

BULLETIN D'INFORMATIQUE APPROFONDIE ET APPLICATIONS

SCIENCE DE L'INFORMATION

COMITE SCIENTIFIQUE

N° 54 - DECEMBRE 1999

*Patrick Abellard
Françoise Adreit
Jalal Almhana
France Chappaz
M'hamed Charifi
Roger Cusin
Bernard Goossens
Patrick Isoardi
Robert Jacquier
Jean - Philippe Lehmann
Nadia Mesli
Patrick Sanchez
Rolland Stutzmann
André Tricot*

DIRECTEUR

Jean - Michel Knippel

REDACTEUR EN CHEF

Edmond Bianco

REDACTEUR ADJOINT

Sami Hilala

REDACTION

Université de Provence
Equipe Hermès. Case 33
3, place Victor Hugo
F - 13331 Marseille Cedex 3
Téléphone: (0)4 91 10 62 30
Télécopie : (0)4 91 50 91 10

DEPOSITAIRE

Université de Provence
Bibliothèque Vniversitaire
3, place Victor Hugo
F - 13331 Marseille Cedex 3
Téléphone: (0)4 91 62 44 16
Télécopie : (0)4 91 95 75 57

1 EDITORIAL.
Printemps, été, automne,

par Edmond Bianco

**3 INFORMATION, SEMIOTIQUE ET
ANALYSE FONCTIONNELLE,**

par France Chappaz

**19 PROPOSITION D'ORGANISATION
PHYSIQUE ET MECANISMES
D'INTERROGATIONS DE THESAURUS,**

par Jean - Michel Knippel

43 VOZZAVEDIBISAR.
Le symbole (suite),

par Edmond Bianco

D'ici quelque temps le bulletin aura ses informations sur WWW:
<http://scamup.univ-mrs.fr> <http://www.up.univ-mrs.fr>

Publication trimestrielle, gratuite, de l'Université de Provence

EDITORIAL,

Printemps, été, automne.

Déjà un vent léger parcourt les hautes herbes et fait frémir les arbres dont les bourgeons d'un vert très tendre éclatent. De petites fleurs timides pointent leur vif éclat coloré dans les endroits abrités. Quelques gouttes de pluie, entre deux nuages joufflus, un rayon de soleil, et toute cette verdure s'agite, verdoie au rythme des chants d'oiseaux.

Mais ce sont d'autres forces qui s'agitent dans les cieux du pouvoir. Des gens qui font semblant de se croire investis de la charge immense du destin des Hommes, veillent sur leurs petits intérêts personnels ou ceux de leur caste.

Ce n'est jamais qu'une année comme une autre, mais quand le jour grandit tout semble absolument nouveau. On oublie un peu ses rhumatismes, une force nouvelle semble couler dans les veines. Le sombre et le froid de l'hiver soulignent de bonheur les tiédeurs revenues.

Après avoir largement semé la pagaille dans l'univers feutré de l'Université, le personnage a prétendu au destin national, il avait fait ses preuves. La science est une dame trop sérieuse pour être confiée aux seuls scientifiques, aussi une bonne mainmise de l'administration par l'intermédiaire des coteries s'est avérée féconde. L'accumulation des commissions, toutes spécialisées, avec va-et-vient, et décision finale à Paris, décentralisation oblige, ne laisse plus filtrer que du politiquement correct.

Parfois le maussade du temps recouvre de froideur et de vent hivernal ce renouveau vibrant, qui se rendort un peu. Ce ne fut pas le cas pour l'année qui s'avance. Si le gel fut discret, Mai et Juin lumineux, avec juste assez de pluie pour que la floraison soit bonne, dès Juillet la chaleur envahit son domaine.

Il y a mille manières de résoudre un problème. Prenons un exemple, prenons en même deux: l'emploi et les retraites. Notre Jospin national a toutes les solutions, il ne les a pas inventées mais il les applique, coupant l'herbe sous les pieds d'une droite qui en reste encore plus stupide que d'habitude. Si c'est possible. Pour l'emploi il suffit de créer l'emploi précaire, la conjoncture est bénie, le chômeur débutant ou le chômeur de métier baisera les pieds de qui leur procurera de quoi s'acheter quelques cigarettes. Et il y a mille manières de privatiser le bien de l'état. C'est ainsi que, prenons le cas d'EDF, nombre de travaux qui pouvaient se faire avec une main d'œuvre qualifiée, mais revendicative, se font maintenant par appel à des sociétés esclavagistes, qui sont prêtes à travailler à des tarifs défiant toute concurrence. Qualité des travaux ? La vraie question c'est: dividendes à actionnaires ? Qu'importe que des centrales nucléaires aient été construites tant bien que mal par des

ouvriers non qualifiés, mais corvéables à merci. S'il y a des fuites, on avisera, peu importe. Que le grand porte-avions du Grand Général ait failli couler dès sa mise à l'eau, qu'importe, le contribuable était là pour le renflouer.

Les oiseaux font la sieste, et le chant des cigales du haut des résineux accompagne la brise qui fait vibrer au loin la campagne endormie. Des ondes de chaleur coiffent les champs de blé, qui achèvent de se dorer. Seuls les pampres verts des vignes donnent leur note fraîche dans la fournaise ambiante.

Lorsque notre ami Yoyo aura réussi la performance des quatre vingt pour cent de travailleurs en contrats à durée déterminée, tous ses soucis s'effaceront comme par miracle. Tout le monde aura du travail ...

« Quel genre de travail ? »

« Qu'est ce que ça peut vous faire, pourvu qu'ils en aient un peu ! »

... et il n'y aura plus de problème de financement des retraites, puisque quatre vingt pour cent des travailleurs travailleront jusqu'à leur mort, qui ne saurait tarder d'ailleurs, vu ce qu'ils vont gagner.

L'eau va se faire rare, l'herbe va jaunir puis sécher, les cigales vont chanter de plus en plus fort dans les pins, le lisier va sentir de plus en plus fort dans les prés, annonçant l'odeur de l'Europe, les vaches vont devenir de plus en plus folles, par l'atteinte spongiforme, puis par la soif. Strauss Kahn créera un impôt sécheresse, Messier de Vivendi, vous savez le marchand d'eau, augmentera le prix de son liquide, vu la richesse biologique et chimique de ce qui coule aux robinets, rien de plus naturel. Et tout continuera à être normal dès les premières pluies d'automne. Nous vivons là finalement une année bien banale. Si ce n'est que nous attendons avec impatience l'arrivée du grand "Bug", le bogue comme disent les journalistes qui ignorent ce délicieux poisson argenté, pour une fois on va un peu rire à voir qui va se casser la figure... Mais on sait déjà qui va payer les violons du bal.

Edmond Bianco

INFORMATION, SEMIOTIQUE ET ANALYSE FONCTIONNELLE

France Chappaz (*)

I. INTRODUCTION

La question des rapports entre information et signification est au coeur de débats épistémologiques vigoureux dans le domaine des sciences cognitives comme dans celui de la linguistique ou de la didactique; mais qu'en est-il de cette question lorsqu'elle s'applique au concept même d'information ! ou plus exactement quelle(s) signification(s) véhiculent les actes sociaux de l'Information comme Fonction de production. Un point de départ basé sur une pragmatique nous permet d'obtenir un corpus de référence à partir duquel nous pourrions tenter d'une part, une recherche d'unités minimales de significations, d'autre part, une analyse plus rigoureuse donc une compréhension mieux fondée de l'évolution des métiers centrés sur la Fonction Information.

Examinant successivement, l'information comme système et sa résolution, à la fois sur le plan de la substance et de la forme, puis l'information comme procès et sa décomposition en "fonctions élémentaires", nous tenterons de mettre en évidence une structuration quasi-stable de l'analyse du poste de travail d'agent d'information. C'est le nouveau métier de Technicien d'Information Médicale qui nous servira d'exemple tout au long de cette étude et de contrôle de pertinence.

Nous allons tenter une double articulation :

- entre un niveau profond de sèmes et les diverses manifestations de l'information en tant que terme-objet au plan superficiel du contenu,
- sur le plan syntaxique, entre les rôles actanciels des divers actants dans une structure et les rôles thématiques des agents d'informations dans une structure que nous montrerons de nature polémique au sein des institutions.

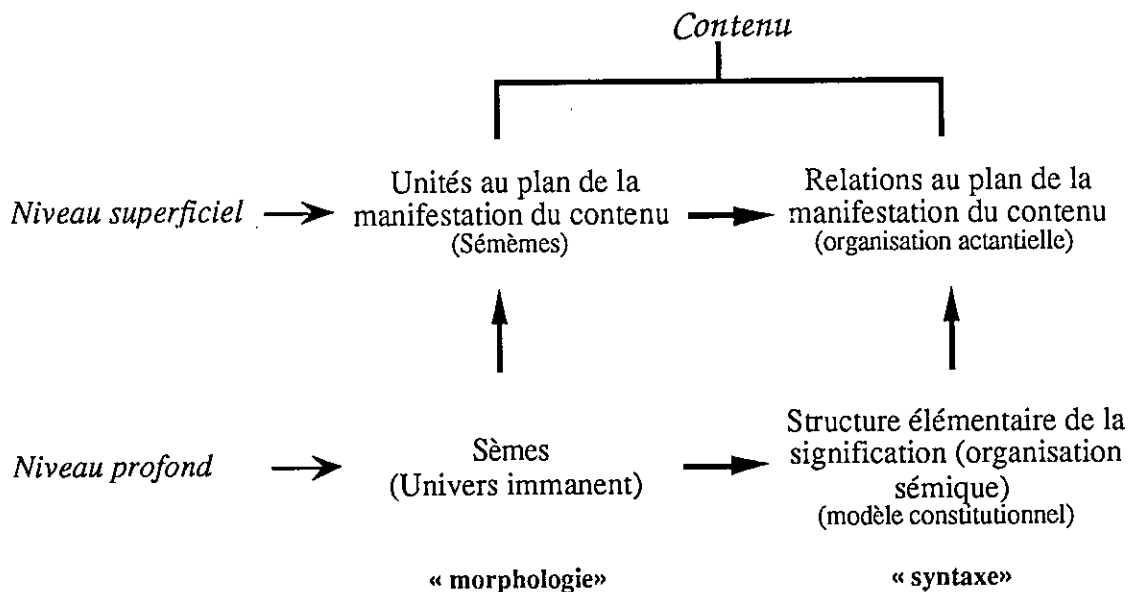


Schéma ci-dessus d'après COURTES

(*) HERMES - Représentation et transfert des connaissances - Université de Provence

Si l'on imagine une articulation de l'Univers sémantique en unités de signification minimales ou sèmes (pour un niveau d'analyse donné bien entendu), on rencontre deux types de sèmes:

- les sèmes nucléaires entrant dans la composition des lexèmes, (c'est-à-dire une unité de contenu, de nature virtuelle, et qui recouvre plusieurs possibilités de sens) ;

l'ensemble de tous les sèmes nucléaires définit le niveau sémiologique du langage;

- les classèmes qui se manifestent dans des jonctions de lexèmes et qui sont déterminés et repérables par le contexte;

l'ensemble des classèmes constitue le niveau sémantique;

- l'ensemble des sèmes nucléaires et des classèmes forme la manifestation du contenu .

- les sémèmes sont des combinaisons obligées de sèmes nucléaires et de classèmes.

On dit souvent que:"Le lexème semble être au sémème ce qu'est une entrée de dictionnaire au mot en contexte).

Un exemple célèbre emprunté à GREIMAS: le lexème "tête" peut s'envisager comme la partie d'un tout qui serait par exemple à une extrémité, supérieure ou antérieure :ex: la «tête d'un arbre» ou la« tête d'un canal»; ainsi, les sèmes /extrémité/ , /supériorité/ /antériorité/ donnent, combinés à d'autres sèmes, les deux sémèmes différents envisagés:

- /extrémité/+ /supériorité/+ /verticalité/ (ex la tête de l'arbre)

- /extrémité/+ /antériorité/+ /horizontalité/ + /continuité/ (ex:la tête d'un réseau)

Le noyau sémique, combinaison de sèmes, (unités pertinentes de signification), commun à ces deux sémèmes est composé du trait pertinent /extrémité/ ainsi que du trait de /supériorité/ qui se décline soit en /verticalité/ ou /antériorité/. Le classème ou sème contextuel est lié à l'insertion contextuelle du sémème dans le discours : la /verticalité/ ou l'/horizontalité sont des classèmes d'ordre contextuels.

(in COURTES (4)(5))

Quelques définitions de sémiotique

II. L'INFORMATION COMME TERME-OBJET (LEXEME) DOTE DE QUALITES (SEMES)

II.1. Les trois représentations du lexème "Information"

Une étude antérieure des paradigmes à l'oeuvre dans la Fonction Information,⁽³⁾ avait fait émerger l'information comme gisement, et l'information comme fluide en écoulement. Nous analyserons un troisième paradigme: l'information comme valeur.

Quelle que soit la façon dont on envisage le lexème "information", nous allons tenter dans ce paragraphe, d'extraire un certain nombre d'invariants ce qui devrait nous permettre de dégager un noyau de sèmes binaires opposés sous-jacents. Il nous faudra bien entendu mettre ensuite en évidence l'articulation des diverses manifestations du mot information avec ces sèmes binaires ainsi proposés. Pour éclairer notre propos, nous proposons un tableau à double entrée "sémèmes-sèmes" : chaque sème se croise ainsi avec chaque sémème, en étant à chaque fois modulé par sa propre spécification.(Fig. 1)

Les trois catégories sémiques que nous retiendrons sont la temporalité, la spatialité et la modalisation (comme mode de production d'actions). Elles vont respectivement se décliner par les sèmes binaires: mémoire/oubli, conservation/perte, tradition/création.

II.2. Premier sémème: le gisement

L'âge industriel a produit le modèle énergétique et qu'il s'agisse de ressource pétrolière, de puits de science ou de banque de données, la perfection à l'œuvre est un réservoir infini dans lequel gît une information morte, ou du moins éternelle. Ce stock est un champ scalaire d'indices, la seule intervention qu'il autorise étant de forer au bon endroit, tomber "juste", en somme!

Les problèmes soulevés sont alors ceux déjà pointés par G.VARET, le couple saisie-recherche qui ne peut bien fonctionner qu'en préservant le stock, en l'archivant dans un espace sûr et en assurant la pertinence, au prix de la redondance et de la structuration du champ par le code.

	Catégorie sémique	Temporalité	Spatialité	Modalisation du faire
	sèmes	Mémoire / Oubli	Conservation / Perte	Tradition /Création
	sémèmes			
INFO	- Gisement	Temps immobile Stock éternel oubli = absence	Espace de stockage infini	Codage-Décodage Redondance/Relevance
RMA	- Fluide s'écoulant	Temps irréversible oubli = ignorance	Écoulement laminaire Situation binaire de fluence: •Circulation ou Blocage •Confidentialité ou Diffusion	Transformation du flux par trans-position, trans-codage, mais toujours avec con-formité du contenu
ITION	- Valeur	• Mémoire et oubli assurent la protection oubli = secret • Mise à jour	Vitesse du fluide infinie. L'expansion et la striction sont multivaluées : accumulation → abondance →sélection → rareté Pertinence	• Médiation avec création d'intermédiaires • Production de Mesures et Moyennes • Sujet fiduciaire
			Plus	Plus

Figure 1. Tableau à double entrée "sémèmes - sèmes"

II.3. Deuxième sémème: le fluide

Mais l'âge moderne vectorise le champ, fluidifie la ressource, fait circuler le flot. Le temps immobile dans le cas précédent, devient temps courant; la fluence parfaite est l'écoulement laminaire.

Sur la spatialité, la conservation se manifeste en nature du débit; qualité d'abord: circulation ou rétention, quantité ensuite: confidentiel ou diffusé. La catégorie sémique est alors l'extensivité pouvant donner l'expansion ou la striction.

Comme dans le précédent sémème, la morphologie du fluide débouche sur le couple codé/clair, mais la question de la trans-position, du trans-codage pose un nouveau couple d'opposition : tradition/création; tradition au sens propre de simple livraison, de transmission strictement con-forme ou création intermédiaire, production d'informations. En tout état de cause, l'intégrité de l'information est de rigueur.

Jusque là, les modèles étaient de physique, le regard, le jugement, ne faisaient pas partie du modèle. Mais une nouvelle représentation de l'information émerge, information-valeur fondée sur l'utilité et qui s'élabore à la convergence du besoin et de la pertinence, valeur fiduciaire aussi, fondée sur la confiance accordée à celui qui l'émet.

II.4. Troisième sémème: la valeur

La valeur mettant en jeu à la fois le désir et le jugement, introduit de fait le sujet fiduciaire qui croit en la valeur de la valeur et pose la nature d'un contrat axiologique entre les protagonistes ; s'il y a contrat, il peut y en avoir rupture et commence à poindre une distribution polémique des divers actants.

II.5. Première catégorie sémique: temporalité

Le jeu du couple mémoire/oubli sur l'information-valeur relève d'une nouvelle appréciation :

- qualitative: l'"oubli", lorsqu'il s'agit du nom d'un malade lui assure en même temps la protection des données le concernant;

- quantitative, l'information valuée, par le jeu de la mesure, des statistiques, "efface" par intégration un certain nombre de données.

Le secret n'est plus forcément un obstacle : il devient source de garantie. En tous cas, il n'est plus dû au chiffrement du message, mais à la conscience du sujet fiduciaire.

La vitesse de propagation du fluide est infinie: l'espace se réduit à l'instant : le monde entier à l'oreille. On en retrouve presque l'âge classique du bouche à oreille des moines...

II.6. Deuxième catégorie sémique: spatialité

La spatialité se déploie en plusieurs termes : depuis la conservation, nous rencontrons et dans ce sens pour raison de cause à effet, l'abondance qui impose la sélection laquelle produit la rareté. Au passage, nous retrouvons la valeur de l'information au sens de SHANNON

Cette opération de sélection, de réduction, ne s'effectue plus au nom seulement de la vérité (qu'une information soit vraie, nous dirons plus volontiers référencée, est la moindre des choses!); cette nécessaire adéquation au référent devient secondaire devant la primauté de la pertinence: temporalité et spatialité se combinent pour la production d'informations valides mais également constamment actualisées, à-jour.

Dépassée, l'idée de vitesse de réponse à une demande: c'est au temps des rois que le message s'appelait "dépêche". Aujourd'hui, c'est l'accélération qui provoque l'anticipation de la question, donc la force d'innovation. L'autre pôle, ne fait que confirmer qu'il s'agit bien d'une dynamique : de la perte de l'information à celle des marchés : le moteur s'arrête en pénurie de fluide.

II.7. Troisième catégorie sémique: modalisation du faire

La modalisation tradition/création, recouvre à la fois la médiation comme création d'intermédiaire, la commensurabilité comme fondement de la prospective ou argument de la déraison. De surcroît, lorsque le sujet croît en la valeur de sa valeur, il en assure la valorisation et la promotion; quelquefois, s'il est plutôt orienté vers la tradition, il manifeste des réticences et des craintes face à cette nouvelle forme signifiante de l'information.

II.8. Tentative

Nous avons donc tenté, à partir d'une triple manifestation du contenu "Information"(les sémèmes Gisement, Fluide, Valeur), de dégager un "noyau sémique" réduit à trois couples de sèmes opposés. Le tableau (Voir Fig. 1) ci-dessus décrit cette organisation "morphologique" du contenu qui ne saurait s'interpréter comme une génétique : selon les contextes, les trois représentations coexistent avec une pondération différente ou se métamorphosent et influencent par la même occasion, les organisations actantielle et fonctionnelle.

Dans le Cas étudié du Technicien d'Information Médicale, la répartition des fonctions selon les trois sémèmes montre bien leur coexistence et leur pondération induite par celle des fonctions (Voir Fig. 3).

Il nous faut maintenant aborder l'étude des relations entre ces entités au niveau "syntaxique", en reprenant les termes du schéma décrit par COURTES dans Sémiotique narrative et discursive (5). Au niveau superficiel de la manifestation du contenu, le corpus est donné par la description des différentes fonctions de l'"Information" dans les divers contextes énoncés dans les études de cas. En effet, l'analyse d'un poste de travail par l'Analyse Fonctionnelle correspond généralement à la réduction d'un ensemble de fonctions qu'un agent donné subsume(*). Notre but est de réduire cet inventaire à un nombre fini et quasi stable aboutissant, à l'aide du modèle actantiel, à la production d'une structure simple de relations entre actants.

III. LE MODELE ACTANTIEL

Le modèle actantiel, tel qu'il a été dégagé à partir des travaux de PROPP, SOURIAU, GREIMAS, se présente au plan systématique comme un ensemble de trois couples d'actants, organisés autour de la relation sujet-objet que constitue le verbe ou la fonction. (d'après GREIMAS A.J. (7)).

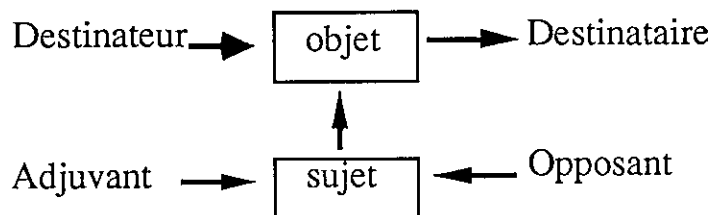


Figure 2. Modèle actantiel

Rappelons que l'actant désigne une unité syntaxique - être ou chose - qui, à un titre quelconque, actif ou passif, participe au procès.

III.1. Le couple sujet/objet:

Le sujet apparaît comme un actant dont la nature dépend de la fonction dans laquelle il s'inscrit; l'actant-sujet = l'être voulant, l'actant-objet est l'être voulu. Sans entrer ici dans le détail nous devons signaler la distinction entre les deux sortes de sujets correspondant aux énoncés d'état et aux énoncés de faire :

- a) les sujets d'états produisent une jonction avec l'objet voulu; ils ne sont sujets qu'en ce qu'ils sont dépositaires d'objets voulus, désirés, donc doués de valeur.
- b) les sujets de faire opèrent des transformations entre les états : le faire transformateur exercé par un sujet de faire vise un énoncé d'état qu'il faut transformer; il peut arriver que sujet d'état et sujet de faire soit pris en charge par un seul et même acteur: COURTES nous donne l'exemple du voleur bénéficiaire de sa propre action(4): le faire y est réflexif; sinon, il est transitif. La relation est toujours orientée du sujet vers l'objet.

III.2. Le couple Destinateur/Destinataire:

Le Destinateur est celui qui communique au Destinataire-sujet les éléments de la compétence mais aussi l'ensemble des valeurs en jeu; c'est aussi celui à qui est communiqué le résultat de la performance du Destinataire-sujet qu'il lui revient de sanctionner; on sera donc amené à distinguer le Destinateur initial "manipulateur" du Destinateur final "judicateur". Ce couple se justifie par rapport à l'objet qui s'inscrit sur l'axe du désir et de la communication : la relation de présupposition est unilatérale du Destinateur vers le Destinataire, au bénéfice de qui se fait l'action.

(*) subsumer: 1897; lat. mod. subsumare, de sub, et sumere "prendre"

III.3. Le couple Adjuvant/Opposant:

Par rapport à la relation sujet/objet, on distingue deux sortes de fonctions (5):

- celles qui agissent dans le sens du désir et qui sont attribuées à un auxiliaire positif: l'Adjuvant

- celles qui créent des obstacles à la réalisation ou à la communication de l'objet et qui sont attribuées à l'auxiliaire négatif: l'Opposant.

Selon la nature des matériaux soumis à l'analyse sémiotique, d'autres types d'actants sont nécessaires tels anti-sujet, anti-destinateur...

Le modèle actantiel comme procès instaure comme nous l'avons déjà signalé, le passage d'une relation d'état à une autre par un recours à un faire-transformateur. L'état est descriptible, le faire est modal.

Ce qui caractérise une action, c'est qu'il se passe quelque chose: c'est une transformation entre deux états successifs et différenciés. Ce changement d'état est effectué par un sujet quelconque affectant un autre sujet quelconque.

Si on appelle "acte" ce qui fait être: l'agir est alors identifiable au "faire-être" et correspond au passage de la potentialité à l'existence. Cet agir présuppose l'existence d'un sujet de faire qui transforme un sujet d'état. Ce faire-transformateur opère sur des états successifs, les sujets en présence pouvant être distincts ou confondus.

Mais le sujet de faire ne peut réaliser la transformation que s'il possède au préalable la compétence, condition nécessaire à la réalisation de l'acte qui en est la performance.; la compétence est le Ce qui fait être.

Appelant S1, le sujet de faire et S2 le sujet d'état, l'action de faire F consistant à ce que le sujet S2 change d'état, par exemple acquière un objet ou une qualité, la formulation symbolique du programme narratif est alors strictement l'expression ci-dessous où la conjonction est symbolisée par « (la forme disjonctive de la jonction étant alors »):

$$F : \{ S1 \rightarrow (S2 \cap O) \} \text{ (in } (4))$$

Les sujets d'état se définissent par leurs propriétés, leurs qualités... et les sujets de faire opèrent des transformations entre les états.

Un acte "performateur" par exemple suppose que S1 et S2 soient confondus et un sujet compétent exige en préalable le mécanisme de formation de la compétence de l'acte transformateur (et non la compétence sur le savoir déclaratif que S2 doit acquérir en aval. Si un faire régit un autre faire, on a la relation factitive de manipulation; dans cette relation factitive (faire-faire), un énoncé de faire régit un autre énoncé de faire: le sujet-manipulateur fait en sorte (F1) que le sujet-manipulé réalise (F2) la conjonction de (ou la disjonction) entre un sujet d'état et un objet de valeur (3 sujets, 1 objet).

La compétence du sujet sera (in (5)) "le vouloir et/ou pouvoir et/ou savoir-faire du sujet (plus la modalité factitive faire-faire) que pré-suppose son faire performantiel".

- le vouloir se rattache à l'axe sujet-objet et instaure le sujet
- le savoir se rattache à l'axe destinateur-destinaire
- le pouvoir se rattache à l'axe adjuvant-opposant.

(in COURTES (4)(5))

Le modèle actantiel comme procès

En tant que procès, il s'agit de mettre en oeuvre les relations précédentes.

Auparavant, il nous faut faire une remarque: notre travail ne porte pas sur la structure narrative d'un discours; il ne peut être question de faire appel à une quelconque chronologie globale d'un discours sur l'Information (consécution temporelle ou logique à rebours) qui mettrait en jeu succession ou simultanéité des fonctions élémentaires, comme cela pourrait se produire dans une étude de processus. Cette étude est essentiellement a-chronique ; les seules relations d'ordre considérées seront locales : par exemple la performance pré-suppose la compétence comme condition nécessaire de l'acte.

IV. L'ANALYSE FONCTIONNELLE

L'Analyse Fonctionnelle est une méthode permettant d'aboutir à l'expression fonctionnelle de besoins satisfaits ou à satisfaire. Elle fait partie intégrante de l'Analyse de la Valeur.

En Analyse fonctionnelle, on convient de différencier :

- une fonction qui satisfait à une raison même d'existence d'un produit telle par exemple la fonction d'éclairage pour une lampe,
- une fonction qui ne concerne pas la satisfaction du besoin essentiel, mais qu'il est cependant nécessaire de réaliser, par exemple la ventilation de la dite lampe.

La première catégorie est qualifiée de fonction de service; la seconde : fonction de contrainte.

Les fonctions de contraintes traduisent généralement les interactions obligées avec l'environnement du produit, qu'il s'agisse de normes de fonctionnement, de spécificités culturelles, ou de modes d'organisation du travail.

(in CHAPPAZ (9))

Les fonctions en Analyse Fonctionnelle

Le déroulement puis l'aboutissement d'une étude de cas, consiste notamment à mettre en évidence une organisation des besoins en termes de "fonctions" élémentaires dont une des caractéristiques est l'expression syntaxiquement homogène, basée presque exclusivement sur le nœud verbal.

Dans ces études de cas, la description de la situation rend compte d'une portion d'univers déterminé, telle que par exemple la redéfinition des missions d'un service de documentation d'une société de presse, ou l'étude d'une cellule de coordination d'information municipale.⁽⁹⁾

Or, cette décomposition que vise (et qu'atteint généralement) l'analyse fonctionnelle, consiste précisément à exhiber des unités d'actes signifiants pour l'agent eu égard aux contextes et aux buts poursuivis par les diverses parties de l'organisation.

Ces fonctions élémentaires de l'Information obtenues à partir d'études de cas, doivent pouvoir se mettre en relation et se combiner au modèle actantiel . C'est ce que nous allons montrer.

V. L'EXEMPLE DE L'"INFORMATION MEDICALE"

V.1. La situation à analyser

Le Programme de Médicalisation des Systèmes d'information (PMSI) a constitué, depuis 1983, une étape essentielle pour la gestion des informations dans le système hospitalier français.⁽⁶⁾ Le Département d'Information Médicale est créé dès 1989, et la loi en définit les missions principales sans préciser vraiment les compétences du personnel affecté : le Médecin responsable et le Technicien d'Information Spécialisée.

Cette étude sur le métier de Technicien en Information Spécialisée en secteur hospitalier, a permis de définir avec précision les fonctions et compétences requises pour ce nouveau poste de travail.

Les principales missions du Département, citées dans les textes officiels, concernent:

- la mesure de l'activité médicale,
- l'aide à la planification,
- l'assurance d'une logistique intégrée de type tertiaire,
- l'approche de la démarche de Qualité totale en milieu hospitalier.

Pour mener à bien ce travail, nous avons mis au point, dans une visée à la fois inter- et trans-disciplinaire, une méthode qui articule Modélisation Systémique, Analyse de la Valeur et Gestion de processus (1).

La définition de concepts dérivés de l'analyse fonctionnelle notamment, nous a permis de transférer la notion de Cahier des Charges fonctionnel à la définition de compétences et celle de réponse au Cahier des charges en programme de formation.

Ce transfert a pour conséquence immédiate que la partition usuelle en Analyse fonctionnelle entre fonctions de contraintes et fonctions de services recouvre de fait la stratification suivante des niveaux de connaissances :

- pour les fonctions de contraintes:
les savoirs de description des nécessités et des contingences, (savoirs-objets décrivant l'actant collectif que constitue le milieu et ses prolongements);
- pour les fonctions de service:
les savoirs de changement d'état (savoir-faire)
et les savoirs de mise à distance (savoir-faire-faire).

Les premières correspondent à des savoirs généralement descriptifs sur les termes d'environnement du processus informationnel considéré. La situation, l'état des choses, sont effectivement différents s'il s'agit d'un Service d'Information en milieu hospitalier ou d'une Cellule d'information municipale destinée à délivrer une information de proximité au grand public. Leur description, correspond plutôt au dévoilement d'un espace de connotation.

Les secondes structurent l'espace de dénotation.

Dans le présent travail, nous nous sommes essentiellement intéressés au noyau sémique de la fonction information; nous n'aborderons donc pas ici la description des contextes, relevant plutôt des sèmes contextuels.

V.2. Les fonctions de service retenues

Le développement de la méthode a permis de définir de façon à la fois complète et minimale les diverses fonctions de services traduisant les tâches du Département d'Information Médicale. Nous en donnons ci-dessous la liste par rapport aux deux postes de travail concernés :

- PROD : Production de documents Médicaux
- RASD : Rassemblement de données
- TRTD : Traitement des données -Etablissement de rapports
- DEMU : Démultiplication, Relai de politique d'Information Médicale
- TRTE : Transfert de méthodes et savoir-faire aux autres services
- GAFI : Garantie - Fiabilité - Protection
- DOCU : Traitement documentaire et bibliométrique
- MEMO : Constitution et gestion de la mémoire des Infos médicales
- SAUV : Sauvegarde et maintenance
- FOCI : Formation initiale, continue, interne et externe au service
- QUAL : Vecteur de l'exigence de qualité dans l'activité hospitalière
- MUOM : Mesure d'unités d'oeuvre médicales
- RPRO : Recherche et Prospective

Ces fonctions ont ensuite été hiérarchisées par postes de travail et ordonnées en trois niveaux: Fonctions Principale, Secondaire, ou Complémentaire, la répartition étant naturellement différente pour chacun des deux postes concernés.

V.3. Répartition des fonctions hiérarchisées par sémème de l'information

Ce que met en évidence la Figure 3 traduit bien l'évolution perçue dans les métiers d'information; primitivement basée sur l'archivage puis sur la documentation, la représentation de l'information rend aujourd'hui secondaire (cercle médian) les fonctions qui proviennent du sème gisement alors que sont considérées comme Principales (Grand cercle) celles qui découlent du sème fluide et presque au même niveau, celles émanant du sème valeur.

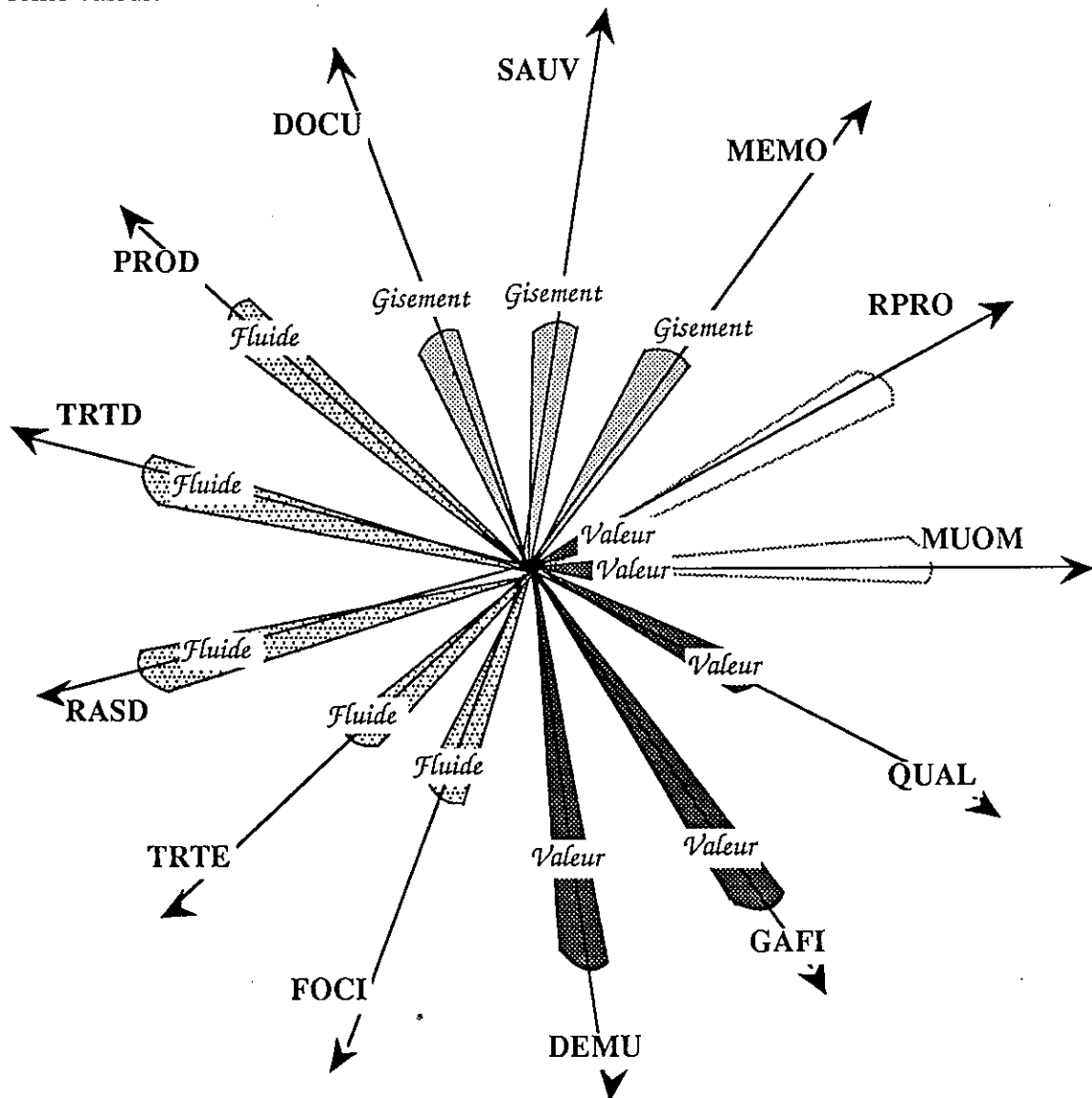


Figure 3. Répartition des fonctions selon les trois sémèmes: gisement, fluide, valeur. (cas du Technicien d'Information Médicale)

V.4. Application du modèle actantiel

V.4.1. Le système

L'ensemble des activités révèle une structure actantielle implicite.

Le Sujet-héros (S1) est naturellement le Technicien d'Information Médicale.

L'objet est constitué principalement:

-de documents ou de données dépourvus de valeur avant le faire transformateur du sujet;

-de savoirs, de méthodes et de savoir-faire que S1 transmet à d'autres sujets: personnels d'autres services hospitaliers ou patients.

Le Destinateur Primaire est le Ministère des Affaires Sociales qui a mis en place le Programme de Médicalisation des Systèmes d'information dans le but notamment d'effectuer une mesure de l'activité hospitalière afin d'en augmenter le rendement et la qualité.

Les destinataires sont souvent à plusieurs degrés: le Service DIM lui-même étant souvent le Destinataire de premier degré, l'Hôpital, l'Institution Médicale, le patient, des Destinataires de second degré.

L'opposant est constituée de l'inertie du milieu, des réticences face à la technique, de l'absence de statut pour le TIM, de l'éclatement physique de certaines structures hospitalières.

Dans un certain nombre de fonctions, le milieu médical, supposé être bénéficiaire à long terme de la politique menée sur l'Information Hospitalière, joue indubitablement le rôle d'anti-destinateur et ne reconnaît pas l'intérêt de ce travail supplémentaire débouchant aisément sur de l'évaluation se détournant aisément en contrôle; une véritable structure polémique est à l'oeuvre, bien caractéristique de l'axiologie sous-tendant le sémème de l'information-valeur.

L'Adjuvant est essentiellement constitué du sentiment d'appartenance à une évolution positive, de la croyance au progrès, de l'adhésion à certains aspects de la politique d'Information Médicale.

V.4.2. Le procès : inventaire et analyse des fonctions

V.4.2.1. Si nous appelons S1 le sujet de faire, S2 le sujet d'état constitué de S1

Lorsqu'il a effectué la transformation F, pour un certain nombre de fonctions, il s'agit simplement d'une opération où S1 et S2, confondus en tant qu'acteurs sont dans deux états différents: conjoints avec un objet O1 dénué de valeur, il se conjoint à un objet doué de valeur O2 après son faire performateur: le faire est alors réflexif et la modalité est réalisante.

Le Programme est alors du type performance:

$$F [(S1 \cap O1) \rightarrow (S2 \cap O2)]$$

C'est le cas des fonctions du type Rassemblement de données, Pratiquer les Sauvegardes, Traiter les données pour effectuer des rapports, Constituer la mémoire des informations médicales, Produire des documents médicaux, Effectuer des recherches documentaires, Pratiquer le codage des dossiers. Le sujet S1 est alors pour la plupart des cas, en syncrétisme avec le Destinataire principal (le Département même d'Information Médicale).

La Médicalisation du Système d'Information en milieu hospitalier étant, encore dans de nombreux cas, en régime transitoire de mise en place, et le personnel de secrétariat médical prévu, a priori non formé aux savoir et savoir-faire sur l'information et la documentation, il est souvent nécessaire de prendre en compte comme actant opposant, des incompétences sur les techniques informatiques et documentaires; dans une analyse plus fine de ces diverses fonctions, il faudrait prendre en compte la nécessité d'un programme d'usage correspondant à l'acquisition préalable de la compétence antérieure à la performance visée.

V.4.2.2. Dans une seconde distribution actantielle

S1 (toujours sujet de faire) effectue une transformation d'état sur un autre sujet S3, constitué d'agents hospitaliers relevant des autres services ; cette fois, l'actant S3 est forcément différent de S1 et l'objet est unique : le faire transformateur F est alors de type transitif : faire communicatif ou formatif; naturellement S1 qui possède l'objet O (Savoirs et méthodes) avant la transformation, ne le perd pas après

Ce faire est CE qui fait être le sujet S3 performant, il est donc une compétence comme le montre COURTES (in ⁽⁴⁾ p.94).

On trouve dans ce cas les fonctions de Transfert de méthodes et techniques, de Formation initiale et continue dont il nous faut bien signaler qu'elles ne faisaient pas explicitement partie des missions initiales définies par le Destinateur. Mais l'Information engendrant des effets dus à sa propre signification (le sémème "fluide"), les agents d'information Médicale ont été placés devant la nécessité de "fluidifier" le processus.

La troisième fonction de compétence : Protection des données médicales intervient au confluent de la "valeur" et de la "temporalité" : de préjudice, l'oubli devient remédiation.

Pour la seule et unique fois dans le processus lié à la Médicalisation de l'Information Hospitalière, l'actant "Patient" S4 arrive en scène. Notre sujet S1 agit sur son état et de non-protégé qu'il était avant la mise en place des procédures de protection, le met au secret, le garantit, le "Fait-être Destinataire-sujet protégé" S5.

V.4.2.3. Troisième modalité du "Faire" dans le corpus considéré, la manipulation

Dans l'exemple pris ici de l'Information Médicale, c'est le sujet S1 qui sera manipulé ; le sujet manipulateur qui est le Destinateur Principal S6 (le Ministère) fait en sorte (F1) que S1 (héros selon le savoir) effectue F2 la conjonction entre le sujet d'état S3 déjà défini (Autres personnels hospitaliers) et l'objet de valeur qui sera selon les cas, la Qualité en milieu hospitalier, la mise en place de la Politique d'Information Médicale.

Naturellement, la manipulation ne peut réussir que si S1 est non seulement apte à réaliser F2 mais également s'il est un sujet fiduciaire, lui-même conjoint de la croyance en la valeur de la valeur. Ces deux fonctions: Vecteur de l'exigence de qualité dans l'activité hospitalière et Démultiplication de la Politique d'Information Médicale, voient émerger un nouveau rôle actantiel: celui de l'anti-destinateur qui, par l'actant collectif du milieu médical, joue un rôle d'opposant régulier et instaure la structure polémique déjà signalée. En simplifiant, nous trouvons bien la formulation symbolique de la manipulation donnée par Courtès ⁽⁴⁾

$$F1\{S6 \rightarrow F2(S1 \rightarrow (S3 \cap O))\}$$

Le dernier ensemble de fonctions caractérisées par la modalité factitive relève des missions fondamentales édictées par le Destinateur S6, en syncrétisme actoriel avec le Destinataire principal. Mais un nouveau sujet de faire S7, (Le Responsable du Service) va jouer le rôle de sujet-manipulé primaire et en même temps de Destinataire: S6 fait (F1) en sorte que S7 fasse (F2) transformer (F3) par S1 des données (O1) en leurs mesures (O2)

Lors de notre étude de cas, nous avons pu mettre en évidence, par la construction du profil fonctionnel du Technicien d'Information Médicale que ces fonctions "Mesure d'unités d'oeuvres Médicales" et "Recherche et Prospective", posaient le problème du statut professionnel du TIM, oscillant au gré des études et du poids de la technique entre secrétaire médicale et chercheur....

Dans notre modèle actantiel, cela se traduit notamment par le syncrétisme actoriel du sujet S1 et de l'actant opposant.

V.4.2.4. Trois cruelles absences

Il était clairement apparu, lors de l'étude de terrain, certaines revendications émanant du personnel agent d'information spécialisée, concernant d'une part, des déficits de formation, d'autre part, et consécutivement à la formation, la reconnaissance statutaire de savoirs professionnels acquis.

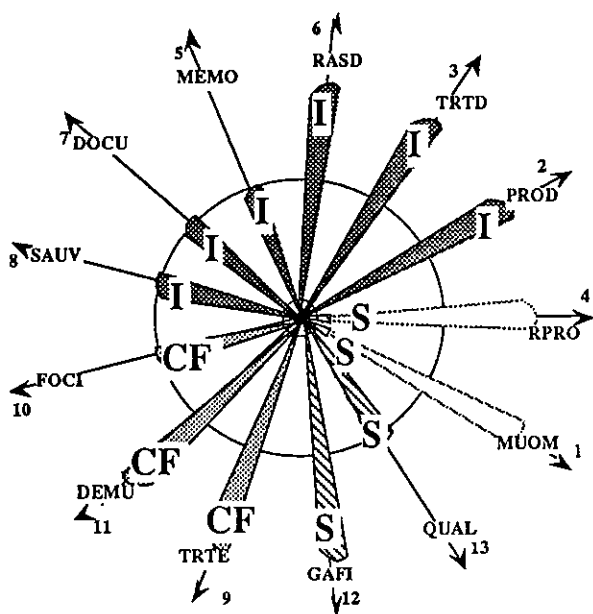


Figure 4a.

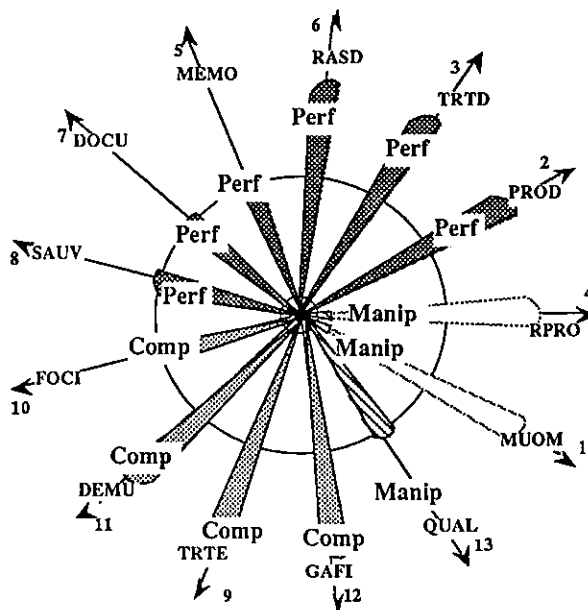


Figure 4b.

Figure 4a: le Technicien d'Information dans ses trois rôles thématiques:
I:l'Information, C-F:la Communication-Formation, S:son rôle socio-économique.

Figure 4b: le Technicien d'Information dans ses divers rôles actantiels et ses trois modalités du faire.

Dans la modalité du faire-performateur, l'opérateur-sujet de faire est confondu avec le sujet d'état; il s'agit d'un faire réflexif.

Dans le faire-être de compétence, l'opérateur-sujet est forcément disjoint du sujet d'état: c'est un faire transitif

Dans le faire manipulateur, il intervient au moins trois sujets dont l'un est forcément en syncrétisme actoriel

Or, si nous poursuivons l'analyse à son niveau anthropomorphique, nous voyons bien, à l'aide de la grille actantielle, comment se traduisent les manques : pas de Destinateur-judicateur, pas de contrat traduisant le niveau axiologique, pas d'épreuve glorifiante. Il ne reste vraiment comme moteur pour continuer à "faire" que la valeur en la croyance de la valeur, qui comme toute croyance, est soumise, in fine, au bon vouloir et au désir du sujet! Du coup, l'absence de formation qualifiante équivaut à /l'empêchement de faire/. Cependant que (S1) est sujet fiduciaire, que d'autres opérateurs hospitaliers (S3) sont en appétit de savoir, du moins jusqu'à un certain point(!), celui de la mise en place du contrat axiologique...., il émerge des anti-sujets du fait de l'impossibilité de se former. Par ailleurs, l'anti-Destinateur (milieu hospitalier) se confond avec l'opposant essentiellement dans les activités de manipulation relatives au lexème de l'information-valeur, notamment sur les missions d'évaluation. Un rapport polémique s'établit donc doublement: entre le Destinateur principal (Le Ministère) et l'anti-Destinateur (le milieu hospitalier), mais aussi entre le sujet S1 et l'anti-sujet (qui ne peut/veut se former).

Le profil professionnel, une fois résolu en ses fonctions élémentaires va pouvoir se déployer en ses trois modalités d'action que sont:

- le faire réflexif de performance,
- le faire transitif de compétence,
- le faire transitif manipulateur.

Or, les trois composantes reconnues du métier en son état actuel sont:

- une composante traditionnelle bien que fortement renouvelée, centrée sur des activités de documentation: notée I dans 4a)
- une composante plus récente, liée notamment à l'émergence des nouvelles technologies, sur la formation et la communication, notée C-F
- une composante visant à l'évaluation et l'amélioration de la qualité de l'ensemble des activités hospitalières, relevant d'aspects socio-économiques, notée S.

Nous constatons, à l'aide des figures 4a) et 4b), qu'à une fonction près (GAFI) dont le rôle thématique est autant à caractériser comme socio-économique que comme gestionnaire des flux puisqu'il peut notamment l'interrompre en en assurant la protection, les modalités du faire recouvrent ces composantes, fondant et confondant ainsi les rôles thématiques représentés par les trois composantes et les rôles actantiels.

VI. EMERGENCE D'UN MODELE CONSTITUTIONNEL

Lorsque nous cherchons enfin à élaborer une organisation encore plus profonde de la signification afin de tenter la mise en place d'une forme canonique, nous avons rencontré au cours de notre étude, sur un axe des contraires, les sèmes /mémoire/oubli/, /conservation/perte/ liés aux catégories sémiques de la temporalité et de la spatialité.

Mais ces deux sèmes relèvent d'une opposition encore plus fondamentale qui est le couple d'éléments en relation d'opposition: /inaccessible/ versus /accessible/ (Figure 5a)) traduisant les diverses modalités possibles pour l'actant-sujet de se conjindre ou non avec l'objet convoité que constitue à un moment donné l'information recherchée.

Cet objet du désir peut être inaccessible soit parce qu'il est ignoré, ou perdu, ou crypté, ou oublié. L'opération de mise en accessibilité recouvre les opposés : référencé, trouvé, décodé, mémorisé.

Chaque couple constitue en fait les états initiaux et finaux de l'objet dont le changement d'état est précisément obtenu par une fonction élémentaire de notre analyse fonctionnelle, revisitée par l'organisation actantielle. Nous trouvons ici quelques fonctions bien connues des documentalistes quelles que soient leurs spécialisations et leurs contextes de travail.

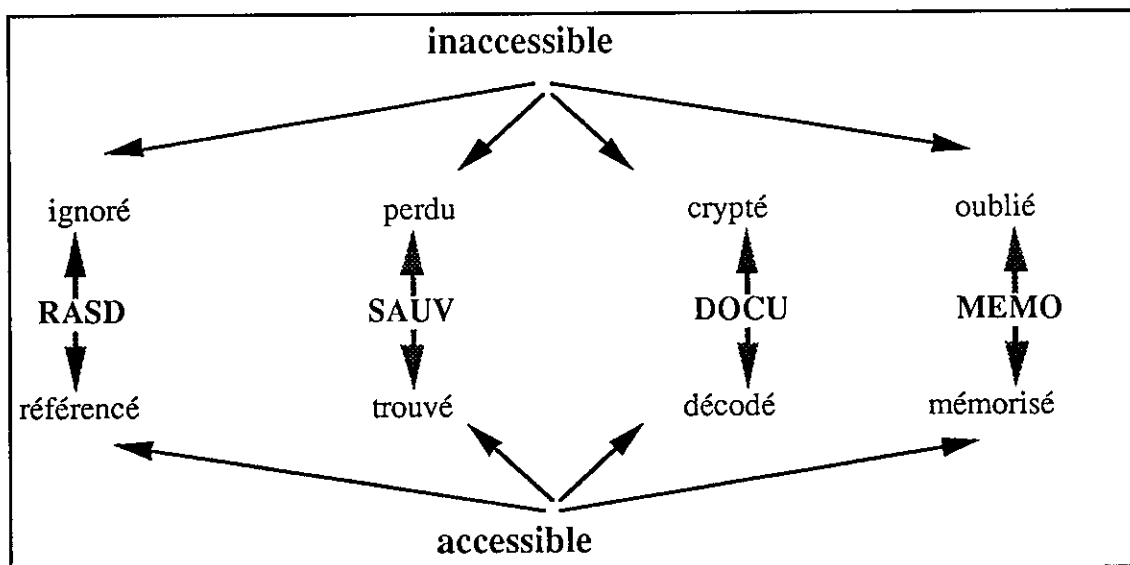


Figure 5a. Couple d'éléments en relation d'opposition

Passons maintenant au troisième couple d'opposés tradition/création et détaillons un peu, non sans rappeler auparavant que tradition est à prendre aussi au sens de simple livraison, de transmission sans "plus-value". Les données sur lesquelles travaillent l'agent d'information se devaient hier d'être intégrales (pas de censure), pérennes (du moins non périmées), abondantes, transparentes pour tous. Le déploiement du sémème de l'information comme valeur s'appuie notamment sur de nouvelles exigences sur les données: elles doivent aboutir à des évaluations globales (donc être intégrées), constamment mises à jour -actualisées- et validées, enfin, elles doivent être sélectionnées pour les objectifs de l'institution, et dans certains cas, protégées de la divulgation. Toutes ces qualités des données ne peuvent s'obtenir sans l'exercice de facultés de jugement de la part de l'agent d'information, facultés débouchant sur les concepts de garantie et de pertinence (Figure 5b)).

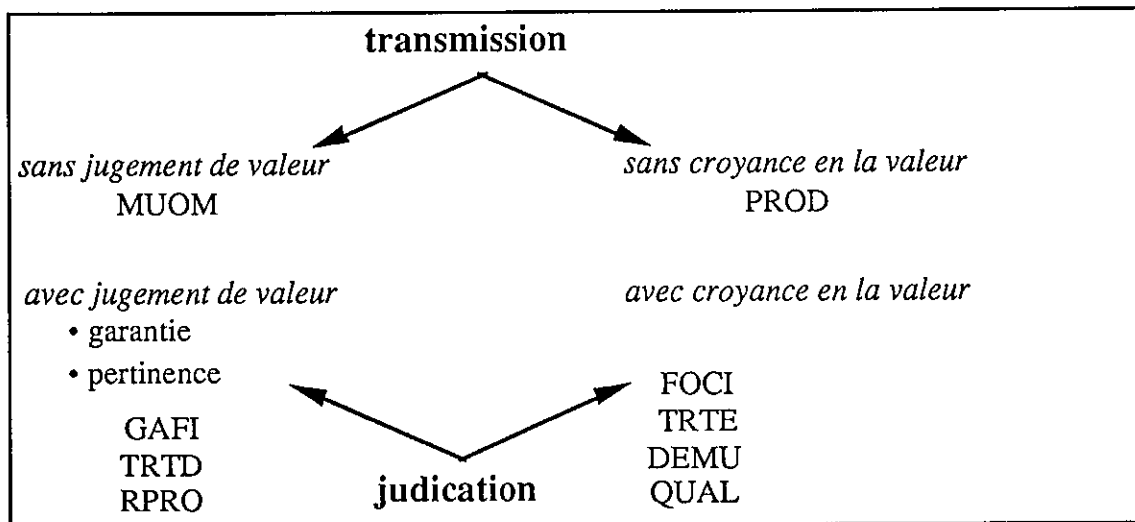


Figure 5b: Vers les concepts de garantie et de pertinence

La garantie qui assure à la fois la fiabilité et la protection, ne peut s'exercer sans faculté de jugement: la mise en secret, comme la certification d'une information, nécessitent la judication. La pertinence qui recouvre à la fois l'actualisation (vérification de la non-péremption) et la sélection en fonction d'objectifs (sélection qui relève de l'évaluation de degrés de pertinence), nécessitent des jugements de valeur et au delà, l'exercice également du processus de judication.

D'autre part, comme nous l'avons déjà dit, l'agent d'informations devient de plus en plus, porteur d'un credo. Il est animé d'une croyance en la valeur de l'information lorsqu'elle est valeur! Et c'est ce qui le pousse à effectuer actions de sensibilisation, de formation, de démultiplication d'une politique d'information...

La production de documents (médicaux ou autres) se passe de la croyance en la valeur; la mesure d'unités d'oeuvres médicales n'a besoin d'aucune croyance ni de jugement liés à la valeur.

Nous voyons donc émerger une organisation profonde de l'activité du TIM qui nous permet de retrouver l'ensemble des fonctions élémentaires repérées dans notre étude de cas.

VII. CONCLUSION

Chaque contextualisation particulière du poste d'agent d'information porte naturellement des éléments spécifiques donnant lieu à des fonctions liées à la situation à prendre en compte. Nous avons voulu essayer de montrer que l'analyse sémiotique assure un fondement et, de ce fait, autorise une nouvelle méthodologie à l'analyse fonctionnelle dans le secteur tertiaire.

Par là, nous espérons apporter une pierre à l'édifice déjà solide des travaux des consultants et conseils en Analyse de la Valeur et Analyse Fonctionnelle.

Sur un plan plus général, le changement de représentation de l'information qui se traduit dans les profils fonctionnels des postes d'agents en information nous semble dépasser ce cadre.

Si nous prenons par exemple le véritable phénomène micro-sociétal que constitue l'engouement récent pour Internet en Europe, nous voyons bien que l'idéologie du "village planétaire" empruntant les notions de transparence totale et de fluidité absolue engendrent tout autant les débordements du fleuve par le trop-plein (de données) que par le détournement, imposant de fait une évolution vers l'endiguement, la protection, le verrouillage. Nous ne sommes pas loin pour les utilisateurs d'Internet, de la nécessité d'un véritable code moral de l'utilisateur.

L'information comme valeur est bien le prix de ce code.

VIII. BIBLIOGRAPHIE

- (1) CHAPPAZ F.
Le métier de Technicien d'Information Médicale: des compétences fonctionnelles au cursus de formation
IDT 94 Gérer l'information pour l'excellence de l'entreprise
Juin 1994 - Paris
- (2) CHAPPAZ F.
Un outil pour concevoir les formations: application au cas de l'information stratégique
Actes du colloque de Génie Industriel
Tours - Mars 1991
- (3) CHAPPAZ F.
L'Information Scientifique et Technique ou le conflit des générations
Colloque DBMIST - Villeurbanne - 1986
- (4) COURTES J.
Analyse Sémiotique du discours
Hachette Sup - 1991
- (5) COURTES J.
Sémiotique Narrative et Discursive
Hachette Sup - 1993
- (6) Dossier collectif
Le PMSI
Gestions hospitalières - n°307 - juin/juillet 1991
- (7) GREIMAS A.J.
Sémiotique structurale
PUF - 1986
- (8) MICHEL J., (coll. SUTTER E.)
Pratique du management de l'information
ADBS - 1992
- (9) ROUX B.,(coll CHAPPAZ F.)
Vers une typologie de l'Information de proximité
Mémoire de Maîtrise Information-Communication - Juin 1994

PROPOSITION D'ORGANISATION PHYSIQUE ET MECANISMES D'INTERROGATIONS DE THESAURUS

Jean - Michel Knippel

(Article suivant celui paru dans le numéro 50 de juin 1998)

1. INTRODUCTION

D'abord, il est souhaitable de ne pas négliger la partie implantation des données, leur structure et le code des différentes opérations. Certes l'utilisateur ne verra pas cette partie cachée, lui qui ne manipule l'objet qu'à travers les méthodes qui lui sont associées. Il serait en effet important de se définir une sorte de continuité discrète qui lie, du moins dans certaines circonstances le code de l'algorithme à la configuration objet. On sépare le fini-borné du fini-illimité. Le premier est le domaine des compilateurs (ici "traducteurs" de symptômes en remèdes). Le compilateur travaille sur le répertoire. Le fini-illimité est le domaine du système ouvert qui assurera la liaison conversationnelle et l'approvisionnement du thésaurus existant par l'accès à d'autres documents fondamentaux : la matière médicale pure de Hahnemann, les matières médicales de Allen et Herwing, le "Synoptic Key" de Boger etc...

Tout utilisateur potentiel souhaitera, lui, rapprocher le langage d'interrogation du langage courant, ou langage "quasi-naturel". Ceci fera l'objet de l'étude des mécanismes d'interrogation de thésaurus au paragraphe trois [SOULE-DUPUY C. 90].

Une des préoccupations face à cette composante de représentation de la base d'informations est le rapprochement avec les langages d'interrogation, permettant la définition des besoins en information. Nous utilisons les graphes, types réseaux de Petri, car ceux-ci sont un outil de modélisation, ou un langage de spécification pour les bases de données, ou encore un langage de représentation des connaissances et ayant une représentation graphique "naturelle". Le paragraphe quatre abordera nos propositions dans ce domaine.

2. PROPOSITON D'ORGANISATION PHYSIQUE DE LA BASE D'INFORMATIONS

Notre approche est basée sur les héritages des travaux de Alan Turing, qui dans son essai "On computable numbers" [TURING A. 48] invente le concept de machine universelle où codes "programme" et codes "données" sont mis au même niveau dans la mémoire [BIANCO E. 92].

Nous désignons par les abréviations suivantes :

- MU : machine universelle,
- P : code de l'algorithme,
- C : configuration de l'algorithme.

Mais il nous faut partir de la machine à cases adressables, afin de nous donner une capacité à faire de l'arithmétique. Ceci a été démontré. La notion de procédure nous donnera ensuite une économie pour la construction des algorithmes [BIANCO E. 92].

Dès lors intervient le véritable ordinateur par la machine formelle ou universelle. Nous utiliserons le processeur de la procédure formelle. Le processeur fonctionne en statique, ce qui signifie qu'on peut le ralentir tant qu'on veut et même l'arrêter, cela ne détruit pas l'information qu'il traite. Nous ne cherchons pas la performance en vitesse pure, qui ne nous semble pas présenter d'intérêt fondamental. Il nous paraît plus important de faire porter l'effort sur la qualité de base du logiciel. Ici le choix du langage nous paraît plus important dans la mesure où il contient des notions de l'informatique qu'on rencontre partout, et à des niveaux divers [BIANCO E. 92]. Par exemple il nous paraît décisif de disposer d'une véritable instruction d'insertion de procédure, car c'est en ce point que se prépare la gestion de la place en mémoire, c'est également là que se transfèrent les informations indispensables. Nous avons d'ailleurs prévu de compléter à brève échéance cette notion, par la notion d'insertion fractionnée, qui introduit la notion de vecteur information, et celle d'insertion différée. Ce sont ces notions de base qui permettent de gérer de manière systématique à la fois le temps et l'espace, par application de l'autojectivité.

De ces bases apparaîtra le langage de programmation minimum qui sera traduit pour le rendre compréhensible à la machine par la compilation qui a déjà fait l'objet de développements par notre équipe [BISIERE CH. 89].

2.1. ORGANISATION DE BASE: MACHINE UNIVERSELLE

Pour expliciter cette organisation de base, nous avons besoin de définir un langage compatible avec l'architecture que nous avons choisie. Le langage choisi est la procédure formelle [BIANCO E. 92] héritier des langages de la machine de Nolin et de la procédure formelle et ses extensions.

Le langage procédure formelle comprend une partie déclarative de paramètres, de variables et d'index, caractérisée par les mots réservés: proc, para, var, index. La "lettre" début annonce le début des instructions et la fin des déclarations. Les instructions se terminent par une lettre " ". La fin des instructions d'une procédure est donnée par la "lettre" fin. La fin de programme est marquée par la "lettre" fin.

Chaque opération utilise au maximum trois opérandes avec:

- 4 opérateurs arithmétiques: '+', '*', '-', '/';
- 5 opérateurs logiques avec saut conditionnel:
 - ' si = vers < étiquette > ',
 - ' si < vers < étiquette > ',
 - ' si <= vers < étiquette > ',
 - ' si > vers < étiquette > ',
 - ' si >= vers < étiquette > ';
- l'affectation ' := ';
- le saut incondtionnel soit ' vers < étiquette > '.

Le ":" permet de ponctuer les étiquettes. On peut également insérer des procédures grâce à l'instruction insere suivie du nom de la procédure à insérer et de la liste des paramètres.

Nous rappelons brièvement la forme des instructions de la procédure formelle. Nous nous limitons pour l'instant au calcul en mémoire centrale.

Les formes:

les opérandes:

x := y		C(a)
x := y Ak y		
<u>si</u> y R y <u>vers</u> < étiquette >		C(a,b)
<u>vers</u> < étiquette >	x =	
<u>insere</u> (n ₁ , n ₂ ,n _i)		C(w)
<u>proc</u> P		
<u>para</u> q ₁	y =	C(I)
<u>index</u> q ₂		
<u>var</u> s; v ₁ , v ₂ ,.....v _s		< constante >

A = opérateur arithmétique

R = opérateur de relation

Les indices: n₁, n₂,n_i, q₁, q₂, s; v₁, v₂,.....v_s sont des entiers.
Le support mémoire centrale se présente de la façon classique:

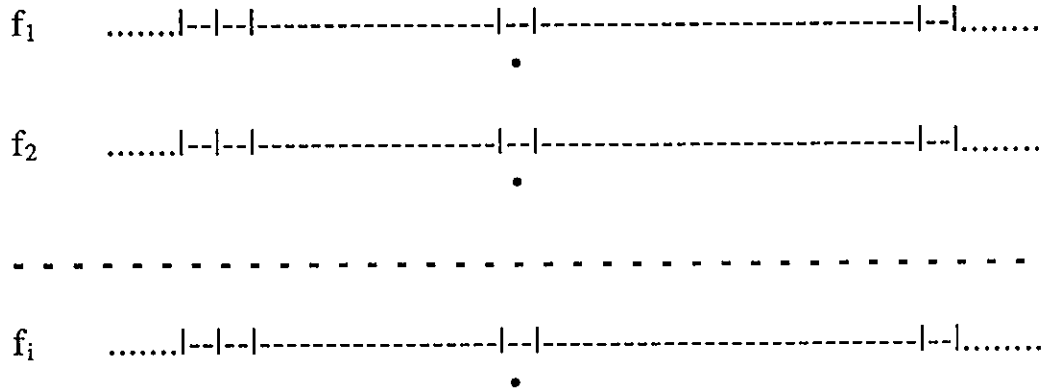
```

|--| |--| |--|-----|-----|-----|-----|-----|-----|-----|
I  w  0 1                                     i                                     k-1

```

Nous avons rajouté la prise en compte de données alphabétiques aux entrées-sorties du compilateur, ainsi que de caractères spéciaux jugés nécessaires, par exemple ' #' pour le traitement des files.

Il nous faut introduire les files externes support de notre organisation de base. Ces files sont constituées de cases en nombre fini illimité. On peut disposer d'autant de cases que le besoin s'en fait sentir. Les files sont constituées de cases toutes identiques entre elles, comme pour la mémoire centrale. A chaque file correspond une tête de lecture-écriture capable d'être déplacée et de lire ou écrire dans la case pointée par cette tête de lecture-écriture notée ' \cdot '. Nous représenterons les files de la façon suivante:



Les instructions d'échange d'information entre mémoire centrale et files sont les suivantes:

- $C(a) := Cf_n$ (transfert de la file à la mémoire centrale),
- $Cf_n := C(a)$ (transfert de la mémoire centrale à la file),
- Avf_n (avance de la tête d'une case à droite),
- Arf_n (arrière de la tête d'une case à gauche),
- $si Cf_n := "$ vers $ez,$
- $Cf_n := '#'$ (valeur '#' marqueur de file).

Les valeurs k et K définissent la machine:

- k est le nombre de cases dans la mémoire centrale,
- K détermine le contenu de cette mémoire.

Pour K valeurs dans la mémoire centrale, nous aurons $K+1$ valeurs dans les files car nous avons une autre valeur '#'.

Compte tenu de l'évolution de la capacité des mémoires "RAM" dynamiques à plusieurs méga-octets; il nous apparaît possible dans un avenir proche de concevoir notre schéma physique, que nous abordons dans le paragraphe suivant, en mémoire centrale spécialisée au traitement du thésaurus qui nous intéresse. Les propositions que nous donnons sont, dans la suite, argumentées sur des supports de files externes.

2.2. SCHEMA PHYSIQUE

Nous donnons ici la possibilité d'intégration et représentation des relations dans le thesaurus de façon très générale. L'administrateur de la base, averti par les concepteurs du fonds documentaire, est totalement maître des relations sémantiques qu'il va introduire au sein du thesaurus pour offrir une meilleure qualité des réponses lors de la recherche rétrospective des informations.

Le système proposé n'utilise aucun "S.G.B.D." standard et travaille exclusivement avec des fichiers classiques organisés de façon séquentielle ou relative. L'organisation définie permet également de n'avoir aucune contrainte majeure en ce qui concerne le nombre de titres, de pages et d'associations.

2.2.1. INDEXATION

Avant de passer à l'architecture, elle même, il nous a semblé bon d'introduire une définition. Le caractère adjacent est celui ci. En effet, le but est de reconnaître facilement les groupes de mots, mots composés et syntagmes. Cette méthode est basée sur le contrôle du caractère adjacent de certains mots à partir d'un thesaurus étendu: le caractère adjacent strict (mots contigus) ou le caractère adjacent large (mots non contigus).

La distance entre les mots significatifs (mots-clés) est exprimée à l'aide d'opérateurs de caractère adjacent:

l'opérateur "-" distance=0

l'opérateur "^" distance \geq 1.

Un groupe de mots est défini comme suit:

- tout mot (ou radical) du thesaurus est considéré comme un groupe de mots élémentaire;

- si g_1 et g_2 sont deux groupes de mots, alors

$g = g_1 - g_2$ est un groupe de mots constitué de g_1 directement suivi de g_2 ;

- si g_1 et g_2 sont deux groupes de mots, alors

$g = g_1 ^ g_2$ est un groupe de mots constitué de g_1 suivi d'un nombre quelconque de mots outils puis de g_2 .

Exemples:

$g_1 = \text{Sulphur}^{\text{iodatum}}$ pour Sulphur iodatum

$g_2 = \text{Barita}^{\text{carbonica}}$ pour Barita carbonica

$g_3 = \text{eau-froide}$ pour eau froide

Dans la figure 4. qui donne l'organisation globale du thesaurus, nous traiterons, à titre d'exemple, les groupes de mots:

$$g_1 = R_k, g_2 = R_k \wedge R_p \text{ et } g_3 = R_l.$$

Les R_z sont des mots ou radicaux comme défini ci-dessus.

L'ensemble physique que nous allons définir permet de mettre en oeuvre la base d'informations en utilisant des fichiers classiques et en respectant les caractéristiques suivantes:

- rapidité d'accès aux informations;
- pas de limitations du nombre de mots-clés, de titres et de pages stockées.

L'architecture globale et assez détaillée est donnée à la figure 4. Les chaînages schématisés ici représentent le cheminement dans la base d'informations à partir d'un radical donné. Ce cheminement est à la base des stratégies mises en oeuvre lors de la recherche des informations pertinentes en réponse à une requête.

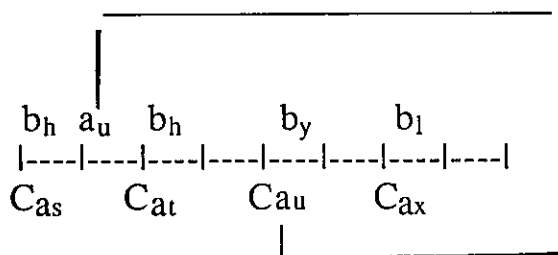


figure 2. Chaînage schématisé par le mécanisme $C_{a,b}$

$C_{as+1,0}$ permettra d'atteindre la case C_{au} et son contenu, lui même une adresse. Il faut naturellement une initialisation convenable de l'index. L'exemple est tiré de la file f2 support des adresses des groupes-mots.

2.2.2. INTEGRATION ET REPRESENTATION DES RELATIONS DANS LE THESAURUS

Chaque fichier principal est proposé sur une file externe. Nous rappelons que les cases mémoires des files ne sont pas numérotées. Toutefois, nous avons pris la liberté de les "numéroter", afin que le lecteur attentif puisse suivre l'explication du mieux possible.

L'organisation globale est illustrée sur un exemple qui est défini comme suit, en faisant les suppositions suivantes:

R_l synonyme de $R_k \wedge R_p$

R_d générique de la classe $\{ R_k \wedge R_p, R_l \}$ et de R_k

$\Rightarrow \{ R_k \wedge R_p, R_l \}$ spécifique de R_d

R_k spécifique de $R_k \wedge R_p$

R_j générique de la classe $\{ R_k \wedge R_p, R_l \}$

$\Rightarrow \{ R_k \wedge R_p, R_l \}$ spécifique de R_j ;

d'où le graphe de la figure suivante:

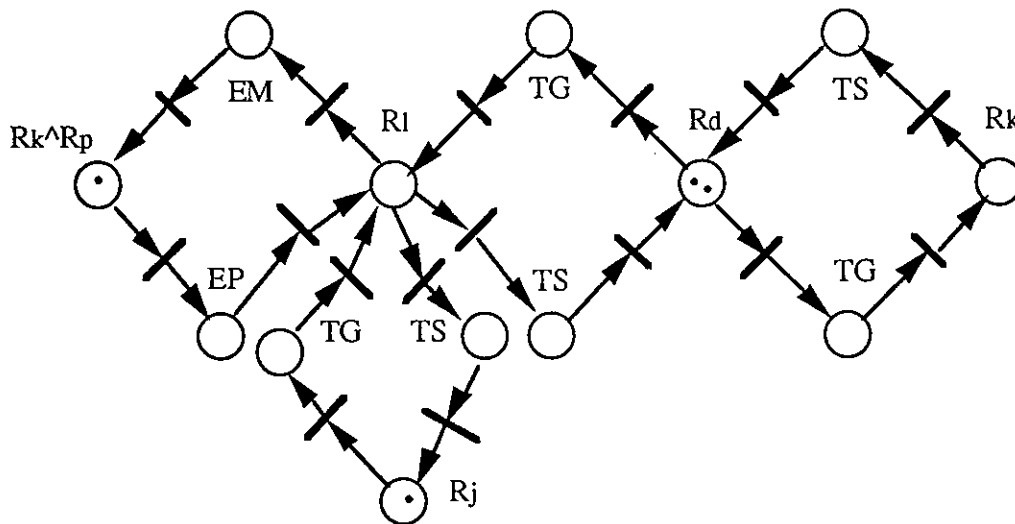


figure 3. Graphe des hypothèses de l'organisation du thesaurus

Les principaux fichiers de type texte utilisés dans la figure 4 sont détaillés ici pour une meilleure compréhension.

Fichier des entrées du thesaurus THESAURUS-ENTREES:

Ce fichier comporte, à terme, tous les radicaux des unités syntaxiques (C_T) extraites par la procédure d'indexation automatique, mais également les radicaux des groupes de mots introduits par l'administrateur au titre des relations sémantiques. Ainsi, chaque radical correspondra à un mot simple ou à un mot composant un (ou plusieurs) groupe(s) de mots qui seront repérés grâce aux deux autres fichiers ADRESSES-GROUPES DE MOTS (C_{T+3}) et GROUPES-MOTS.

Fichier d'adressage des groupes de mots ADRESSES-GROUPES DE MOTS:

Pour un radical donné (C_G), il permet de chaîner les adresses physiques dans le fichier GROUPES-MOTS des groupes de mots comprenant ce radical (C_{G+1}).

Fichier des groupes de mots GROUPES-MOTS:

Il représente les groupes de mots, ou de radicaux (C_M), ainsi que les opérateurs qui relient chaque radical du groupe de mots (C_{M+1}).

Rappelons que nous appelons "groupe de mots" un radical simple (mot simple) ou un ensemble de radicaux ordonnés et connectés par des opérateurs (mot composé ou groupe de mots). De ce fait, à tout point d'entrée du thesaurus correspond au moins un "groupe de mots" dans ce fichier.

Ce fichier permet également de chaîner les groupes mots synonymes au sens défini, et comprend deux pointeurs vers le fichier des relations générique (C_{M+6}) -spécifique (C_{M+7}).

C_{M+2} , adresse physique dans le même fichier du radical relié au radical(i) par l'opérateur(i); C_{M+3} , nombre total d'apparitions du groupe de mots i dans la base; C_{M+4} , adresse physique du synonyme suivant dans la classe de synonymie dans le même fichier et C_{M+5} , adresse physique dans le fichier ADRESSES-RUBRIQUES de la liste des rubriques dans lesquelles ce groupe de mots i apparaît, ne sont valués que pour le premier radical d'un groupe de mots.

Volontairement, cette structure n'est pas optimale au sens où, à terme, des champs ne seront pas valués mais elle a l'avantage de permettre un traitement homogène des groupes de mots et de limiter le nombre de "fichiers" et donc le nombre d'accès à réaliser.

Fichier de la relation GÉNÉRIQUE-SPECIFIQUE:

Ce fichier permet de mettre en oeuvre la relation de hiérarchie entre groupes de mots et de retrouver, à partir d'un groupe de mots donné, la liste de ses termes génériques et spécifiques.

A partir des origines (têtes de chaînages ou d'anneaux) enregistrés dans le fichier GROUPES-MOTS, ce fichier permet de représenter les listes chaînées associant un groupe de mots et la liste des arcs qui partent de ce groupe de mots: liaison générique (C_S, C_{S+1}); et un groupe de mots et la liste des arcs qui arrivent vers ce groupe de mots: liaison spécifique (C_{S+2}, C_{S+3}).

Fichier d'adressage des rubriques ADRESSES-RUBRIQUES:

Il permet d'établir la liaison entre un mot-clé ou groupe de mots du thesaurus et les titres des rubriques où ce groupe apparaît. Il contient alors les adresses physiques des titres des rubriques référencés (C_{A+1}).

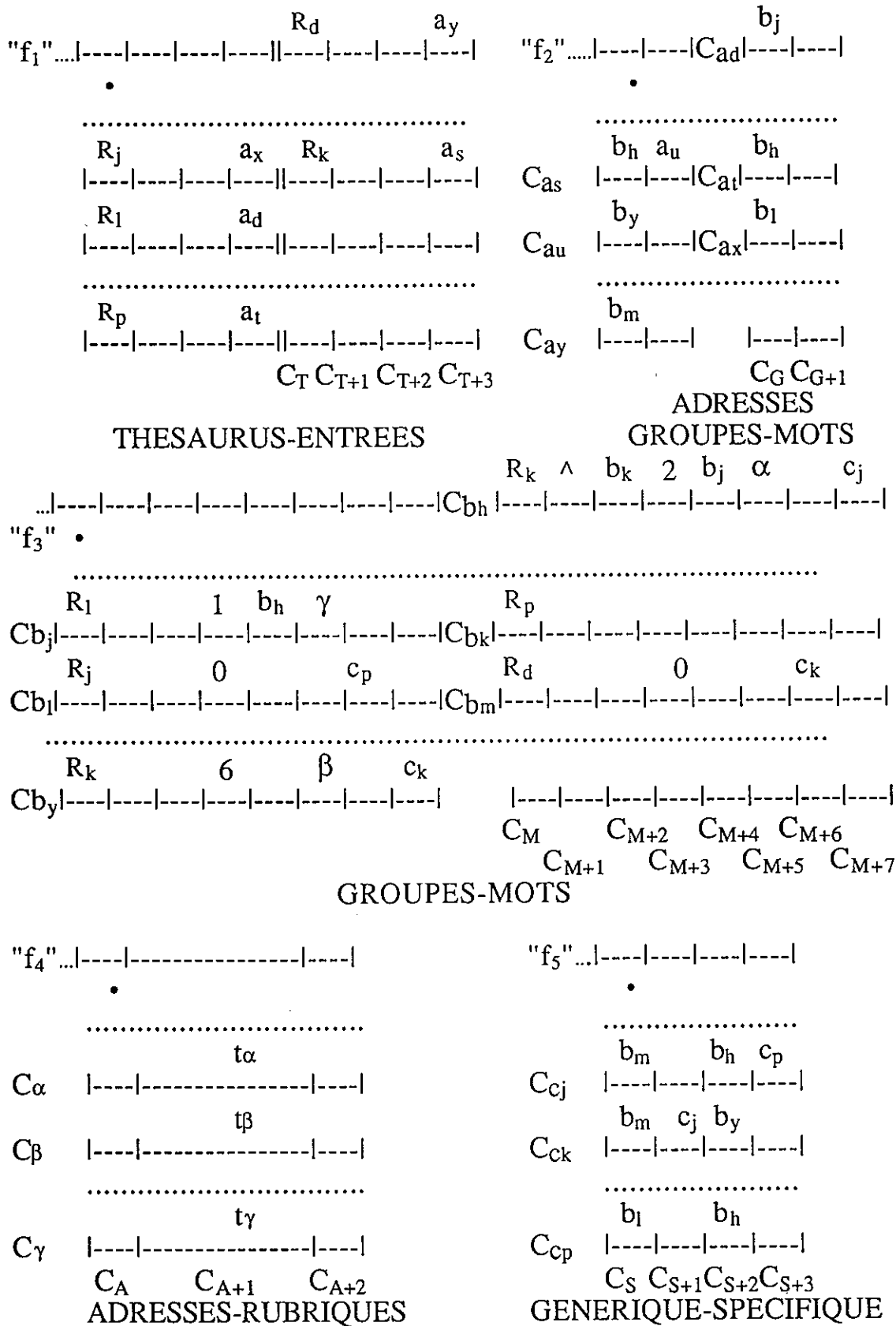


figure 4. Organisation globale du thesaurus.

3. MECANISMES D'INTERROGATION DU THESAURUS

Dans cette partie, consacrée plus précisément aux mécanismes d'interrogation du système de recherche d'information, nous n'aborderons pas toutes les différentes notions liées à la possibilité d'interrogation en langage courant et à la prise en compte des caractéristiques du thesaurus [SOULE-DUPUY CH. 90]. Nous utiliserons les éléments définis dans les articles précédents pour le traitement des groupes de mots et des relations. Nous précisons que dans notre cas du répertoire de Kent, dans un premier temps, les modalités d'intégration de ces notions dans l'analyse et le traitement de la requête en langage courant ne sont pas intégrées.

Pourquoi? L'application n'est pas dite "grand public", au sens où il n'y a pas formulation des requêtes dans un langage proche du langage naturel. Ce principe se justifie d'autant que:

- le vocabulaire met en oeuvre des termes familiers aux utilisateurs homéopathes,
- la consultation du fonds documentaire s'effectue à partir de postes de travail "haut de gamme", dont les claviers sont de bonne qualité et permettent une frappe aisée.

Les problèmes de tolérance orthographique soit de catégorie des fautes d'orthographe, soit de catégorie fautes typographiques ne font pas l'objet de notre étude ici et ont déjà été développées [LASKRI M.T. 87].

Nous allons, avant de voir la structuration libre à priori de la requête au paragraphe 3.2, proposer quelques remarques sur la méthode d'analyse d'un texte en la replaçant dans son cadre général, c'est à dire la construction de thesaurus.

3.1. ANALYSE PRELIMINAIRE

Cette analyse préliminaire, nous le rappelons, se situe dans le contexte de l'analyse de première approche, que nous avons présentée dans le paragraphe 3.2 de l'article précédent. Nous avons opté pour un analyseur rassemblant la connaissance syntaxique. L'analyse consiste à identifier les composants de la phrase et à remplacer les termes du composant par un noeud correspondant dans le schéma intermédiaire (arbre syntaxique). La gestion d'une telle analyse, pour la construction du schéma intermédiaire,

est contrôlée par des règles d'analyse condition-action. Ainsi, nous disposons de connaissances lexicales et de connaissances syntaxiques:

- les connaissances lexicales sont consignées dans un dictionnaire appelé thesaurus primaire;
- les connaissances syntaxiques reflètent la structure générale de la phrase à l'aide de laquelle nous pouvons unifier une grande variété de phrases.

Cette structure est composée donc de:

- NOEUDS représentant des catégories syntaxiques qui peuvent être optionnelles et/ou répétitives et engendrant d'autres catégories syntaxiques,
- FEUILLES qui sont des primitives représentant une catégorie syntaxique de base, qui contrairement aux noeuds n'engendrent plus de catégories syntaxiques, mais des catégories lexicales dont l'ensemble constitue le thesaurus primaire.

Nous donnons ici les notations, qui seront utilisées dans le paragraphe suivant, pour les différentes catégories syntaxiques:

-GN	: groupe nominal,
-GVC	: groupe verbal-complément,
-AD	: adverbe,
-GVC _i	: ième groupe verbal-complément,
-GNR	: groupe nominal réel,
-GV _i	: ième groupe verbal,
-PIP _N	: phrase incidente précédant le GNR,
-GC _i	: ième groupe complément associé à GV _i ,
-PIP _S	: phrase incidente succédant au GNR,
-GC _{ij}	: jème groupe complément d'un même GC _i ,
-GCR _{ij}	: jème groupe complément réel de GC _i ,
-PIP _{C_{ij}}	: jème phrase incidente précédant le GCR _{ij} ,
-PISC _{ij}	: jème phrase incidente succédant au GCR _{ij} .

Analyser correctement une phrase revient donc à générer sans ambiguïté son schéma intermédiaire. Ainsi lors de cette analyse, il est raisonnable de prévoir que l'automate que nous construisons peut se tromper de temps en temps. Mais, nous le construisons de manière à ce qu'il puisse se corriger; pour cela, il faut lui apporter une quantité fantastique d'informations.

Cet automate communique avec le thesaurus primaire pour définir le sens des termes de la phrase à analyser pour spécifier qu'il s'agit d'un terme propre au domaine d'application, d'un verbe, d'un pronom personnel, d'un adjectif possessif, etc... Cette analyse nous permet donc de savoir si nous traitons un terme du groupe nominal, du groupe verbal, ou du groupe complément, afin de générer correctement le schéma intermédiaire.

Enfin, pour fournir une extensibilité au système, l'automate réalisant l'analyse de première approche, permet d'enregistrer les termes qui n'existent pas encore dans le thesaurus primaire et dans le thesaurus secondaire, et de reprendre ensuite l'analyse; cette fonction se réalise avec l'accord de l'utilisateur.

Le système général est résumé dans la figure 5. Il se compose de trois parties essentielles afin d'assurer convenablement les fonctions énumérées, ci dessous:

- la partie PRE-ANALYSEUR permet de transformer principalement la phrase en une suite de catégories syntaxiques;
- la partie MOTEUR D'INFERENCE (MI) traite le résultat du PRE-ANALYSEUR en puisant les connaissances dans la base de connaissances (BC) pour donner la structure complète de la phrase.
- la partie CODIFICATION DES REGLES permet d'obtenir une forme interne facilement gérable par le moteur d'inférence.

Naturellement pour chaque modification de la base de connaissances (ajout de nouvelles règles, suppression de règles ou correction de certaines règles), il faudra re-exécuter le programme de codification pour avoir la nouvelle forme interne de la BC sur laquelle travaillera le MI [DEWEZE A. 93].

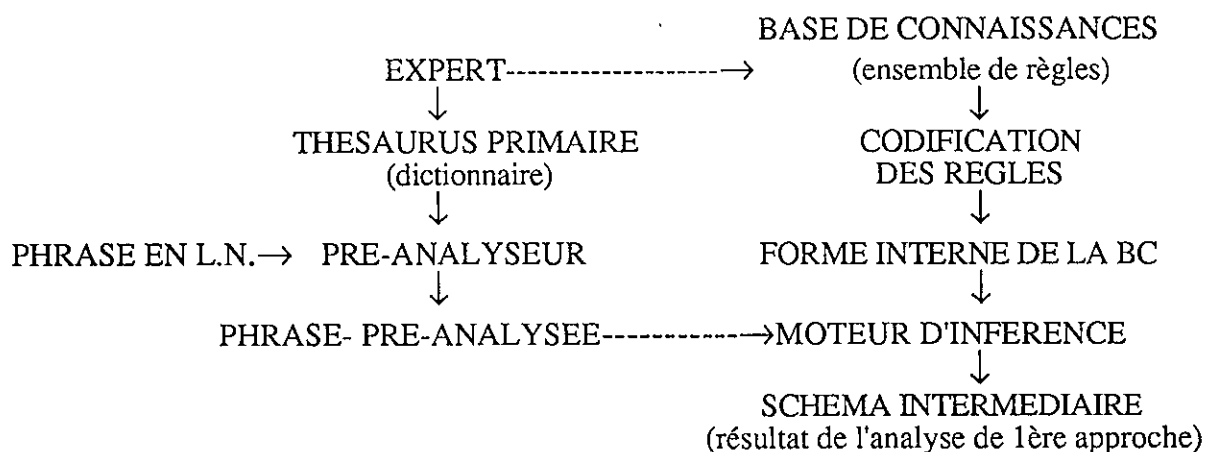


figure 5. Système général assurant l'analyse de première approche

3.2. ANALYSE DE PREMIERE APPROCHE DU LANGAGE NATUREL SUR UN EXEMPLE

Pour voir comment le système fonctionne, nous prenons quelques règles et nous analysons une phrase avec celles-ci. Afin de rendre l'exemple d'analyse évident au suivi et à la compréhension de la méthode, l'ensemble des règles retenues ne concerne que quelques formes de phrases par rapport au schéma intermédiaire général.

- R₀: Si C(noeud_m) = <nil>
Alors attacher 'PH'
- R₁: Si C(noeud_m) = 'PH'
Alors attacher 'GN'
- R₂: Si C(noeud_m) = 'GN' et C(Ptr) = 'GNR'
Alors attacher 'GNR'
- R₃: Si C(noeud_m) = 'GNR' et C(Ptr) = 'GV_i'
Alors | attacher 'GVC_i' à 'PH'
| attacher 'GV_i'
- R₄: Si C(noeud_m) = 'GC_i' et C(Ptr) = 'GV_k'
Alors | attacher 'GVC_k' à 'PH'
| attacher 'GV_k'
- R₅: Si C(noeud_m) = 'GV_k' et C(Ptr) = 'GC_{ki}'
Alors | attacher 'GC_k' à 'GVC_k'
| attacher 'GC_{ki}'
- R₆: Si C(noeud_m) = 'GC_{ki}' et C(Ptr) = 'GC_{ki}'
Alors | attacher 'GC_{ki}' à 'GC_k'
- R_n: Si C(Ptr) = <nil>
Alors FIN.

Soit la phrase suivante à analyser:

" LEDUM est le meilleur antidote de la piqûre de moustique, c'est aussi un excellent remède préventif, et est aussi efficace contre les piqûres de tiques . "

Le pré-analyseur donnerait la suite des catégories:

GNR = LEDUM

GV₁ = est

GC₁₁ = le meilleur antidote de la piqûre de moustique

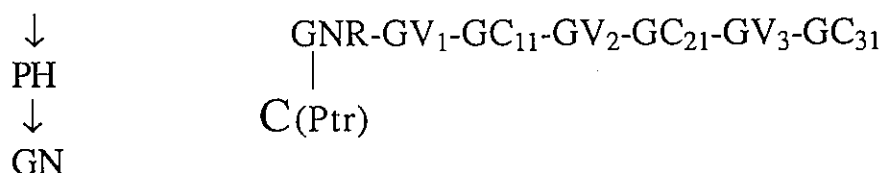
GV₂ = c'est aussi

GC₂₁ = un excellent remède préventif

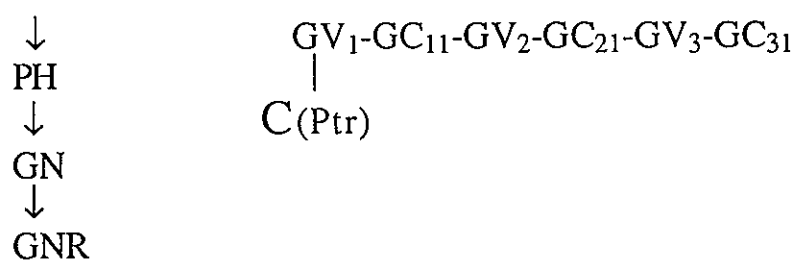
GV₃ = et est aussi efficace

GC₃₁ = contre les piqûres de tiques

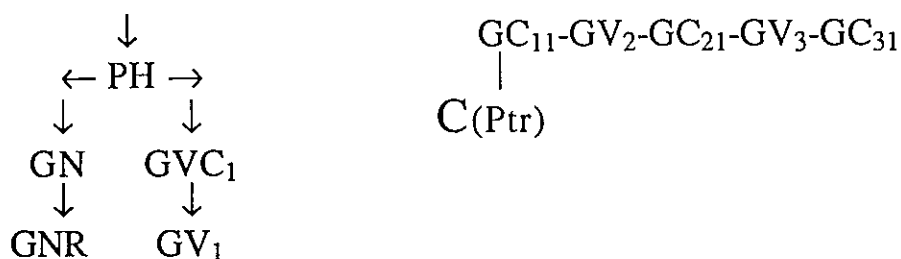
L'application de la règle R₀, ensuite de la règle R₁ donnerait:



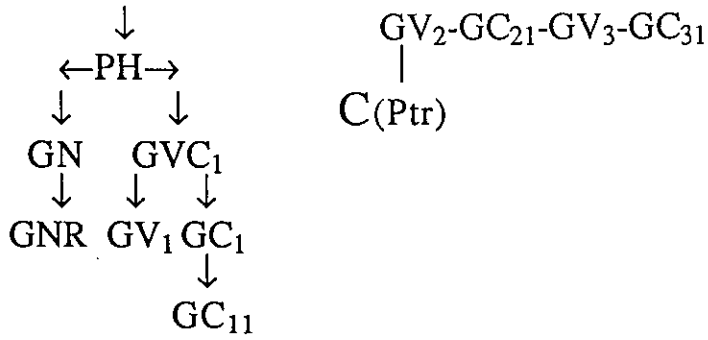
L'application de la règle R₂ donnerait:



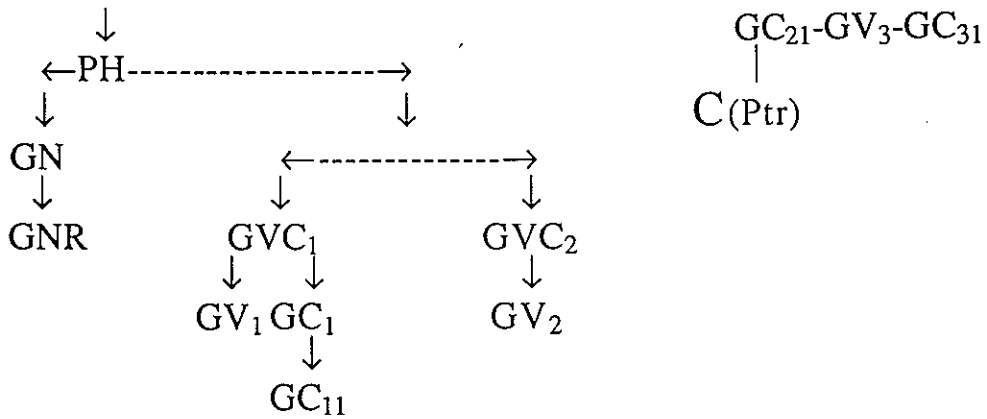
L'application de la règle R₃ donnerait:



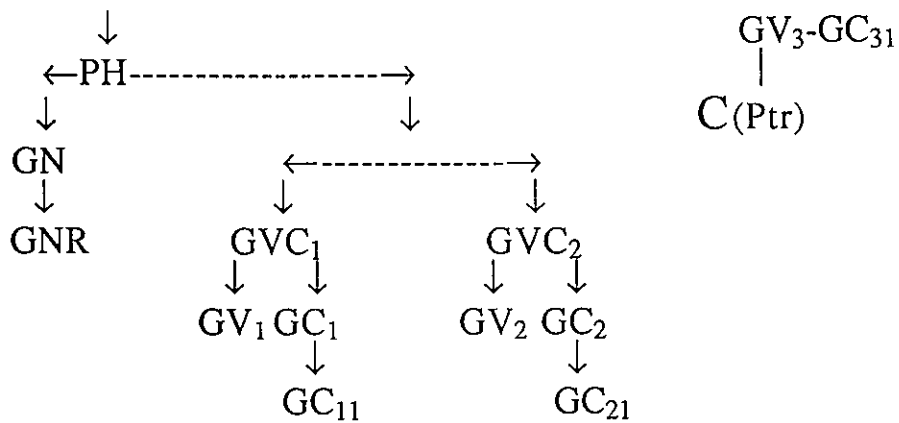
L'application de la règle R₅ donnerait:



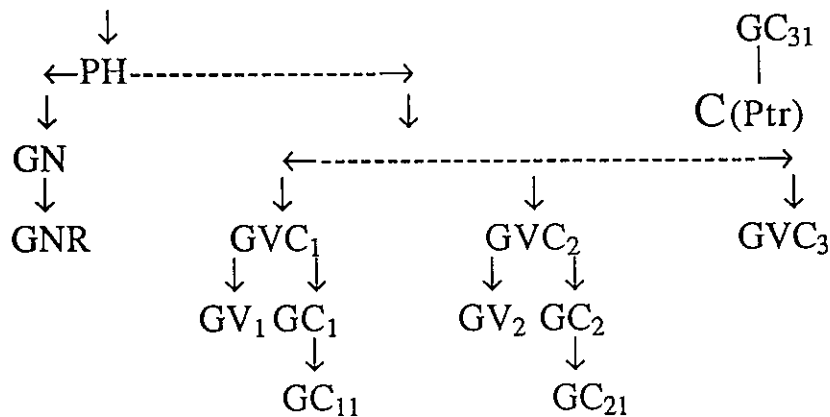
L'application de la règle R₄ donnerait:



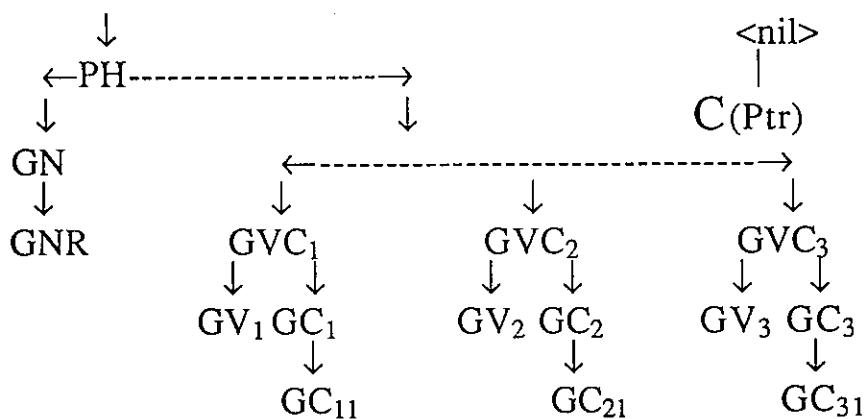
L'application de la règle R₅ donnerait:



L'application de la règle R₄ donnerait:



L'application de la règle R₅ donnerait:



L'application de la règle R_n arrête l'analyse.

Ainsi peut se traiter l'analyse de première approche à base de connaissances par production du schéma intermédiaire, résultat de l'analyse de première approche d'une phrase à analyser.

Un nombre important de recherches en traitement automatique du langage naturel se distingue par l'utilisation des analyseurs hors contexte, voir les travaux d'Ichbiah, et les analyseurs ATN. Cependant, il est à constater que se pose le problème de l'intégration du contexte pour guider l'analyse; et les analyseurs rassemblant la connaissance syntaxique, " Wait And See Parser " offrent une alternative. En effet, le traitement syntaxique est contrôlé par des règles d'analyse condition-action permettant la génération de la représentation interne en fonction de la forme de la phrase prévue dans la base de connaissances.

4. BILAN SUR LE LANGAGE D'INTERROGATION DU THESAURUS

Le paragraphe précédent considère que l'utilisateur formule sa requête en langage libre. Or, il a été constaté dans des expériences d'expérimentation de systèmes d'interrogation que patients ou praticiens, qui interrogent une base assez fréquemment ont tendance à réduire leur requête à une juxtaposition de mots ou de syntagmes.[SOULE-DUPUY CH. 90]. Ainsi, un utilisateur expérimenté, c'est à dire qui a une certaine pratique de l'interrogation du système, pourra limiter sa requête aux seuls groupes de mots significatifs [VIALLET F. 92].

Nous pouvons alors envisager d'étendre les possibilités du langage d'interrogation pour lui donner la possibilité d'affiner sa requête en la structurant à sa guise à l'aide de certaines règles mises à disposition. Ces règles lui donnent la possibilité d'exprimer des associations d'idées.

4.1. STRUCTURATION DE LA REQUETE

Nous rappelons qu'une requête est formée d'une partie principale et d'une partie secondaire, qui peut être vide ou formée d'une séquence de composantes élémentaires, et que l'on met après la partie principale [SIMONET M. 90].

syntaxe: <PARTIE PRINCIPALE> ; <PARTIE SECONDAIRE>

Un point virgule sépare la partie principale de la partie secondaire. La partie secondaire est une séquence de composantes élémentaires séparées par un point virgule et précédées de plus ou moins.

Une composante élémentaire correspond à ce qui a été vu jusqu'à présent. C'est une séquence de mots, pouvant inclure un nom de chapitre, des noms de remèdes et de concepts mais sans virgule ni plus (à moins de mettre entre parenthèses les termes du choix séparés par plus).

4.1.1. EXPLICATION DU DECOUPAGE DE LA REQUETE

Par insertion de signes de ponctuation (une virgule, par exemple) entre deux syntagmes, l'utilisateur peut juxtaposer les mots ou groupes de mots significatifs exprimant ses besoins et uniquement ceux-ci: "..., g₁, g₂,...".

Tous les mots ne jouant qu'un rôle syntaxique sont ainsi éliminés d'office, le système repère alors exactement les groupes de mots à rechercher et l'analyse de première approche s'en trouve optimisée. Il n'aura pas besoin de raisonner par le caractère adjacent implicite, le groupe de mots est explicité, il apparaît ou non dans le thesaurus.

Par exemple: " toux, sang "

au lieu de: "Je voudrais m'informer de toutes les rubriques contenant les mots toux et sang, ou l'un des mots de leurs familles respectives."

4.1.2. OPERATEUR OU LOGIQUE

Par l'emploi du "parenthésage", l'utilisateur peut exprimer, dans sa requête, l'équivalence entre deux ou plusieurs termes. Nous parlerons de synonymie locale, par opposition à la synonymie globale qui est définie au niveau du thesaurus et qui est vraie quelle que soit la requête sur un cas donné. Le système considère tous les groupes de mots entre parenthèses, associés ou non par des "ou", éventuellement séparés par des virgules, comme étant synonymes, et ceci uniquement pendant le temps de traitement de cette requête. La notion de concept permettra, si l'utilisateur le souhaite de mémoriser cette requête de façon définitive et non pas temporaire.

Par exemple: " humide + brouillard + temps brumeux + air des caves " traduit la notion d'humidité. Humidité peut par cette définition être conservé comme concept.

4.1.3. OPERATEUR ET LOGIQUE

En utilisant un opérateur et entre deux mots ou groupe de mots, il est possible de représenter un et logique et d'imposer la présence simultanée, dans une même information, des concepts ainsi associés pour que celle-ci puisse être sélectionnée. La notation de l'opérateur et est la virgule ou l'espace.

Par la requête "...g₁ , g₂ ...", on ne veut que tout ce qui concerne à la fois g₁ et g₂. On rejette toute information contenant seulement g₁ ou g₂.

Par exemple: "humide mouillé", cerne la notion d'humidité précisée dans le paragraphe précédent.

Par la combinaison des règles et des opérateurs définis, un utilisateur pourra alors formuler des requêtes du type:

" (gums bleeding + mucous membrane excoriation) ".

Les règles que nous fournissons permettent d'optimiser le traitement de la requête et d'accroître la pertinence du système (et des réponses), tout en conservant les mêmes principes mis en oeuvre.

4.1.4. OPERATEUR SAUF

L'usage de l'opérateur sauf suppose qu'il soit précédé d'une requête initiale, qu'il va servir à corriger en retranchant certaines rubriques du résultat initial. Il est noté par le caractère '-'.

Exemple: "gums bleeding - detached"

retranchera du résultat la rubrique "Detached" de "teeth", "gums", et "bleed easily". Cet exemple est extrait de la définition du concept Gingivite.

Les noms des remèdes peuvent être utilisés à la place des mots.

Exemple: " *Puls.* - *Lyc.* "

effectue la différence des remèdes, retenant les rubriques contenant *Puls.* mais non *Lyc.*

4.2. TRADUCTION DES REQUETES EN GRAPHERS

Notre objectif dans ce paragraphe est d'exprimer précisément la traduction de la requête en graphes. Ceci est dans le droit fil des propositions de Sowa [PRC-GDR 92]. Cette traduction s'effectue de la manière suivante:

- chaque donnée intervenant dans la requête est traduite par une place $^{\circ}ti$, place Donnée du RDPFD,

- chaque "relation calculée" est l'opérateur Or de la place P_{ij} ,
Remarque: une relation calculée a toujours un et un seul concept résultat.

A titre d'exemple montrons les graphes traduisants les requêtes Q_1 et Q_2 .

Q_1 "Dans quelles rubriques Arnica montana est il présent ?".

Matière médicale du remède: Arn.

Recherche: dans tout le répertoire.

Affichage des rubriques: toutes.

Affichage des rubriques: remède au degré quelconque.

La matière médicale d'un remède est constituée par l'ensemble des rubriques où apparaît ce remède. Le même ensemble de rubriques serait fourni par une requête constituée du seul nom de ce remède. La différence est que l'option matière médicale d'un remède offre des possibilités de sélection et de présentation spécifiques aux remèdes [SAPHIR 91]

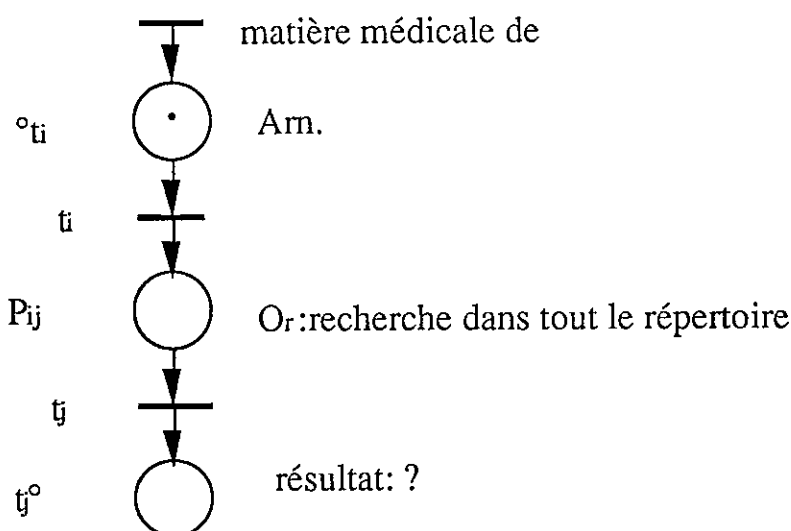


figure 6. Recherche de la matière médicale du remède Arnica montana

Q_2 "Quelles sont les rubriques et leur nombre de remèdes dont les symptômes sont communs à exactement : ars., ars-h., ars-i., ars-m., ars-s-f. ?".

Communauté de symptômes: ars., ars-h., ars-i., ars-m., ars-s-f.

Recherche: dans tout le répertoire.

Symptômes communs à: exactement 5 remèdes.

Affichage des rubriques: toutes.

La présentation des rubriques communes à un ensemble de remèdes permet de dégager les rubriques communes à cet ensemble de remèdes. La recherche peut se faire dans tout le répertoire, dans un chapitre, ou dans la requête courante (exemple: " Puls. - Lyc. ").

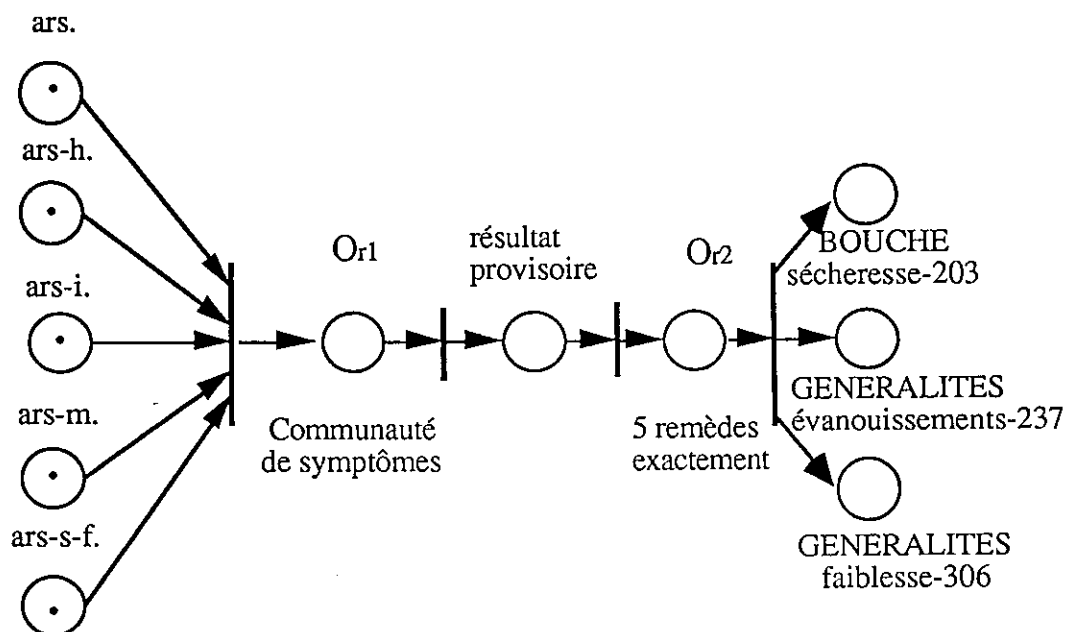


figure 7. Recherche de rubriques communes à un ensemble de remèdes

Volontairement nous avons traduit ces requêtes du langage d'interrogation, en requête de SAPHIR afin de rester familiarisés avec plusieurs formes.

Les travaux menés sur la représentation des concepts et des requêtes donnent deux résultats majeurs, la description sous forme de graphes d'un certain nombre d'objets et l'émergence d'une méthode, dite de "paraphrasage" canonique, permettant à partir d'un corpus la mise en évidence des concepts et des relations.

Nous avons simplement montré ici qu'il était possible de formaliser les requêtes de façon unifiée permettant d'étendre les capacités descriptives du modèle de base. Le travail restant dans cette partie peut s'orienter sur la description formelle de langages; l'objectif principal est de proposer une grammaire décrivant une représentation des requêtes. Ensuite une phase d'apprentissage nécessitera de faire le point sur les travaux concernant l'apprentissage à partir d'exemples, c'est à dire construire un graphe qui soit une généralisation d'un ensemble de graphes donnés. Alors seulement pourra être abordée la phase de développement d'interface comprenant l'intégration d'éditeurs graphiques et linéaires, des traducteurs de ces représentations vers la représentation physique de la représentation standard mentionnée au paragraphe 2.

5. DISCUSSION

En ce qui concerne la proposition d'organisation physique, la partie compilateur et dérouleur de la machine universelle est opérationnelle sur divers calculateurs de type APPLE, IBM, etc...[BISIÈRE CH. 89]. Elle sera ensuite implantée sur le processeur de la procédure formelle, MAC1, actuellement en fin de montage et tests [BIANCO E. 92].

Il restera maintenant un travail non négligeable de vérification à entreprendre sur notre répertoire. En effet, la phase de développement réaliste était limitée par le nombre des états pris en compte par les logiciels de vérifications. Les résultats actuels permettent de traiter des systèmes de plusieurs millions d'états, alors qu'on en était à 50 000 états il y a quelques années. Rappelons que l'homéopathie différencie seize à dix huit mille symptômes, avec une "banque de remèdes" évalués déjà à six cent quarante à l'époque de Kent.

L'originalité de notre approche réside dans la coopération de deux modes:

- un langage graphique, qui a inspiré des systèmes hypertextes qui facilite la manipulation d'objets complexes à partir d'un graphe visualisant un schéma,
- un mode d'interrogation en langage quasi-naturel, inspiré des systèmes de recherche d'informations documentaires qui permet une recherche portant sur le contenu sémantique du document [LN-GRTC 93].

Indépendamment de tout impact d'un tel travail, il paraît important dans une théorie qui se veut homogène, et capable de recouvrir tout le champ de l'informatique fondamentale, de montrer que les propriétés avancées, sont vraies quelles que soient les formes possibles adoptées.

6. REFERENCES BIBLIOGRAPHIQUES

Publications sur le projet

- [BIANCO E. 92]

Le processeur de la procédure formelle.
Bulletin d'informatique approfondie et applications.
ISSN 0291-5413. N°31. Mars 1992. Marseille.

- [BISIÈRE CH. 89]

Bisière Ch., Iwanenko D., Knippel J.M., Massat J.L.
Compilateur de machine universelle: MU.
Bulletin d'informatique approfondie et applications.
ISSN 0291-5413. N°23. Juin 1989. Marseille.

Autres références

- [DEGOULET P. 92]
Degoulet P., Fieschi M.
Collection Informatique et Santé. N°5.
Nouvelles Méthodes de Traitement de l'Information Médicale.
Springer-Verlag. 1992. Paris.
- [DEWEZE A. 93]
Informatique documentaire. 4ème édition.
Masson. 1993. Paris.
- [GOOSSENS B. 93]
Goossens B., Akil M.
MT: A multithreaded 64 bits RISC CPU.
Parallel Computing Technologies (PaCT-93).
Edt. by V.Malyshkin. 1993. Obninsk. Russie.
- [PRC-GDR 92]
Actes des 4èmes Journées nationales PRC-GDR Intelligence Artificielle.
TEKNEA. 1992. Marseille.
- [SOULE-DUPUY CH. 90]
Systèmes de recherche d'informations. Le système videotex INFOLAB.
Thèse pour l'obtention du grade de docteur de l'Université.
Spécialité: Informatique. IRIT / CERFIA. UA CNRS 824.
Université Paul Sabatier. 1990. Toulouse.
- [TURING A. 48]
Collected Works of A.M.Turing.
Mechanical Intelligence.
Ince D.C. Editor. North-Holland. 1992. Pays Bas.
- [VIALLET F.92]
Garraud J., De Graeve J., Viallet F.
Construction et maintenance d'un thesaurus dans les télécommunications.
Conférence: Bases de données et intelligence artificielle.
Saint Denis. 1992. France.

Note interne

- [LN-GRTC 93]
Rapport scientifique et technique 1993 de l'équipe langage naturel (LN).
Groupe représentation et traitement des connaissances. CNRS.
GRTC - CNRS. 1993. Marseille.

VOUZZAVEDIBISAR,

Le symbole (suite).

Mais il est un symbole encore plus puissant que celui de l'argent bien qu'il lui ait été et qu'il lui soit encore étroitement lié. Le symbole est par essence abstrait, et tout symbole mystique est une abstraction d'abstraction qui accroît encore sa puissance par action sur les imaginations pusillanimes, et aussi sur les autres, mais c'est alors question d'intérêts et de partage de pouvoir.

Il fut une époque romantique où le langage des amoureux s'enrichissait de symboles délicats et merveilleux, dont l'odeur envoûtante et la richesse des couleurs soulignaient avec force les mots d'amour. On a oublié le langage des fleurs. La simplicité des œillets et des roses a été un peu délaissée au profit de fleurs étranges d'origine lointaine qui ne sont pas toujours très à l'aise chez nous. Ces symboles d'une vie agréable, qui sont les marques des relations harmonieuses entre les gens, perdent graduellement de leur intérêt. Ce qui devrait être considéré comme des symboles forts de la société, le véritable ciment qui assure le plaisir dans la relation, est largement déprécié au profit d'autres symboles, moins conviviaux mais tellement plus efficaces pour l'agrégation du pouvoir. Le "profit", puisqu'il faut le nommer. Voilà un symbole réaliste, moteur de toute bonne "économie". Qui n'est pas parfaitement convaincu que seule une bonne économie est capable de faire bien fonctionner un pays? Même si, pour cela, on doit passer de trois à quatre millions de chômeurs.

Dont tout le monde sait, tout aussi parfaitement, qu'il ne faut pas trop les indemniser pour ne pas leur donner le goût de ne rien faire, le plaisir d'être assisté. De toute façon il est clair que si tout marchait mal c'est de leur faute, puisqu'après les avoir virés, tout est tellement allé mieux !

On peut se demander comment il se fait que de tels symboles n'aient pas été remis en cause. Le plus grave est qu'ils l'ont été, et sérieusement. Mais les solutions alternatives envisagées non seulement n'ont rien résolu mais se sont révélées pires que le mal. En conséquence le symbole "idéologie" a endossé une très mauvaise réputation, et le symbole "économie" est sorti renforcé. D'autres symboles sous-jacents ont également profité de l'aubaine.

Au cours des âges les hommes ont ressenti un pressant besoin de se réfugier à l'ombre des croix, et des croix, on en a connu de toutes sortes, toutes plus ou moins sanglantes. L'une des plus vieilles, la croix du Christ se révèle inusable à l'usage. Au départ, simple instrument de torture des Carthaginois, dégoulinante de sang, elle a envahi le monde chrétien, et le monde qui a été envahi par le monde chrétien.

D'autres croix ont eu un succès plus discret ou plus éphémère. La croix de Malte, la croix de fer, la croix d'occident, qui ressort toujours quand on parle d'ordre nouveau, la croix de Lorraine, symbole d'un grand Général puis d'une confrérie de féaux (+) amis d'ycelui (*), toutes les croix de guerre, distribuées au poids de viande hachée, et enfin la svastika, la croix gammée, symbole du héros d'un abominable thriller qui n'a pas fini de faire des adeptes.

La croix chrétienne, un temps parfumée à l'opium, a pu voir son étoile décliner dans le ciel. Mais depuis, l'ange exterminateur est venu rétablir l'ordre sur terre. Le sabre et le goupillon relèvent la tête dans un même sursaut jaculatoire (ô mânes de Pierre Dac).

Mais tout danger n'est pas écarté car l'œil dans le triangle a un concurrent dangereux qui date de l'hégire. Le nouveau venu de six cents ans plus jeune ne semble guère s'embarrasser de scrupules. Il suffit d'observer les talibans en pleine action pour s'en convaincre.

C'est comme si l'humanité grimpait à une sorte d'échelle dont les barreaux seraient constitués de tous ces symboles à plusieurs faces, et qui, rongés par la vermine du temps, se brisent périodiquement, faisant chuter le monde dans le chaos.

Quand je reviens à la naïve machine de Nolin, qui traite certes des suites de symboles, en fait une poignée de petits dessins figés, j'ai peine à l'imaginer en train de triturer cette bouillie malsaine destinée à combler les vides imaginatifs des populations en manque d'oppressions.

Edmond Bianco

(+) féal, e, aux. Adjectif (de foi). Fidèle.

(*) relire Ronsard.

Université de Provence
Atelier de Reprographie
Centre Saint Charles
3, place Victor Hugo
F - 13331 Marseille Cedex 3

