochastique. Chaque neurone peut être dans un état actif ou passif (états représentés la plupart du temps par les valeurs 1 et 0) et est relié à ses voisins par des liaisons synaptiques (représentées par des nombres réels quelconques). Une

liaison synaptique entre deux neurones représente en quelque sorte le degré de compatibilité ou d'incompatibilité entre ces neurones. Si la liaison est positive, les états actifs des neurones sont compatibles; si la liaison est négative, la position la plus harmonieuse entre les deux neurones sera atteinte lorsque l'un des neurones sera actif et l'autre passif. Si deux neurones sont actifs avec une liaison négative entre eux, une *tension* en résulte. Une *énergie*, fonction des états de tous les neurones et des valeurs des liaisons synaptiques, peut être calculée pour l'ensemble du réseau (Hopfield, 1982). Cette énergie sera d'autant plus élevée qu'existeront des états de tension entre les neurones. Si maintenant, les états des neurones représentent les valeurs possibles d'un certain système (physique, biologique, psychologique, etc.), et si les liaisons synaptiques représentent des contraintes de fonctionnement du système, la question que l'on cherchera à résoudre sera de trouver l'état du réseau (donné par l'état individuel de chacun de ses neurones) qui satisfera au mieux l'ensemble des contraintes engrammées dans les liaisons synaptiques. A partir d'un état initial des neurones, un processus dit de *relaxation* permet d'atteindre le premier état énergétique minimal, comme la bille de la figure 1 atteint le premier *fond de vallée*. Il existe cependant de nombreux états énergétiques minimaux correspondant à des minima locaux de la fonction énergie. L'atteinte du minimum local le plus proche de l'état initial du réseaux est suffisant dans le cas où le réseau est utilisé comme *mémoire associative* ou pour faire de la classification (on associe alors à l'information représentée par l'état initial du réseau, *une autre information* qui est l'état du réseau dans le minimum local le plus proche). Il n'en va pas de même lorsqu'on cherche la meilleure solution d'un problème, solution représentée par le minimum énergétique absolu du réseau. Il s'agit alors de transcender les minima locaux dont la technique actuellement la plus éprouvée est celle du *recuit-simulé*.

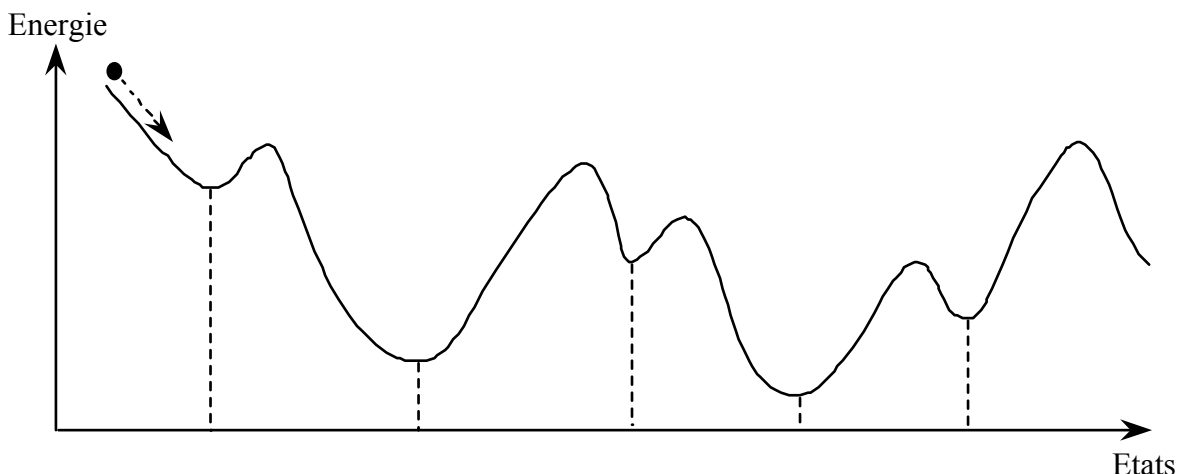


Fig. 1: Minima énergétiques d'un réseau à connexions symétriques. La bille symbolise un état du réseau à un instant donné, son mouvement représente le phénomène de relaxation.

Minima énergétiques et recuit-simulé

Le recuit-simulé (Metropolis *et al.*, 1953, Kirkpatrick, 1983) est une procédure mathématique stochastique de recherche de solutions dont le principe général est analogue à une méthode pratique de solidification de métaux. Elle constitue l'une des plus anciennes techniques employées pour transcender l'optimalité locale dans le cas de problèmes à solutions discrètes (Glover, 1986). La dynamique stochastique de l'algorithme autorise en effet la sortie de minima locaux et la convergence vers le minimum absolu, sans toutefois la certitude de l'atteindre. La procédure est itérative sur une quantité qui représente en quelque sorte l'*état d'agitation* ou de *désordre* du système, quantité appelée *température* en référence à ses sources d'inspiration. Celle-ci, élevée au départ pour éviter les minima locaux, est graduellement abaissée selon un protocole de refroidissement précis, à l'image d'un métal que l'on doit refroidir graduellement afin d'éviter des structures internes instables (tensions) et obtenir ainsi une structure très stable, proche d'un cristal. Elle consiste en une quantité aléatoire introduite dans le calcul de l'état du neurone en fonction de ses liaisons synaptiques. Lorsque cette quantité (température) est nulle, l'état du neurone est

entièrement déterminé par les valeurs de ces liaisons. Lorsqu'elle est non nulle, elle introduit un aléa dans la détermination de l'état du neurone. Plus elle augmente, plus cet état devient aléatoire et indépendant des connexions synaptiques. L'état du réseau complet est alors d'autant plus chaotique et imprévisible que la température est élevée, et d'autant plus stable et déterministe que la température est basse. A température nulle, l'énergie du réseau ne peut que décroître ou rester stationnaire, alors qu'une température qui s'élève donne au réseau la liberté d'accroître son énergie. Ce qui explique la possibilité de sortir de minima locaux, conférée au réseau par une température suffisante.

II. Modèles computationnels en psychiatrie

L'intelligence artificielle en psychiatrie

L'intelligence artificielle par nature est étroitement liée au champ des sciences cognitives (Winograd 1987, Demongeot *et al.*, 1989). Dans cette optique, la plupart des approches théoriques et pratiques du traitement automatique de l'information ont été reprises dans des tentatives de démonstration de leur pertinence en tant que modèle de référence du fonctionnement cérébral. Kydd et Wright (1986) par exemple argumentent sur la modélisation du phénomène mental par un *automate d'états finis*. J.R. Anderson (1983) propose une théorie des *Modèles Mentaux* qu'il concrétise dans son modèle ACT* (Adaptative Control of Thought) de la mémoire et du fonctionnement cérébral.

L'intelligence artificielle a par ailleurs été concrètement mise en pratique dans le champ de la psychiatrie, avec pas moins de 15 réalisations de systèmes à bases de connaissances (Binik, 1988, Morelli, 1989, Ohayon, 1990, Rialle et Ohayon, 1991). Les bases de connaissances des systèmes experts sont le lieu par excellence des modèles computationnels du raisonnement et de l'inférence en psychiatrie. Elles sont dévolues dans la majorité des cas à la simulation du raisonnement diagnostique, pronostique ou thérapeutique et, plus rarement mais de manière non moins réelle, à l'aide directe au patient dans son processus psychothérapeutique. Nous en citerons ici quelques-unes des plus marquantes. Les premiers systèmes à bases de connaissances en psychiatrie sont apparus dès la fin des années 70. Heiser et Brooks (1978), proposent les premiers un système expert d'aide à la prescription thérapeutique dans le domaine de la psychose et de la dépression. Quelques années plus tard, Mulsant et Servan-Schreiber (1986) proposent le système BLUE BOX dans le domaine diagnostique de la dépression. Feinberg et Carrol (1982) construisent pour leur part un système d'aide au diagnostic des dépressions endogènes. Un système expert en Géro-psycho-Psychiatrie — METHUSELAH — est proposé par G. Werner (1987). Le système expert JPSY de Portelli *et al.* (1989), fondé sur la nosographie psychiatrique française, est consacré au diagnostic des troubles mentaux chez l'enfant. Binik *et al.* (1988) expérimentent un système expert pour l'aide à la psychothérapie des dysfonctionnements sexuels. Plugge *et al.* (1990) produisent le système EVINCE-I, pour le diagnostic différentiel de la démence. A partir de 1982, M. Ohayon (1990) développe le système ADINFER, dont la base de connaissances se veut être une fidèle translation du DSM-III-R et dont l'un des buts principaux est de répondre aux nécessités des études multicentriques de psychopharmacologie.

Par ailleurs, le test de Turing (1950), longtemps considéré comme le moyen idéal pour déterminer le caractère intelligent d'un programme d'ordinateur, a suscité une réflexion approfondie chez un certain nombre de psychiatres. Ses limitations ont notamment été démontrées par Heiser *et al.* (1979) et Ohayon (1990).

La possibilité d'un déficit du traitement de l'information dans la schizophrénie a encouragé la modélisation (Gallaway et Naghdi, 1982) : selon eux, les schizophrènes sont déficitaires dans le traitement de l'information caractérisé par des processus conscients, sériels et "limited-channel-capacity". Les processus parallèles sont quant à eux, plus automatiques, et semblent opérer en dehors de la conscience. Ces derniers semblent normaux, voire super-normaux dans la

schizophrénie. L'idée principale de ce modèle est donc que les schizophrènes présentent un déficit des processus sériels contrôlés mais non des traitements parallèles automatiques.

Modèles connexionnistes

La psychiatrie, fidèle à ses précédents rendez-vous avec l'intelligence artificielle, tente de plus en plus d'utiliser les réseaux neuronaux formels pour reproduire et analyser des déficits cognitifs rencontrés chez des malades psychiatriques (Farah et McClelland, 1992, Cohen et Servan-Schreiber, 1992a et 1992b). Malgré la distance qui sépare les modèles computationnels en général du système cérébral, quelques chercheurs risquent régulièrement la démonstration que le cerveau fonctionne en partie, jusqu'à un certain point ou dans certaines de ses composantes, comme l'un des modèles proposés par le connexionnisme. Nous présenterons succinctement les principaux travaux conduits dans cette perspective. Ces travaux sont d'une inégale importance : si la plupart font appels à des outils théoriques complexes et présentent des simulations mettant en œuvre des moyens importants, d'autres en restent à l'établissement d'analogies entre certains modèles mathématiques et des comportements neuropsychologiques rencontrés chez le sujet humain.

Modélisation de comportements pathologiques et RNF

La modélisation de certains processus mentaux par réseaux de neurones correspond à une interrogation de plus en plus prégnante en neuropsychologie et en psychiatrie (Seron, 1990, Caramazza, 1990) : le substrat neuronal dans lequel prennent naissance les phénomènes mentaux peut-il être mimé par des réseaux de neurones formels ? Si oui, dans quelle mesure l'observation du "comportement" du réseau, lorsque l'on fait varier ses paramètres, peut-elle apporter des indications sur certaines pathologies ?

Dans cette optique, R.E. Hoffman (1987) a établi une série de parallèles entre les phénomènes observables dans un réseau de Hopfield et ceux qui sont observables chez un malade psychiatrique. Il rapproche notamment le comportement chaotique du réseau formel, à température élevée, avec le caractère imprévisible et sans suite logique du discours chez le maniaque. En outre, le réseau hopfieldien est caractérisé par une limite sévère du nombre d'éléments mémorisables, limite au-delà de laquelle un phénomène dit d'*oubli catastrophique* survient très rapidement. Le rappel associatif des contenus mémorisés est alors complètement perturbé, donnant au réseau un comportement de type *délinant*, analogue à certains symptômes schizophréniques.

Modélisation connexionniste de la dyslexie acquise

Ackley, Hinton et Sejnowsky (1985) ont proposé un premier système d'apprentissage du sens des mots d'un langage. Ce système est fondé sur un modèle connexionniste stochastique, qu'ils ont eux-mêmes élaboré et auquel ils ont donné le nom de *Machine de BOLTZMANN*. Plus récemment, Hinton et Shallice (1991) ont proposé un autre système, dont le principe est différent mais qui opère toujours dans un domaine de décisions probabilistes. Ils ont plus particulièrement consacré ce travail à la modélisation de la *dyslexie acquise*. C'est ce second système qui est résumé ici. L'intention première de Hinton et Shallice était de démontrer comment un réseau connexionniste pouvait directement associer des sémèmes à des graphèmes, sans avoir à dédier un neurone particulier à chaque mot. Ce n'est qu'après la réalisation du réseau, et au cours de son étude approfondie, qu'ils se sont aperçu de ses étonnantes capacités à reproduire des phénomènes neuropsychologiques. Ceux-ci sont observés lors de la suppression de certains neurones et liens synaptiques. Le réseau commet alors des erreurs très similaires à celles que l'on rencontre chez des sujets atteints de dyslexie acquise (dyslexie profonde et dyslexie sémantique). Chaque unité à l'entrée du réseau (unités de *graphème*) code deux informations : un caractère et sa position dans le mot. Les sorties du réseau sont des unités de sens (*sémèmes*). Chaque mot présenté au réseau est censé mettre en activité les sémèmes qui lui sont associés et dont l'ensemble en représente le sens. Pour restreindre la complexité de l'expérimentation, les mots sont limités à 3 ou 4 caractères, le vocabulaire est limité à une quarantaine de mots, et les sémèmes sont au nombre de 68, répartis en plusieurs sous-groupes. Le réseau est à apprentissage supervisé, utilisant une version itérative de la

procédure de rétropropagation du gradient de l'erreur telle qu'expliquée dans Rumelhart *et al.* (1986). La connectivité du réseau, au demeurant assez complexe, est destinée à *capturer* les connexions de renforcement ou d'inhibition tant au niveau des entrées (*i.e.* entre caractères alphabétiques) qu'au niveau des sorties (*i.e.* entre unités de sens). La caractéristique fondamentale du réseau est la présence d'états attracteurs dans sa dynamique de relaxation (Amit, 1989).

Les *lésions* pratiquées sur le réseau sont de deux types : suppression de neurones ou suppression de connexions synaptiques entre neurones. Elles conduisent à deux types de réponses incorrectes : les omissions (*i.e.* tous les sémèmes activés sont justes, mais il en manque quelques-uns), et les erreurs (*i.e.* parmi les sémèmes activés, certains le sont à tort). Parmi ces dernières, les auteurs en distinguent 4 sortes : les erreurs sémantiques (des mots différents sont associés au même ensemble de sémèmes), les erreurs visuelles (mots associés à tort à des sémèmes en raison de leur proximité orthographique), les erreurs mixtes (sémantiques et visuelles) et les erreurs autres. Différents niveaux de sévérité des lésions ont été pratiqués sur le réseau, et ont été comparés à différents types de patients. Les deux principales similarités de comportement entre un sujet dyslexique et le réseau concernent d'une part la classe d'erreurs commises et d'autre part les performances en situation de choix forcés (*i.e.* lorsque le choix sémantique est pratiqué en dépit du trop grand manque d'informations causé par les lésions). Ces modélisations ont en eux-mêmes une référence à une représentation de ce que l'on désigne de lexique mental. Stip et Lecours (1992) ont montré que l'accès au lexique mental était modifié dans la dépression majeure lors d'une épreuve de décision lexicale vis à vis de mots dépressogènes. Nous voyons là encore que l'affect en tant que tel est à considérer dans un système cognitif global. La lésion est remplacée dans ce genre de représentation par une perturbation de l'humeur. Il reste à faire des progrès dans l'analyse du langage des schizophrènes et en particulier dans la production des néologismes et de certaines variétés de discours psychotiques comme la schizophasie glossomaniaque et glossolalique (Lecours, Stip & Tremblay 1992). Selon eux, la modélisation devrait rendre compte non seulement de déficits mais encore d'hyperfonctionnements lexicaux.

Modélisation connexionniste de symptômes biologiques et comportementaux de la schizophrénie

Dans un récent travail, J. Cohen et D. Servan-Schreiber (1992a) proposent trois simulations connexionnistes de phénomènes biologiques et comportementaux associés à la schizophrénie correspondant à trois tests cognitifs : le test de Stroop, le test de performance continue et la tâche de désambiguïsation lexicale. Au niveau comportemental, Cohen et Servan-Schreiber s'intéressent aux désordres de l'attention sélective et à ceux du langage; tandis qu'au niveau biologique, ils tentent d'établir un parallèle entre l'effet d'un paramètre essentiel dans la fonction non-linéaire du neurone formel, le *gain d'activation*, et celui d'un neuro-modulateur particulièrement lié aux déficits cognitifs de la schizophrénie: la concentration de dopamine dans le cortex pré-frontal. Ils montrent notamment comment certaines variations du gain, dans le module connexionniste simulant le cortex pré-frontal, introduisent des anomalies fonctionnelles quantitativement similaires à celles que l'on observe chez des schizophrènes, au cours des mêmes tâches. Les RNF utilisés sont à apprentissage supervisé au moyen de la règle de rétropropagation du gradient de l'erreur. Les auteurs partagent l'idée que cette règle n'est pas vraiment représentative d'un fonctionnement biologique. Ils n'en considèrent pas moins qu'elle reste une bonne formule de rétroaction pour l'ajustement des poids synaptiques.

Le test de Stroop, objet de la première expérimentation, consiste en deux tâches: dans la première, les sujets nomment la couleur de l'encre avec laquelle sont écrits certains mots; dans la seconde, ils doivent lire les mots en ignorant la couleur avec laquelle ils sont écrits. L'expérience subjective de ce test est qu'il est plus facile de lire un mot sans tenir compte de la couleur de l'encre que de citer la couleur d'un mot sans tenir compte de son sens. Le modèle de cette première expérimentation est composé d'un module de représentation des stimuli-mots, d'un module de représentation du contexte (*i.e.* la "demande": lire le mot ou citer sa couleur) et d'un module de réponses. Ce modèle connexionniste de l'attention sélective est repris d'un travail antérieur conduit par Cohen, Dunbar et McClelland (1990).

Cohen et Servan-Schreiber établissent dans un premier temps la fidélité du réseau à reproduire, l'effet du contexte (associé au lobe pré-frontal) sur la réponse du sujet. Ils perturbent ensuite le réseau pour voir si le comportement du sujet reste fidèle à une hypothèse d'explication des déficits attentionnels dans la schizophrénie. L'hypothèse des auteurs étant que si le lobe pré-frontal est le siège des représentations contextuelles et si, d'autre part, la schizophrénie implique un dysfonctionnement de ce lobe, alors un déficit spécifique de la schizophrénie devrait être obtenu en perturbant le lobe pré-frontal.

Cohen et Servan-Schreiber concluent en effet qu'une réduction du gain dans le module de contexte introduit l'effet attendu, à savoir une augmentation des délais de réponse et une augmentation disproportionnée des conflits de choix de couleur. Il est intéressant de constater que ces conclusions sont en rapport avec les études cliniques de l'effet Stroop dans la schizophrénie (Everett *et al.*, 1989).

Méthode du recuit-simulé et psychanalyse

Galatzer-Levy (1988) tente de montrer qu'un modèle d'état énergétique tel que celui qui a été évoqué dans la partie I (§ *Notions d'états énergétiques*) possède une dynamique interne et des propriétés propres à mimer un comportement humain de recherche de solutions en situation de conflits émotionnels. Le déterminisme psychique peut être traduit, jusqu'à un certain point, en termes de décisions dictées par des lois quantitatives. Ce n'est pas contredire Freud qui proposait justement un modèle économique, énergétique et topique du psychisme humain. Des notions typiquement quantitatives, d'intensité ou de fréquence d'affects, dictent le comportement et l'état psychique. Cet état dans lequel se trouve une personne, en cure psychanalytique, à un instant donné est plus ou moins stable. A toute action ou décision est associée une grandeur particulière qui est l'énergie psychique investie par le sujet dans cette action ou cette décision.

Galatzer-Levy trouve alors frappante l'analogie qui peut être établie entre la technique du recuit-simulé et le processus psychanalytique qui consiste à amener le sujet dans un état énergétique plus élevé. Cet état lui permet alors de retrouver des sensations et des émotions psychologiques antérieurement vécues, mais transformées en tensions du fait d'une recherche brutale d'état stable. La cure conduit le sujet dans des états énergétiques suffisamment élevés, puis le maintient suffisamment longtemps pour que des états énergétiques plus stables soient atteints. L'intensité de l'affect, de l'investissement psychique, diminue ensuite graduellement, au cours de la cure, jusqu'à ce que le sujet ait atteint un état de stabilité optimale correspondant à une certaine résolution des conflits.

Galatzer-Levy rappelle également que l'un des atouts de l'intelligence artificielle est de ne poser pratiquement aucune distance entre les hypothèses théoriques qu'elle produit et les vérifications expérimentales de ces hypothèses dans des programmes informatiques. Cette proximité entre des hypothèses et leur mise en œuvre concrète à des fins de vérification et de validation n'est pas l'apanage de toutes les disciplines des sciences humaines.

Modélisation connexionniste de l'hypnose

Dixon *et al.* (1990) proposent une réflexion sur les possibilités d'appliquer le modèle connexionniste de l'effet Stroop, tel que celui de Cohen et Servan-Schreiber, pour la modélisation de la réponse hypnotique. Leur étude, portant sur 27 sujets repartis en trois catégories d'hypnotisabilité, confirme l'hypothèse d'une relation significative entre la sensibilité à l'hypnose et la relation mot-couleur du test de Stroop. Les sujets à forte sensibilité à l'hypnose révèlent, par le test de Stroop, une plus grande sensibilité à l'expression verbale qu'à l'expression de la couleur. Les auteurs suggèrent donc que la différence significative des délais de réponse entre tests avec congruence et tests avec non-congruence mot/couleur, observés chez les sujet hypnotiques, peuvent être expliquées par des connexions synaptiques anormalement renforcées dans la partie verbale du modèle.

Dans un autre exemple de modélisation, des chercheurs ont réalisé un désordre dissociatif (Li et Spiegel, 1992) dans lequel le traumatisme est représenté. Ce traumatisme fixe dans une position rigide une certaine architecture du réseau neuronal. Il représente une "contrainte" pour la forme du

réseau. Cette contrainte n'est pas un lieu ou une mémoire en tant que tel mais plutôt une fixation-limite des interconnexions du réseau.

Processus du rêve et connexionnisme

La modélisation des processus du rêve par une approche connexionniste a montré qu'en augmentant le bruit et en diminuant l'inhibition durant le sommeil-REM, le réseau avait tendance à changer brutalement d'état, avec moins d'harmonie et la formation d'attracteurs parasites, qui expliquerait les caractéristiques oniriques faits de changements thématiques soudains, de déplacement et de condensation (Globus, 1989). Dans une autre proposition de modélisation du sommeil, les auteurs ont représentés dans un réseau neuronal des facteurs modulateurs aminergiques et cholinergiques: sous différents états de modulation simulée aminergique et cholinergique, les séquences de sortie du réseau sont très différentes. Dans l'état de veille simulée, les séquences sont progressives, avec des changements d'une séquence (A1,A2,A3...) à une autre (B1,B2,B3,...) induits par des stimuli extérieurs au réseau. Dans le sommeil non-REM simulé, les séquences de sortie tendent à persévérer. Enfin dans le sommeil REM simulé, les séquences sont discontinues et mélangées (Sutton *et al.* 1992). C'est bien dans cette dernière phase que se produit l'activité onirique.

Modélisation connexionniste de la cognition des autistes de haut niveau.

Le fonctionnement cognitif de certains autistes, "idiots-savants" pose des questions très pertinentes (Mottron *et al.* 1991) à la neuropsychologie. La modélisation de leur capacités spéciales est semblait-il plus connexionniste que modulaire tant est présente la faillite des interprétations sérielles des hyperfonctionnements: calculateur prodige, hypermnésique visuel, génie musical etc. Un modèle connexionniste pour la modélisation d'un sujet calculateur de date a notamment été proposé par Norris (1990). L'apprentissage des calculs de calendrier est de type supervisé et utilise la classique règle de rétropropagation du gradient de l'erreur.

Limitations des modèles séquentiels et potentialités du parallélisme

Parmi les nombreuses mutations technologiques qui sont à l'œuvre aujourd'hui, celle de l'ordinateur parallèle en est l'une des plus significatives (Hillis 1992, Deng 1992). La pression qui préside au développement croissant des ordinateurs parallèles est de nature essentiellement technologique : on cherche avant tout à réduire les durées de traitement, compte tenu de la quantité de plus en plus importante de calculs demandés à l'ordinateur. Il existe cependant une raison plus épistémologique en faveur de cette nouvelle génération d'ordinateurs : c'est celle du constat d'insuffisance sinon d'échec du modèle classique d'ordinateur: la *machine de Von NEUMANN* (Backus, 1978). La *programmation massivement parallèle*, permise par les ordinateurs parallèles, possède quelques aspects d'une révolution dans le traitement automatique de l'information et la modélisation cognitive (Bailey, 1992).

L'ordinateur classique propose un mode strictement séquentiel de fonctionnement et se résume à un processeur unique (le microprocesseur dans le cas des micro-ordinateurs) dont la tâche est d'exécuter, dans l'ordre où elles lui sont présentées, les instructions qui constituent un programme. Sa mémoire peut être schématisée par une série de *cases* accessibles par une *adresse* et contenant une *valeur*. Elle est découpée en *mots* de 8, 16, 32,..., bits, et une information est codée dans chacun de ces *mots* de mémoire. Le processeur possède un registre d'adresse, un registre d'instruction et un compteur ordinal qui lui indique, après chaque instruction, l'adresse de la prochaine zone de mémoire à charger dans ses registres. On retrouve donc, inscrite au cœur de l'ordinateur, une dichotomie entre adresse et valeur qui se retrouve naturellement dans la plupart des langages de programmation et dans les grands principes d'algorithmique. Il devient alors difficile, après une telle omniprésence de la dichotomie variable-valeur, de concevoir un modèle autrement qu'en ces termes. D'une manière plus ou moins consciente, le modèle d'ordinateur de Von Neumann oblige donc à concevoir les modèles computationnels à la fois en termes de couples (variable, valeur) ou

de triplets (objet, attribut, valeur) — ce qui revient au même — et de suites ordonnées d'opérations arithmétiques et logiques sur ces termes.

Lorsqu'il s'agit de construire un modèle d'un phénomène comportemental ou cognitif, les notions de variables et de valeurs de ces variables à un instant donné deviennent difficilement déterminables de manière catégorique. La grande variabilité d'expression des phénomènes dont on cherche une représentation est donc difficilement compatible avec la rigueur imposée par une représentation computationnelle en programmation classique. Or, on demande de plus en plus à l'informatique, et particulièrement à l'intelligence artificielle, de reproduire des fonctions qui ne peuvent pas, sauf au prix d'un réductionnisme excessif, être modélisées dans la stricte séquentialité et la dualité variable-valeur. Il en résulte une inadéquation entre l'outil et l'objectif qui lui est assigné, le premier étant séquentiel et le second ne l'étant pas *a priori*. Cette inadéquation résulte en fait d'une *différence de complexité* entre d'une part le mode opératoire pas à pas et ordonné du processeur électronique, et les modalités multiples, asynchrones, concurrentielles, ..., propres aux processus psychologiques (pensées, représentations mentales, intentionnalité, déductions, analogies...). Représenter de tels processus par une liste d'instructions bien précises et ordonnées ne peut alors se faire qu'au prix d'un réductionnisme drastique.

D'un point de vue pratique, les simulations de tels processus sur ordinateur séquentiel, outre le réductionnisme important auquel elles doivent se plier, se heurtent à des délais d'exécution sans proportion avec ceux qui sont raisonnablement admis (des durées de traitement de plusieurs heures à plusieurs journées ont ainsi fréquemment été rapportées dans des domaines tels que le connexionnisme ou l'analyse linguistique et textuelle). Ces temps de traitement constituent donc un obstacle pratique majeur à l'expérimentation de modèles.

Il faut cependant se garder d'établir un rapprochement trop rapide entre machine parallèle et système cérébral : le cerveau ne fonctionne pas comme un ordinateur parallèle! Bien que la compréhension du fonctionnement cérébral soit encore très imparfaite, ce dont on est davantage certain, c'est que le cerveau traite les informations qui lui parviennent des récepteurs sensoriels lorsqu'elles se présentent, et elles se présentent le plus souvent de manière simultanée. Par conséquent, la sélection des informations pertinentes dans le flot d'informations ressemble à du parallélisme (Sokolowski, 1992).

Discussion et Conclusions

Les modèles de dysfonctionnement cognitif qui ont été rapportés ici permettent d'affirmer la place importante qu'ils tiennent dans la problématique générale de la modélisation cognitive. Ils expriment en effet l'idée que, pour être vraisemblable, un modèle cognitif doit être "lésionnable". La difficulté en psychiatrie est d'abord d'accepter et de se représenter la "lésion". A la différence d'un ordinateur conventionnel, un être vivant — particulièrement l'humain — peut *fonctionner* avec et malgré une certaine quantité de lésions dans son organisation structurale et fonctionnelle. Il possède non seulement la capacité de palier jusqu'à un certain point ses déficits fonctionnels, mais aussi celle de rétablir progressivement, dans certaines conditions, son fonctionnement normal. De sorte que, pour être réaliste, un modèle doit pouvoir tolérer et rendre compte d'un certain nombre de perturbations. Les réseaux connexionnistes dotés d'une dynamique à attracteurs (Amit, 1989) semblent constituer, actuellement, l'une des meilleures voies de modélisation de ces perturbations.

Les méthodes connexionnistes sont souvent investies du rôle de concurrent ou de complément des méthodes de l'intelligence artificielle "classique". Une ambiguïté se développe d'ailleurs quand à la définition et aux contours précis de l'intelligence artificielle: les spécialistes de longue date de cette discipline ont tendance à restreindre ce vocable aux méthodes purement symboliques et à revendiquer pleinement le champ de l'apprentissage symbolique automatique. Les non-spécialistes de l'IA ont tendance à assimiler ce terme aux méthodes pratiques avancées de réalisation des

modèles cognitifs, quelles que soient ces méthodes (Perez, 1989). Il reste que, au-delà des questions de définitions et de spécificités, les deux paradigmes apparaissent aussi étroitement liés que le haut et le bas d'un même objet : le tout neuronal ne semble possible que dans certaines fonctions cognitives (perceptions, localisation, reconnaissance de formes ou de sons,...), alors que le tout symbolique n'est concevable qu'à partir d'un certain niveau d'abstraction. Si l'on veut concevoir des systèmes doués d'un minimum d'intelligence et d'autonomie, la jonction entre les deux niveaux doit nécessairement être réalisée. Le problème du passage du neuronal au symbolique, et *vice versa*, est un thème majeur de préoccupation en modélisation cognitive.

La référence au connexionnisme pour l'explication des phénomènes mentaux et des comportements humains reçoit régulièrement de sévères critiques. Massaro (1988) souligne, en même temps que Fodor et Pylyshyn (1988), la faiblesse des modèles connexionnistes quant à l'objectif de modélisation de la réalité psycho-physique et biologique. L'intérêt pour les modèles parallèles et distribués ne cesse néanmoins de se développer et d'exciter la curiosité des chercheurs en psychiatrie (Cohen et Servan-Schreiber, 1992b). Le secret de cet engouement réside certainement dans les étonnantes capacités des réseaux neuronaux formels à épouser des comportements neuro-biologiquement et psychologiquement vraisemblables. Leur apparente supériorité par rapport aux techniques de modélisation symbolique, proposées par l'intelligence artificielle classique, est due à plusieurs facteurs dont le principal est peut-être la possibilité qu'ils offrent d'expérimenter aisément des modifications de paramètres, des perturbations, des "lésions" structurelles et fonctionnelles et d'en mesurer les conséquences dans le comportement général du modèle. Les perturbations induites par ces manipulations paramétriques sont ensuite rapprochées des dysfonctionnements cognitifs observés sur des malades psychiatriques. Pour l'heure les recherches se situent essentiellement au niveau de la mesure de l'adéquation entre le modèle et la réalité : jusqu'à quel point le RNF donne-t-il les réponses attendues et se comporte-t-il comme l'original vivant ? L'enthousiasme actuellement suscité par les RNF se situe essentiellement dans la recherche de cette adéquation et des mécanismes d'apprentissage et d'adaptation qui la rendent possible. C'est ainsi par exemple que le réseau de Hopfield est utilisé comme mémoire associative, modèle d'optimisation combinatoire ou modèle de certains phénomènes mentaux. Le modèle génétique apporte des solutions dans de nombreux problèmes combinatoires (Davis, 1991) et offre un cadre conceptuel à la modélisation en neuropsychologie (Dehaene et Changeux, 1989). L'induction de connaissances concernant le sujet humain à partir de l'observation du réseau et de la manipulation paramétrique reste néanmoins une question centrale. Que peut-on observer sur un modèle formel qui puisse informer ou répondre à une question concernant l'Homme ? Dans quelles conditions peut-on effectuer le mouvement inverse d'induction de phénomènes biologiques ou neuro-psychologiques, non directement observables chez le sujet humain, à partir des réponses et des comportements observables sur le RNF ? En bref, pourra-t-on élaborer des modèles mathématiques de la *psyché* et les utiliser pour inférer des comportements humains ? Les premiers modèles neuro-psycho-mimétiques qui fleurissent actuellement dans le domaine de la psychiatrie, s'ils ne permettent pas encore de répondre à toutes ces questions, les placent néanmoins à l'avant scène des questions épistémologiques posées par les sciences cognitives. Il serait intéressant que l'hypothèse neuro-développementale puisse être modélisable c'est-à-dire que le réseau puisse se modifier en profondeur à une période précise de son fonctionnement. Ce serait fondamental pour la recherche sur la schizophrénie où les perturbations cognitivo-comportementales ne s'expriment qu'après un laps de temps comme l'adolescence. Dans le champ des relations entre la psychiatrie et les neurosciences, la modélisation occupe une place grandissante.

Remerciements. Cette étude a pu être menée à bien grâce à la subvention N 031 77 91 08 et à la *bourse d'excellence* attribuées à V. Rialle par le CONSEIL RÉGIONAL RHÔNE-ALPES, ainsi qu'à la subvention 23-1992 du Fond CAFIR, Université de Montréal, Département de Psychiatrie et à la Bourse de Centre du Centre Fernand Seguin, attribuées à E. Stip.

Bibliographie

Ackley DH, Hinton GE, Sejnowski TJ (1985) A learning algorithm for Boltzmann machines. *Cognitive Science*, 9, pp 147-169.

- Amit DJ (1989) *Modeling Brain Function - The world of attractor neural networks*. New York: Cambridge University Press.
- Amy B (1989) Contextual cognitive machines. In: Tiberghien G (ed), *Advances in cognitive sciences*. Vol. 2 : theory and applications. Ellis Horwood.
- Amy B, Giacometti A, Gut A (1990) Modèles connexionnistes de l'expertise. *Proceedings NEURO NIMES '90*, Nanterre, France: EC2, pp 99-119.
- Anderson JR (1983) *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Andreasen N (1986) Can schizophrenia be localized in the brain? Progress in psychiatry. American Psychiatric Press. Washington DC.
- Backus J (1978) Can programming be liberated from the von Neumann style ? A functional style and its algebra of programs. *Communications of the ACM*, 21, pp 613-641.
- Bailey J (1992) First We Reshape Our Computers, Then Our Computers Reshape Us: The Broader Intellectual Impact of Parallelism. *Dædalus*, 121(1), pp 67-86.
- Bayes T (1763) An essay towards solving a problem in the doctrine of chances, *Phil. Trans. B.*, **53**, 370. In *Biometrika*, 1958, **45**, p 296.
- Biniak YM, Servan-Schreiber D, Freiwald S, Hall KSK (1988) Intelligent Computer-Based Assessment in Psychotherapy: An Expert System for Sexual Dysfunction. *Journal of Nervous and Mental disease*, 176(7), pp 387-400.
- Caramazza A (1990) Des déficits causés par les lésions cérébrales aux systèmes cognitifs du sujet normal. In: Psychologie et Cerveau, Presse Universitaire de France, Paris, pp 177-194.
- Cohen JD, Dunbar K, McClelland JL (1990) On the control of automatic processes: a parallel distributed processing model of the stroop effect. *Psychological Review*, 97 (3), pp 332-361.
- Cohen J, Servan-Schreiber D (1992a) Context, cortex, and dopamine: A connectionist approach to behavior and biology in schizophrenia. *Psychological Review*, 99(1), pp 45-77.
- Cohen JD, Servan-schreiber D (1992b) Introduction to Neural Network Models in Psychiatry. *Psychiatric Annals*, 22(3), pp 113-118.
- Davis (1991) *The handbook of Genetic Algorithms*. Van Nostrand Reinhold.
- Dehaene S, Changeux JP (1989) A simple Model of Prefrontal Cortex Function in Delayed-Response Tasks. *Journal of Cognitive Neuroscience*, 1(3), pp 244-261.
- Demongeot J, Hervé T, Berthomier F, François O (1989) Neural networks : From neuro-computing to neuro-modelling. In : A. Goldbeter (ed), *Theoretical models for cell to cell signalling*. New York: Academic Press, pp 109-121.
- Demongeot J, Hervé T, Rialle V, Roche C (eds) (1988) *Artificial Intelligence and Cognitive Sciences* (Manchester University Press, New York).
- Deng Y (1992) Perspectives on Parallel Computing. *Dædalus*, 121(1), pp 31-52.
- Dixon M, Brunet A, Laurence JR (1990) Hypnotisability et Automaticity: Toward a Parallel Distributed Processing Model of Hypnotic Responding. *Journal of Abnormal Psychology*, 99(4), pp 336-343.
- Dreyfus HL (1984) *Intelligence Artificielle : Mythes et limites*. Trad. RM Vassalo-Villaneau, Ed. Flammarion, Paris.
- Everett J, Laplante L, Thomas J,(1989),The selective attention deficit in schizophrenia: limited resources or cognitive fatigue. *Journal of nervous and mental disease*, 177, pp 735-738.
- Farah M, McClelland JL (1992) Neural Network Models and Cognitive Neuropsychology. *Psychiatric Annals*, 22(3), pp 148-153.
- Feinberg M, Carrol B J (1982) Separation of subtypes of depression using discriminant analysis I: separation of unipolar endogenous depression from non endogenous depression, *British Journal of Psychiatry*, 140, pp 384-391.
- Feldman JA (1989) Connectionist Representation of Concepts. In: R. Pfeifer, Z. Schreter, F. Fogelman-Soulié, L. Steels (eds), *Connectionism in perspective*. Amsterdam, The Netherland: North-Holland Publ., pp 25-45.
- Fodor JA, Pylyshyn ZW (1988) Connectionism and cognitive architecture: a critical analysis. *Cognition*, 28 (1-2), pp 3-71.
- Galatzer-Levy RM (1988) On working through: a model from artificial intelligence. *Journal of the American Psychoanal. Association*, 36(1), pp 125-151.
- Galloway E, Naghdi S (1982) An information processing model for schizophrenia. *Archives of General Psychiatry*, 39, pp 339-347.
- Giacometti A, Amy B et Grumbach A (1992), Theory and Experiments in Connectionist AI : a Tightly-coupled Hybrid System. In: *Proc. ICANN'92*, Artificial Neural Networks 2, Elsevier Science Publishers.
- Globus GG (1989) Connectionism and the Dreaming Mind. *Journal of Mind and Behavior*, 10 (2), pp 179-195.
- Goldberg DE (1989) *Genetic algorithms in search, optimisation and machine learning*. Reading, Massachusetts: Addison-Wesley.
- Hebb DO (1949) *The Organization of Behavior*. New York: John Wiley, Sons.
- Heiser JF, Colby KM, Faijght WS, Parkinson RC (1979) Can psychiatrists distinguish a computer simulation of paranoia from the real thing ? The limitations of Turing-like tests as measures of the adequacy of simulations, *J. Psychia. Res.*, 15(3), pp 149-162.
- Heiser J-F, Brooks R E (1978) Design considerations for a clinical psychopharmacology advisor, *Proc., 2nd Ann. Symp. On Computer Applications in Medical Care*, New-York,IEEE, pp 278-285.
- Hendler JA (1989) On the need of hybrid systems. *Connection Science* (special issue on Hybrid Connectionist / Symbolic Systems), 1 (3).
- Hillis WD (1992) What Is Massively Parallel Computing, and Why Is It Important? *Dædalus*, 121(1), pp 1-15.
- Hinton GE, Sejnowski TJ (1986) Learning and Relearning in Boltzmann Machines. In: Rumelhart, D. E., McClelland, J L, the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 1. Cambridge, MA: MIT Press/Bradford Books, pp 282-317.
- Hinton GE et Shallice T (1991) Lesioning an Attractor Network: Investigation of Acquired Dyslexia. *Psychological Review*, 8(1), pp 74-95.
- Hoffman RE (1987) Computer Simulation of Neural Information Processing and the Schizophrenia-Mania Dichotomy. *Archives of General Psychiatry*, 44, pp 178-188.
- Holland JH (1975) *Adaptation in Natural and Artificial Systems*. Ann Arbor, Michigan : University of Michigan Press.
- Holland JH (1986) Escaping brittleness. *Machine Learning*, 2.
- Hopfield JJ, Tank DW (1985) "Neural" computation of décisions in optimisation problems, *Biol. Cybern.*, 52, 141-152.
- Hopfield JJ (1982) Neural Networks and physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Science* (79), USA, 2554-2558.
- Imbert M (1987) *Neurosciences et sciences cognitives*. Le Débat, 47, Paris: Gallimard.
- Kaplan S, Weaver M, French R (1990) Active Symbols and Internal Models: Towards a Cognitive Connectionism. *Artificial Intelligence and Society*, 4(1), pp 51-71.
- Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. *Science*, 220, pp 671-680.
- Kodratoff Y (1989) IA et approche connexionniste, *Bulletin de liaison de la recherche en informatique et en automatique*. INRIA, 124, pp 16-21.
- Kohonen T (1988) *Self-organization and associative memory*. Berlin: Springer-Verlag.
- Kydd RR, Wright JJ (1986) Mental phenomena as change of state in a finite-state machine. *Australian and New Zealand Journal of Psychiatry*, 20(2), pp 158-165.
- Lecours AR, Stip E, Tremblay M (1992) La schizophasie et le discours du schizophrène. *Sémiotiques.*, CNRS, 3, pp 9-22.
- Li D, Spiegel D (1992) A Neural Network Model od Dissociative Disorders, *Psychiatric Annals*, 22(3), pp 144-147.
- MacCarthy J (1977) Epistemological problems of artificial intelligence. *Fifth Int Joint Conf on Art Int*, MA: Cambridge.
- Massaro DW (1988) Some criticisms of connectionist models of human performance. *Journal of Memory and Language*, 27(2), pp 213-234.
- McClelland J L, Rumelhart D E, the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 2: Psychological and Biological Models, Cambridge, Ma: MIT Press/Bradford Books.
- McCulloch W, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Physics*, 5, pp 115-133.
- Memmi D (1989) Connectionism and Artificial Intelligence. *Proceedings NEURO-NIMES '89*. Nanterre, France: EC2, pp 17-34.

- Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E (1953) Equation of state calculations by fast computing machines. *Journal of Chem. Physics*, 21, pp 1087-1092.
- Morelli R A (1989) Artificial intelligence in psychiatry : issues and questions, in: Y. Kim and Spelman FA, *Proc.11th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, IEEE, New-York, USA, vol 6, pp 1812-1813.
- Mottron L, Belleville S, Stip E (1991) Etude des stratégies d'encodage chez un autiste hypermnésique de haut niveau. *L'Encéphale*, 17, p 549.
- Musant B, Servan-Schreiber D (1984) Knowledge engineering: a daily activity on an hospital information ward. *Computers and Biomedical Research*, 17, pp 71-91.
- Newell A (1980) Physical Symbol Systems. *Cognitive Science*, 4, pp 135-183.
- Norris D (1990) How to built a connectionnist idiot-savant ? *Cognition* 35, pp 277-291.
- Ohayon M (1990) *Rapport de psychiatrie - Intelligence artificielle et Psychiatrie*, tome 1. Paris: Masson.
- Perez JC (1989) *De nouvelles voies vers l'intelligence artificielle*. Paris: Masson.
- Pinkas G (1992) Represented Unrestricted First-Order Logic Formulas in Connectionist Networks. In : Proc. of the AAAI-92 Workshop "Integrating Neural and Symbolic Processes - The Cognitive Dimension", Sun R, Bookman LA (eds), San Jose, California, pp 23-30.
- Plugge LA, Verhey FRJ, Jolles J (1990) A desktop expert system for the differential diagnosis of dementia : an evaluation study, *International Journal of Technology Assessment in Health Care*, 6, pp 147-156.
- Pollack JB (1990) Recursive Distributed Representations. *Artificial Intelligence*, 46, pp 77-105.
- Portelli C, Frydman F, Mises R (1989) Intelligence artificielle en pédo-psychiatrie: JPSY, un système expert d'aide au diagnostic. In: M Ohayon, *Intelligence Artificielle et Psychiatrie*, Paris:Masson, pp 372-388.
- Proulx R, Bégin J (1990) A new learning algorithm for the BSB model. *Proceedings of the International Joint Conference on Neural Networks*, Vol. 1 : Neural and Cognitive Science Track, Hillsdale, NJ: Laurence Erlbaum, pp 704-706.
- Rialle V, Ohayon M (1991) Computers and Artificial Intelligence in Psychiatry: Brief History and State of the Art. In JH Nagel, WM Smith (eds), *13th An. Int. Conf. IEEE Engin. in Medicine and Biology Society*. 13(3), Piscataway, New Jersey: IEEE, pp 1280-1281.
- Rialle V (1992) Notes sur le connexionnisme, l'analyse de textes et le parallélisme. *Cahiers du Centre de Recherche en Cognition et Information ATO•CI*. Université du Québec à Montréal, 1(1).
- Rialle V., Ohayon M, Amy B, Bessière P (1991) Medical Knowledge Modeling in a Symbolic-Connectionist Perspective. In JH Nagel, WM Smith (eds), *13th An. Int. Conf. IEEE Engin. in Medicine and Biol. Society*. 13(3), Piscataway, New Jersey: IEEE, pp 1109-1110.
- Rosenblatt F (1961) *Principles of Neurodynamics: Perceptrons and the theory of Brain Mechanisms*. Washington D.C: Spartan.
- Rumelhart D E, McClelland, J L, the PDP Research Group (1986) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1: Foundations. Cambridge, Ma: MIT Press/Bradford Books.
- Seron X (1990) La neuropsychologie cognitive: Quelques réflexions sur ses fondements théoriques et méthodologiques. In: *Les modèles expérimentaux et la clinique psychiatrique*, Confrontations Psychiatriques, No 30.253-270. Paris: Editions médicales SPECIA.
- Servan-Schreiber D (1986) Artificial Intelligence and Psychiatry, *The Journal of Nervous and Mental Disease*, 174 (4), pp 191-202 .
- Sokolowski R (1992) Parallelism in Conscious Experience. *Daedalus*, 121(1), pp 87-104.
- Stip E, Lecours AR (1992): Fonctionnement neuropsychologique du déprimé, Epreuve de décision lexicale dans la dépression majeure. *L'Encéphale*, 18, pp 575-583.
- Sun R (1991) Integrating Rules and Connectionism for Robust Processing. PhD. Thesis, Brandeis University, Computer Science Dept, Waltham, MA.
- Sun R, Bookman LA (1992) Introduction. In : Proc. of the AAAI-92 Workshop "Integrating Neural and Symbolic Processes - The Cognitive Dimension", Sun R, Bookman LA (eds), San Jose, California, pp 3-7.
- Tiberghien G (1986) Psychologie cognitive, sciences cognitives et technologie de la connaissance. In: JL Le Moigne (ed), *Intelligence des mécanismes, Mécanismes de l'intelligence*. Paris: Fayard.
- Tiberghien G (1992) Questions de modélisation et de simulation cognitives. In: JF Le Ny (ed), *Intelligence Naturelle et Intelligence Artificielle*, Symposium de l'Association de Psychologie Scientifique de Langue Française (APSLF), Rome, 23-25 sept. 1991.
- Touretzky DS (1989) BoltzCONS : Dynamic Symbol Structures in a Connectionist Network. Tech. Rep. No. CMU-CS-89-172, Carnegie Mellon University.
- Turing AM (1983) *Pensée et Machine*. Traduction française, Paris: Edition du Champ du Vallon.
- Varela FJ (1989) *Connaître les sciences cognitives: tendances et perspectives*. Paris, France: Le Seuil (traduit de *Cognitive Science. A cartography of Current Ideas*, 1988).
- Volkow ND, Wolf AP (eds) (1991) *Positron-Emission Tomography in Schizophrenia Research*. American Psychiatric Press, Washington, USA.
- Walker SF (1990) A Brief History of Connectionism and its Psychological Implications. *Artificial Intelligence and Society*, 4(1), pp 17-38.
- Werner G (1987) Methuselah—An expert system for diagnosis in geriatric psychiatry, *Computers and Biomedical Research*, 20(5), pp 477-478.
- Winograd T (1987) *Thinking machines : can there be? are we?* Stanford Report CS-87-1161, Stanford University.

II. PUBLICATIONS

II.1. Revues nationales

1. BUSQUET G., DUBOIS F., GUERIN B., RIALLE V. & SOTTO J.J., Enquête sur les Travaux Dirigés au cours du Deuxième Cycle des Etudes Médicales. *Alpes-Médecine*, 1982, 2, 630-635.
2. RIALLE V. & SEIGNEURIN D., Logiciel d'Exploitation Informatique de la Correction des Q.R.O.C. et Q.C.M. - Application aux épreuves d'hématologie en D.C.E.M.3.. *Revue d'Education Médicale*, 6 (5), 1983.
3. RIALLE V. & VICENT C., Contribution à la Réalisation d'une Banque Nationale Informatisée de Q.R.O.C. et Q.C.M.. *Revue d'Education Médicale*, 6 (9), 1983.
4. RIALLE V., Logiciel d'Evaluation des Connaissances par Q.C.M.. *Revue d'Education Médicale*, 7 (3), 1984.
5. RIALLE V., DUBOIS F., BUSQUET G. & DEMONGEOT J., Une Expérience d'Introduction de l'Informatique en P.C.E.M.2. *Revue d'Education Médicale*, 7 (6), 1984.
6. RIALLE V., Analyse Informatique des Questions d'Examen ou Concours, de type Q.R.O.C. ou Q.C.M.. *Alpes Médecine*, 1985, V, 615-619.
7. RIALLE V., Ergonomie Cognitive et Bases de Connaissances. *Journal de l'Intelligence Artificielle*, 9, Paris, 1988, 4-9. (texte de la communication faite au congrès FORUM IA 88, Paris, 11-13 Septembre 1988).

II.2. Revues internationales

1. RIALLE V., VILA A. & BESNARD Y., Heterogeneous knowledge representation using a finite automaton and first order logic: a case study in electromyography. *Artificial Intelligence in Medicine*, 3 (2), 1991, 65-74.
2. Beltrami E. & Rialle V. (1993), La problématique du profil type des abuseurs sexuels. *Revue Sexologique*, 1(2), 55-78.
3. RIALLE V. & STIP E., Modélisation cognitive en psychiatrie : des modèles symboliques aux parallèles et distribués. *Journal of Psychiatry and Neuroscience*, 19(3), 1994. sous presse.
4. RIALLE V., STIP E. & O'CONNOR K., Computer Mediated Psychotherapy: Ethical issues and difficulties in Implementation. *Humane Medicine: A Journal of the Art and Science of Medicine*. 1994. sous presse.

II.3. Congrès nationaux et principales communications locales

1. RIALLE V., Système Expert de Diagnostic en Génétique Médicale. 3^o Journées Médico-Chirurgicales de Annaba (Algérie), 18 - 19 septembre 1985.
2. FOUILLET P., REYNES J.F., VILA A., REYMOND F. & RIALLE V., Protocole d'Etude Electrophysiologique des Neuropathies. *Actes des IVèmes Journées Francophones d'Electromyographie*, Nantes, 18-20 juin 1986.
3. RIALLE V., Ergonomie Cognitive et Bases de Connaissances. *FORUM IA '88*, Paris, 11-13 Septembre 1988.
4. AUGERAT P., BACONNIER P., FAUVAGE B., GRIMBERT F., MALEK H., PEYRIN J.C., RIALLE V. & VERMONT J., Aide au diagnostic d'urgence au moyen de l'intelligence artificielle. *Actes de la CONVENTION IA 88-89*, Paris, éditions Hermes, Paris, 575-590.
5. RIALLE V., PAGONIS D., VERMONT J., AUGERAT P. & GIRARDET P., Symbolic simulation of emergency care reasoning : toward an intelligent problem-solving assistance. *Actes des Journées Intelligence Artificielle en Simulation Numérique et Symbolique*, 22-23-24 mars 1989, Editions Aléas, Lyon, 1991.
6. RIALLE V., BESNARD Y. & VILA A., NEUROP : Un système expert clinique pour l'exploration neuro-musculaire et le diagnostic des neuropathies. *Actes des Journées d'Informatique Médicale de Toulouse*, Congrès F.N.A.I.M., Toulouse, 24-25-26 Mai 1989, 432-434.
7. RIALLE V., Systèmes à bases de connaissances en milieu médical: Problématique et pratiques. *Séminaires de sciences cognitives*, Pôle Régional Rhône-Alpes de Sciences Cognitives, 23 mars 1990 (mentionné dans Neurosciences, Informatique & Sciences cognitives - Rapport Final de Recherche, Responsable G. Tiberghien, Laboratoire de Psychologie Cognitive, Université Pierre Mendès France, Grenoble, p. 51).
8. RIALLE V., Problématique de l'aide à la décision médicale assistée par ordinateur: applications à l'électromyographie et à la réanimation post-chirurgicale. *Séminaire dans le cadre des Jeudis de Neurinfo*, 28 avril 1990, Laboratoire Neurinfo, Institut Méditerranéen de Technologie, Marseille.
9. RIALLE V., SAURA C., VILA A., BESNARD Y., REYMOND F. & ABAOUB L., NEUROP: aide au diagnostic électrophysiologique des neuropathies. *Actes du FORUM DE MICRO-INFORMATIQUE MEDICALE*, Faculté de Médecine de Toulouse, novembre 1990.
10. RIALLE V., VILA A. & BESNARD Y., Représentation hétérogène de connaissances : utilisation conjointe d'un automate fini et de la logique du premier ordre pour le diagnostic électrophysiologique des neuropathies. *Actes de la CONVENTION IA '91: 3^o conférence européenne sur les techniques et les applications de l'intelligence artificielle en milieu industriel et de service*, 15-17 Janvier 1991, Paris, éditions Hermes, 295-304.
11. CAVERNI J.P., PERIS J.L., RIALLE V. & VERONIQUE D., Formalisation des processus mentaux dans une activité d'expertise: l'évaluation des connaissances. *COGNISUD: Assises Régionales des Sciences de la Cognition*, C.N.R.S., Marseille, 24-25-26 janvier 1991.
12. VILA A., ZIEBELIN D., RIALLE V., REYMOND F., ABAOUB L., BESNARD Y., LE GOFF V., TAREL V. & FARRE M.T., Systèmes d'aide au diagnostic électromyographique

des affections neuromusculaires: quel système pour quel diagnostic? 3ème Journée Franco-romande de Neurologie, Grenoble, 17 mai 1991.

II.4. Congrès internationaux

1. RIALLE V., Génération de Connaissances Expertes - Utilisation de l'Analyse des Données. *Proceedings Fifth International Workshop on Expert Systems and Their Applications*, Avignon, 13-15 mai 1985, Ed. EC2, Nanterre, France, Vol II, 901-915.
2. RIALLE V., Du Langage Naturel Spécialisé au Système Expert. *Proceedings COGNITIVA 1985*, Paris, 3-7 juin 1985, 371-375.
3. VILA A., KRESS L., RIALLE V., ROBERT C. & ZIEBELIN D., Are expert systems an aid for diagnosing neuropathies ? *Electroencephalography and Clinical Neurophysiology* 66 (1987) S109, (reprint abstract, *VIII International Congress of Electromyography and Related Clinical Neurophysiology*, Sorrento, Italie, 24-29 mai 1987).
4. RIALLE V., Aide au diagnostic et l'apprentissage dans un domaine médical incertain, incomplet et évolutif. *Proc. IA-BIOMED 86, Première Conf. Int. sur Les Impacts de l'Intelligence Artificielle en Biologie et Santé*, 29 sept. - 2 oct. 1986, Montpellier, pp. 224 - 233. (Egalement dans: Pr. Joly (Ed.) : *Informatique et Médecine*, éditeur, EC2, Nanterre, France, 1988, 30-40).
5. ROBERT C., DUHAMEL A., RIALLE V. & VILA A., Using Clinical Datafiles in Building Expert Systems. International Conference on Artificial Intelligence and Cognitive Sciences, Grenoble, 17-20 février 1987. In: *Artificial Intelligence and Cognitive Sciences*, J. Demongeot, T. Hervé, V. Rialle et C. Roche (Eds), Manchester University Press, New York, 1988, 387-393.
6. KRESS L., RIALLE V. & VILA A., Methodology of knowledge base building for an expert system in electromyography. *Proceedings of the International Conference MEDICAL INFORMATICS: COMPUTERS IN CLINICAL MEDICINE* (Nottingham, 13 - 15 septembre 1988), British Medical Informatics Society (Eds), London, 1988, 83-88.
7. RIALLE V., Pour une intégration des systèmes experts en milieu clinique : quelques réflexions sur la problématique de l'utilisation des systèmes à bases de connaissances en pratique clinique courante. *Actes des Journées Francophones d'Informatique Médicale de Montpellier*, 14-16 mars 1989, édition de l'Ecole Nationale de la Santé Publique, Rennes, 1989, 279-285.
8. VILA A., REYMOND F. & RIALLE V., Evaluation clinique des techniques d'analyse automatique de l'EMG, *Actes des VIIèmes Journées d'EMG de Langue Française, et IIIèmes Journées Languedociennes d'Electromyographie*, 3-5 mai 1990, Montpellier, France, 55.
9. RIALLE V., Contribution à la définition du profil de cognaticien: caractéristiques fonctionnelles, savoir et savoir-faire dans l'élaboration d'une base de connaissances. *Proceedings du 2nd Congrès Européen Multi-média, Intelligence Artificielle et Formation*, 24-26 septembre 1990, Lille, France.
10. VILA A., REYMOND F. & RIALLE V., Diagnostic value of different computerized methods of automatic quantitative electromyography. *Muscle and Nerve*, 1990, 13 (10), 972, (reprint abstract *Proceedings of the 4th International Conference on Computerized and Quantitative EMG*, and : *9th International Single Fiber and Quantitative EMG Course*, 11-13 septembre 1990, Mainz, Allemagne Fédérale).

11. RIALLE V. & M. OHAYON M., Computers and Artificial Intelligence in Psychiatry: Brief History and State of the Art. In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1280-1281.
12. RIALLE V., M. OHAYON M., AMY B. & BESSIÈRE P., Medical Knowledge Modeling in a Symbolic-Connectionist Perspective. In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1109-1110.
13. BESNARD Y., RIALLE V., VILA A., REYMOND F., ABAOUB L. & DAHOU M., NEUROP: An expert system in electromyography based on a multilevel knowledge representation. In J.H. Nagel and W.M. Smith (Eds), *13th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 13, Part 3/5, 1991, IEEE Service Center, Piscataway, New Jersey, 1302-1303.
14. GIACOMETTI A., IORDANOVA I., AMY B., VILA A., REYMOND F., ABAOUB L., DAHOU M. & RIALLE V., A Hybrid Approach to Computer Aided Diagnosis in Electromyography. *Proc. 14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 14, 1992, 1012-1013.
15. IORDANOVA I., RIALLE V. & VILA A., Use of Unsupervised Neural Networks for Classification Tasks in Electromyography. *14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 14, 1992, 1014-1015.
16. IORDANOVA I., GIACOMETTI A., VILA A., AMY B., REYMOND F., ABAOUB L., DAHOU M. & RIALLE V., SHADE: A hybrid expert system in electromyography. *9th International Congress of Electromyography and Clinical Neurophysiology*, Jérusalem, Israël, 8-12 juin 1992.
17. RIALLE V., Modélisation Cognitive: Présentation, éléments d'analyse et discussion. *2nd Congrès Annuel de la Société Canadienne pour l'Etude du Cerveau, du Comportement et des Sciences Cognitives*, Université Laval, Québec, 13-14 juin 1992.
18. LANGEVIN E., BAIROCH A., APPEL R.D., RIALLE V., SCHERRER J.R. & HOCHSTRASSER D.F., ExPASy : Un système expert pour l'analyse des séquences protéiques. Quatrièmes Journées Francophones d'Informatique Médicale, Bruxelles, 3-4 juin 1993.

II.5. Posters de congrès

1. RIALLE V., Logiciel de Test de Connaissances au moyen de Questions à Choix Multiple. *Troisièmes Journées Universitaires Francophones de Pédagogie Médicale*, Lyon, 14-15 septembre 1984.
2. JOANNARD A. & RIALLE V., Aide au diagnostic, Génétique et Système Expert. *Journées de la Recherche Médicale*, Faculté de Médecine de Grenoble, avril 1986.
3. VILA A., FOUILLET P., REYNES J.F., REYMOND F., RIALLE V. & ROBERT C., Diagnostic électrophysiologique des neuropathies. *Journées de la Recherche Médicale*, Faculté

de Médecine de Grenoble, 3-4 Avril 1987. In: *Actualités 1986-1987*, Collection Grenoble Sciences, Université J. Fourier, Grenoble, Février 1988, 83.

4. AUGERAT P., BACONNIER P., DEMONGEOT J., FAUVAGE B., GRIMBERT F., MALEK H., PEYRIN J.C., RIALLE V., VERMONT J. & GELAS P., Système expert d'analyse des gaz du sang. *Journées de la Recherche Médicale*, Faculté de Médecine de Grenoble, 22-23 Avril 1988. In: *Actualités 1987-1988*, Collection Grenoble Sciences, Université J. Fourier, Grenoble, Nov. 1988, 137.
5. BESNARD Y., RIALLE V. & VILA A., NEUROP : Un système expert pour l'exploration électrophysiologique neuro-musculaire et le diagnostic des neuropathies. *Journées de la Recherche Médicale*, Faculté de Médecine de Grenoble, 28-29 Avril 1989. In: *Actualités 1988-1989*, Collection Grenoble Sciences, Université J. Fourier, Grenoble, Juillet 1990, 85.
6. AUGERAT P., BACONNIER P., DEMONGEOT J., FAUVAGE B., GRIMBERT F., MALEK H., PEYRIN J.C., RIALLE V., VERMONT J. & GELAS P., Un système expert pour l'interprétation de résultats de gaz du sang. *Journées de la Recherche Médicale*, Faculté de Médecine de Grenoble, 28-29 Avril 1989. In: *Actualités 1988-1989*, Collection Grenoble Sciences, Université J. Fourier, Grenoble, Juillet 1990, 100.
7. RIALLE V., VILA A. & BESNARD Y., Computer Aided Electrophysiological Diagnosis: Application to Neuropathies. *First Europeen Conference on Biomedical Engineering*, Nice, France, 17-20 février 1991.
8. RIALLE V., VILA A., SAURA C., BESNARD Y., REYMOND F. & ABAOUB L., NEUROP: aide au diagnostic électrophysiologique des neuropathies. *Journées de la Recherche Médicale*, Faculté de Médecine de Grenoble, 5-6 Avril 1991.

II.6. Mémoires, rapports, thèse

1. BIGALLET C. & RIALLE V., *L'espace communautaire*. Mémoire du Diplôme d'Etudes Approfondies de Prospective et Politique Scientifiques, Institut de Recherche Economique et de Planification (I.R.E.P), Université des Sciences Sociales Pierre Mendès-France - Grenoble, 1975.
2. RIALLE V., *Les périphériques audiovisuels: une assistance aux utilisateurs de systèmes informatiques*. Mémoire du Diplôme d'Etudes Approfondies de Mathématiques Appliquées aux Sciences Sociales, Institut de Mathématiques en Sciences Sociales (I.M.S.S), Université des Sciences Sociales Pierre Mendès-France - Grenoble, 1983.
3. CINQUIN P., RIALLE V. & DEMONGEOT J., *Propositions pour une structure d'imagerie médicale intégrée*. Rapport interne du Département de Mathématiques, Statistiques et Informatique Médicale, Faculté de Médecine de Grenoble, novembre 1984, 33 pages.
4. RIALLE V., *Aide au diagnostic et à l'apprentissage dans un domaine médical incertain, incomplet et évolutif - Etude des méthodes existantes et proposition d'une méthodologie nouvelle*. Thèse de Doctorat (nouveau régime) de Génie Biologique et Médical, Université Joseph Fourier, Grenoble, UFR de Médecine. Soutenue le 20 février 1987.
5. RIALLE V., Notes sur le connexionnisme, l'analyse de textes et le parallélisme. *Cahiers du Centre de Recherche en Cognition et Information ATO-CI*. Université du Québec à Montréal, 1(1), 1992.

6. RIALLE V., Rapport final du projet de recherche : Intelligence Artificielle en Psychiatrie. Présenté à la Région Rhône-Alpes. Octobre 1992.
7. RIALLE ., Rapport final de bourse d'excellence pour le séjour d'une année de recherche au Québec. Présenté à la Région Rhône-Alpes. Octobre 1992.
8. AMY B., GIACOMETTI A., ORSIER B., RIALLE V. et VILA A., *NORIA : Application des systèmes hybrides symboliques et connexionnistes à la modélisation de l'expertise*. Rapport d'activité, Action concertée "Sciences de la cognition 91", Ministère de l'Education Nationale et Ministère de la Recherche et de l'Espace. Juillet 1993, 114 pages.

II.7. Contributions à des ouvrages

1. RIALLE V., Aide au diagnostic et l'apprentissage dans un domaine médical incertain, incomplet et évolutif. In: *Informatique et Médecine*, éditeur EC2, Nanterre, France, 1988, 30-40.
2. ROBERT C., DUHAMEL A., RIALLE V. & VILA A., Using Clinical Datafiles in Building Expert Systems. In: *Artificial Intelligence and Cognitive Sciences*, J. Demongeot, T. Hervé, V. Rialle et C. Roche (Eds), Manchester University Press, New York, 1988, 387-393.
3. RIALLE V., Data Analysis as an aid to Learning and Knowledge Base Making in a medical field. In *Artificial Intelligence and Cognitive Sciences*, J. Demongeot, T. Hervé, V. Rialle et C. Roche (Eds), Manchester University Press, New York, 1988, 375-386.
4. RIALLE V., De la Représentation à l'Emergence: Décision et Cognition en Intelligence Artificielle Biomédicale. In: Dr H. Joly (éd.), *Biomédecine 2000*, Nanterre: EC2 (ou Masson), 1994. sous presse (18 pages).

II.8. Edition d'ouvrages

1. J. DEMONGEOT, T. HERVÉ, V. RIALLE ET C. ROCHE (Eds). *Artificial Intelligence and Cognitive Sciences*. Manchester University Press, New York, 1988.
2. RIALLE V. & PAYETTE D. (Eds.) (1994), *Modèles de la cognition : vers une science de l'esprit*. Numéro spécial de la série LEKTON. Montréal: Presses Universitaires du Québec. Sous presse.

II.9. Soumission en cours

ORSIER B., IORDANOVA I., RIALLE V., GIACOMETTI A. & VILA A., Hybrid systems for expertise modeling : from concepts to a medical application in electromyography. *Computers and Artificial Intelligence*, soumis.

III. PERSPECTIVES ET PROJETS DE DEVELOPPEMENT

Les développements que vont prendre dans les prochaines années mes activités de recherches s'inscrivent dans l'axe principal défini au chapitre I.

Si l'on inclut, comme il se doit, les travaux de Turing, Von Neumann, McCulloch et Pits, et certainement Wiener, l'intelligence artificielle est maintenant forte de plus de 50 années de recherche. Elle a conduit, particulièrement dans les quinze dernières années, à un nombre considérable de méthodes de raisonnement, de représentation de concepts et de savoirs, et d'apprentissage de connaissances à partir d'expériences. Mes travaux se situent dans une optique de développement et d'intégration de ces recherches.

Au confluent de l'informatique et de la médecine, ils participent :

- d'une part de la recherche pure, en particulier dans le domaine de la modélisation symbolique et connexionniste et de l'autonomie artificielle.
- d'autre part de la recherche appliquée par l'intégration au sein d'architectures logicielles cohérentes, de différentes méthodes d'intelligence artificielle au sens le plus large du terme, au bénéfice de la recherche médicale.

Les deux tendances complémentaires se retrouvent dans les divers projets qui vont être présentés, à des degrés variables selon les projets.

Ces projets, d'envergures différentes, revêtent une dimension locale ou internationale. Ceux de dimension locale seront présentés en premier, suivis des projets inscrits dans des coopérations internationales. Soit, dans l'ordre de présentation :

- Projet européen MIX - programme BRA ESPRIT-III de la CEE.
- Projet COGNIMED : intégration et évaluation de principes d'apprentissage et de construction automatique de connaissances (algorithmes génétiques, réseaux neuronaux, induction automatique...).
- Projets en coopération scientifique avec le Québec.

III.1. Projet MIX, ESPRIT-III

III.1.1 Présentation

Nous avons déposé en 1993 un projet de recherche sur le thème des systèmes hybrides symboliques et connexionnistes dans le cadre du Programme ESPRIT-III de la Communauté Economique Européenne. Ce projet intitulé MIX, pour *Modular Integration of Connectionist and Symbolic Processing in Knowledge-Based Systems*, a été accepté par la CEE. Il débute donc officiellement le 1er février 1994 et durera 3 ans.

Il s'agit d'un BRA (Basic Research Action) dont l'élaboration a été coordonnée par le Centre Universitaire d'Informatique de Genève (CUI, Pr. C. Pellegrini) en collaboration avec 5 autres partenaires européens: Le Centre de Recherche en Informatique de Nancy (INRIA-Lorraine, CRIN-CNRS, Pr. J.P. Haton), l'IMAG (Bernard Amy, LIFIA-IMAG, et Vincent Rialle, TIMC-IMAG), Kratzer Automatisierung (Dr. D. Butz, Germany), Fakultät Für Informatik (Munich, Pr. B. Fronhofer) et l'ETSI Telecommunicacion (Universidad Polytechnica de Madrid, Dr. J.C. Gonzalez).

Comme indiqué dans l'introduction, ce projet a pour objectif de clarifier l'impact théorique de l'intégration des méthodes symboliques et des méthodes connexionnistes d'apprentissage automatique. Il se propose à cet effet de produire un environnement logiciel modulaire (selon une orientation d'IA distribuée) d'intégration de ces méthodes et d'effectuer quelques applications dans le domaine de l'industrie, principalement, et de la médecine.

L'un des principaux domaines de validation des travaux sera constitué par les soins intensifs, avec une importante contribution de Grenoble. D'où la place centrale que devraient prendre, dans ce projet européen, le sujet de recherche constitué par le domaine de la toxicologie clinique présenté au point suivant.

Le thème du projet MIX relève des thèmes prioritaires 6.1 (Theories and models for the design of heterogeneous systems) et 6.2 (Basic aspects of multiple computing agents) de l'appel d'offre des Projet ESPRIT-III. Les deux paradigmes, IA symbolique et connexionnisme, englobent une large variété de principes théoriques et algorithmiques. L'IA symbolique recouvre notamment le *Raisonnement Fondé sur des Cas* (RFC) à propos duquel l'équipe de Grenoble a la responsabilité de l'intégration RFC-connexionnisme au sein du projet.

III.1.2. Contributions spécifiques de l'équipe grenobloise

Outre une contribution globale aux travaux de recherche prévus dans le projet, l'équipe grenobloise a deux contributions spécifiques dans ce projet : d'une part l'étude de l'intégration entre RFC et réseaux connexionnistes (étude dans laquelle intervient également l'équipe de Nancy), d'autre part une application médicale des travaux du consortium MIX (prévue durant la troisième année du projet).

Dans l'architecture modulaire d'intégration des paradigmes symbolique et connexionniste, le Raisonnement Fondé sur des Cas occupe donc une place importante.

Présentation succincte du RFC

La notion de RFC recouvre un ensemble de techniques de traitement de l'information dont le but commun est d'associer — par ressemblance, similarité de composition, de structure, de comportement, etc. — des exemples de problèmes, de cas, de situations diagnostiques, etc., qui alimentent quotidiennement l'activité de l'expert. Le RFC est une forme de raisonnement par analogie. Il se distingue néanmoins de la théorie proprement dite du raisonnement analogique en ce que les relations de similarités du RFC sont calculées au sein d'un même domaine, alors que le raisonnement par analogie établit des similarités structurelles et fonctionnelles entre domaines différents. Le RFC est considéré comme une méthode de résolution de problèmes fondé une mémoire sélective d'expériences bien assimilées, plutôt que sur une connaissance explicite et figée. L'inférence est alors conduite à partir de cas et non plus à partir de règles. Résoudre le problème revient donc premièrement à rechercher dans la mémoire le cas qui évoque le plus le cas en cours d'étude; deuxièmement à adapter, suivant un certain nombre de techniques, le cas sélectionné au cas en cours. Le problème de l'adaptation ne se pose de toute évidence que lorsque le cas sélectionné n'est pas suffisamment semblable au cas en cours. Le RFC s'appuie donc sur deux principes de base : a) le principe de similarité de cas; b) le principe d'adaptation de réponse. Ces deux principes dirigent l'architecture générale d'un système de RFC.

L'intérêt du RFC dans le domaine médical en particulier réside dans sa capacité à prendre en compte, dans un processus de décision, l'immense réservoir de connaissances et d'expériences que

représente en l'occurrence une base de dossiers médicaux. Il possède néanmoins ses propres limites (tout raisonnement ne saurait notamment être ramené à des calculs de proximité entre cas), mais n'en constitue pas moins un paradigme intéressant qui suscite depuis quelques années des travaux de recherche novateurs.

RFC et connexionnisme à Grenoble

Le raisonnement fondé sur des cas est depuis plusieurs mois étudié à Grenoble dans le cadre des systèmes hybrides symboliques et connexionnistes. C'est dans le domaine de l'électromyographie qu'une première réflexion a été conduite : un projet de DEA d'informatique, réalisé par Melle Maria Malek et intitulé "Conception d'un Système de Raisonnement Par Analogie Appliqué au Diagnostic Electro-physiologique des Neuropathies" a été conduit sous la direction de V. Rialle et en collaboration avec B. Amy, en 1992-1993. Dans ce travail, Melle Malek a analysé la problématique de l'intégration complète de la base de connaissance du diagnostic électrophysiologique des neuropathies (contenue dans le système NEURO, cf. I.2.4, I.2.19 et I.3.2.) dans un système hybride, utilisant à la fois des représentations symboliques, des mécanismes connexionnistes et un raisonnement par cas. Elle a notamment proposé une mesure de ressemblance entre cas sous contrainte de variation considérable du nombre de variables électrophysiologiques (le choix et le nombre de nerfs explorés pouvant varier considérablement d'un malade à l'autre). Melle Malek poursuit actuellement ces recherches sur l'intégration du RFC et des réseaux neuronaux dans le cadre d'une thèse d'informatique en relation avec le projet MIX.

Application médicale

Un projet de thèse de Génie Biologique et Médical, consistant en la conception et la réalisation d'un SYSTÈME D'AIDE AU DIAGNOSTIC DES COMAS TOXIQUES CHEZ L'ADULTE a été formulé fin 1993 par le Docteur Vincent Danel, responsable du service de toxicologie clinique du CHU de Grenoble.

Ce projet est à visée diagnostique et pédagogique et met en œuvre une base de données exploitable selon deux démarches distinctes :

- une première démarche qui consiste à *produire* une connaissance — par analyse de la base et opérations d'induction de connaissances, d'apprentissage, de généralisation, etc. De nombreuses méthodes peuvent concourir à produire ce type de connaissances. Ces méthodes empruntent à la statistique, à l'analyse des données, au connexionnisme et à l'apprentissage automatique (*Machine learning*).

- une démarche qui consiste à utiliser directement la base pour y pratiquer un raisonnement fondé sur les cas, selon les principes du RFC évoqués précédemment.

Le système à concevoir devra mettre en œuvre ces différentes méthodes et combiner les connaissances propres du médecin avec les connaissances issues de la base de dossiers médicaux.

Ce domaine de recherche devrait constituer l'application médicale du projet MIX dans sa troisième année.

III.2. Projet COGNIMED

COGNIMED est un projet soutenu par la REGION RHONE-ALPES dans le cadre de son programme pluriannuel "Intelligence Artificielle, Systèmes Experts et Applications", appel d'offre 1993 (titre complet du projet : "Théorie des classifieurs et Expérimentations : le projet COGNIMED"). Il a pour principal objectif d'intégrer au sein d'une même structure logicielle la

plupart des modèles de classification automatique et d'apprentissage de règles et d'arbre de décision et se trouve en cela en *synergie* avec le projet MIX précédemment exposé. Il a pour but à moyen terme de mettre au service du corps médical du CHU un certain nombre d'outils décisionnels issus de quelques avancées pertinentes en matière d'intelligence artificielle au sens le plus large du terme. Ces outils, théoriques et pratiques, ont pour point commun de s'appuyer sur des ensembles de données (appelés *Bases de cas* dans le contexte du raisonnement par cas, et *base d'exemples* ou *bases d'apprentissage* dans les contextes du connexionnisme ou de l'induction automatique de connaissances). Ces données sont en l'occurrence constituées par des fichiers-patients anonymes. L'un des objectifs particulier du système sera d'extraire de ces fichiers une connaissance susceptible d'être utilisée par le médecin, dans ses tâches décisionnelles, ou encore par un système à base de connaissances.

L'objectif concret du projet est d'élaborer une "plate-forme" logicielle permettant l'utilisation des meilleures méthodes de production de connaissances classificatoires et de modèles décisionnels. Ces méthodes incluent notamment l'induction automatique d'arbres ou de règles de décision (type ID3, CART, IN2, classifieur bayésien, ...), les réseaux de neurones formels (perceptron multicouche, réseau de Kohonen, réseaux hybrides...), les algorithmes génétiques (travaux de J. Holland, système NEWBOOLE,...) et le raisonnement par cas. Ce projet complète mais ne remplace pas les méthodes d'analyse des données, telles que l'analyse factorielle discriminante ou l'analyse en composantes principales, disponibles dans de nombreux logiciels (il intégrera néanmoins, si possible l'algorithme des nuées dynamiques et l'algorithme COBWEB).

Cette plate-forme permettra notamment le choix entre divers principes de similarité entre cas, proposera divers algorithmes d'intelligence artificielle pour l'analyse décisionnelle et l'apprentissage automatique de comportements diagnostiques à partir d'exemples, et offrira des facilités pour l'évaluation des méthodes et leur comparaison.

Le logiciel sera conçu selon une **programmation par objets**. Il sera modulaire, de type boîte à outils évoluée, avec facilités d'ajouts de modules (*i.e.* des "classifieurs automatiques", des principes de similarité,...), présentation graphique des résultats et interface de communication convivial.

Cette plate-forme logicielle sera conçue pour s'insérer, dès que les conditions s'y prêteront, dans le Système d'Information Hospitalier Intégré de l'hôpital A. Michallon afin de s'ajouter aux outils logistiques qui seront mis au service du corps médical. A cette fin, tous les programmes seront réalisés en C et transportables sur station de travail DEC. La machine massivement parallèle MASPARE sera utilisée dans les divers traitements connexionnistes envisagés sur les bases de données.

L'une des expérimentations majeures de COGNIMED se fera dans le cadre du projet de coopération franco-québécois développé au point III.6.

En pratique :

Les premiers jalons de ce projet ont été posés par une étude fonctionnelle et une première maquette réalisées par 4 étudiants du DESS Informatique Double Compétence de l'Université Joseph Fourier (BOURAS Hassan, DALLA-VALE Laurent, GAULD Gérard, MEARY Stéphane, voir Annexe 1). Il constitue actuellement le sujet de travail de M. E. Duriez (ingénieur, titulaire du DEA de sciences cognitives, voir Annexe 1) dans le cadre de son service national civil.

III.3. Projets franco-québécois

III.3.1 Textes et réseaux neuronaux

Il s'agit, comme introduit dans les parties I.1, I.2.16 et I.3.5, d'une coopération scientifique franco-québécoise avec le Centre de Recherche en Cognition et Information ATO-CI, de l'Université du Québec à Montréal - UQAM, dirigé par le Professeur Jean-Guy Meunier.

Thème des travaux : traitement parallèle et distribué des grands fichiers de textes.

Contexte : Ce projet constitue la poursuite d'une action de coopération (mentionnée en I.3.3) entre le laboratoire TIMC-IMAG et le Centre de Recherche en Cognition et Information ATO-CI. Cette coopération a été initiée l'année dernière par le séjour de longue durée que j'ai effectué à Montréal en qualité de chercheur invité par le partenaire québécois et dans le cadre d'une *délégation*. Ce séjour a été l'occasion d'établir des liens solides centrés sur des intérêts de recherche convergents et complémentaires (voir l'ouvrage : RIALLE V., Notes sur le connexionnisme, l'analyse de textes et le parallélisme. *Cahiers du Centre de Recherche en Cognition et Information ATO-CI*. Université du Québec à Montréal, 1(1), 1992, 55 pages)).

Par ailleurs, le Professeur Meunier conduit une action d'envergure internationale pour le développement de la recherche en matière de traitement textuel.

Problématique générale : Les intérêts communs mentionnés précédemment s'enracinent dans l'idée qu'une modélisation de processus de raisonnements et de décisions est étroitement liée à une modélisation de la pensée sous sa forme linguistique. Cependant, cette forme linguistique reste encore largement à définir. On constate en effet que la linguistique computationnelle, telle qu'elle a été pratiquée durant ces dernières décennies, n'a pas débouché sur un modèle opérationnel dans des applications de taille ou d'enjeux importants. Or l'utilisation du langage, particulièrement sous sa forme écrite, pose des problèmes dont on mesure chaque jour un peu plus l'importance. Il s'agit par exemple de la surcharge informationnelle des organisations, du besoin de recherche rapide d'informations, de l'analyse de textes longs, de la traduction automatique, pour n'en citer que quelques uns. Problèmes qui constituent autant de thèmes de recherche pour l'intelligence artificielle. On sait de plus que la médecine, dans ses formes actuelles d'organisation institutionnelle (Centres hospitaliers en particulier) traite des volumes d'information toujours plus importants dans lesquels l'usage du texte libre reste une nécessité.

Orientation de recherche : L'orientation générale des recherches que je conduis dans ce domaine s'inscrit dans la lignée de la sémiotique cognitive développée par F. Rastier (auteur de *Sémantique et recherche cognitive*, PUF, 1991) et de la conception de la linguistique développée par P. Jorion (auteur de *Principes des systèmes intelligents*, Masson, 1990; et de *Aristote aurait-il pu faire la machine pensante ?*, preprint). Il s'agit, pour ce dernier auteur, d'effectuer une sorte de retour aux sources de la pensée aristotélicienne, qui prévalait avant que la linguistique ne consacre des séparations très marquées entre syntaxe et sémantique. L'orientation de modélisation encouragée par P. Jorion s'appuie sur le principe de subsomption selon lequel le sens émerge dès l'association de deux termes et se construit ainsi de proche en proche. L'analyse syntaxique, vouée à l'échec dans le cas des langages naturels, n'apparaît donc plus comme le passage obligé entre le vocabulaire et le sens.

Objectifs pratiques et moyens: Le volume de documents textuels manipulé par les grandes organisations (dont certaines organisations de Santé) tend à dépasser les capacités en personnel et les délais socialement admissibles de ces organisations. Or, des méthodes automatisées de lecture de textes électroniques, de recherches de concepts, de "fouille" systématique, de résumé automatique, d'extraction de connaissances, etc., émergent peu à peu des recherches en Intelligence Artificielle, Sciences Cognitives, Connexionnisme et Linguistique Computationnelle.

D'un point de vue pratique en matière de traitement textuel on constate actuellement :

- d'une part l'existence de logiciels déjà assez anciens (capables d'effectuer des tâches classiques d'analyse syntaxique, d'indexation automatique, analyse statistique de contenu, etc.) et limités (dans leurs capacités propres, dans la taille des documents qu'ils peuvent analyser, etc.);
- d'autre part l'arrivée de méthodes théoriques et de technologies (ordinateurs parallèles, disques lasers, etc.) avancées, potentiellement capables de répondre aux demandes croissantes et parfois urgentes de traitement intelligemment assisté de l'information.

Le projet de *Gestion Intelligemment Assistée de Documents*, conduit par le Pr. J.-G. Meunier, a pour objectif de répondre à ce besoin de méthodes et d'outils de traitement textuel (une part du projet, qui ne nous intéresse pas directement ici, concerne également la fouille de documents graphiques).

Ce projet, multi-national et réunissant divers partenaires institutionnels et commerciaux, a récemment reçu l'agrément du Conseil pour la Recherche en Sciences Humaine (CRSH) du Canada, avec un important budget.

L'aspect du projet dans lequel le laboratoire TIMC-IMAG intervient est l'application des méthodes avancées de modélisation des raisonnements et des processus cognitifs que nous développons, pour l'analyse et la gestion des documents textuels dans les grandes organisations (en particulier dans le domaine de la santé).

L'objectif spécifique à la partie grenobloise est double :

- sur le plan théorique, il consiste à expérimenter un certain nombre de réseaux connexionnistes sur des textes.

Une plate-forme logicielle permettant l'implémentation rapide et l'expérimentation de modèles connexionnistes sur la machine "massivement parallèle" MASPARE est en cours de développement. Les perspectives concernent l'utilisation de machines massivement parallèles pour le traitement de volumes très considérables de données textuelles dans des laps de temps raisonnables, à l'aide de méthodes informatiques issues des récentes avancées en matière de parallélisme et de connexionnisme.

Les premiers jalons grenoblois de ce projet sont constitués par le travail de recherche de M. O. Nérot que j'encadre cette année dans le DEA de Sciences Cognitives de l'Université Joseph Fourier, et dont le titre est : Apprentissage sémantique cognitif dans les bases de données textuelles (travail sur MASPARE).

Par ailleurs, une demande de subvention a été adressée en février 1993 au Centre Jacques Cartier de Lyon en réponse à son 8ème appel d'offre.

III.3.2 Applications en psychiatrie

Il s'agit, comme introduit en I.2.17, d'une coopération franco-québécoise avec le Professeur Edouard Beltrami, psychiatre au Département de Sexologie de l'Université du Québec à Montréal et chercheur au Centre ATO-CI (précédemment mentionné). Ici encore, cette coopération a été initiée lors de mon séjour de longue durée au Québec.

Thème des travaux : Classification et recherche de profils types d'abuseurs sexuels par réseaux de neurones formels.

Résumé : La typologie de l'abuseur sexuel répond à un besoin social important pour les cours de justice et afin de permettre un diagnostic adéquat. On constate que les classifications psychiatriques traditionnelles, codifiées par exemple dans le DSM-III-R, permettent un diagnostic scientifique valide en cour de justice. Par contre, tous les efforts de produire un ou des profils d'abuseur ont jusqu'à présent échoué. Nous postulons que les biais théoriques, voire certains préjugés culturels, ont masqué quelques éléments cruciaux pour le diagnostic psychiatrique dans ce domaine. Une procédure informatique non entachée de subjectivité pourrait apporter des éléments déterminants dans ce domaine sensible du diagnostic. Cette recherche est donc orientée vers la construction d'un système qui pourrait inférer lui-même des catégories ou des profils scientifiquement valides à partir des données descriptives de sujets témoins et d'abuseurs avoués. Une base de données déjà existante et représentative des abuseurs sexuels est en cours de perfectionnement. Les sujets abuseurs proviennent de plusieurs thérapeutes traitant en externe, dans plusieurs villes du Québec qui nous ont déjà donné leur collaboration. Nous prenons comme témoins des dysfonctionnels sexuels qui n'ont jamais abusé sexuellement, mais qui ont été soumis aux mêmes tests et dans les mêmes conditions (les études dans ce domaine, avec des cas témoins sont très rares).

Les travaux en cours et en projet consistent à explorer les capacités de classification de quelques méthodes, incluant les réseaux de neurones formels et divers autres algorithmes, afin de produire un modèle valide de classification. La carte topographique fournie par le réseau neuronal auto-organisant de Kohonen est en cours d'expérimentation pour la résolution de notre problème de classification.

Un article vient d'être produit sur les premiers résultats obtenus. Il s'agit de :

Beltrami E. & Rialle V. (1993), La problématique du profil type des abuseurs sexuels. *Revue Sexologique*, 1(2), 55-78.

En pratique :

Une demande de subvention est actuellement adressée par le Docteur Beltrami au Fond pour la Recherche en Santé du Québec.

Un certain nombre d'expérimentations de classifieurs seront entreprises dès la rentrée universitaire 1993. Ces expérimentations feront appel aux outils développés dans le cadre du projet COGNIMED.

IV. CONCLUSION

Ce rapport a tenté de présenter l'ensemble de mes travaux de recherche, en relation avec l'enseignement et l'encadrement d'étudiants.

Le fil conducteur des divers champs d'investigation qui ont été présentés est la modélisation "computationnelle" du raisonnement et de la décision.

Plusieurs axes de progression peuvent être observés, parmi lesquels:

- un axe d'intégration de méthodes historiquement extérieures au domaine de l'intelligence artificielle (tel que l'on concevait ce domaine il y a encore quelques années, et que l'on nomme quelquefois *I.A. classique* ou *IA symbolique*) dans le but de réaliser concrètement des systèmes capables d'aider un médecin dans ses tâches décisionnelles. Cette intégration passait notamment, dès les premières années de recherche, par l'analyse des données et l'apprentissage automatique de connaissances (thème de ma thèse, 1987);
- un axe de progression depuis l'étude de méthodes très formelles développées par l'I.A. classique, vers l'étude et l'application de méthodes réputées plus *neurologiquement* et *psychologiquement* vraisemblables en matière de modélisation du raisonnement décisionnel humain; Il s'agit notamment des méthodes fondées sur les réseaux de neurones formels.
- un axe d'intégration des diverses dimensions de l'intelligence artificielle, incluant le domaine de l'aide à la décision et celui de la modélisation de la pensée sous sa forme linguistique.
- un axe de progression depuis le travail de développement au niveau de la machine (système expert NEUROPOP par exemple) jusqu'au travail de réflexion sur les tenants et aboutissants de la modélisation des processus cognitifs (relations avec le Département de philosophie de l'UQAM notamment);
- un axe de progression depuis le travail individuel jusqu'à l'encadrement d'un nombre en progression régulière d'étudiants de troisième cycle;
- un axe de progression vers des coopérations de recherche inter-équipes d'une part (TIMC-LIFIA), internationales d'autre part (TIMAC-UQAM et TIMC-consortium MIX).

La recherche, l'enseignement et la direction d'étudiants dans les domaines qui ont été présentés dans ce rapport, constituent donc le cœur de mes activités présentes et à venir dans l'Université et le champ de l'informatique.

ANNEXE - Activité d'encadrement d'étudiants de 3ème cycle

A. Présentation générale

L'activité d'encadrement d'étudiants de troisième cycle a débuté dès 1983. Au total, 33 étudiants de troisième cycle ont été encadrés, dont :

- 19 projets de DESS Informatique Double Compétence (Université J. Fourier),
- 3 projets de DEA de Génie Biologique et Médical (Universités J. Fourier et Lyon I)
- 2 projets de DEA de Sciences Cognitives (Université J. Fourier),
- 4 étudiants répartis dans 2 projets de troisième année ingénieur ENSIMAG (Institut National Polytechnique de Grenoble),
- 1 projet de DEA d'Informatique (Université J. Fourier),
- 1 projet de DESS Informatique et Applications aux Sciences de la vie (Université Paris V),
- 3 thèses de médecine consacrée à la réalisation de systèmes d'aide à la décision — en génétique médicale et électromyographie — (Faculté de Médecine de Grenoble, Université J. Fourier).

Il faut mentionner également l'encadrement de 6 stages d'été, en fin de cursus de troisième cycle.

Actuellement, cette activité d'encadrement se poursuit par le suivi de deux étudiants en thèse :

Maria Malek, thèse d'Informatique.

Dr. Vincent Danel, thèse de Génie Biologique et Médical.

B. Liste des étudiants

CHALANSONNEX Alice, FORTIN Pierre, LUNG Lai-man, MENOUD Michel, PRETAT Virginie, RECORD François, SAIMMAINE Isabelle, TESTE Laurent, VINTROU Jean-Paul, VOIRON Christine, YAN Jing-Wu,

Les 11 étudiants précédents étaient répartis en cinq groupes spécialisés chacun dans une partie de la conception et réalisation d'un gros système d'aide à la décision en génétique médicale, comprenant : une partie base de données, une partie interface en langage naturel et une partie aide à la décision.

Mémoires de projet du DESS Informatique Double Compétence, Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble, 1985.

KRESS Lionel

"Choix de systèmes et méthodologie de construction d'une base de connaissances - Application à l'Electromyographie",

Mémoire de Diplôme d'Etudes Approfondies de Génie Biologique et Médical, année 1986-1987.

BUCHARLES Christophe

"MENINGE: un système expert utilisant des objets et des règles de production - Possibilités de traitement automatique de cas à partir d'une base de données clinique",
Mémoire de Diplôme d'Etudes Approfondies de Génie Biologique et Médical, année 1986-1987.

BAUDINO Nathalie et GOMES Christine

"Contribution à l'élaboration d'un système informatique d'aide au diagnostic des syndromes malformatifs — Réalisation d'une base de données à partir de syndromes comprenant une fente labiale et/ou palatine", Thèse de médecine, Faculté de Médecine de Grenoble, soutenue le 30 mai 1985.

GARCIA Constanza, GODEL Christian, GOARDOU Christian, VARIGNY François

"Contribution à la réalisation d'un serveur de résultats d'analyse et d'un système expert pour l'analyse des gaz du sang"

Mémoire de projet du DESS Informatique Double Compétence, Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble, 1988.

BESNARD Yves

"Réalisation d'un Système-Expert en Electromyographie avec le générateur NEXPERT"

Mémoire de stage-projet du DESS Informatique et Applications aux Sciences de la vie
Université Paris V, lieu du stage : Faculté de Médecine de Grenoble, 1988.

CADDOUX-EUDES Monique

"Système expert d'aide au diagnostic des troubles de la transmission neuro-musculaire"
thèse de médecine soutenue le 20 décembre 1991.

Jury : Pr J. Perret, Pr J. Demongeot, Pr R. Sarrazin, Dr V. Rialle, Dr A. Vila

Mention *Très honorable*, avec proposition d'échange avec d'autres universités.

Programmation entièrement en Prolog avec interface utilisateur convivial, sur ordinateur Macintosh.

DAHOU Mostéfa

"Représentation de connaissances en logique du premier ordre et évaluation d'un système à base de connaissances pour le diagnostic électrophysiologique des neuropathies"

Mémoire de DEA de Génie Biologique et Médical. Année 1991-92.

MALEK Maria

Projet du D.E.A d'Informatique de l'Université Joseph Fourier, année 1992-1993

Sujet : Conception d'un système de raisonnement par analogie appliqué au diagnostic électrophysiologique des neuropathies

Stage effectué au laboratoire IMAG-TIMB,

en collaboration avec le laboratoire IMAG-LIFIA (B. Amy)

Actuellement en thèse d'Informatique sur les systèmes hybrides symbolique et connexionnistes (cf. III.1)

NERO Olivier

Projet du D.E.A. de Sciences Cognitives de Grenoble, année 1992-1993

Sujet : Apprentissage sémantique cognitif dans les bases de données textuelles

Stage effectué au laboratoire IMAG-TIMB, en coopération scientifique avec le Centre de Recherche en Cognition et Information ATO-CI, Université du Québec à Montréal

(Pr. J. G. Meunier).

DURIEZ Emmanuel

Projet du D.E.A. de Sciences Cognitives de Grenoble, année 1992-1993

Sujet : Systèmes hybrides et modélisation du rappel et de la reconnaissance en psychologie cognitive

Stage effectué au laboratoire IMAG-TIMB, en collaboration avec le laboratoire IMAG-LIFIA (B. Amy) et le Laboratoire de Psychologie Expérimentale (Pr. G. Tiberghien).

BOURAS Hassan, DALLA-VALE Laurent, GAULD Gérard, MEARY Stéphane
Projet du DESS Informatique Double Compétence, Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble, 1992.

DANEL Vincent

Thèse de Génie Biologique et Médical

Sujet : Aide au diagnostic des comas toxiques chez l'adulte
1993-...

C. Stages d'été

1985 VIAL André

Logiciel d'aide au diagnostic des maladies génétiques : Conception, réalisation et maintenance

de la base de données associée.

Mémoire de stage de DESS Informatique Double Compétence,

Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble.

1988 CHAIX Corinne

Contribution à l'achèvement de la première version d'un système expert d'analyse des gaz du sang

Mémoire de stage de DESS Informatique Double Compétence,

Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble.

Utilisation du générateur de systèmes expert NEXPERT, sur Macintosh.

1988 PERRIN Catherine

Contribution à la réalisation d'un système expert en Electromyographie

Mémoire de stage de DESS Informatique Double Compétence,

Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble.

Utilisation du générateur de systèmes expert NEXPERT, sur Macintosh.

1989 BOULE Bernard

Contribution à la réalisation d'un système à base de connaissance en Prolog pour l'analyse des gaz du sang et le conseil d'urgence en réanimation.

Mémoire de stage de DESS Informatique Double Compétence,

Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble.

Programmation en Prolog, sur Macintosh

1990 SAURA Claude

Contribution à la réalisation du système expert NEUROP pour le diagnostic des neuropathies

Mémoire de stage de DESS Informatique Double Compétence

Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble.

Programmation en Prolog, sur Macintosh

1991 GACON Véronique

Contribution à la mise en œuvre opérationnelle du système expert NEUROP.

Mémoire de stage de DESS Informatique Double Compétence

Université J. Fourier, Institut Grenoblois d'Etudes Informatiques, UFR IMA - Grenoble.

Programmation en Prolog, sur Macintosh

1993 MALEK Maria

Contribution à la mise au point du système NEUROP dans le cadre du projet ESTEEM

Résumé

La thématique centrale de ce mémoire concerne l'Intelligence Artificielle appliquée au diagnostic Médical (IAM). Son orientation repose sur des concepts d'apprentissage automatique de décisions plus que de connaissances pré-définies et figées, d'adaptation à de nouvelles situations plus que de comportements répétitifs et stéréotypés. Au goulot d'étranglement que constitue la construction d'une base de connaissances déclaratives, les systèmes préconisés opposent une utilisation intensive des *bases de cas*, véritables prolongements sur support informatique de la mémoire à long terme du médecin, et conduisent à l'*émergence* de capacités classificatoires automatisées au moyen d'algorithmes appropriés.

L'entreprise consistant à programmer une machine afin qu'elle produise des raisonnements habituellement attendus d'un spécialiste met au premier plan quelques difficiles questions relatives au *pourquoi* et au *comment* de tels systèmes. Le constat sinon d'échec du moins d'immenses difficultés de mise en œuvre des systèmes à bases de connaissances déclaratives, ouvrent la porte à un retour en force de l'*expérience* et de la mémoire cumulée de l'activité décisionnelle du praticien au cours des mois et des années de pratique. Dans cette optique, l'expérience emmagasinée sous forme de bases de cas dans la mémoire de l'ordinateur prend le pas sur la connaissance figée et laborieusement élaborée dans une base de connaissances pour la construction d'un système d'IAM. Des systèmes *hybrides* — permettant d'associer des connaissances déclarées par le spécialiste et des connaissances apprises automatiquement — constituent en quelque sorte un idéal que l'on s'efforce d'atteindre, notamment dans notre projet ESPRIT-III : MIX.

L'introduction de ce mémoire tente de préciser ce passage de la représentation à l'émergence de connaissances qui consacre en quelque sorte l'immersion de l'IAM dans les sciences de la cognition (et qui correspond en gros à mon parcours de chercheur depuis une douzaine d'années). Divers aspects de la problématique générale de la construction de classifieurs y sont abordés et une présentation succincte des diverses approches de l'émergence est proposée (connexionnisme, algorithmique génétique, induction...) et illustrée par le projet COGNIMED.

Autour et parfois en marge de cette problématique centrale, s'ordonnent un certain nombre de travaux que j'ai pu conduire ces dernières années dans les domaines de la psychiatrie-psychologie, de l'analyse textuelle et de la "philosophie de l'esprit". Ces travaux sont également évoqués.

Le mémoire est structuré en quatre parties principales et une annexe :

- Une partie introductive présentant de manière synthétique la thématique générale de recherche ainsi qu'un bref état de l'art du domaine dans lequel se placent mes travaux. La présentation des problèmes et des orientations de ce domaine sera appuyée par une bibliographie propre à l'introduction.
- La première partie est consacrée à la description des recherches qui ont été effectuées depuis une dizaine d'années. Outre la description des thèmes de recherche, cette partie inclut la présentation de quelques articles et résumés de travaux.
- La deuxième partie donne une liste complète et structurée des publications, communications, posters, etc.
- La troisième partie, orientée vers le futur, est consacrée aux projets imminents et aux perspectives de recherche et de développement à moyen terme.
- L'annexe présente l'activité d'encadrement d'étudiants de troisième cycle, directement liée aux activités d'enseignement et de recherche.

Mots-clés : Médecine, intelligence artificielle, systèmes experts, apprentissage automatique, connexionnisme, systèmes hybrides, modélisation psycho- et neuro-cognitive.