



Une forme d'expression du diagnostic : contribution à l'étude ergonomique de la gestion automatique du dossier médical

Elisabeth Defour

► To cite this version:

Elisabeth Defour. Une forme d'expression du diagnostic : contribution à l'étude ergonomique de la gestion automatique du dossier médical. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1970. Français. NNT : . tel-00282249

HAL Id: tel-00282249

<https://theses.hal.science/tel-00282249>

Submitted on 27 May 2008

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.

FACULTE DE MEDECINE
ET DE PHARMACIE DE GRENOBLE

Année 1970

N° D'ordre :

UNE FORME D'EXPRESSION DU DIAGNOSTIC :
CONTRIBUTION A L'ETUDE ERGONOMIQUE
DE LA GESTION AUTOMATIQUE DU DOSSIER MEDICAL

T H E S E

présentée

à la Faculté de Médecine et de Pharmacie
de Grenoble
pour obtenir

le grade de DOCTEUR en MEDECINE

(Diplôme d'Etat)

par

*Madame DEFOUR Elisabeth née PECCOUD
Externe des Hôpitaux de LYON
Née le 5 Juin 1944
A CHAMBERY (73)*

Cette thèse sera soutenue publiquement le 27 JUIN 1970, devant :
Monsieur le Professeur G. CABANEL, Président du Jury et Messieurs les Professeurs
J. KUNTZMANN, J. BONNET-EYMARD, M. GEINDRE
Monsieur le Docteur J. VALOIS.

FACULTE DE MEDECINE
ET DE PHARMACIE DE GRENOBLE

Année 1970

N° D'ordre :

UNE FORME D'EXPRESSION DU DIAGNOSTIC :
CONTRIBUTION A L'ETUDE ERGONOMIQUE
DE LA GESTION AUTOMATIQUE DU DOSSIER MEDICAL

T H E S E

présentée

à la Faculté de Médecine et de Pharmacie
de Grenoble
pour obtenir

le grade de DOCTEUR en MEDECINE

(Diplôme d'Etat)

par

*Madame DEFOUR Elisabeth née PECCOUD
Externe des Hôpitaux de LYON
Née le 5 Juin 1944
A CHAMBERY (73)*

Cette thèse sera soutenue publiquement le 27 JUIN 1970, devant :
Monsieur le Professeur G. CABANEL, Président du Jury et Messieurs les Professeurs
J. KUNTZMANN, J. BONNET-EYMARD, M. GEINDRE
Monsieur le Docteur J. VALOIS.

TABLE DES MATIERES

<u>Chapitre I : 1'EXPRESSION DU DIAGNOSTIC</u>	8
I - ANALYSE DU DIAGNOSTIC	9
1. Les éléments constitutifs du diagnostic	9
2. La structure du diagnostic	12
II - GRAMMAIRE CONTEXT-FREE D'EXPRESSION D'UN DIAGNOSTIC	13
1. Abrégé d'une définition de grammaire pour la compréhension de ce qui suit	13
2. Définition d'une grammaire formelle de description du diagnostic	16
III - DESCRIPTION DU SYSTEME	22
1. Mode "Définition de données"	23
2. Mode "manipulation de données"	26
3. Mode "récupération de données"	28
4. Quelques détails techniques	31
IV - STRUCTURE-TYPE DES DONNEES POUR LE DIAGNOSTIC	36
 <u>Chapitre II : <u>ERGONOMIE DE L'UTILISATION D'UN SYSTEME INFORMATIQUE</u></u>	
<u>DE GESTION DES DOSSIERS MEDICAUX</u>	39
I - CARACTERES GENERAUX D'UN SYSTEME AUTOMATISE DE GESTION D'UN FICHIER MEDICAL	40
1. Fonctions	40
2. Interrogation du fichier	42
3. L'interrogation des données	44

II - ETUDE DE DIFFERENTS SYSTEMES.....	45
1. Système de textes	45
2. Mots clés	46
3. Codes	47
4. Les questionnaires	48
5. Le langage normalisé	49
 CHAPITRE III : <u>CARACTERISTIQUES ERGONOMIQUES DU SYSTEME DEFINI A L'IMAG</u> .	52
 I - STRUCTURE ACTUELLE	58
1. Définition de données	58
2. Création de données	58
3. Interrogation des données	62
4. Conclusions	62
 II - EVOLUTION DU SYSTEME	63
1. Modification de la structure	63
2. Introduction des synonymes	63
3. Contrôle d'erreurs	64
4. Modification du terminal	65
 CONCLUSION	66
 BIBLIOGRAPHIE	67

FACULTE MIXTE DE MEDECINE
ET DE
PHARMACIE DE GRENOBLE

PERSONNEL DE LA FACULTE

DOYEN HONORAIRE : M.J. ROGET
DOYEN : M.G. CABANEL
ASSESSEURS : MM.A. BOUCHERLE
M. TANCHE

PROFESSEURS HONORAIRES

Melle R. BARRIER
M. L. ABONNENC
M. L. BETHOUX
M. B. JOYEUX
M. J. JAUDEL

MAITRE DE CONFERENCES HONORAIRE

M. P. FLANDRIN

PROFESSEURS

MM.	G. ARNAUD	Clinique des maladies infectieuses
	J. BARRIE	Clinique chirurgicale
J.L.	BONNET	Clinique Ophtalmologique
	J. BONNET-EYMARD	Pathologie Médicale
	A. BOUCHERLE	Chimie et Toxicologie
	J. CABANAC	Pathologie Chirurgicale
	G. CABANEL	Hydrologie et Climatologie
	F. CALAS	Anatomie
	G. CARRAZ	Biologie Animale et Pharmacodynamie
	G. CAU	Médecine Légale et du Travail
	R. CHATEAU	Thérapeutique (Neuro-Psychiatrie)
	A. COEUR	Pharmacie Chimique et Chimie Analytique
	R. CONTAMIN	Clinique Gynécologique
	P. COUDERC	Anatomie Pathologique
Mme	A. DEBELMAS	Matière Médicale
MM.	P. DUGOIS	Clinique de Dermatologie et Syphiligraphie
	R. FAU	Clinique de Neurologie et Psychiatrie
	J. LACHARME	Biologie Végétale

.../...

J. LATURAZE	Chimie Biologique Pharmaceutique
J. LEDRU	Clinique Médicale B
Y. MALINAS	Clinique Obstétricale
Y. MAZARE	Clinique Médicale A
C. MOURIQUAND	Histologie
F. PIAGET	Clinique O.R.L.
R. RINALDI	Physique
J. ROGET	Clinique de Pédiatrie et Puériculture
R. SEIGNEURIN	Microbiologie et Hygiène
P. VIGNAIS	Biochimie Médicale

PROFESSEURS A TITRE PERSONNEL

MM.	H. BEZES	Chirurgie générale
	M. TANCHE	Physiologie
	A. VERAÏN	Physique

PROFESSEURS SANS CHAIRE

MM.	A. BEAUDOING	Pédiatrie
	M. GEINDRE	Electroradiologie
	R. LATREILLE	Chirurgie générale
	P. MARTIN NOEL	Médecine générale
	M. REVOL	Urologie
	J. REYMOND	Chirurgie générale
Mme	A. VERAÏN	Pharmacie Galénique

MAITRES DE CONFERENCES AGREGES

Mlle	C. AGNIUS DELORD	Physique Pharmacie
Mlle	J. ALARY	Chimie analytique
MM.	P. AMBLARD	Dermatologie
	P. AMBROISE THOMAS	Parasitologie
	Y. BOUCHET	Anatomie
	J. BUTEL	Orthopédie
	J. CHAMPETIER	Anatomie et organogénèse
	M. COLOMB	Biochimie médicale
	R. CHARACHON	O.R.L.
	J. FAURE	Médecine légale
J.P.	GAUTRAY	Gynécologie-obstétrique
	M. GAVEND	Pharmacologie
	J. GROULADE	Biochimie médicale
	D. HOLLARD	Hématologie
	R. HUGONOT	Hygiène
	P. JALBERT	Histologie
	C. LUU DUC	Chimie organique
	R. MAGNIN	Hygiène et médecine préventive
	M. MICOUD	Maladies infectieuses
J.M.	MULLER	Thérapeutique

MM.	J. PERRET	Neurologie
	L. PERRIN	Pathologie expérimentale
	M. RACHAIL	Médecine interne
	J. DE ROUGEMONT	Neuro-chirurgie
	R. SARRAZIN	Anatomie
	L. SIROT	Chirurgie générale
	P. STIEGLITZ	Anesthésiologie
	R. VOOG	Médecine interne
	C. VROUSSOS	Radiologie

CHARGEES DE FONCTIONS

Mme	H. BERIEL	Physiologie animale
-----	-----------	---------------------

EXAMINATEURS DE LA THESE

MM.	G. CABANEL	PRESIDENT
	J. KUNTZMANN	
	M. GEINDRE	
	J. VALOIS	

A Monsieur le Professeur G. CABANEL,
Doyen de la Faculté de Médecine

qui nous a fait l'honneur de présider
au jugement de cette thèse

Nous le prions de croire à notre déférente gratitude.

A Monsieur le Professeur J. KUNTZMANN
 Monsieur le Professeur M. GEINDRE

qui ont bien voulu accepter de juger notre
travail

qu'ils veuillent bien croire à l'expression de
nos sincères remerciements.

A Monsieur le Docteur J. VALOIS

qui nous a guidée constamment dans notre étude.
Qu'il veuille bien croire à notre très vive reconnaissance.
L'enthousiasme dont il témoigne pour l'informatique
médicale et la confiance qu'il nous a accordée
en acceptant de nous accueillir sans réserves dans son
équipe, dont pourtant l'éloignement géographique
aurait pu nous séparer, furent pour nous une expérience
dont nous nous souviendrons.

Que Monsieur J.R. ABRIAL
et tous ses collaborateurs de l'IMAG

veuillent croire que nous mesurons pleinement
l'importance du concours qu'ils nous ont apporté.
Nous avons pu apprécier à quel point l'idée n'est
rien si l'outil pour la mettre en oeuvre n'existe
pas et nous leur sommes reconnaissante de nous
avoir permis de reprendre pour notre propre compte
les termes qu'ils utilisaient il y a quelques mois,
pour nous faire comprendre les caractéristiques
essentiels de leur système informatique.

A nos Maîtres dans l'Externat,

Monsieur le Professeur P. MOUNIER-KUHN

Monsieur le Professeur A. TRILLAT

Monsieur le Docteur E. ROCHET

Monsieur le Professeur M. BETHENOD

Monsieur le Professeur P. P-RAVAULT

Monsieur le Professeur M. LEVRAT

qui ont toujours été pour nous des guides dévoués et sûrs
dans notre découverte de la pathologie et notre connaissance du
malade.

Qu'ils trouvent ici l'expression de notre profonde reconnaissance.

A Monsieur le Docteur BARJHOUX
 Monsieur le Docteur LANCON
 Monsieur le Docteur STEFANINI
 médecins au Centre Hospitalier de Chambéry

Qu'ils veuillent bien accepter mes très sincères remerciements.

A mon Mari
 mes Parents

et à François
 dont l'aide me fut très précieuse tout au
 long de mes études médicales.

I N T R O D U C T I O N

L'aide que peut apporter l'ordinateur dans la gestion du dossier médical devient une des préoccupations croissantes du corps médical.

Au Centre Hospitalier Régional de Grenoble et à l'Institut de Mathématiques Appliquées (IMAG) de cette ville, les travaux sur ce sujet ont débuté en 1967, avec l'étude d'une nouvelle forme de langage : un langage normalisé. Elle fut conduite par le Docteur J. Valois et Messieurs Ph. Duquesnel et F. Peccoud. Une première application fut réalisée pour la description des demandes d'examens radiologiques. Ce travail, si limité soit-il, permettait d'envisager l'extension du langage à d'autres domaines du dossier médical, par la constitution d'un fichier hybride (1) .

En fait, une nouvelle direction de recherche fut prise dans le cadre de deux contrats de recherche passés entre le Ministère des Affaires Sociales, la DGRST, et l'IMAG. Un système conversationnel de définition et d'interrogation de données fut conçu et réalisé par l'équipe informatique de Monsieur J.R. Abrial, et l'étude de son application à la gestion automatique de l'ensemble du dossier médical fut entreprise.

Notre participation s'attachait à définir les modalités d'expression du diagnostic. Après avoir analysé les possibilités éventuelles d'expression du diagnostic selon le langage normalisé, nous avons utilisé la structure d'information définie à l'IMAG, pour décrire les différents éléments constituant un diagnostic. Cette étude nous a conduits à nous demander quels sont les caractères ergonomiques d'un système informatique de gestion du dossier médical pour un médecin. C'est selon ces caractères que nous avons tenté de juger le système proposé.

P L A N

Chapitre I : L'EXPRESSION DU DIAGNOSTIC

I - ANALYSE DU DIAGNOSTIC

- Constituants du diagnostic
- Structure du diagnostic

II - GRAMMAIRE CONTEXT-FREE D'EXPRESSION D'UN DIAGNOSTIC

- Abrégé d'une définition de grammaire
- Définition d'une grammaire formelle de description du diagnostic
- Remarques

III - DESCRIPTION DU SYSTEME D'INTERROGATION ET DE DEFINITION DE DONNEES DEFINI A L'IMAG

- Mode "définition de données"
- Mode "Manipulation de données"
- Mode "Récupération de données"
- Quelques détails techniques

IV - STRUCTURE TYPE DES DONNEES POUR LE DIAGNOSTIC

Chapitre II : ERGONOMIE DE L'UTILISATION D'UN SYSTEME INFORMATIQUE DE GESTION DES DOSSIERS MEDICAUX

I - CARACTERES GENERAUX D'UN SYSTEME AUTOMATISE DE GESTION D'UN FICHIER MEDICAL :

- Fonctions : structures algébriques disponibles
création des données
interrogation du fichier

- Caractéristiques ergonomiques de ces fonctions pour un médecin lors de :
 - la définition de données
 - la création de données
 - l'interrogation de données

II - ETUDE DE DIFFERENTS SYSTEMES

- texte
- mots-clés
- codes
- questionnaires
- langage normalisé

Chapitre III : CARACTERISTIQUES ERGONOMIQUES DU SYSTEME DEFINI A L'IMAG

I - STRUCTURE ACTUELLE

- définition de données
- création de données
 - nature sémantique des caractéristiques
 - forme du dialogue entre médecin et ordinateur
 - structure-type du fichier
 - conclusion
- interrogation de données
 - travail clinique
 - recherche
- conclusion

ii- EVOLUTION DU SYSTEME

- Modification de la structure
- Introduction des synonymes
- Contrôle d'erreurs
- Modification du terminal

CONCLUSION

CHAPITRE I - L'EXPRESSION DU DIAGNOSTIC

I - ANALYSE DU DIAGNOSTIC

Avant d'adapter l'expression du diagnostic à une structure algébrique, il était nécessaire d'analyser ce diagnostic pour préciser la nature de ses constituants.

1 - LES ELEMENTS CONSTITUTIFS DU DIAGNOSTIC

Un diagnostic est composé de trois concepts :

- l'affection
- la topographie de cette affection
- l'étiologie de cette affection.

1-1. L'affection

C'est le processus morbide considéré dans ses manifestations actuelles, abstraction faite de ses causes. On peut en distinguer trois types :

- les maladies générales, qui affectent l'ensemble de l'organisme

exemple : DIABETE

- les lésions qui regroupent toutes les affections dont la localisation est précise et limitée

exemple : FRACTURE, ABCES, TUMEUR.....

Dans ce groupe sont incluses les interventions chirurgicales, que l'on considère comme des lésions expérimentales.

- les symptômes, et les syndromes (ensemble de symptômes), c'est-à-dire les caractères évoquant un état morbide, sans être par eux-même la maladie

exemple : AMAIGRISSEMENT, HYPERTHERMIE, NEURALGIE.

.../...

Elle peut être étudiée à différents niveaux :

- la région anatomique, ensemble d'organes situés en un même point de l'organisme
exemple : TÊTE, COU, THORAX.
- le système, ou appareil, constitué d'organes ayant le même rôle physiologique, quelque soit leur localisation
exemple : l'appareil digestif, dont les organes, situés aussi bien dans la tête, que dans le cou, le thorax et l'abdomen, remplissent tous la fonction de la digestion.
- l'organe, qui est l'unité de base physiologique et anatomique
exemple : ESTOMAC, RECTUM, FEMUR.
- le tissu, ensemble de cellules ayant une même morphologie et des potentialités identiques
exemple : OS, MUSCLE, TISSU ADIPEUX
- la cellule, unité fondamentale histologique.

Nous nous arrêtons là, mais il serait possible de décrire aussi les constituants de la cellule....

1-3. L'origine ou étiologie :

Elle peut être extérieure ou intérieure à l'organisme.

- Dans les étiologies externes nous classons :

- les agents physiques tels que rayons X, chaleur, lumière... et les traumatismes (compression, perforation, torsion...)

.../...

- les toxiques (produits chimiques, toxiques d'origine végétale et animale, médicaments...)
- les agents infectieux qui regroupent d'une part les microbes (virus, bactéries, rickettsies), d'autre part les parasites animaux et végétaux.
- Dans les étiologies internes, nous regroupons les autres étiologies, parmi lesquelles nous distinguons un certain nombre de catégories :
 - étiologie congénitale, qu'il s'agisse de troubles génétiques, c'est-à-dire liés au patrimoine héréditaire, ou de troubles dus à la vie foetale.
 - dysfonctionnements (troubles métaboliques, circulatoires, nutritionnels, neurovégétatifs, allergiques, etc...)
 - étiologie psychique
 - tumeurs bénignes et tumeurs malignes (dans ces dernières se retrouvent tous les cancers, même les cancers généralisés pour lesquels le terme "tumeur" peut sembler impropre. Nous les classons dans les étiologies internes... en attendant que cette étiologie soit précisée !)
- Enfin, l'étiologie peut être inconnue. On parlera de maladie cryptogénétique.

La classification que nous venons de présenter peut, sans nul doute, prêter à discussion. Nous voulons simplement montrer par là que chaque concept fondamental du diagnostic peut être lui-même divisé en de nouveaux concepts qui forment un certain nombre de classes.

Par exemple, si l'étiologie d'une affection est "STREPTOCOQUE", le clinicien peut désirer classer cette affection dans les BACTERIES elles-mêmes classées dans les INFECTIONS. Cela représente déjà trois classes étiologiques. Mais le

.../...

biologiste peut vouloir plus de précision et définir pour le streptocoque considéré, sa famille, sa tribu, son genre, son espèce.

2 - LA STRUCTURE DU DIAGNOSTIC :

- Un diagnostic, au niveau le plus élémentaire, peut donc contenir les trois concepts fondamentaux que nous avons définis. En fait, seule l'affection est toujours exprimée. Topographie et étiologie peuvent manquer. Chacun de ces concepts peut être traduit par un ou plusieurs mots français (certains mots pouvant exprimer deux ou trois concepts : exemple : GASTRITE ; STREPTOCOCCEMIE).
- Un diagnostic peut être aussi une construction plus complexe, mais il s'agit alors de plusieurs diagnostics élémentaires, reliés entre eux par un mot, que nous avons nommé opérateur logique : ce sont les prépositions et conjonctions françaises CHEZ, ET, AVEC, PAR, OU, POUR.
 - L'opérateur "ET" sépare deux diagnostics indépendants l'un de l'autre, qui coexistent chez un même malade.
exemple : ULCERE D'ESTOMAC ET TUBERCULOSE PULMONAIRE.
 - L'opérateur "CHEZ" précède un diagnostic qui ajoute une notion de "terrain" au diagnostic qu'il suit :
exemple : PERITONITE CHEZ UN ETHYLIQUE
 - L'opérateur "PAR" précède un diagnostic qui est l'étiologie (ou la complication) du diagnostic qu'il suit :
exemple : PERITONITE PAR ULCERE PERFORE
 - L'opérateur "AVEC" a le rôle inverse. Le diagnostic qu'il suit est l'étiologie du diagnostic qu'il précède (ou sa complication) :
L'exemple précédent peut s'écrire : ULCERE PERFORE AVEC
PERITONITE
 - L'opérateur "POUR" suit un diagnostic dont l'affection est une
.../...

intervention chirurgicale. Il introduit un deuxième diagnostic qui est la cause de cette intervention.

exemple : GASTRECTOMIE POUR ULCERE D'ESTOMAC

- Enfin, l'opérateur "OU" exprime le doute entre deux diagnostics

exemple : ULCERE OU CANCER D'ESTOMAC.

Pour tenter de sauvegarder la souplesse nécessaire à l'expression du diagnostic, nous avons été amenés à étudier, sur proposition du Docteur Valois, une structure d'expression du diagnostic en utilisant comme structure algébrique une grammaire context-free de reconnaissance des formes.

II - GRAMMAIRE CONTEXT-FREE D'EXPRESSION D'UN DIAGNOSTIC

Nous avons été guidés dans cette étude par le travail réalisé par Ph. Duquesnel (2).

L'idée générale exprimée dans la référence citée consiste à exploiter une structure algébrique de grammaire context-free pour relier, d'une manière plus souple que celle qu'autorisent les structures d'arbres fixes utilisées dans les méthodes actuellement disponibles, les concepts associés dans l'expression de fait d'une discipline scientifique bien délimitée.

1 - ABREGE D'UNE DEFINITION DE GRAMMAIRE POUR LA COMPREHENSION DE CE QUI SUIT.

Tout langage est la construction, suivant un ensemble de règles précises appelé syntaxe, de phrases comprenant un certain nombre de mots. La nature algébrique de ces structures syntaxiques a été mise en évidence récemment ; il ne nous appartient pas, avec notre incompetence, d'en discuter, mais simplement de présenter des concepts fondamentaux de théorie algébrique des langages que nous avons pu comprendre.

.../...

Définition : une grammaire est une suite de règles syntaxiques de production, chaque règle permettant de remplacer une variable syntaxique par un groupe ou plusieurs groupes de variables syntaxiques équivalents.

Nous noterons chaque variable syntaxique par un nom encadré de crochets. Chaque règle de grammaire sera exprimée par une partie gauche composée d'une variable syntaxique, suivie d'une flèche, suivie d'une partie droite composée d'une liste de variables syntaxiques. Le signe | séparant ces variables signifie "ou".

A titre d'exemple, nous pourrions donner la structure de grammaire équivalente à la structure du diagnostic exposée dans le paragraphe précédent :

1ère règle :

$$\begin{aligned} \langle \text{diagnostic} \rangle \rightarrow & \langle \text{diagnostic élémentaire} \rangle | \\ & \langle \text{diagnostic élémentaire} \rangle \langle \text{opérateur logique} \rangle | \\ & \langle \text{diagnostic élémentaire} \rangle \end{aligned}$$

Le symbolisme signifie qu'un diagnostic se décomposera soit en un diagnostic élémentaire, soit en un diagnostic élémentaire suivi d'un opérateur logique suivi d'un diagnostic élémentaire.

2ème règle : $\langle \text{opérateur logique} \rangle \rightarrow \text{et} \mid \text{ou} \mid \text{chez} \mid \text{par}$

Ceci signifie simplement que les opérateurs utilisés dans notre langage de description du diagnostic sont les quatre opérateurs exprimés dans la partie droite de la règle ci-dessus.

3ème règle : $\langle \text{diagnostic élémentaire} \rangle \rightarrow \langle \text{lésion} \rangle \langle \text{topographie} \rangle \langle \text{étiologie} \rangle$

Nous indiquons par là que nous décomposons un diagnostic élémentaire en trois parties.

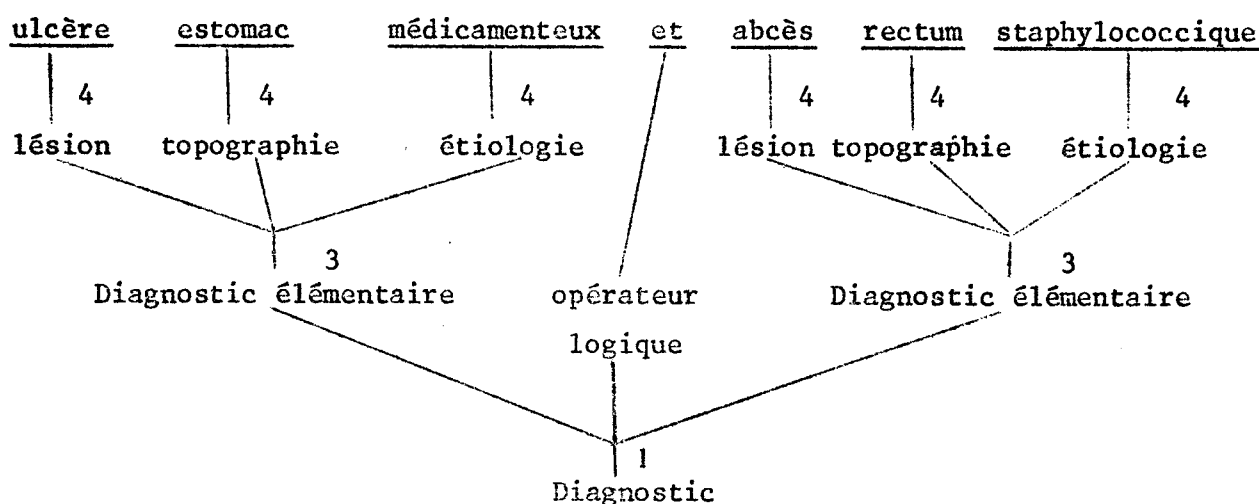
4ème règle annexe, à titre d'exemple :

.../...

<lésion> → ulcère | abcès | fracture | tumeur (1)
 <topographie> → estomac | rectum | fémur
 <étiologie> → médicamenteux | staphylococcique | traumatique

Dans tout langage, il faut préciser quels sont les mots (appelés en théorie des langages 'variable terminale' du langage) qui peuvent être utilisés dans ce langage. Nous avons ici défini un langage acceptant 4 mots de description de lésion, 3 mots de description de topographie et 3 mots de description d'étiologie.

exemple 1 :

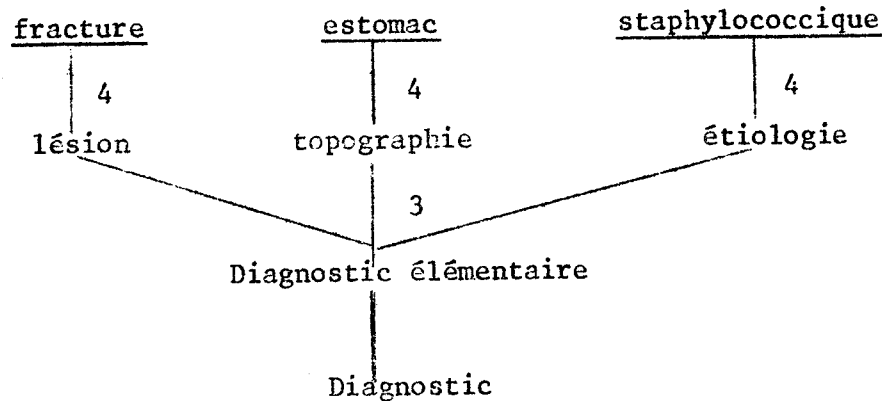


(les chiffres indiquent le numéro de la règle utilisée)

(1) Nous avons dans le diagnostic inscrit 'tumeur' comme terme de lésion et 'cancer' comme terme d'étiologie.

.../...

exemple 2 :



Dans ces exemples nous avons représenté l'arbre d'analyse des phrases, c'est-à-dire la vérification, à partir des mots de la phrase, d'une structure correspondant aux règles.

On peut cependant noter que, si la première phrase est bien construite, elle a aussi un sens médical, tandis que la seconde, bien que syntaxiquement correcte, est parfaitement absurde du point de vue médical.

Cette définition, plus empirique qu'axiomatique, de ce qu'est une grammaire formelle, peut nous permettre d'aborder l'utilisation d'un tel outil pour structurer un domaine d'information médicale.

2 - DEFINITION D'UNE GRAMMAIRE FORMELLE DE DESCRIPTION DU DIAGNOSTIC.

Monsieur Duquesnel s'était attaché, avec le docteur Valois, à définir une grammaire de description des demandes d'examens radiologiques puis à écrire les programmes permettant de réaliser l'analyse syntaxique des demandes exprimées dans ce langage.

Notre but, plus modeste, est de trouver une grammaire d'un langage de description du diagnostic en laissant le soin aux informaticiens de juger s'il est opportun d'écrire vraiment les programmes qui permettraient d'analyser les phrases de ce langage. L'exercice peut sembler à priori stérile;

.../...

Règle 1 : <Diagnostic> → <diagnostic élémentaire> | <diagnostic élémentaire>
<opérateur logique> <diagnostic élémentaire>

Règle 2 : <opérateur logique> et | ou | par | chez

Règle 3 : <diagnostic élémentaire> → <lésion> | <lésion> <topographie> | <lésion> <étio-
logie> | <lésion> <topographie> <étiologie> |
<lésion> <étiologie> <topographie>

- Groupe de règles n° 2.

Règle 4 : <lésion> → mot de lésion | mot de lésion-étio | mot de lésion-topo |
mot de lésion-topo-étio | mot de lésion <groupe qualificatif> |
mot de lésion-étio <groupe qualificatif> | mot de lésion-topo
<groupe qualificatif> |
mot de lésion-topo-étio <groupe qualificatif> | <groupe qualifi-
catif> mot de lésion | <groupe qualificatif> mot de lésion-étio |
<groupe qualificatif> mot de lésion topo | <groupe qualificatif>
mot de lésion-étio-topo

Cette règle signifie ceci : la lésion peut être décrite par un seul mot (variable terminale). La description de ce mot peut être affinée par un groupe de qualificatifs qui peut le précéder ou le suivre.

Le mot n'est pas uniquement un nom commun français. Il peut aussi être un adjectif, ou un nom propre.

exemple : HODGKIN

On a distingué quatre sortes de mots :

le 'mot de lésion' ne contient qu'un seul sens correspondant au seul concept

lésion : exemple ULCERE

le 'mot de lésion-topo' contient les deux concepts : lésion et topographie :

exemple

G A S T R I T E
└───┘ └───┘
↓ ↓

topographie=estomac lésion=inflammation

de même dans le 'mot de lésion-étio' on retrouve la lésion et l'étiologie :

exemple : STAPHYLOCCOCIE = infection à staphylocoque

enfin le 'mot de lésion-topo-étio' contient les 3 concepts fondamentaux :

exemple : STAPHYLOCCOCEMIE = infection générale par présence de staphylocoques
dans le sang.

.../...

bles terminales. Nous les récapitulons ci-dessous en indiquant pour chacune le nombre de mots qu'il avait retenu. Dans notre grammaire, c'est par plusieurs milliers que se compteraient les mots correspondant aux variables terminales 'mot de topographie' (on y retrouverait tous les termes anatomiques) et 'mot de lésion' (tous les termes de pathologie y seraient rassemblés). C'est donc à un inventaire représentant un travail comparable à la création du code SNOD (3) que nous aurions dû procéder, pour pouvoir vraiment dire que la grammaire était définie. Cette remarque permet de constater que la souplesse acquise par la création d'une telle grammaire, ne le serait en fait qu'au prix d'un travail considérable de classement et de normalisation de tous les concepts médicaux. C'est certainement la restriction la plus importante que nous pouvons faire à l'utilisation d'un langage normalisé en médecine.

2) Une autre restriction est liée au fait que la souplesse d'expression n'est jamais totale. En effet, quelle que soit la complexité introduite dans les règles d'une grammaire, il apparaît impossible d'en créer une qui permette d'accepter toutes les formes d'expression imaginées par des médecins non prévenus de son contenu. Autrement dit, à une souplesse d'expression croissante, correspond une difficulté croissante pour le médecin dans l'analyse des erreurs syntaxiques ; à la limite, il faudrait que le médecin devienne un spécialiste de théorie des langages pour utiliser correctement ce genre de grammaire : telle n'est pas sa vocation !

A ces difficultés, s'ajoutent celles, non négligeables, de la constitution d'un dictionnaire (voir chapitre 2, p.49) qui permettrait l'expression du diagnostic selon une structure algébrique de grammaire context-free.

Nous avons donc étudié comment pourrait être utilisé le système décrit à l'IMAG par J.R. Abrial et son équipe.

III - DESCRIPTION DU SYSTEME.

Ce système (3), qui est en cours de réalisation à l'Université de Grenoble sur du matériel IBM, est une banque de données générales destinée à fonctionner sur un système en "time sharing".

.../...

- En mode "définition de données", il donne la possibilité de décrire des modèles de structure de données complexes de même que leur représentation interne.
- En mode "manipulation de données" on peut créer, supprimer ou mettre à jour des données "vives" correspondant à un modèle précédemment introduit.
- En mode "récupération de données" le système répond à des questions cohérentes avec ce qui a été fait lors des phases précédentes.

Exemple : le système peut être utilisé pour :

- définir la structure d'un dossier médical
- stocker et mettre à jour des informations au sujet des malades d'un hôpital
- retrouver les informations contenues dans les dossiers médicaux.

1 - MODE "DEFINITION DE DONNEES"

La description d'un modèle de structure de données complexes se fait au moyen d'un langage. Ce langage n'a pas d'instructions mais seulement des déclarations. Il contient toutes les caractéristiques puissantes que l'on trouve dans la partie "instruction" des langages de programmation actuels (condition, récursion, boucle, procédures, structure de blocs, etc...) pour les mettre dans ce langage de déclaration.

Ceci a d'ailleurs été déjà fait partiellement en d'autres occasions (Cobol, PL/I, Algol,...).

Les données sont décrites d'un point de vue sémantique indépendant de la machine.

Exemple : Commentaire texte

Nom mot

Groupe-sanguin (A+ AB+ A- AB- B+ O+ B- O- Autres)

Pouls de 0 à 300

.../...

Nous venons de décrire ici 4 objets en donnant au système leur nature sémantique, c'est-à-dire un texte, un mot, une liste de valeurs exclusives, une plage de valeurs : ce sont des caractéristiques simples.

Très souvent il n'est pas suffisant de décrire un objet par une seule caractéristique ; on les groupe alors sous un nom générique ; elles forment un bloc délimité par les mots 'Début' et 'Fin'. Les blocs peuvent s'imbriquer les uns dans les autres.

MALADE

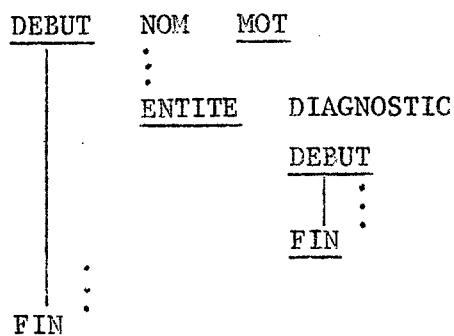
Quelquefois une caractéristique d'un objet n'a qu'une existence conditionnelle (conditionnelle à la valeur d'une ou plusieurs autres).

MALADE

On veut pouvoir exprimer le fait qu'une donnée peut être répétée un nombre indéfini de fois = c'est une entité (l'entité peut être une caractéristique simple ou un bloc).

Exemple : une personne ne "possède" d'habitude qu'un seul nom mais peut-être plusieurs diagnostics. Nous écrivons alors :

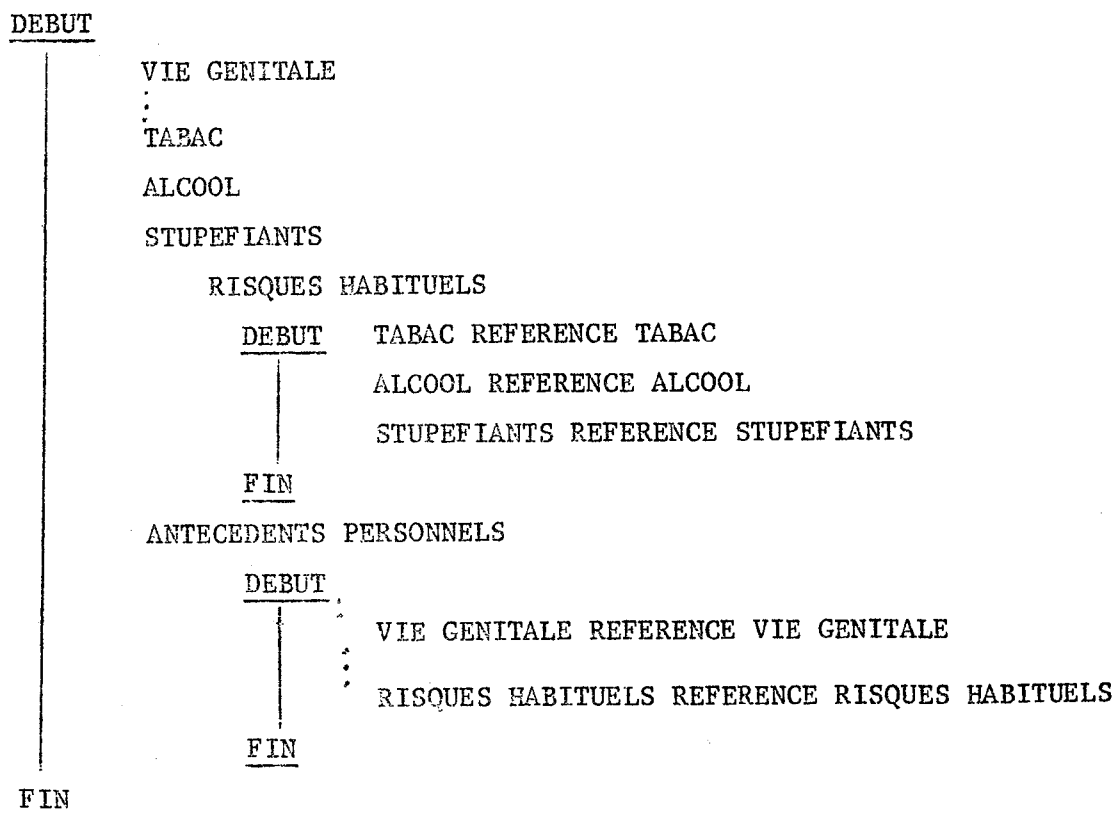
MALADE



Nous voudrions maintenant pouvoir exprimer des relations entre les objets (ou ensembles d'objets). Jusqu'à maintenant la seule relation que nous ayons implicitement introduite est celle d'appartenance.

Suivant la terminologie d'ALGOL W, nous introduisons la REFERENCE.

Exemple :



.../...

N'importe quel niveau de référence en chaîne est permis.

Nous avons introduit d'autres possibilités comme IDEM (référence au niveau du modèle) qui permet des définitions récursives de données, et même un modèle autodescriptif du système (système universel). Voir paragraphe 4 pour d'autres détails sur le langage.

En mode "description de données", le "programmeur" rentre dans le système une définition complète décrite dans le langage que nous venons de définir. Une sorte de compilateur traduit ce programme en une description interne qui consiste principalement en un dictionnaire et une structure complexe chaînée, qui ressemble d'assez près à celle que l'on peut trouver dans une machine après compilation de la partie déclarative d'un programme écrit en ALGOL ou PL/1.

2 - MODE "MANIPULATION DE DONNEES"

Ce mode peut s'utiliser de différentes manières. Celle que nous allons décrire correspond à un emploi conversationnel.

Pour créer des données correspondant à un modèle précédemment rentré, l'utilisateur est conduit par la machine qui lui pose des questions d'après le modèle et ses conditions. Ainsi l'utilisateur rentre les unes après les autres les différentes valeurs (ou la valeur 'indéfini') des caractéristiques, et le système, avant de poser la question suivante, contrôle leur validité sémantique d'après le modèle. Lorsque le système rencontre une entité, il boucle en demandant à l'utilisateur de bien vouloir créer tous les éléments de l'ensemble correspondant avant de continuer.

Exemple : Nous donnons un exemple de création de diagnostic, dont le modèle est décrit p. 36.

La séquence de création est la suivante :

VOUS POUVEZ DEMANDER UNE DES FONCTIONS SUIVANTES (K, I, M, N, S) (1)

K

.../...

(1) K signifie création, I interrogation, M mise à jour, N dénombrement, S surveillance.

QUE VOULEZ-VOUS CREER ?

un malade

NOM ?

dupont

PRENOM ?

jean

AGE ?

566

ERREUR. BORNES DE 0 A 110

AGE ?

56

SEXE ?

m

ERREUR . CECI N'APPARTIENT PAS A LA LISTE DE VALEURS.

LA VOULEZ-VOUS ?

oui

MASCULIN . FEMININ

SEXE ?

masculin

CREATION DE DIAGNOSTIC ?

oui

LESION ?

ulcère

Q-LESION ?

u (1)

TOPOGRAPHIE ?

duodénum

Q-TOPOGRAPHIE ?

bulbe

ETIOLOGIE ?

u

Q-ETIOLOGIE ?

u

.../...

(1) u signifie indéfini.

COMMENTAIRE ?

non

AVEZ-VOUS D'AUTRES DIAGNOSTIC A CREER ?

non

AVEZ-VOUS D'AUTRES MALADE A CREER ?

non

La mise à jour de l'observation se fait de la même façon que la création.
Ces deux fonctions modifient les éléments d'un fichier physique.

3 - MODE "RECUPERATION DE DONNEES"

Ce mode comprend les fonctions d'Interrogation et de Dénombrement.

Interrogation.

La représentation interne du modèle fabriqué lors de la première phase est la clé du fichier créé ou mis à jour durant la phase de manipulation de données. Les questions sont fournies au système grâce à un langage de questions qui permet de désigner des objets (ou des ensembles d'objets) en suivant une méthode associative.

Exemples :

DIAGNOSTIC DE MALADE AYANT NOM = 'DUPONT' ET PRENOM = 'JEAN';?

est une question cohérente avec l'exemple précédent.

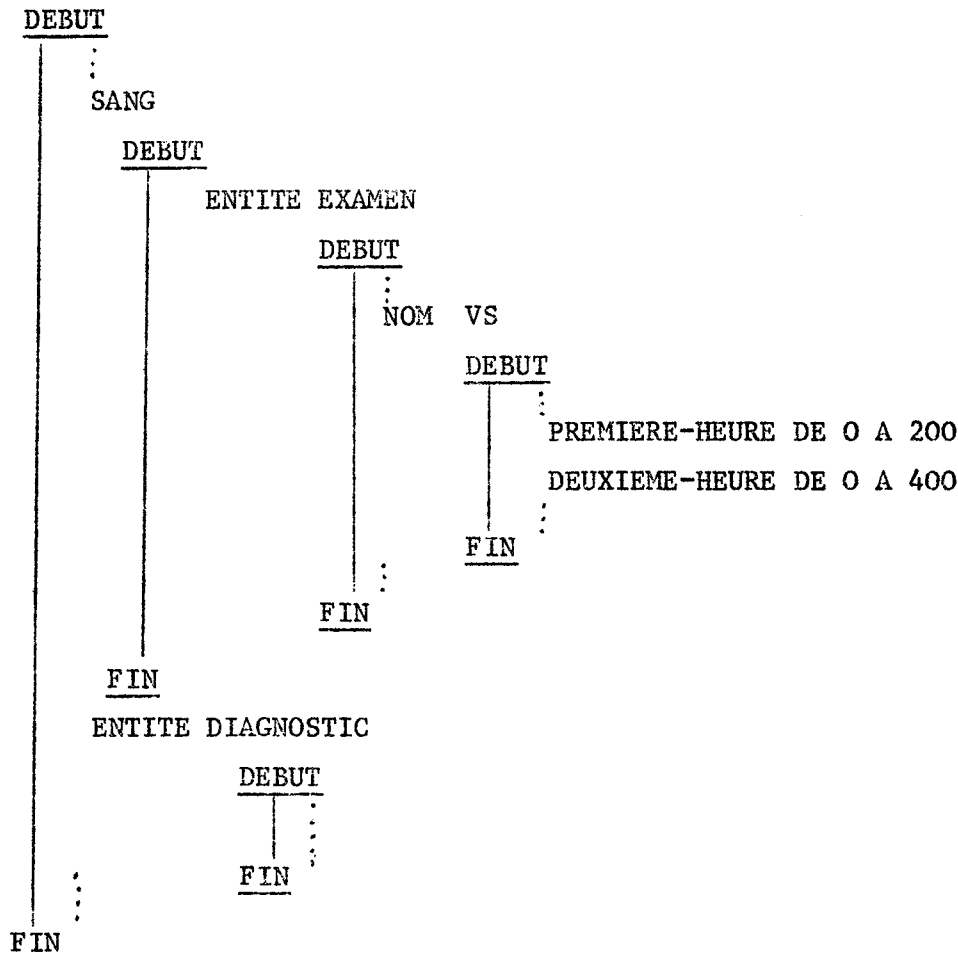
La question "quels sont les diagnostics des malades dont la VS est supérieure à 100 à la 2ème heure ?" sera posée de la façon suivante :

DIAGNOSTIC DE MALADE AYANT EXISTE EXAMEN AYANT NOM = 'VS' ET DEUXIEME HEURE
> 100 ; DE SANG ; ?

c'est une question cohérente avec le modèle suivant :

.../...

ENTITE MALADE



La syntaxe de ce langage de questions ressemble à celle qui est utilisée pour désigner des objets en COBOL. De plus l'opérateur AYANT permet des recherches associatives sur les ensembles. N'importe quelle imbrication d'opérateurs AYANT cohérente avec le modèle correspondant est permise. D'autres opérateurs sont aussi disponibles comme :

ET, OU, NON, EXISTE, x, +, -, /, PREMIER, DERNIER, etc...

Exemples :

NOM DE MALADE AYANT EXISTE DIAGNOSTIC AYANT LESION =

'ULCERE' ET TOPOGRAPHIE = 'ESTOMAC'; ;?

PREMIER MALADE AYANT EXISTE EXAMEN DE SANG = 'ANTICORPS.ANTIDNA'; ?

DERNIER EXAMEN AYANT NOM = 'VS' DE SANG DE MALADE AYANT NOM = 'DUPONT'; ?

EXAMEN AYANT NOM = 'BIOPSIE' OU NOM = 'ENDOSCOPIE'....

.../...

(dans cet exemple l'opérateur OU est non exclusif. Cette question signifie en fait = donnez-moi les biopsies et les endoscopies...)

Les programmes qui répondent aux questions sont essentiellement :

- un petit contrôleur de syntaxe et de sémantique qui traduit la forme externe de la question en une structure d'arbres.
- un interpréteur qui se déplace à travers le fichier, d'après le modèle et la question.
- un éditeur pour présenter les résultats.

En fait, le langage de description (paragraphe 2) et le langage de question ne font qu'un.

Dénombrement

Cette fonction permet de dénombrer les données du fichier en utilisant les mêmes critères logiques que la fonction d'interrogation.

Exemple : on peut demander le nombre de malades admis à l'hôpital par l'intermédiaire de la consultation externe en Mai 1970. Après avoir demandé la fonction N on pose la question suivante :

MALADE AYANT ADMISSION = 'CONSULTATION' ET MOIS DE DATE = 5
ET ANNEE DE DATE = 1970 DE SEJOUR ; ?

Surveillance

Cette fonction, accessoire, permet de simplifier l'écriture des mises à jour et des questions. En effet, si cette fonction n'existait pas, lorsque l'on poserait les questions concernant un même malade, il faudrait chaque fois répéter ... DE MALADE AYANT NOM = 'X'. Lorsque l'on interroge le système sous surveillance, les questions posées sont automatiquement rapportées au malade surveillé. Il en est de même pour la mise à jour.

Exemple : QUE VOULEZ-VOUS SURVEILLER ?

- un malade ayant nom = 'dupont' et prénom = 'jean'; ?

VOUS POUVEZ INTERROGER OU METTRE A JOUR ?

- i

QUE VOULEZ-VOUS SAVOIR ?

- diagnostic ?

.../...

On obtiendra ainsi le diagnostic de Monsieur Dupont.

Suppression

Cette fonction permet de supprimer lorsqu'on le désire n'importe quel enregistrement, sans obligation de le remplacer par un nouvel enregistrement.

4 - QUELQUES DETAILS TECHNIQUES

L'organisation générale du système est décrite dans le schéma ci-après.

Nous allons maintenant décrire rapidement la structure interne des modèles, les fichiers, le programme de réponse, l'organisation de la mémoire et l'implémentation des références.

4-1. Structure interne

Le système s'occupe d'objets identifiés par un nom. Par exemple, résumé, nom, température, sexe, état-civil, pouls, etc...

La structure interne est donc d'abord un dictionnaire de noms. Ceux-ci réfèrent à des méta-objets qui peuvent être : des entités, des blocs, des caractéristiques, des valeurs, des conditions. Chaque métaobjet possède à son tour des caractéristiques.

Exemple : Une entité possède d'autres entités, une entité a un nom, un bloc externe, etc...

La seconde partie de la structure interne est donc l'extension de celui-ci, qui donne pour chaque métaobjet les éléments dont il est composé, et les relations qu'il entretient avec les autres.

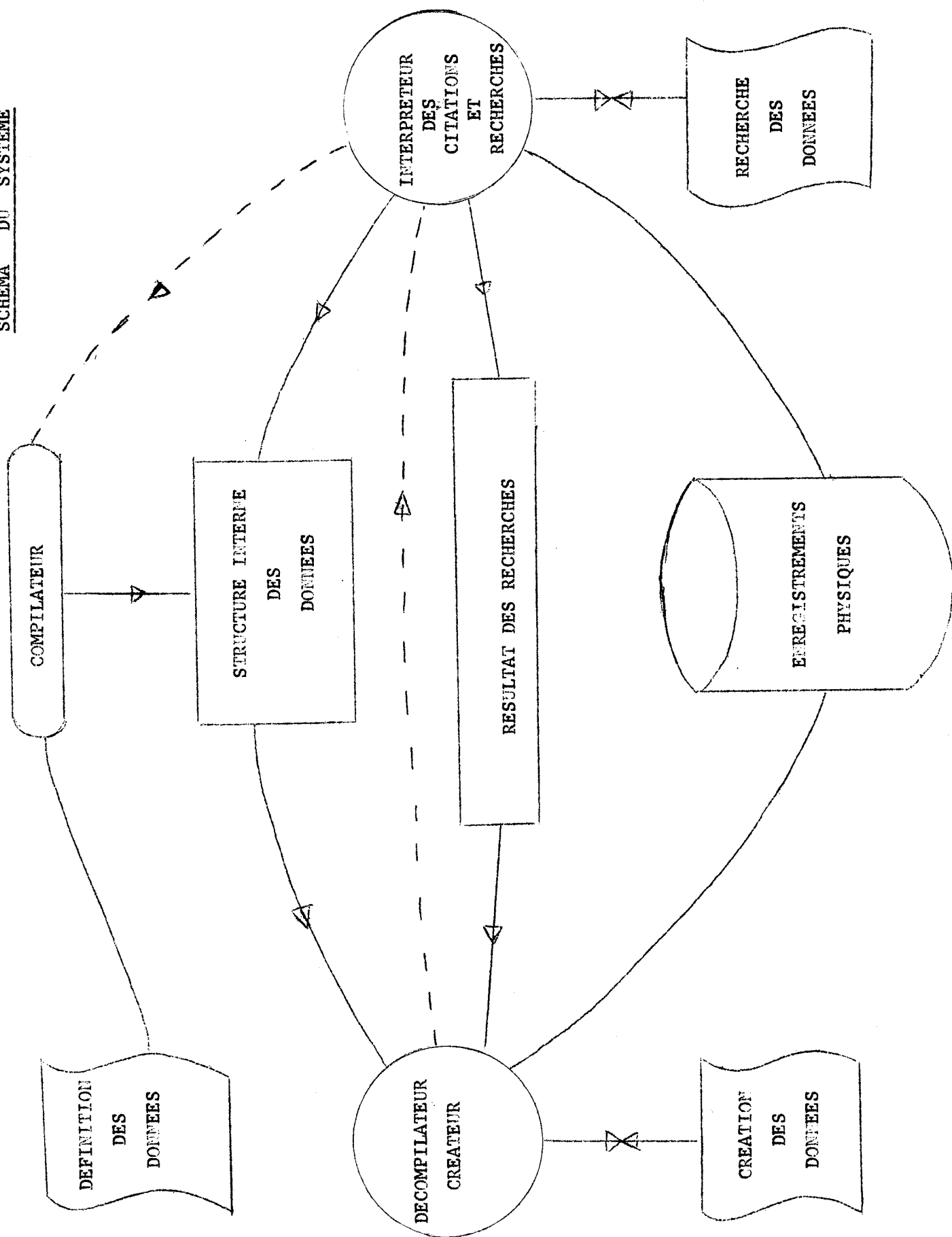
4-2. Le fichier

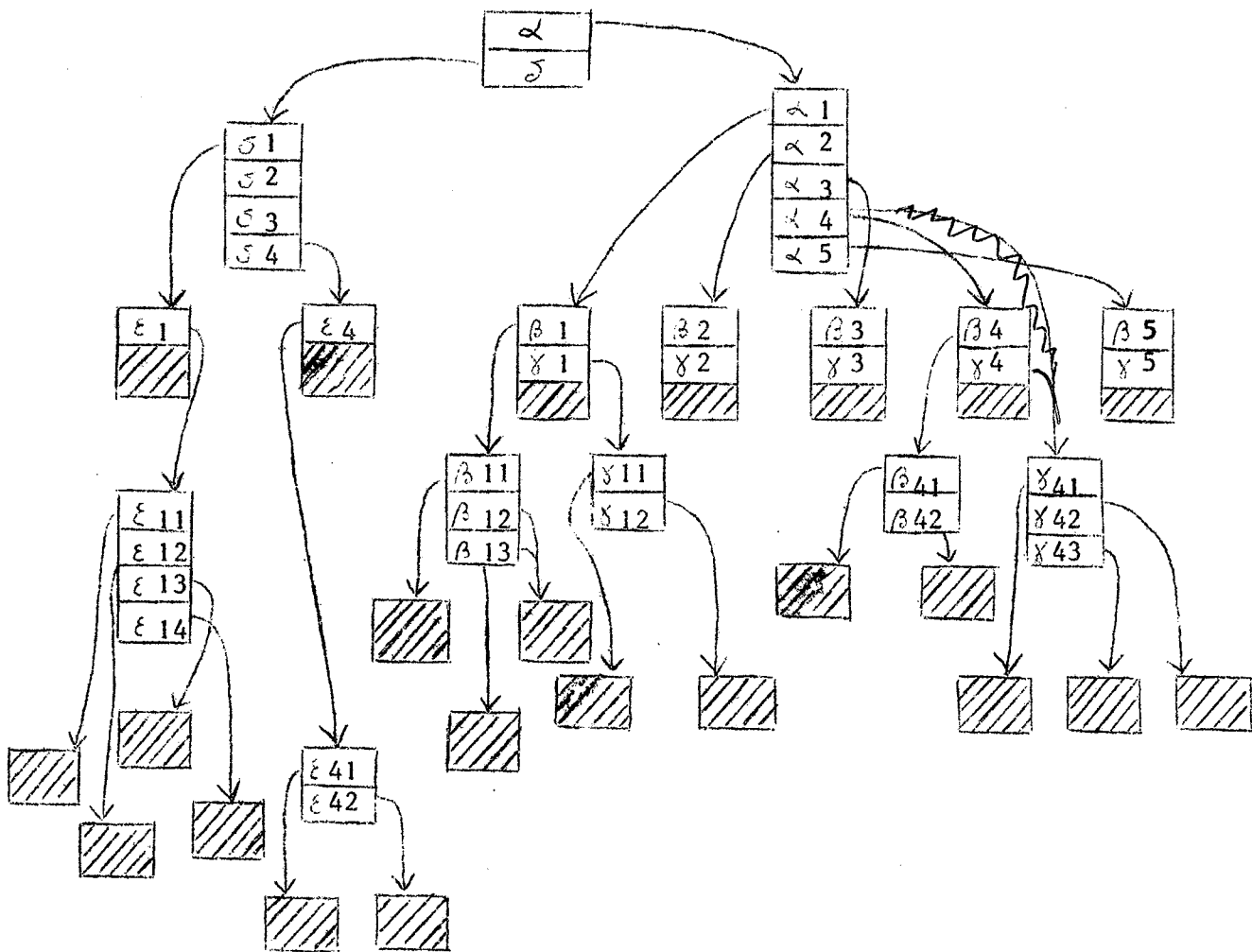
Le fichier est composé d'enregistrements. Le niveau des enregistrements est le même que le niveau des éléments d'ensemble.

Supposons que nous ayons le modèle suivant :

.../...

SCHEMA DU SYSTEME





Un élément d'ensemble est constitué par :

- une liste de classes d'entités qu'il possède
- une liste des blocs présents
- une liste des valeurs de ses caractéristiques

Les éléments d'ensemble appartenant à la même classe sont disséminés sur la mémoire (disque) mais chacun peut être référencé par un dictionnaire. Les dictionnaires eux-même sont disséminés sur la mémoire et sont référencés par les éléments de la liste des classes d'entités de leur élément-père.

.../...

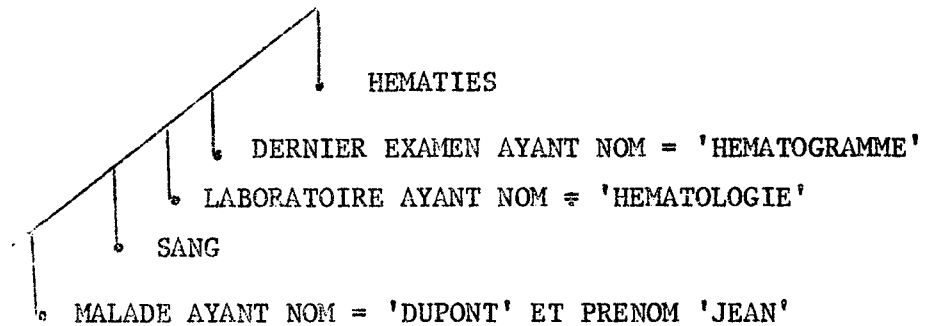
4-3. Programme de récupération

La partie principale de ce programme est un algorithme d'interprétation dont le rôle consiste à donner une réponse à la question posée.

Exemple : à la question suivante :

HEMATIE DE DERNIER EXAMEN AYANT NOM = 'HEMOGRAMME' ; DE LABORATOIRE AYANT
NOM = 'HEMATOLOGIE' ; DE SANG DE MALADE AYANT NOM = 'DUPONT' ET PRENOM = 'JEAN' ; ?

correspond l'arbre suivant



L'interprétation a lieu comme suit :

elle sélectionne les malades qui s'appellent JEAN DUPONT, puis le sang de ces malades, etc...

Cet algorithme est donc un simple algorithme de déplacement dans un arbre.

A chaque noeud on peut restreindre le nombre de successeurs en interprétant les conditions "AYANT".

4-4. Organisation de la mémoire

L'organisation de la mémoire est un système indépendant fait de deux espaces :

- 1'espace virtuel
- 1'espace réel

L'espace virtuel est segmenté.

Un segment est un élément de stockage de taille variable mais composé d'un nombre entier de pavés. Un pavé a un nombre fixe de mots par exemple 64.

.../...

Entre deux segments on trouve 0 ou 1 trou.

Chaque segment est référencé par un segment seulement, mais il peut en référencer plusieurs. L'organisation des segments est donc un arbre. Evidemment, les éléments d'ensemble et les dictionnaires que nous avons décrit au § 4-2 sont des segments.

L'espace réel est fait de deux sous-ensembles :

- le plan primaire
- le plan secondaire

L'extension du plan secondaire et de l'espace virtuel sont les mêmes.

L'espace réel est paginé. Chaque page consiste en un nombre fixe de mots, par exemple 1024.

Pour les 2 espaces, nous avons plusieurs opérations primaires qui peuvent créer, mettre à jour, ou supprimer des segments ou des trous, déplacer des pages du plan primaire ou du plan secondaire et vice-versa et enfin assurer le passage de l'espace virtuel à l'espace réel.

4-5. Implémentation des références

Elles ont été introduites au §1. Elles permettent à la structure formelle (donc à la structure physique) d'être plus générale qu'un arbre en donnant la possibilité d'établir n'importe quelles relations entre éléments et leurs caractéristiques.

Comme nous l'avons souligné au §4-4 , l'organisation de la mémoire est dynamique. Il est donc impossible d'implémenter une référence sous la forme d'une adresse absolue. Ce que nous avons fait consiste à stocker un "programme d'adresses" qui plutôt que l'adresse de l'élément donne le moyen de l'atteindre.

Exemple : si l'objet référencé est la 1ère caractéristique du 3ème élément de la 4ème classe d'entités du 14ème élément de la 13ème classe d'entités du système, le "programme d'adresses" correspondant sera (1 3 4 14 13).

.../...

6 - Conclusion

Ce système est un système très général, flexible. Son utilisation en mode conversationnel en fait une structure de choix pour la gestion automatique du dossier médical. Ce problème a été étudié par S. Cohen (4).

Nous allons voir maintenant une de ses applications : l'expression du diagnostic.

IV - STRUCTURE-TYPE DES DONNEES POUR LE DIAGNOSTIC.

La définition est la suivante :

DEBUT ENTITE MALADE

DEBUT NOM MOT

PRENOM MOT

AGE DE 0 A 150

SEXE (MASCULIN FEMININ)

ENTITE DIAGNOSTIC

DEBUT SI EXISTE DIAGNOSTIC DE MALADE EN COURS

ALORS OPL (PLUS CHEZ PAR OU-BIEN)

LESION MOT

Q-LESION MOT

TOPOGRAPHIE MOT

Q-TOPO MOT

ETIOLOGIE MOT

Q-ETIO MOT

COMMENTAIRE TEXTE

FIN

FIN

FIN

Le diagnostic élémentaire est une entité : pour un même malade, on peut donc, à la création, le répéter autant de fois qu'on le désire.

Cette entité contient sept caractéristiques : les six premières décrivent, deux à deux, les concepts fondamentaux : affection, topographie et

.../...

étiologie. Chacune de ces caractéristiques a pour valeur sémantique un mot : c'est, lors de la création une chaîne de caractères sans blancs, contenant 16 caractères au maximum. Notons que le terme Q- signifie 'qualificatif de' et que le nom 'LESION' a été préféré à celui d' "affection" car ce dernier contient un nombre supérieur de caractères.

La septième caractéristique 'COMMENTAIRE' permet de décrire, sous forme de texte, toutes les informations concernant le diagnostic, qui n'ont pu être créées à l'intérieur des six caractéristiques précédentes.

La condition SI EXISTE DIAGNOSTIC DE MALADE EN COURS ALORS OPL (PLUS CHEZ PAR OU-BIEN) signifie ceci : lorsque, pour un malade donné, on désire créer plusieurs diagnostics élémentaires, tous ces diagnostics, en dehors du premier, sont forcément précédés par un opérateur logique. Nous n'avons donné à cet opérateur que quatre des six valeurs précédemment définies. En effet nous avons vu qu'il existait une équivalence entre PAR et AVEC à condition d'inverser l'ordre des diagnostics qu'ils relient ("Diagnostic 1 AVEC Diagnostic 2" est équivalent à "Diagnostic 2 PAR Diagnostic 1").

L'opérateur POUR ne nous a pas semblé indispensable dans cette première version. Nous utilisons à sa place l'opérateur CHEZ. Notons que pour les opérateurs ET et OU, nous avons dû employer les mots PLUS et OU-BIEN, les mots ET et OU étant interdits à l'utilisateur comme valeur d'une caractéristique car ils font partie du langage de définition du système.

Nous réservons pour la troisième partie de notre propos les conclusions pratiques que nous avons pu tirer des expériences de création de diagnostic selon la structure présentée.

La définition formelle d'un langage de description du diagnostic conformément à la méthodologie présentée par PH. Duquesnel, et l'utilisation du système conversationnel mis au point à l'IMAG nous ont posé un certain nombre de questions sur l'évaluation ergonomique d'un système de gestion du dossier médical. Sans prétendre être exhaustif, ne serait-ce que par notre manque de connaissances sur les contraintes technologiques de l'informatique, nous pensons pouvoir présenter les différents paramètres qui permettent d'évaluer, du point de vue médical, un tel système.

<u>Symbole</u>	<u>Mot</u>	<u>nombre de mots</u>
MI	Mot d'incidence	3
MPR	Mot de position radiologique	5
MTA	Mot de Topographie Anatomique	228
MTD	Mot de Topographie descriptive	61
MTER	Mot de Technique d'Examen radiologique	30
MTOR	Mot de Technique opératoire radiologique	25
NR	Nom propre de radiologue	15
OPL1	Opérateur logique 1	1
OPL2	Opérateur logique 2	1
OPL3	Opérateur logique 3	1
Q	Qualificatif	30
QG	Qualificatif gauche	15
CS	Opérateur de centrage	1
QTOR	Qualificatif de Technique Opération radiologique	10
NOMB	Nombre	
OPM	Opérateur de moyen	2

Figure I - Nombre de valeurs terminales utilisées
dans le langage normalisé en radiologie -

CHAPITRE II - ERGONOMIE DE L'UTILISATION D'UN SYSTEME INFORMATIQUE
DE GESTION DES DOSSIERS MEDICAUX

I - CARACTERES GENERAUX D'UN SYSTEME AUTOMATISE DE GESTION D'UN FICHIER MEDICAL.

1 - FONCTIONS

Dans la description de ces fonctions nous reprendrons les termes utilisés pour le système de l'IMAG, chaque fois que possible.

1-1. Structures algébriques à la disposition du médecin :

Pour enregistrer les données qu'il possède, le médecin doit choisir une structure algébrique et définir la structure sémantique des données médicales en fonction de cette structure algébrique. Nous ne décrirons pas la structure interne de ces différentes structures algébriques, mais uniquement l'aspect qu'elles revêtent aux yeux de l'utilisateur.

a - Systèmes "à chaînes de caractères"

- Le plus simple est la chaîne de caractères c'est-à-dire le texte.
L'utilisateur peut enregistrer le texte qu'il désire.
- Un certain nombre de caractères peut être isolé, ils formeront des mots appelés 'mots-clés' ou descripteurs. C'est par ces mots-clés que l'information stockée est caractérisée. Une partie de l'information peut être enregistrée sous une autre forme, par exemple un texte. Le mot-clé correspondant à l'information constitue l'étiquette qui permet de repérer la totalité de l'information à l'intérieur du fichier. L'ensemble des mots-clés constitue un dictionnaire (ou thésaurus).

Pour augmenter les possibilités d'utilisation de ces mots-clés, on peut définir des synonymes pour chaque mot-clé. Ces mots-clés peuvent être reliés entre eux par une syntaxe plus ou moins complexe.

.../...

b - les codes :

Chaque information reçoit un code qui peut être numérique, ou alphabétique ou alphanumérique. Chaque code correspond à un terme défini, l'ensemble des termes avec leur code constitue un dictionnaire.

Un certain nombre de règles peuvent établir des relations entre des codes (par exemple, dans le code SNOD (8) les 3 premiers chiffres d'un code représente la topographie du diagnostic, les 3 autres, séparés par un trait d'union des premiers représentent l'étiologie).

c - les questionnaires :

L'information est stockée sur des fichiers qui supportent des listes finies de termes. Seuls les termes de ces listes peuvent être utilisés. Ils constituent un dictionnaire.

d - Le système décrit à l'IMAG ne sera pas repris ici.

e - Le langage normalisé :

C'est le langage étudié par Monsieur Duquesnel. Dans ce langage, les mots, que l'utilisateur emploie tels quels, sont affectés d'une valeur syntaxique et d'une valeur lexicale. La valeur syntaxique correspond aux règles de la grammaire de reconnaissance des formes définie pour ce langage. La valeur lexicale représente en quelque sorte le sens du mot, elle découle de l'analyse morphologique du mot.

L'ensemble des mots constitue le dictionnaire, qui est variable.

1-2. Création des données :

Cette fonction consiste en l'initialisation puis la mise à jour du fichier dossier médical : initialisation, c'est-à-dire entrée dans l'ordinateur de nouveaux dossiers médicaux, mise à jour c'est-à-dire enregistrement de nouvelles informations pour un dossier déjà créé (par exemple un nouvel examen de laboratoire, l'apparition d'un nouveau symptôme) ou bien rectification d'une information erronée ou caduque (par exemple modification du diagnostic).

1-3. Interrogation du fichier

L'extraction de données du fichier peut se faire dans deux buts très différents :

- la surveillance clinique : l'extraction porte sur un enregistrement bien déterminé, le fichier médical remplace alors le dossier du malade traditionnel avec ses nombreuses pages d'observation, les différents résultats des examens paracliniques, le tout souvent en désordre !
- la recherche : c'est le domaine de la documentation automatique qui nécessite un accès à l'ensemble du fichier.

2 - CARACTERISTIQUES ERGONOMIQUES DE CES FONCTIONS POUR UN MEDECIN

2-1. Dans la définition des données :

Nous retenons trois caractéristiques essentielles :

a - L'indépendance entre la structure logique des données et la structure des données dans le fichier :

la logique de la structure algébrique choisie doit permettre de représenter l'enchaînement, du point de vue médical, des différents aspects d'une maladie, en laissant complètement dans l'ombre la manière dont les informations sont stockées en mémoire. Par exemple, le médecin ne doit pas se préoccuper davantage de la structure de stockage des informations sur un support magnétique, qu'une femme au volant ne doit se soucier du nombre et de la forme des pistons de la voiture qu'elle conduit !

b - La possibilité de modifier la structure des données sans détruire le fichier créé : en effet, la définition d'une structure de données ne peut être envisagée par le médecin comme un acte définitif. L'observation varie en fonction des découvertes médicales et des changements du mode de vie des patients. Tel médecin, qui se préoccupe actuellement du rôle de l'armature des

.../...

soutiens-gorge dans l'étude des affections mammaires (19), devra peut-être se demander dans quelques années quelle est l'importance de la composition des textiles utilisés. Il serait très regrettable que le rajout d'une nouvelle rubrique dans la structure-type rende inutilisable les enregistrements déjà stockés.

c - La simplicité dans la définition de données :

avant de définir une structure-type, l'utilisateur récapitule les problèmes qu'il étudiera avec le fichier qu'il va créer. Si l'on désire, en médecine, disposer d'une structure-type qui réponde aux vœux de chaque spécialiste, on conçoit aisément combien les questions posées au système seront complexes et variées. Il est donc souhaitable que la structure-type permette de répondre à toutes ces questions. On sera tenté de définir une structure très vaste : la définition et le maniement de cette structure risquent d'être malaisés.

2-2. La création des données :

Pour ce mode d'utilisation de l'ordinateur, il nous semble important d'envisager les points suivants : le contrôle de la validité des informations, la simplicité d'utilisation du système, l'absence d'ambiguïté dans les informations enregistrées.

a - Le contrôle d'erreurs : ce contrôle peut avoir lieu à différents niveaux. Au niveau du vocabulaire utilisé, il faut pouvoir vérifier que les termes employés sont bien des termes médicaux. Au niveau de l'exactitude d'une information, il faut contrôler si la donnée effectivement enregistrée est bien celle que l'on désire stocker, ou si une erreur d'enregistrement a entraîné le stockage d'une donnée erronée.

b - L'absence d'ambiguïté : à un terme, deux médecins peuvent donner un sens différent, inversement il arrive que l'on désigne la même information par deux termes distincts. Ce risque d'ambiguïté, s'il existe dans le système utilisé, peut fausser le travail de recherche. Par exemple, l'estimation de l'intensité d'un ictère, dans un dossier de gastroentérologie peut se faire de la façon suivante : "Faible = 1, Moyen = 2, Fort = 3". (10)

Les médecins noteront-ils tous la même intensité pour un malade donné ?

.../...

On peut en douter !

Ce problème pose celui de la standardisation du langage médical : il serait souhaitable que les médecins soient tous d'accord sur le sens précis des termes qu'ils emploient, et que les critères d'appréciation subjectifs tels que "faible, moyen, fort" soient remplacés par des valeurs objectives, autant que possible mesurables. Cette question est, hélas, loin d'être résolue !

c - La facilité et la souplesse de création :

tout en répondant aux caractéristiques précédentes, il faut qu'un système informatique soit d'une utilisation facile, sinon le médecin aura vite tendance à trouver trop lourd cet auxiliaire de son travail. La manière de créer les données prend ainsi beaucoup d'importance. Que le médecin se trouve devant une machine à écrire, qu'il ait un questionnaire à cocher ou des cartes à perforer, qu'il doive écrire sur un écran cathodique avec un crayon lumineux, son souhait reste toujours le même : une création rapide avec des contraintes d'utilisation du terminal faciles à apprendre et à retenir, et peu de contraintes de langage.

3 - L'INTERROGATION DES DONNEES

a - S'il s'agit du travail clinique la donnée recherchée doit être obtenue avec rapidité et être complète ; on peut souhaiter par exemple obtenir le résultat d'un examen radiologique : celui-ci comporte non seulement le diagnostic résumé en une phrase, mais aussi le commentaire descriptif de l'image. Cela impose un système conversationnel, c'est-à-dire un système grâce auquel médecin et ordinateur peuvent dialoguer, et qui apporte une réponse rapide. Il faut que le temps nécessaire à l'obtention d'une réponse soit presque immédiat : si l'on acceptait sans mot dire de passer des heures dans un service d'archives traditionnel à la recherche d'un renseignement, on trouve le temps long en attendant une minute ce même renseignement devant un terminal !

b - Dans la documentation automatique le problème est bien différent : le temps de réponse a peu d'importance, il est possible de travailler en "batch processing" (cela signifie que l'on pose en même temps une série de questions

.../...

sur cartes perforées à l'ordinateur, qui donne les réponses au bout de plusieurs heures). Mais on s'attachera à obtenir des réponses aussi précises que possible. Cette finesse d'interrogation dépend du nombre d'étiquettes choisies dans la définition de la structure (par étiquette, nous entendons ici aussi bien les mots-clés, que les termes d'un questionnaire, ou les numéros d'un code...) Enfin, l'absence d'ambiguïtés est très importante. Celles-là entraînent dans la recherche soit des bruits, c'est-à-dire des réponses ne correspondant pas à la question (erreurs par excès) soit des silences lorsque des réponses correctes manquent (erreur par défaut).

II - ETUDE DE DIFFERENTS SYSTEMES :

En fonction des caractéristiques que nous venons de définir, nous avons étudié les différents systèmes possibles pour exprimer un diagnostic.

1 - SYSTEMES DE TEXTES :

Le texte libre existe dans un certain nombre de dossiers médicaux (De Heaulme et coll. en psychiatrie(5)). Il donne au médecin une liberté d'expression totale: celui-ci peut enregistrer les données formulées comme il le désire, sans limitation de longueur. On conçoit l'intérêt d'un tel système pour le diagnostic. En effet, un diagnostic radiologique par exemple, peut être très incomplet si l'on se contente d'une seule phrase telle que 'ulcère du bulbe duodénal'. Le compte-rendu du radiologue précisant la taille de cet ulcère, les déformations adjacentes, le retentissement sur le transit digestif, etc... peut être utile à connaître. L'emploi d'un texte peut sembler très séduisant.

Cependant, à une telle facilité d'utilisation correspond une très grande difficulté d'exploitation. D'une part, il n'existe évidemment aucun contrôle d'erreurs, puisque le texte dépend totalement de celui qui le crée. D'autre part, on ne peut s'en servir, en mode "interrogation de données", que s'il porte une étiquette permettant de le repérer. Cette exploitation reste d'ailleurs strictement limitée à un domaine clinique, où l'on se contente de la restitution ad integrum du texte enregistré. Mais toute documentation automatique est impossible, puisque rien ne permet de faire un tri à l'intérieur d'un texte, et à

.../...

plus forte raison, de comparer deux textes. Ce système ne permet donc pas à lui seul, une exploitation satisfaisante du diagnostic par ordinateur. Il reste cependant une structure d'appoint, qui apporte les précisions que l'on ne peut enfermer dans une structure limitée.

2 - MOTS-CLES

Ce système a été expérimenté pour les dossiers médicaux des malades hospitalisés traités par l'Assistance Publique à Paris (6) et sert de base à la bibliographie médicale américaine (7).

Le médecin choisit, dans les textes constituant le dossier médical, les mots qui lui semblent significatifs. L'ensemble de ces mots est rassemblé dans un dictionnaire, qui augmente à chaque nouveau mot-clé et n'a pas de limites imposées. Les mots synonymes sont notés, les termes de signification voisine sont regroupés dans une liste thématique .

Ce système est évidemment d'un maniement très facile pour le médecin. Celui-ci écrit les mots-clés en français, il n'a besoin de connaître aucun code. Lorsqu'il ignore si le terme qu'il emploie existe dans le dictionnaire des mots-clés, il peut consulter le dictionnaire, mais cette consultation n'est pas obligatoire à chaque création de données. Le médecin dispose ainsi d'un mode d'expression simple, et très souple, puisque le choix des termes retenus ne dépend que de lui.

En contrepartie de cette souplesse, il faut souligner une absence totale de contrôle d'erreurs. Non seulement le médecin peut faire des erreurs de vocabulaire, mais il peut encore négliger de créer telle ou telle donnée indispensable à la bonne compréhension des informations. Par exemple si un diagnostic comporte uniquement "ulcère", il n'aura aucun sens ; mais le médecin peut oublier de préciser sa topographie, si rien ne l'y incite. Qu'un certain nombre de mots aient le même sens, peu importe, puisque les synonymes sont répertoriés dans le dictionnaire. Mais certains termes peuvent désigner différentes choses, tels que 'coronarité', antrite, médullaire... Un manque de précision risque de fausser le travail clinique. Surtout, la recherche peut perdre toute valeur, si les "bruits" et les "silences" entraînés par les confusions de termes, sont

.../...

nombreux. Le risque est encore augmenté par le fait que chaque médecin choisit comme mot-clés le terme qui lui semble, personnellement, significatif. En somme, ce système simple et souple, perd beaucoup de son intérêt en raison des gros risques d'erreurs qu'il contient.

3 - CODES

Les codes sont très répandus aux USA : citons par exemple l'OMS (13), le SNOP (9), le SNOD (8) qui traitent du diagnostic. Ils ont été repris en France par Chevallier (6), par Meary en orthopédie (12). H. Sarles et son équipe ont créé un code alphabétique (10).

Les mots, ou plutôt les ensembles de mots, sont affectés d'un nombre dont les chiffres expriment certaines relations. Exemple : dans le SNOD, un diagnostic est coté de la façon suivante : le code est composé de deux nombres de trois chiffres séparés par un tiret ; les trois premiers chiffres indiquent la topographie, les trois suivants l'étiologie du diagnostic considéré.

L'ensemble de ces mots est regroupé dans un dictionnaire. La création de ce dictionnaire se fait à priori et non au fur et à mesure de l'utilisation, comme dans le cas précédent. Ce n'est pas cependant un dictionnaire fermé : de nouvelles rubriques peuvent être ajoutées en fonction des besoins.

On voit tout de suite les obstacles que l'utilisateur va rencontrer : la création du dictionnaire est longue si le code se veut complet. Il est pourtant évident qu'il ne pourra contenir tout le vocabulaire médical avec ses nuances. La rubrique "autre" permet d'enregistrer tout ce qui n'est pas codé. Dans cette rubrique se trouvent des affections qui, pour être rares, n'en sont pas moins essentiellement différentes.

Si la création est achevée, le médecin qui désire se servir de ce code devra, soit le connaître par coeur, ce qui peut sembler difficile, soit consulter le dictionnaire. Cela rend la création très lourde, même si un certain nombre de dictionnaires possède plusieurs entrées.

.../...

On pourrait penser qu'au prix de cette lourdeur dans la création on obtienne un contrôle d'erreurs assez strict. Le contrôle existe en effet au niveau du vocabulaire médical. Un code faux, c'est à dire ne correspondant pas à un terme du dictionnaire, sera refusé. Mais il n'y a pas toujours un contrôle dans l'exactitude des données à l'intérieur des termes admis. Une erreur de chiffre peut entraîner une erreur importante dans un enregistrement qui sera accepté. (Par exemple, dans le code SNOD, le code 350-100 signifie bronchite aiguë, le code 450-100 correspond à l'endocardite bactérienne). Dans certains cas cependant, il existe un système de contrôle pour l'exactitude de chaque donnée : c'est la paraphrase. Le médecin tape le code qu'il désire, la machine répond en donnant le terme auquel correspond le code inscrit. Cette méthode est employée à Grenoble, dans l'aide au diagnostic en toxicologie (11) : si le médecin tape "35", l'ordinateur répond "vomissements" ; le médecin peut donc contrôler l'information qu'il a créée. Ce contrôle d'erreurs qui n'est d'ailleurs possible que si le système est utilisé en conversationnel augmente encore la longueur de la création. De même l'interrogation, qui se fait en "batch", le plus souvent, est beaucoup plus malaisée qu'avec des termes en langage naturel. Le code n'est donc pas un instrument de choix dans la clinique, mais il peut rendre de grands services dans la recherche.

4 - LES QUESTIONNAIRES

Ils connaissent une grande vogue en Europe. Citons le dossier médical à l'Institut Gustave Roussy (17), les travaux de Bernier (14), Broustet (16), en France, ceux de Besson à Lausanne (15), les questionnaires suédois (18). La définition de tels systèmes n'a jamais été faite pour un domaine très général ; chacun des auteurs les a expérimentés uniquement pour sa spécialité, en fonction de sa région géographique (le choix des ethnies, dans les antécédents personnels et familiaux, est significatif !). Si l'on envisageait l'utilisation d'un tel système pour l'expression du diagnostic, sa définition en serait très longue. En outre, la plupart des questionnaires se présentent sous forme de listes finies. L'évolution de la pathologie obligerait à une refonte complète du système. Lors de la création de données, la manipulation du système est très lourde. Quel que soit le système d'entrée (cartes perforées, feuilles de marque

lues par un lecteur optique...), le médecin doit utiliser le cahier regroupant l'ensemble du questionnaire. Dès que le problème traité est étendu, ce cahier devient très encombrant.

La méthode n'est pas exempte d'erreurs. Il est, bien sûr, impossible d'utiliser des termes qui ne font pas partie du vocabulaire médical, mais à l'intérieur de celui-ci, les erreurs sont faciles. Elles sont de deux ordres. D'une part, si la rubrique que l'on désire noter n'existe pas telle quelle dans le questionnaire, on peut être tenté ou de ne pas créer l'information, ou de créer une information fausse par confusion des termes dont on dispose. D'autre part, les erreurs lors de la perforation ou du marquage des feuilles ne sont pas détectées. Les limites du système rendent son utilisation très difficile en clinique. Le fait d'utiliser des listes finies enlève beaucoup de souplesse à l'expression, même si l'on dispose de dictionnaires volumineux. En outre, dans un certain nombre de systèmes, le nombre de termes à choisir est limité (un pour une liste par exemple), et le nombre de résumés peut lui-même être limité (on dispose par exemple de huit ou dix résumés par malade, en fonction de la place disponible en mémoire). L'évolution de l'observation ne peut donc pas être rendue de façon satisfaisante. On voit donc bien que de tels systèmes ne sont pas satisfaisants pour le travail clinique car ils ne peuvent jamais se substituer au dossier traditionnel. Dans la recherche, ce seraient, selon leurs auteurs, de bons outils de travail. Cependant, le risque important d'erreurs, l'absence d'évolutivité du système, permettent de se demander quelle est la valeur d'une recherche fondée sur de telles bases !

5 - LE LANGAGE NORMALISE

Dans l'esprit de son auteur (2), ce langage, étudié pour les demandes d'examen radiologiques, devait pouvoir être étendu à l'expression du diagnostic. Il devait être intégré dans un fichier hybride : le langage normalisé permettrait la rédaction d'étiquettes, qui introduiraient un enregistrement plus précis et plus complet en langage clair.

Nous avons vu (Chapitre I) que la définition de la grammaire n'offrait pas de trop grandes difficultés. Mais la constitution du dictionnaire doit être préalable à l'utilisation du système (précisons cependant qu'il existe des possibilités de mise à jour). En effet ce dictionnaire contient un arbre des bases (ce qui correspond à une liste des mots) et une table de traduction qui comporte pour chaque mot des renseignements de nature syntaxique, des renseignements de nature morphologique, l'adresse du synonyme de référence, et enfin des rensei-

gnements permettant l'étude d'inclusions du niveau du mot (par exemple : acromion est inclus dans épaule, estomac est inclus dans abdomen).

Cela implique que chaque mot ait une valeur syntaxique précise, qu'il n'y ait aucune confusion possible sur son sens : par exemple dans les demandes d'examen radiologique, un mot de technique d'examen radiologique ne peut être que cela. Un mot ne doit correspondre en principe qu'à un seul concept. Dans les demandes d'examens radiologiques, les exceptions étaient rares : par exemple, angiographie contient à la fois un mot de technique radiologique (graphie) et un mot de topographie anatomique (angio). Ces cas étaient résolus par des règles de programmation particulières, ce qui était possible dans ces cas très limités.

Pour l'expression du diagnostic, de tels cas sont légion : il suffit de penser à tous les mots composés avec le suffixe "ite" comme "gastrite", "entérocolite", etc...

Une deuxième difficulté dans la définition du système est la structure de la grammaire. Celle-ci est définie une fois pour toutes. Si l'on se rend compte à posteriori que certaines règles sont inutiles ou mal conçues, il faut redéfinir tout le système.

On peut regretter ces difficultés de définition, car c'est certainement, pour l'utilisateur, un système très performant tant pour la création que pour l'interrogation des données.

Pendant la création, le contrôle d'erreurs est très strict ; il existe bien sûr au niveau du vocabulaire médical, puisque seuls les termes contenus dans le dictionnaire sont admis. Mais il existe aussi au niveau de chaque mot qui n'est accepté que si l'on respecte sa valeur syntaxique. D'autre part, les fautes d'orthographe sont "corrigées" puisqu'un mot mal écrit est refusé, comme un mot inconnu. Cependant le contrôle n'est pas total : des mots syntaxiquement corrects, dans un ordre correct, sont acceptés même si le sens de la phrase qu'ils forment est absurde. Le risque de confusion entre termes de même valeur syntaxique persiste.

.../...

La création elle-même est simple : le langage est très proche du français, et correspond à la logique médicale. Il est assez souple, étant donné l'existence des synonymes, et la possibilité d'exprimer les relations entre les différents concepts grâce aux règles de grammaire.

A l'interrogation, c'est également un système très satisfaisant. Le fichier hybride correspond bien aux nécessités de la clinique. Pour la documentation automatique, ce système présente un gros avantage : il existe un algorithme de comparaison qui permet de comparer deux phrases contenant un certain nombre d'informations communes, même si elles ne sont pas identiques. Par exemple, "radiographie du bras" et "radiographie du bras gauche" sont deux phrases comparables, bien que syntaxiquement différentes. De même "radiographie de l'hypocondre droit" et "radiographie de l'abdomen" sont deux phrases comparables. La comparaison a lieu aussi bien pour la syntaxe que pour la sémantique. On voit donc les nombreux avantages de ce système pour la création et la manipulation de données. Il est regrettable que les difficultés inhérentes à sa définition soient un obstacle sérieux à son exploitation.

De cette étude de différents systèmes, nous pouvons tirer quelques conclusions : la plupart de ces systèmes (codes - questionnaires - mots-clés - langage normalisé) sont entièrement dépendants de ceux qui les définissent. Cela peut sembler une lapalissade, mais on ne peut négliger cet aspect en médecine : en effet, la précision, l'exactitude, la logique des données dépendent de la définition du système. A l'inverse, s'il n'existe pas de structure définie (textes) la fiabilité des données enregistrées est très discutable. Enfin, aucun de ces systèmes ne permet un contrôle d'erreurs total. Dans une dernière partie, nous allons étudier comment le système en cours d'étude à l'IMAG répond aux caractéristiques ergonomiques que nous avons définies.

CHAPITRE III - CARACTERISTIQUES ERGONOMIQUES

DU SYSTEME DEFINI A L'IMAG

La structure que nous avons définie (cf. Chapitre I) a été expérimentée par un très petit nombre de malades (50). Cela nous a permis d'appréhender un certain nombre de difficultés. Dans une étude des possibilités évolutives du système, nous verrons comment répondre à ces critiques.

Y A-T-IL D'AUTRES MALADES	A CREER ?
oui	
NOM	
be...	
PRENOM	
solange	
AGE	
36	
SEXE	
feminin	
CREATION DE DIAGNOSTIC	?
oui	
LESION	
pneumopathie	
Q-LESION	
aigue	
TOPOGRAPHIE	
poumon	
Q-TOPO	
gauche	
ETIOLOGIE	
grippe	
Q-ETIO	
probable	
TEXTE:COMMENTAIRE	?
non	
Y A-T-IL D'AUTRES DIAGNOSTICS	A CREER ?

Figure 2 - Exemple de création -

L'intitulé du diagnostic est le suivant : "Pneumopathie aigüe grippale probable -
(Poumon gauche"

BA
AUGUSTINE
75
FEMININ
CARCINOME
MALPIGHIEN
UTERUS
COL
CANCER

Figure 3 - Exemple d'interrogation -

Interrogation portant sur une malade dont le diagnostic est : "carcionome malpighien du col utérin" - Notons que "col" est ici un "qualificatif topographique" et "cancer" une "etiologie".

AB
MARIE-LOUISE
80
FEMININ
HERNIE
HIATALE
AND
DIVERTICULOSE
DUODENUM
AND
LITHIASE
BILIAIRE
AND
ANEMIE
HYPOCHROME

* WEBER POSITIF **

Figure 4 - Exemple d'interrogation -

Interrogation d'un dossier dont le diagnostic est "Hernie hiatale, diverticulose duodénale, lithiase biliaire, anémie hypochrome - Weber positifs".

malade ayant existe diagnostic ayant lesion ='cirrhose' ET

_etilogie ='alcool' ; ; ?

REPONSE :

4 MALADES

QUE VOULEZ VOUS DENOMBRER ?

_malade ayant lesion de un diagnostic ='cirrhose' ; ?

REPONSE :

5 MALADES

QUE VOULEZ VOUS DENOMBRER ?

_malade ayant existe diagnostic ayant lesion ='cirrhose' et

_etilogie ≠ 'alcool' ; ; ?

REPONSE :

2 MALADES

QUE VOULEZ VOUS DENOMBRER ?

_malade ayant existe diagnostic ayant lesion ='cirrhose' ; ; ?

REPONSE :

6 MALADES

QUE VOULEZ VOUS DENOMBRER ?

i
_QUE VOULEZ VOUS SAVOIR ?
_etilogie de diagnostic ayant lesion ='cirrhose' et etilogie =
'alcool' ; de malade ?

REPONSE :

ETIOLOGIE	POST-HEPATITE
ETIOLOGIE	NON-ALCOOLIQUE

QUE VOULEZ VOUS SAVOIR ?
_non

Figure 5 - Exemple d'interrogation -

L'interrogation porte sur l'ensemble du fichier.

Les questions doivent se comprendre de la façon suivante :

- 1 - Combien de malades ont pour diagnostic "cirrhose éthylique" ?
- 2 - Combien de malades ont pour premier diagnostic une "cirrhose" ?
- 3 - Combien de malades ont une cirrhose non éthylique ?
(le signe ≠ signifie "différent de")
- 4 - Combien de malades ont une cirrhose ? (question dont la réponse est différente de celle de la seconde question, car certains malades ont comme premier diagnostic un diagnostic autre que cirrhose.)

Ces quatre questions sont posées par l'intermédiaire de la fonction "dénombrement".

La cinquième est la suivante : quelle est l'étiologie des cirrhoses non éthyliques ?

QUE VOULEZ VOUS SAVOIR ?

_etiologie de diagnostic de malade ayant age <45 et lesion

_de diagnostic ='cirrhose' ;?

REPONSE :

ETIOLOGIE

ALCOOL
LITHIASE

ETIOLOGIE

ALCOOL

QUE VOULEZ VOUS SAVOIR ?

_sexe de malade ayant age <45 et lesion de diagnostic =

_ 'cirrhose' et etiologie de diagnostic ='alcool' ;?

REPONSE :

SEXE

MASCULIN

SEXE

MASCULIN

QUE VOULEZ VOUS SAVOIR ?

_diagnostic de malade ayant age <45 et lesion de

_diagnostic ='cirrhose' et etiologie de diagnostic =

_ 'alcool' ;?

REPONSE :

DIAGNOSTIC 1

LESION

TOPOGRAPHIE

ETIOLOGIE

CIRRHOSE

FOIE

ALCOOL

DIAGNOSTIC 2

OPL

LESION

TOPOGRAPHIE

AND

LITHIASE

PANCREAS

DIAGNOSTIC 3

OPL

LESION

TOPOGRAPHIE

ETIOLOGIE

AND

PANCREATITE

PANCREAS

LITHIASE

DIAGNOSTIC 1

LESION

TOPOGRAPHIE

ETIOLOGIE

CIRRHOSE

FOIE

ALCOOL

DIAGNOSTIC 2

OPL

LESION

TOPOGRAPHIE

AND

PSORIASIS

PEAU

DIAGNOSTIC 3

OPL

LESION

topographie

Q-TOPO

AND

URETHRITE

urethre

CERVICAL

QUE VOULEZ VOUS SAVOIR ?

_etiologie de diagnostic ayant lesion ='cirrhose';de

_malade ayant age <45 ;?

REPONSE :

ETIOLOGIE

ALCOOL

ETIOLOGIE

ALCOOL

QUE VOULEZ VOUS SAVOIR ?

_non

Figure 6 - Exemple d'interrogation -

- 1ère question : quelle est l'étiologie des diagnostics des malades de moins de 45 ans ayant une cirrhose ?
(la réponse comporte trois étiologies)
- 2ème question : quel est le sexe des malades de moins de 45 ans ayant une cirrhose éthylique ?
- dans la 3ème question, on demande de donner intégralement les diagnostics de ces malades de moins de 45 ans porteurs d'une cirrhose éthylique : la réponse montre qu'il existe bien deux malades, l'un d'eux ayant deux diagnostics où l'étiologie soit précisée ("alcool" et "lithiase") ce qui correspond aux deux premières étiologies de la question 1. Cette première question était mal posée : on pose alors la question suivante : quelle est l'étiologie des cirrhoses chez les malades de moins de 45 ans ?

On voit, par cet exemple, les difficultés de l'interrogation pour le médecin qui doit rigoureusement suivre la logique du système, et les possibilités de contrôle des erreurs d'interrogation.

I - STRUCTURE ACTUELLE

1 - DEFINITION DE DONNEES

La structure-type (cf. chapitre I) est extrêmement simple. Il n'a été défini qu'une entité, l'entité diagnostic : il est bien évident que pour un malade, on peut exprimer plusieurs diagnostics. Les concepts fondamentaux sont sous forme de caractéristiques. Donc une entité diagnostic ne peut comporter qu'une seule lésion, une seule topographie, une seule étiologie. On verra les conséquences de ce fait sur la création de données. La structure logique des données est complètement indépendante de leur structure dans le fichier. Le médecin ne se préoccupe ni de la structure interne, ni de la taille des mémoires. Lors de la définition de structure, il doit simplement respecter les règles d'écriture du langage, peu nombreuses et faciles à apprendre. Il faut noter cependant que cette structure une fois définie, aucune possibilité de modification n'existait sans détruire le fichier.

2 - CREATION DE DONNEES

La création se fait sur un terminal type machine à écrire. L'ordinateur pose les questions correspondant à chaque caractéristique. Le médecin tape la réponse selon la nature sémantique qu'il a définie ; par exemple pour "lésion = mot" : le médecin doit écrire une chaîne de caractères sans blanc comportant 16 caractères au plus (pour des raisons techniques). Nous avons dû choisir certaines abréviations pour des expressions composées trop longues. Exemple : "syndrome oedémato ascitique" a été enregistré sous la forme "sd - oedematoascit."

Nous voyons ici un des écueils de la structure actuelle : comme nous n'avons défini, comme valeurs sémantiques, que des "mots" et un "texte", l'utilisateur est totalement libre d'enregistrer ce qu'il désire. Il n'existe aucun contrôle de la validité du langage médical et de l'exactitude des informations.

D'autres avantages et inconvénients sont inhérents à la forme de dialogue entre médecin et ordinateur. Ce n'est pas le médecin qui impose le résultat de ses observations à la machine, c'est la machine qui interroge et le médecin doit répondre à la question posée, pour que la séquence des questions continue à se dérouler.

C'est un double avantage :

- si le médecin avait tendance à oublier certains aspects de l'information (par exemple une étiologie), l'ordinateur lui rappelle toutes les données qu'il est utile de rechercher et de connaître.

- il est possible d'indiquer comme réponse l'absence d'une caractéristique (par exemple : "étiologie = inconnue"), ce qui peut être une information aussi utile en médecine que la présence d'un signe.

Mais c'est aussi un inconvénient :

- le médecin doit répondre à toute la séquence de questions, en écrivant "indéfini" pour les caractéristiques inconnues. Cela augmente considérablement le temps de création. *

- or ce temps de création est déjà long puisque l'on doit dactylographier les réponses, et qu'un médecin n'a pas la formation d'une bonne secrétaire ! Cette dactylographie permet un certain contrôle d'erreurs : on risque moins d'écrire en toutes lettres "endocardite" pour "bronchite" que de frapper un 4 à la place d'un 3 ! Mais si le nombre des données à créer est élevé, on risque, à la longue, de multiplier les abréviations abusives ou les fautes d'orthographe. Cela aura de grosses conséquences sur la recherche.

* notons cependant que l'on tape 'u', ce que l'ordinateur traduit par 'indéfini'.

.../...

Un dernier aspect de la création dépend de la structure-type du fichier diagnostique. Les concepts étant sous forme de caractéristiques, si l'on veut enregistrer un diagnostic contenant plusieurs concepts de même valeur syntaxique, on doit créer autant de diagnostics qu'il existe de concepts. Par exemple "polypose duodéno-jéjunale" devient à la création "polypose duodénale, puis, polypose jéjunale".

S'agit-il d'une qualité ou d'un défaut de la structure actuelle ? On a renoncé aux entités pour chaque concept, car à chacune d'elle, il aurait été nécessaire de "boucler" l'interrogatoire avant de poursuivre. La séquence aurait été la suivante :

CREATION DE DIAGNOSTIC ?

oui

CREATION DE LESION ?

oui

-ulcère

Y A-T-IL D'AUTRES LESIONS A CREER ?

non

CREATION DE Q-LESION ?

oui

-perforé

Y A-T-IL D'AUTRES Q-LESION A CREER ?

non

:
:
:
:
:
:
:

Pour chaque diagnostic, le temps de création serait ainsi très allongé. Mais cela permettrait de créer autant de concepts qu'il en existe dans une phrase de diagnostic. Par exemple, il arrive qu'une lésion ait plusieurs qualificatifs: "ulcère perforé bouché". Pour la création, on se trouve devant l'alternative suivante : enregistrer "perforé-bouché" comme un seul qualificatif, à la rubrique Q-LESION, ou bien enregistrer le premier qualificatif "perforé" comme

.../...

tel, et le second "bouché" dans le commentaire. Du choix de la forme d'enregistrement dépendra la documentation automatique.

Cette structure-type raccourcit le temps de création au prix d'une limite des possibilités d'expression (même si le commentaire permet d'enregistrer tout ce que les autres caractéristiques ne pouvaient contenir). L'étendue du fichier actuel ne permet pas d'affirmer s'il est préférable de conserver cette structure telle quelle ou bien d'y multiplier les entités. Seule une utilisation à plus grande échelle, au cours de laquelle on pourrait mesurer la fréquence d'apparition des concepts répétés dans un même diagnostic, peut apporter une réponse à ce problème.

En conclusion, à la création, cette structure-type, ne contient pratiquement pas de contrôle d'erreurs. Si la dactylographie entraîne le médecin à mieux contrôler l'enregistrement de chaque terme du diagnostic, celui-là ne dépend que du créateur : chaque médecin enregistre ce qu'il désire, sous la rubrique de son choix. A l'usage, cela se traduit de la manière suivante : le médecin analyse le contenu du diagnostic selon ses connaissances personnelles, si bien que tel terme, enregistré par l'un comme qualificatif topographique, peut être considéré par un autre comme topographie. Exemple : "carcinome malpighien du col utérin" a été enregistré comme sur la figure (3).

- col est un qualificatif topographique (analogue à la topographie descriptive selon Ph. Duquesnel)
- utérus est la topographie (topographie anatomique selon Ph. Duquesnel).

Un médecin non prévenu aurait pu inverser l'enregistrement de ces deux termes.

Cette absence de contrôle tant au niveau du vocabulaire médical en général, que de l'analyse, concept par concept, d'un diagnostic particulier, a de grosses conséquences dans le domaine de la recherche.

C'est le reproche essentiel que l'on peut formuler à l'égard de la structure actuelle, dont l'utilisation, lors de la phase de création, reste très simple, quoiqu'un peu lente.

3 - INTERROGATION DES DONNEES

- Pour le travail clinique, cette structure semble satisfaisante :

le médecin peut obtenir rapidement, sous une forme claire, tout ou partie d'un diagnostic concernant un malade donné. Cette interrogation, qui peut être suivie d'une mise à jour, permet de suivre pas à pas l'évolution du malade et de modifier en conséquence l'intitulé du diagnostic (voir fig. (3) (4) (5) (6)....). Notons simplement que les questions doivent être posées conformément à la logique du système.

- La recherche est beaucoup plus difficile : elle est même, pour l'instant, presque impossible.

En effet, dans cette première version, nous ne disposons pas de synonymes. Comme la documentation automatique ne peut se faire que sur les caractéristiques telles qu'elles ont été enregistrées, une erreur orthographique, une modification dans un terme (adjectif employé au lieu du nom, pluriel au lieu du singulier) entraînent un silence.

Par exemple si pour l'étiologie on enregistre une fois "estomac", une autre fois "gastrique", on ne pourra rechercher tous les "ulcères gastriques" qu'en posant la question suivante :

DIAGNOSTIC AYANT LESION = "ULCERE" ET TOPOGRAPHIE = "GASTRIQUE" OU TOPOGRAPHIE "ESTOMAC" ; ?

Il est donc nécessaire que ce soit toujours le même médecin qui pose les questions, ou que chaque médecin utilisateur connaisse l'ensemble du fichier ! On se rend compte de la difficulté d'une recherche dans ces conditions.

4 - CONCLUSION

La structure utilisée actuellement est très satisfaisante pour le clinicien : une création simple, une possibilité d'expression étendue, un interrogatoire facile et rapide, une mise à jour correspondant à la variabilité des informations médicales, donnaient au médecin un outil très adapté à ses besoins.

.../...

Mais la recherche, alors qu'il n'existe aucun contrôle d'erreur, aucune liste de termes de référence, ne peut être entreprise pour l'instant. Nous allons voir quelles modifications peuvent être apportées au système pour augmenter ses possibilités d'utilisation.

II - EVOLUTION DU SYSTEME

De l'étude de la structure actuelle, il ressort que le système doit être surtout amélioré pour la recherche : ceci peut se faire par action sur trois points essentiellement :

- modification de la structure
- introduction de synonymes
- contrôle d'erreurs

I - MODIFICATION DE LA STRUCTURE

Cela sera possible dans la version du système actuellement en cours d'expérimentation (on pourra passer le fichier déjà créé dans la nouvelle structure, en le conservant totalement). Nous ne redétaillerons pas quelles modifications seront nécessaires. Disons simplement qu'il s'agit de l'introduction d'un plus ou moins grand nombre d'entités, qui permettrait le transfert d'informations enregistrées à l'intérieur du commentaire (donc non accessibles à la documentation automatique) dans une caractéristique.

Peut-être sera-t-il également nécessaire d'augmenter le nombre des caractéristiques : nous nous sommes contentés d'étudier les trois concepts fondamentaux, mais il serait peut-être utile de les décomposer pour mettre davantage en évidence les classes décrites dans l'analyse du diagnostic.

2 - INTRODUCTION DES SYNONYMES

Cela deviendra désormais possible grâce à un programme appelé "macrogénérateur". Ce programme s'intercale entre l'organe d'entrée-sortie et le reste du système.

.../...

Le système de fonctionnement est le suivant : dans la phase de création, pour un mot dont on veut préciser les synonymes, on définit une "macro". L'"appel" de la macro est composé du terme synonyme, son "expansion", du synonyme de référence. Exemple : si "foie" est le synonyme de référence, "hépatique" le terme utilisé, l'appel de la macro est "hépatique", son expansion est "foie".

Il est donc nécessaire de définir une nouvelle "macro" pour chaque nouveau synonyme.

Lors des mises à jour, chaque fois qu'un synonyme "appel" sera rencontré, le macrogénérateur permettra l'enregistrement dans le fichier du terme pris comme synonyme de référence.

Lors de l'interrogation du fichier, c'est ce dernier terme qui sera édité. Il faut donc bien noter que le macrogénérateur autorise une plus grande liberté dans la création, mais ne modifie pas le contenu du fichier.

Le macrogénérateur sera peut-être utile aussi pour des termes contenant plusieurs concepts, tel 'gastrite'. On pourrait par exemple définir une "macro" ayant pour appel 'gastrite' et pour expansion 'inflammatoire' comme lésion, et 'estomac' comme topographie. Mais dans ce cas, il ne sera ensuite possible d'interroger le système que sur les inflammations de l'estomac, non sur les gastrites, à moins de définir une nouvelle macro dont l'appel serait = 'gastritique' et l'expansion 'malade ayant existe diagnostic ayant lésion = 'inflammation' et topographie = 'estomac' ''.

3 - CONTROLE D'ERREURS

Ce contrôle pourrait dès maintenant se faire par l'introduction, à la place des "mots", de listes de valeurs pour chaque caractéristique. Ces listes sont ouvertes = à la fin de chacune d'elles existe la valeur "autre", avec la condition :

SI VALEUR = 'AUTRE' , ALORS NOUVELLE VALEUR = 'MOT'.

.../...

Ces listes peuvent être consultées par le médecin qui en ignorerait le contenu. L'utilisateur pourrait donc, si de telles listes étaient créées pour le diagnostic, soit utiliser une valeur connue, soit créer une nouvelle valeur, en fonction de ses besoins.

Ceci correspondrait au contrôle qui existe dans les questionnaires, mais avec un avantage considérable sur ceux-ci : l'existence de listes non finies.

On pourrait même concevoir, à l'intérieur de ces listes, une classification. En effet, la consultation d'une liste de plusieurs milliers de termes serait très longue et fastidieuse. On pourrait donc regrouper les termes par classes, en s'inspirant de la classification établie lors de l'analyse du diagnostic.

Par exemple, si l'étiologie à créer est "staphylocoque", cette valeur se trouve dans la liste des "bactéries" qui se trouve dans la liste des "microbes" inclus dans la liste des "agents infectieux", valeur de la liste "étiologies".

Le médecin qui ne saurait où classer une affection staphylococcique demanderait la liste des valeurs des étiologies. Il repérerait le terme "agent infectieux", demanderait la liste correspondante, repérerait le terme "bactérie"... et ainsi de suite.

De cet exemple, on peut déduire la nécessité d'une modification du terminal.

4 - MODIFICATION DU TERMINAL

Cette modification ne concerne pas le système à proprement parler. Mais l'utilisateur, soucieux de son confort, apprécierait beaucoup plus, pour consulter les listes dans l'exemple ci-dessus, un écran cathodique, où l'affichage de ces listes serait immédiat, qu'une imprimante, même rapide, où la lecture est beaucoup moins aisée. Pour la création, le médecin, qui n'a pas encore vu apparaître la dactylographie au nombre des matières enseignées dans les Facultés de Médecine, préférerait certainement la collaboration d'une charmante secrétaire, à l'utilisation de ses propres doigts ! Mais nous abordons là un aspect très différent de l'utilisation du système, dont l'évaluation financière nous entraînerait très loin de notre propos !

IV - CONCLUSION

L'expression du diagnostic au moyen d'un ordinateur nous a conduit, après une analyse des concepts constituant ce diagnostic, à l'étude de deux systèmes :

- d'une part, une grammaire context-free de reconnaissance des formes qui permettrait l'élaboration d'un langage normalisé.
- d'autre part, un système conversationnel de définition et d'interrogation de données.

La difficulté d'élaboration du langage normalisé, malgré la souplesse qu'il apporterait dans l'expression du diagnostic, nous a fait préférer le deuxième système.

Après une définition de la structure-type selon laquelle serait créé le fichier diagnostic, nous avons étudié les questions que nous posait l'utilisation d'un tel système. Pour ce faire, nous avons, dans une seconde partie, tenté de présenter les différentes caractéristiques ergonomiques permettant à un médecin d'évaluer un système de gestion automatique du dossier médical, puis nous avons, à la lumière de ces paramètres, estimé brièvement les différents systèmes utilisables = texte en langage libre, mots-clés, codes, questionnaires, langage normalisé.

Enfin, dans une troisième partie, nous avons tenté de juger le système que nous avons proposé expérimenté dans sa version actuelle, et nous avons proposé, quelques modifications dans son évolution, pour que ce système, déjà très satisfaisant dans le travail clinique, devienne aussi pour le médecin un auxiliaire précieux dans le domaine de la recherche.

Langage normalisé

(1) Ph. DUQUESNEL - F. PECCOUD

Langage artificiel en radiologie - Fichiers hybrides.

Communication A.F.I.R.O. - Nancy - 1967

J. VALOIS - Ph. DUQUESNEL

Langage médical en radiologie -

L'informatique au service de l'homme p.167-184
Grenoble - Mars 1969

(2) Ph. DUQUESNEL

Langage normalisé pour la description des informations médicales en radiologie.

Thèse de doctorat de troisième Cycle - Faculté des Sciences - Grenoble - Juin 1969

Système défini à l'IMAG

(3) J.R. ABRIAL, G. BEAUME, G. HENNERON, B. LEPAPE, R. MORIN, G. VIGLIANO

Système de définition et d'interrogation de données.

Communication interne à l'IMAG - Grenoble - Mars 1969

(4) S. COHEN

Description d'un système de gestion en temps réel et en mode conversationnel du dossier médical.

Thèse de doctorat en médecine - Grenoble - Juin 1970

- (5) R. AMIEL, M. de HEAULME, Cl. LEROY, J.L. ARMAND-LAROCHE, F. LEBIGRE, J. CARON.

Le recueil et le traitement des informations en
psychiatrie. Exposé d'une méthodologie de la
constitution et de l'exploitation des dossiers.

Symposium de Toulouse - Mars 1968 - p. 343-360.

Mots-clés

- (6) J. CHEVALLIER - G. CHEVALLIER

La fiche médicale du malade hospitalisé et son
exploitation en ordinateur.

Concours médical : 91-26 p. 5637-5644

91-27 p. 5799-5803

- (7) MEDLARS

Medical Literature Analysis and Retrieval System
National Library of Medicine - Bethesda - USA

Codes

- (8) American Medical Association

Standard Nomenclature of Diseases and Operations

5 th Edition - The Blakiston Division -
BC Graw-Hill - Book Co - New-York - 1961

- (9) College of American Pathologist

Systematized Nomenclature of Pathology
First Edition - Chicago - 1965

- (10) J.C. GARDIN, M. BORILLO, R. CROS, A. GAUTHIER, P. BERNARD, H. SARLES

Propriétés d'un système descriptif de signes
médicaux pour l'étude de la pathologie digestive
sur ordinateurs.

Symposium de Toulouse - Mars 1968 - P. 293-300

.../...

(11) M. MATTEI Les ordinateurs à l'aide du diagnostic médical
Application en Toxicologie.

Thèse de doctorat en médecine - Grenoble -
Décembre 1969 -

(12) R. MEARY Code de classement intéressant l'orthopédie et
la Traumatologie -

Masson et Cie - Paris - 1966

(13) Organisation mondiale de la Santé :
Classification statistique internationale des
maladies, traumatismes et causes de décès -
Révision 1965 - Volume 1 - Genève - 1968 -

Questionnaires

(14) J.J. BERNIER, N. VIDON J. MASSELOT
Traitement des dossiers médicaux en ordinateur-
Etude des contraintes de divers systèmes d'en-
trée des informations d'après 500 dossiers de
gastroentérologie -

Symposium de Toulouse - Mars 1968 - p. 301-320

(15) A. BESSON Application d'ordinateurs en médecine -

Dunod - Paris - 1969 -

(16) J.P. BROUSTET, Ph. MOTHES, P. BESSE, H. BRICAUD, P. BROUSTET

Classement, mise à jour et interrogation automa-
tique des dossiers médicaux stockés en ordina-
teurs - Contraintes - Solutions envisagées -
Résultats pratiques.

Symposium de Toulouse - Mars 1968 - p. 361-378

.../...

(17) P. DENOIX

L'électronique, instrument de la logistique hospitalière -

Techniques hospitalières - 1967 - 259 p. 45-50

(18) A.M. FONTAN

Notions sur les machines à traiter l'information et quelques unes de leurs applications au domaine médical -

L'exemple suédois -

Thèse de doctorat en médecine - Lille - 1962 -
Tome II - p. 289-318

(19) Ch. GROS, C. VROUSSOS, J.P. WALTER, R. VERGUES avec la collaboration de M. TREILLET et Y. HUMMEL

Exploitation du dossier sénologique par l'ordinateur pour l'étude des affections mammaires -

Symposium de Toulouse - Mars 1968 - p. 321-336

Vu, le Président de Thèse : G. CABANEL

Vu, le Doyen de la Faculté de Médecine : G. CABANEL

Vu et permis d'imprimer,

le Recteur de l'Académie de Grenoble : M. NIVEAU

SERMENT D'HIPPOCRATE

En présence des Maîtres de cette Faculté, de mes chers condisciples, devant l'effigie d'Hippocrate.

Je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la Médecine.

Je donnerai mes soins à l'indigent et je n'exigerai jamais un salaire au-dessus de mon travail. Je ne participerai à aucun partage clandestin d'honoraires.

Admis dans l'intérieur des maisons, mes yeux ne verront pas ce qui s'y passe, ma langue taira les secrets qui me seront confiés et mon état ne servira à corrompre les mœurs ni à favoriser le crime.

Je ne permettrai pas que des considérations de religion, de nation, de partis ou de classe sociale viennent s'interposer entre mon devoir et mon patient.

Je garderai le respect absolu de la vie humaine.

Même sous la menace, je n'admettrai pas de faire usage de mes connaissances médicales contre les lois de l'humanité.

Respectueux et reconnaissant envers mes Maîtres je rendrai à leurs enfants l'instruction que j'ai reçue de leur père.

Que les hommes m'accordent leur estime si je suis fidèle à mes promesses.

Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.

