

89 | 339

Sc N 89 / 412 A

Université de Nancy I
Faculté des Sciences

Centre de Recherche en
Informatique de Nancy (UA n° 262)

**APPRENTISSAGE POUR LA RECHERCHE
INTERACTIVE ET PROGRESSIVE D'IMAGES :
PROCESSUS EXPRIM ET PROTOTYPE RIVAGE**

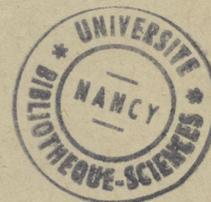
THESE

Présentée et soutenue le 9 octobre 1989
pour l'obtention du

Doctorat de l'Université de Nancy 1
Spécialité INFORMATIQUE

par

Gilles HALIN



Composition du jury :

Président :	Odile FOUCAUT	Professeur à l'Université de Nancy II
Rapporteurs :	Jean Marie PIERREL Yves CHIARAMELLA	Professeur à l'Université de Nancy I Professeur à l'Université de Grenoble
Examineurs :	Marion CREHANGE Maryse QUERE Jean SALLANTIN	Professeur à l'Université de Nancy II, Directeur de thèse. Professeur à l'Université de Nancy II Chargé de recherche CNRS, Centre de Recherche en Informatique de Montpellier
Invités :	Muriel CLUZEAU Michel REBUFFET	Maitre de conférence à l'IUT de Dijon Société Européenne de Propulsion (MS2I)

**APPRENTISSAGE POUR LA RECHERCHE
INTERACTIVE ET PROGRESSIVE D'IMAGES :
PROCESSUS EXPRIM ET PROTOTYPE RIVAGE**

THESE

Présentée et soutenue le 9 octobre 1989
pour l'obtention du

Doctorat de l'Université de Nancy 1
Spécialité INFORMATIQUE

par
Gilles HALIN



Composition du jury :

- | | | |
|----------------------|-------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Président : | Odile FOUCAUT | Professeur à l'Université de Nancy II |
| Rapporteurs : | Jean Marie PIERREL
Yves CHIARAMELLA | Professeur à l'Université de Nancy I
Professeur à l'Université de Grenoble |
| Examineurs : | Marion CREHANGE

Maryse QUERE
Jean SALLANTIN | Professeur à l'Université de Nancy II,
Directeur de thèse.
Professeur à l'Université de Nancy II
Chargé de recherche CNRS, Centre de
Recherche en Informatique de Montpellier |
| Invités : | Muriel CLUZEAU
Michel PERDUEFT | Maitre de conférence à l'UT de Dijon
Société Européenne de Recherche (AESR) |

A mes parents,
Pour l'aide et l'amour que vous m'offrez constamment.

A Sandrine,
Pour l'amour et la tendresse que tu me prodigues, et la confiance que tu m'accordes.

A mon frère,

A toute ma famille,

A mes amis.

Je tiens à remercier ici :

Madame Marion Créhange, Professeur à l'Université de Nancy II, qui m'a accueilli très chaleureusement dans son équipe en me proposant le sujet de cette thèse. Je la remercie pour les conseils très enrichissants qu'elle m'a si amicalement donnés, et en particulier pour l'ambiance à la fois studieuse et décontractée qu'elle sait faire régner dans son équipe.

Madame Odile Foucaut, Professeur à l'Université de Nancy II, qui me fait l'honneur de présider ce Jury. Je tiens à lui exprimer ici toute ma gratitude pour l'aide qu'elle m'a apportée tout au long de cette thèse.

Monsieur Yves Chiaramella, Professeur à l'Université de Grenoble, qui m'a fait le grand plaisir d'être rapporteur de cette thèse. Je le remercie pour la patience dont il a fait preuve pour lire les différentes versions de mon manuscrit et pour les remarques toujours pertinentes qu'il m'a adressées.

Monsieur Jean-Jacques Pirevel, Professeur à l'Université de Nancy I,
pour l'honneur qu'il m'a fait d'avoir accepté de rendre
compte de mon travail. Ses remarques ont été essentielles
quant à la forme à donner à ce manuscrit, je tiens à
vous exprimer ici ma profonde reconnaissance.

Madame Maryse Quéiré, Professeur à l'Université de Nancy II,
qui me fait un grand honneur à siéger dans ce Jury.

Monsieur Jean Sallantin, Chargé de Recherche au CNRS, pour
l'intérêt qu'il a porté à ce travail en acceptant de siéger à
ce Jury.

Madame Michèle Chazeau, Maître de conférence à l'IST de Dijon,
qui a su jouer pleinement son rôle d'expert en recherche
iconographique, je la remercie de m'avoir très amicalement
communiqué une partie de son expérience.

Monsieur Michel Rebuffet, Directeur d'études à la Société
Européenne de Propulsion (SEP), sans qui je n'aurais
jamais obtenu un contrat Cifre avec la SEP, je tiens à le
remercier pour m'avoir soutenu et aidé tout au long de
ce contrat.

Je tiens à remercier aussi tous ceux qui ont contribué
à la réalisation de cette thèse :

François, pour tes lectures méthodiques, j'espère que l'aide
que je pourrai t'apporter dans ton travail sera à la hauteur
de celle que tu m'as donnée.

Sabine, pour les nombreux tests que tu as effectués. J'ai eu un grand plaisir à travailler en ta compagnie.

Sandrine, pour la saisie extraordinairement complète du Thesaurus et des indexations qui m'a permis d'expérimenter en toute confiance.

Le Directeur et tout le personnel des Archives Photographiques du Fort St Cyr qui ont accepté de participer à ces recherches.

Ainsi que tous les membres actuels de l'équipe Express (non encore cités) : Noureddine, Odile, Daniel, David, Pierre et Dominique pour le soutien qu'ils m'ont apporté et la bonne humeur qu'ils font régner dans l'équipe.

En espérant n'avoir oublié personne,

Nancy le 9/9/89

Gilles Halin.



RESUME

La recherche d'images est une recherche d'informations à part entière, mais les modèles utilisés dans les systèmes de recherche d'informations (SRI) existants ont été créés pour rechercher de l'information textuelle. Ces modèles doivent être adaptés pour prendre en compte ce nouveau type d'information qu'est l'image. L'arrivée des supports optiques a permis à l'ordinateur de manipuler l'image en tant que document au même titre qu'un document textuel. Un système de recherche d'images doit prendre en compte les caractéristiques de l'image et utiliser des modèles de recherche d'informations adaptés.

Cette thèse propose un processus interactif et progressif de recherche d'images basé sur le processus EXPRIM et inspiré des techniques d'apprentissage symbolique développées en IA. Le processus EXPRIM itère sur trois phases non réellement distinctes:

- la phase Avant-Visualisation : l'utilisateur formule une demande que le système interprète pour sélectionner un ensemble d'images à partir de leur description se trouvant dans une base descriptive.
- la phase Visualisation : l'utilisateur visualise les images sélectionnées par le système et effectue un choix en composant deux ensembles d'images : les images choisies et les images rejetées.
- la phase Après-Visualisation : le système analyse les choix de l'utilisateur pour essayer de comprendre ses besoins afin de les formuler sous la forme d'une nouvelle demande.

La méthode proposée assimile le processus EXPRIM à un processus d'apprentissage où le concept à apprendre est le besoin de l'utilisateur et où les exemples et contre-exemples du concept sont respectivement les images choisies et les images rejetées. Le concept est représenté dans le thesaurus en attachant des poids aux termes mesurant la capacité des termes à exprimer les besoins de l'utilisateur ; l'ensemble de ces poids forme ce qu'on appelle le **niveau d'expression** de la demande.

Cette méthode contient : un modèle de représentation du concept à l'intérieur du thesaurus, un mécanisme d'évolution de ce concept en fonction des choix de l'utilisateur, une méthode de construction du concept à partir de sa représentation dans le thesaurus, des mécanismes d'interprétation et de déformation du concept appris pour la sélection d'images dans la base.

Le prototype RIVAGE met en œuvre la méthode avec comme champ d'expérimentation une base d'images du Ministère de la Culture. Ce prototype a été réalisé dans un environnement objet et il a été évalué à l'aide de techniques classiques d'évaluation des systèmes de recherche d'informations. Cette réalisation constitue la deuxième expérience de mise œuvre d'un tel système car au cours du projet ESPRIT, une première approche a été proposée et une amorce de système a été réalisée.

Cette thèse se termine en présentant une évolution du système RIVAGE, qui est en cours d'expérimentation. Cette évolution consiste à implanter une structure Hypertexte adaptée à la recherche progressive et interactive d'images, offrant à l'utilisateur tout le confort nécessaire à une telle recherche.

**Apprentissage pour le recherche interactive et progressive
d'images :
processus EXPRIM et prototype RIVAGE**

Plan de la thèse

INTRODUCTION	1
Première partie	
Etat de l'art sur les systèmes de recherche d'informations	
Introduction	5
Chapitre 1	
Notion de base en recherche d'informations classique	7
I Introduction	7
II Problématique	8
III Notion de pertinence	8
III.1 La pertinence globale d'un système	9
III.2 La pertinence utilisateur	11
III.3 La pertinence système	12
IV Les fonctions d'un système de recherche d'informations	12
IV.1 L'indexation	13
IV.1.1 Cadre général	13
IV.1.2 Les modèles de documents	14
IV.2 L'interrogation	15
IV.2.1 Généralités	15
IV.2.2 La formulation booléenne	16
V Conclusion	17
Chapitre 2	
Les principaux modèles et fonctions	20
I L'indexation	20
I.1 Les modèles d'indexation	20
I.1.1 Définition des données de recherche	20
I.1.2 Les langages d'indexation	21
I.1.2.1 Modes d'indexation	21
I.1.2.2 Les vocabulaires des langages libres	22
I.1.2.3 Les vocabulaires des langages contrôlés	22
I.1.3 Conclusion	23
I.2 Structure d'un thesaurus	24
I.2.1 Les unités lexicales	24
I.2.2 Les relations sémantiques	26

I.2.2.1 L'appartenance à un thème.....	26
I.2.2.2 L'équivalence sémantique.....	27
I.2.2.3 La relation de hiérarchie.....	28
I.2.2.4 La relation d'association.....	29
I.3 Conclusion.....	30
II Notion de modèles de correspondance.....	31
II.1 Généralités.....	31
II.2 Problématique.....	31
II.3 Le modèle vectoriel.....	32
II.4 Conclusion.....	34
III Autres formes d'interrogation.....	34
III.1 La formulation avec pondération.....	34
III.2 La formulation avec un thesaurus.....	35
IV Conclusion.....	36
Chapitre 3	
Les tendances possibles.....	38
I Le bouclage de pertinence.....	38
II Les modèles de correspondances.....	39
II.1 Les modèles probabilistes.....	39
II.2 Les modèles à base de logique booléenne	
II.2.1 Mise en correspondance.....	44
II.2.2 Inconvénients des modèles booléens.....	44
II.2.3 Solutions possibles.....	44
II.3 Le langage naturel et les modèles linguistiques.....	46
II.3.1 L'interrogation en langage naturel.....	46
II.3.2 Le langage naturel et ses inconvénients.....	48
III Autres approches de l'interrogation.....	49
III.1 Avantages et inconvénients de la formulation booléenne.....	49
III.2 La pondération dans la formulation booléenne.....	51
III.3 Les solutions intermédiaires.....	52
III.4 Conclusion.....	54
IV Conclusion.....	55
Chapitre 4	
Apport de l'Intelligence Artificielle dans la recherche d'information.....	57
I Introduction.....	57
II L'indexation.....	58
III L'interrogation.....	59
III.1 Les systèmes experts.....	60
III.2 Les systèmes à base de connaissances.....	60
III.2.1 Le système IOTA.....	62
III.2.2 Le système I ³ R.....	64

III.3 L'apprentissage et ses applications en recherche d'informations.....	67
III.3.1 L'apprentissage à court terme.....	67
III.3.2 L'apprentissage à long terme.....	69
IV Conclusion.....	70
Conclusion.....	71
Deuxième partie	
Une recherche d'informations particulière la recherche d'images	
Introduction.....	73
Chapitre 1	
Cadre général.....	75
I Les différences apportées par l'image.....	75
I.1 Le sens et l'image.....	75
I.2 Les images et les mots.....	77
I.2.1 Perception d'un mot.....	77
I.2.2 Le pouvoir d'imagerie.....	78
I.2.3 Application à la recherche d'images.....	79
I.3 L'analyse de l'image.....	80
I.3.1 Système descriptif des représentations.....	80
I.3.2 L'analyse d'une photographie.....	83
I.3.2.1 La morphologie d'une photographie.....	83
I.3.2.2 Le contenu informatif.....	84
I.3.2.3 La connotation.....	86
I.3.3 Conclusion.....	87
I.4 La recherche en elle-même.....	87
I.4.1 Historique.....	87
I.4.2 Les supports optiques.....	88
I.4.3 La visualisation.....	90
I.4.4 Conclusion.....	91
II La recherche d'images et ses applications.....	91
II.1 Les utilisateurs et les utilisations.....	91
II.1.1 Profil des demandeurs d'images.....	92
II.1.2 Les types de demandes.....	93
II.2 Les systèmes existants.....	94
II.2.1 Les systèmes liés à une application.....	94
II.2.2 Les systèmes construits pour la recherche d'images.....	96
III Conclusion.....	98
Chapitre 2	
L'approche EXPRIM.....	101
I Introduction.....	101

II Historique.....	101
III Les idées principales du projet EXPRIM.....	102
III.1. Le processus EXPRIM.....	102
III.2 L'architecture proposée.....	104
III.3 Les connaissances et expertises utilisées.....	105
III.3.1 L'expertise du domaine.....	105
III.3.2 L'expertise de base.....	108
III.4 Conclusion.....	109
IV Les études et réalisations.....	109
IV.1 Première réalisation.....	110
IV.1.1 Architecture du système et description des différents constituants.....	110
IV.1.2 Les connaissances et données utilisées et leur représentation.....	110
IV.1.3 Les fonctions réalisées.....	111
IV.1.4 Le jeu d'essai.....	112
IV.1.5 L'évaluation du système.....	112
IV.1.6 Conclusion.....	113
IV.2 Deuxième réalisation.....	114
IV.2.1 Architecture du système et description des différents constituants.....	114
IV.2.2 Les connaissances et données utilisées et leur représentation.....	114
IV.2.3. Les fonctions réalisées.....	115
IV.2.4 Le jeu d'essai.....	118
IV.2.5 L'évaluation du système.....	118
IV.2.6 Conclusion.....	118
V Conclusion.....	119
Chapitre 3	
Première étude : le projet ESPRIT.....	121
I Introduction.....	121
II Spécification du système.....	122
II.1 Organisation générale.....	122
II.1.1 Le graphe d'enchaînement des processus.....	122
II.1.2 Description générale des processus.....	124
II.2 Présentation des deux principaux processus.....	125
II.2.1 L'évaluation d'une demande (phase "Avant-Visualisation").....	125
II.2.1.1 Position du problème.....	125
II.2.1.2 L'évaluation et la transformation de la requête utilisateur.....	126
II.2.2 La reformulation de la demande (phase "Après-Visualisation").....	128
II.2.2.1 Analyse des choix.....	129
II.2.2.2 Reformulation de la demande.....	132
III Réalisation.....	134
III.1 Introduction.....	134
III.2 L'application pilote.....	134

III.2.1 Les images et descriptions.....	135
III.2.2 Le thesaurus.....	137
III.3 L'expérimentation.....	138
III.3.1 Architecture du système.....	138
III.3.1.1 Présentation générale.....	138
III.3.1.2 Les choix des différents composants.....	138
III.3.1.2.1 L'imageur documentaire.....	138
III.3.1.2.2 Le système documentaire.....	141
III.3.1.2.3 Le moteur d'inférence.....	141
III.3.1.2.4 Le module de dialogue.....	142
III.3.1.2.5 Le module de pilotage.....	142
III.3.2 Les parties réalisées.....	142
IV Conclusion.....	144
Conclusion.....	146
Quatrième partie	
Le processus EXPRIM vu comme un problème d'apprentissage	
Introduction.....	149
Chapitre 1	
Position du problème.....	151
I Introduction.....	151
II La reformulation de requête, un problème de "Relevance feedback".....	152
III Notions générales sur l'apprentissage symbolique.....	154
III.1. Généralités.....	154
III.2 Quelques méthodes.....	156
III.2.1 Généralisation par points de vue et apprentissage de concepts, de Quinqueton & Sallantin.....	156
III.2.2 INDUCE de Michalski.....	158
III.2.3 AGAPE de Kodratoff.....	160
III.2.4 L'espace des versions de Mitchell.....	161
III.3 Conclusion.....	164
IV Conclusion.....	165
Chapitre 2	
La méthode proposée.....	167
I Introduction.....	167
II Présentation intuitive de la méthode.....	167
III Description des données impliquées dans la méthode.....	169
III.1 Le thesaurus, la connaissance sémantique.....	170

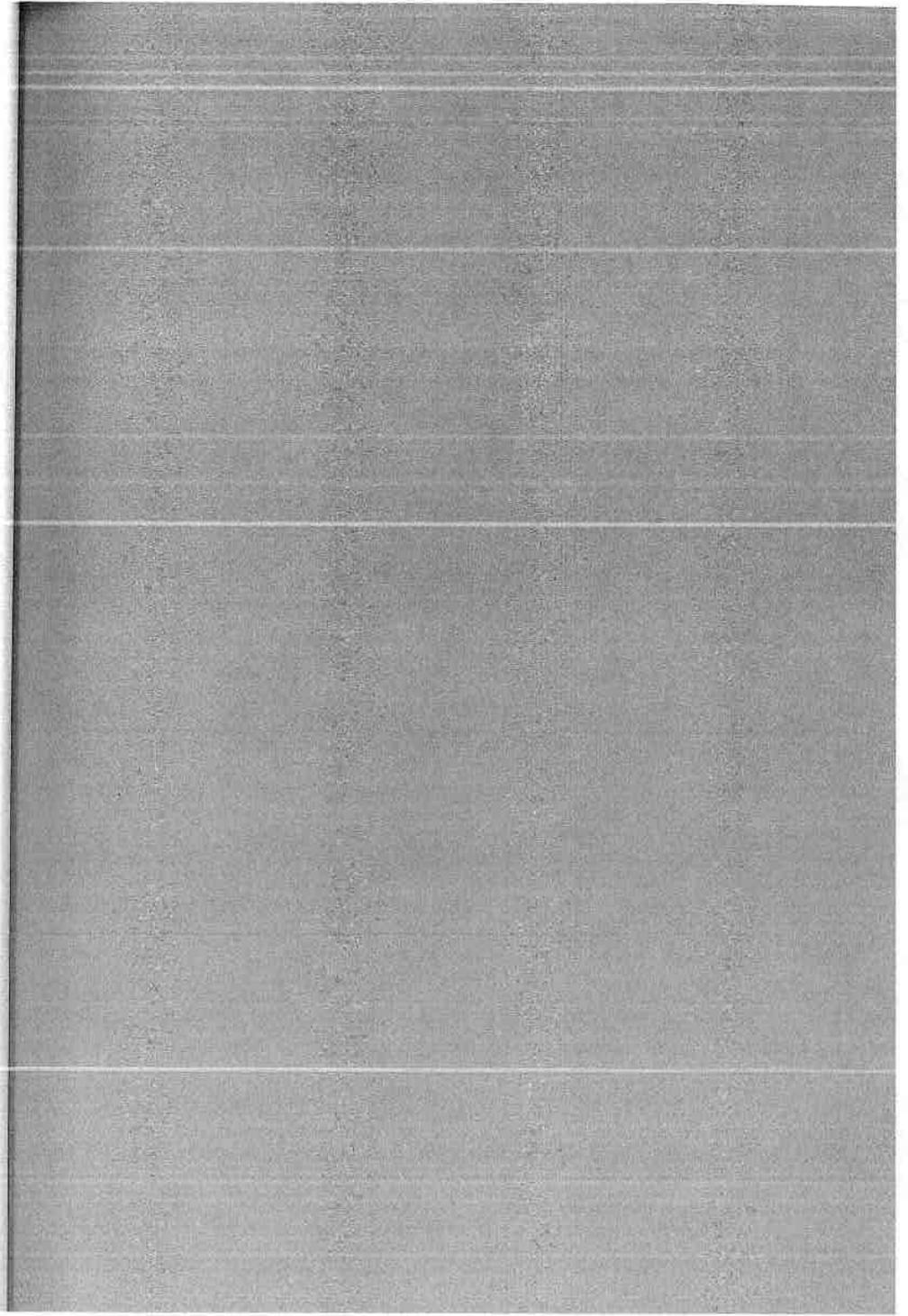
III.1.1 Généralités sur les taxonomies.....	170
III.1.2 Le thesaurus vu comme un ensemble de taxonomies.....	170
III.2 Les descriptions d'images, exemples du concept.....	172
III.3 La demande, le concept à apprendre.....	173
III.4 Niveau d'expression de la demande.....	174
III.4.1 Représentation.....	174
III.4.2 Evolution globale.....	174
IV Evolution du poids d'expressivité pendant la phase "Avant-Visualisation".....	176
IV.1 Mise en place d'un premier niveau.....	176
IV.2 La mise en correspondance entre demande et descriptions d'images.....	177
IV.2.1 Le modèle de mise en correspondance.....	177
IV.2.2 La ventilation et l'interprétation.....	180
IV.2.3 La déformation (3).....	186
V Evolution du niveau d'expression pendant la phase de "Visualisation".....	188
VI Evolution du niveau d'expression dans la phase "Après-visualisation".....	189
VI.1 La pertinence d'un terme (4).....	189
VI.2 Mise à jour du niveau d'expression (5).....	192
VII Formulation du concept appris, construction d'une nouvelle demande.....	193
VII.1 Extraction des termes exprimant le concept (6).....	193
VII.2 Dialogue pour une reformulation (7).....	195
VIII Conclusion.....	197
Chapitre 3	
Expérimentation de la méthode le prototype RIVAGE.....	199
I Introduction.....	199
II Présentation du prototype.....	200
II.1 Architecture.....	200
II.2 Le dialogue et le déroulement d'une étape de recherche.....	202
II.2.1 Le déroulement d'une étape.....	202
II.2.2 L'interface utilisateur en Smalltalk.....	203
II.2.3 Les fenêtres utilisées dans RIVAGE.....	205
II.2.3.1 La formulation de la demande.....	205
II.2.3.2 Visualisation et choix d'images.....	207
II.2.3.3 Les autres fenêtres.....	208
III Présentation des objets.....	211
III.1 Formalisme utilisé.....	211
III.2 Les objets manipulés.....	212
III.2.1 La session.....	213
III.2.2 Le thesaurus.....	216
III.2.3 La base descriptive.....	218

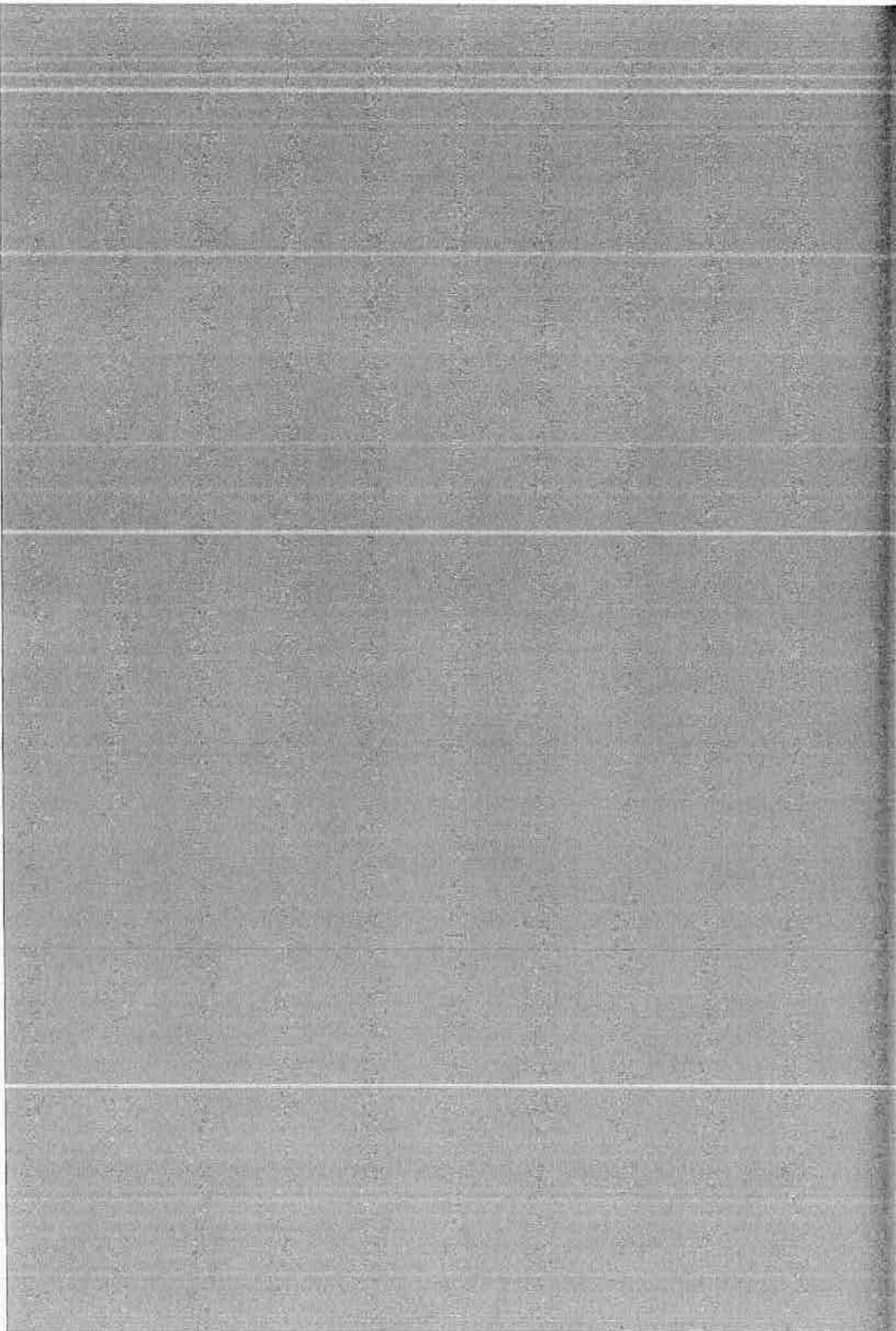
III.2.4 Les facettes.....	219
IV Description des fonctions principales.....	220
IV.1 La ventilation et l'interprétation de la demande.....	220
IV.1.1 Quelques remarques.....	220
IV.1.2 La fonction ProcheSémantiquement.....	222
IV.2 La déformation de la demande.....	225
IV.3 La Visualisation.....	230
IV.4 Analyse des choix	
calcul des poids de pertinence.....	232
IV.5 La propagation des poids.....	233
IV.6 La reformulation.....	235
IV.6.1 Le choix des seuils.....	235
IV.6.2 Parcours d'un champ sémantique en langage objet.....	238
V Evaluation du système.....	241
V.1 Avertissement.....	241
V.2 La démarche choisie.....	241
V.3 Résultats de l'évaluation.....	242
VI Evolution & perspectives.....	247
VI.1 Les évolutions à court terme.....	247
VI.2 Les évolutions à moyen terme.....	249
VI.3 Les évolutions à long terme, perspectives du système RIVAGE.....	249
VI.3.1 Apprentissage à long terme.....	250
VI.3.2 Formalisation du processus EXPRIM.....	251
IV.3.3 Explicabilité associée au processus.....	251
VII Conclusion.....	252

Chapitre 4

Implantation d'une structure hypertexte dans le système RIVAGE.....	255
I Introduction.....	255
II Notion générale d'hypertexte.....	255
II.1 Historique.....	255
II.2 Les principales idées.....	256
II.3 Les solutions apportées.....	256
II.3.1 Le lien.....	257
II.3.2 Le nœud.....	258
II.4 Les problèmes liés à l'utilisation de système hypertexte.....	258
II.5 Les remèdes.....	259
III Définition du concept d'Hypermedia.....	260
IV RIVAGE et les besoins d'un utilisateur recherchant des images.....	260
V Réalisation d'une structure hypertexte dans le système RIVAGE.....	261
V.1 Le réseau de nœuds.....	261
V.2 Le nœud étape.....	263
V.3 Les aides associées au parcours.....	265

VI Conclusion.....	266
Conclusion.....	268
CONCLUSION.....	270





*L'image est péremptoire,
elle a toujours le dernier mot :
aucune connaissance ne peut
la contredire, l'aménager, la sublimer.*

Roland Barthes (L'image, p157).

INTRODUCTION

L'image peut revêtir plusieurs formes allant du reflet dans un miroir à la représentation mentale d'objet ou de sensation, en passant par la peinture, la sculpture, le dessin, la photographie. L'image que nous allons considérer ici, d'une manière générale, est celle qui peut être diffusée sur un support papier.

A l'ère où l'image peut être construite ou plutôt synthétisée par une machine, les outils effectuant d'une manière automatique des recherches rétrospectives sur des images ne prennent pas en considération les caractéristiques intrinsèques de l'information recherchée. L'arrivée de nouveaux supports de l'information : les supports optiques, mémorisant l'image en grand nombre et avec une bonne qualité, a donné la possibilité aux systèmes informatiques de manipuler l'image en tant que résultat d'une recherche. Cet événement a offert à la recherche d'images de nouvelles perspectives à travers de nouvelles applications. Chacune de ces applications représente une concrétisation, à grande échelle, d'une exploitation de l'image à des fins utiles. En effet, la recherche d'images prend toute son importance lorsque l'image recherchée est le témoin d'un événement, présent ou passé, que l'on souhaite illustrer – ce type de besoin est celui d'une agence de presse – mais aussi lorsqu'elle est le support d'un diagnostic en médecine ou en maintenance, ou encore le point de départ d'une phase de création en publicité, dans une agence de mode, ou lorsqu'elle représente le support d'une communication, d'un cours ...

C'est pour donner à toutes ces formes de recherche d'images un outil adapté à leurs besoins, qu'est né le projet EXPRIM, au sein d'une équipe du laboratoire CRIN, il y a plus de quatre ans. Ce projet propose un processus **progressif** de recherche où la visualisation de l'image joue un rôle important et où l'**interactivité** avec l'utilisateur est très forte. Le processus de recherche est composé de trois phases, non disjointes, permettant à l'utilisateur d'exprimer ses besoins dans la phase "Avant-Visualisation", de les illustrer pendant la phase "Visualisation", en choisissant, parmi les images proposées par le système, celles qui s'en approchent et en rejetant celles qui en sont trop éloignées, puis grâce à la phase "Après-Visualisation" d'en voir une expression construite par le système à partir des choix effectués dans la phase précédente. Cette nouvelle expression peut alors constituer le point de départ d'une nouvelle recherche.

Notre travail, à l'origine, était d'étudier les problèmes liés à la phase "Après-Visualisation" afin de proposer une méthode permettant de la réaliser. Mais au fur et à mesure de l'étude, il est

apparu que cette phase ne pouvait être analysée indépendamment des deux autres.

Cette thèse est donc consacrée à l'ensemble du projet, avec comme élément directeur la réalisation de la phase "Après-Visualisation". La solution finale que nous proposons est de considérer l'ensemble du processus EXPRIM comme un processus d'**apprentissage** où le concept à apprendre est le besoin de l'utilisateur et où des exemples et contre-exemples de ce concept sont respectivement les images choisies et les images rejetées lors de la phase de visualisation.

Le projet EXPRIM est à l'intersection de plusieurs domaines de recherche en informatique : en effet, l'image étant maintenant une information manipulable par un système informatique, la recherche d'image peut alors être considéré comme un domaine particulier de la **recherche d'information**. Une recherche progressive et interactive d'images passe par l'instauration d'un dialogue entre le système et l'utilisateur devant permettre à ce dernier de satisfaire progressivement ses besoins qui peuvent être, au départ, mal définis. Nous sommes ici dans le domaine de la **communication homme-machine**, domaine qui fait largement appel à l'**intelligence artificielle**. La manipulation simultanée d'images et de texte place notre étude dans un domaine plus général qu'est la réalisation de systèmes **hypermedia**, problème très lié à celui de la conception de **bases de données multimédia** et plus généralement à celui de la représentation et de la manipulation d'**objets à structure complexe**.

Ce travail nous a permis d'étudier, à des degrés de précision variables, un ou plusieurs composants de ces domaines de l'informatique, afin d'en extraire ceux pouvant être exploités dans une réalisation du projet EXPRIM. Cette thèse retrace alors toutes les étapes de ce travail qui sont présentées sous la forme de quatre parties.

La première partie présente un état de l'art sur les systèmes de recherche d'informations (Information Retrieval Systems) qui ont été conçus, jusqu'à ce jour, pour rechercher des informations de type textuel. En effet, la recherche d'images, qu'elle soit interactive ou non, est soumise aux problèmes classiques de ces systèmes. Cette première partie présente l'environnement et les composants d'un système de recherche d'informations dans un cadre général, en suivant une complexité croissante. Nous présentons, dans le premier chapitre, les notions de base nécessaires à la manipulation d'un système de recherche d'informations. Puis, nous tentons d'exposer, dans un cadre plus général, la problématique de la recherche d'informations par l'étude des composants, des modèles utilisés ainsi que des tendances possibles pour l'amélioration des performances (chapitre 2 et 3). L'apport de l'intelligence artificielle dans la recherche d'informations, ainsi que la présentation de quelques systèmes,

terminent cette première partie (chapitre 4).

La deuxième partie tente de montrer quelles sont les différences qu'apporte l'image dans les systèmes de recherche d'informations et comment elles peuvent être prises en compte d'une manière générale dans ces systèmes. Le premier chapitre expose les caractéristiques de l'image ainsi que l'apport des supports optiques dans la recherche d'images. Le deuxième chapitre présente le projet EXPRIM à travers son histoire, une étude critique des idées émises et des réalisations effectuées avant notre arrivée dans l'équipe. Notre participation à un projet ESPRIT n°901 et l'étude que nous avons menée, en collaboration étroite avec la Société Européenne de Propulsion (SEP), le Bureau Marcel van Dijk de Bruxelles (BMvD) et les Archives Photographiques du Ministère de la Culture, sont détaillées dans le troisième chapitre. Ce chapitre présente la spécification et l'architecture du système que nous devons réaliser dans le cadre de ce projet afin d'étudier la faisabilité technique d'un système "intelligent" de recherche d'images intégrant les idées du projet EXPRIM, l'Imageur Documentaire de la SEP et l'expérience en recherche d'informations du BMvD. Il retrace aussi les différentes étapes de la réalisation du système et met en évidence les erreurs qui ont été commises et qui ont entraîné l'abandon de la réalisation. Quoique cette expérience n'ait pas abouti à une réalisation concrète, il nous a semblé important de la présenter parce qu'elle a été très déterminante dans l'évolution de nos idées et qu'elle a occupé deux années sur les quatre années que nous avons consacrées à cette thèse.

Enfin la troisième partie montre l'aboutissement de notre étude qui consiste à placer le processus EXPRIM dans le cadre assez large de l'**apprentissage symbolique** étudié plus généralement en Intelligence Artificielle. L'originalité de notre approche est de combiner les résultats en recherche d'informations sur les problèmes de bouclage de pertinence (Relevance feedback) aux résultats de techniques d'apprentissage symbolique utilisant des taxonomies. Le premier chapitre de cette partie présente d'une manière informelle la problématique de l'apprentissage symbolique ainsi que les méthodes qui nous ont fortement inspiré. Le deuxième chapitre présente le processus de recherche d'images vu comme un problème d'apprentissage où la demande de l'utilisateur joue le rôle du concept à apprendre, et où les images choisies et rejetées lors de la phase de visualisation jouent le rôle d'exemples et contre-exemples du concept. Le prototype RIVAGE (système de Recherche progressive et interactive d'images), qui met en œuvre la méthode proposée dans le deuxième chapitre, est présenté dans le troisième chapitre. Ce système, que nous avons développé dans un environnement objet à l'aide du langage Smalltalk 80, constitue la première réalisation du processus EXPRIM dans sa totalité. L'ensemble du système est décrit dans le détail ainsi qu'une évaluation de ses performances montrant sa capacité à effectuer des recherches interactives et progressives d'images. Cette

réalisation se veut très évolutive et les nombreuses perspectives qui sont exposées dans ce même chapitre ont de fortes chances d'être réalisées sur ce prototype dans un futur plus ou moins proche. Le dernier chapitre présente l'une de ces évolutions, qui s'est concrétisée récemment par l'implantation sur le système RIVAGE d'une structure Hypertexte associée à l'interrogation d'une base d'images intégrant les fonctionnalités du processus EXPRIM.

Première partie

Etat de l'art sur les systèmes de recherche d'informations

Introduction

Il existe plusieurs formes de systèmes de recherche d'informations qui se différencient essentiellement par l'organisation de l'information. Ainsi, on peut faire la différence entre une base de donnée réalisée sur un SGBD (Système de Gestion de Base de Données) et une base de donnée dite documentaire gérée par un système documentaire (SD). Une base de données réalisée sur un SGBD contient des informations d'une organisation ou d'une entreprise dont la représentation et le contenu ont été définis en fonction des besoins par ses administrateurs (schéma conceptuel) [Rol 87][Del 82]. L'accès à l'information est précis et il dépend de la structuration initiale de la base ; le résultat d'une recherche peut être structuré. Les informations d'une base de donnée documentaire sont assez diversifiées et peuvent intéresser un assez vaste public. Le contenu d'une telle base ne peut être structuré en fonction de traitements spécifiques. L'accès à l'information doit permettre de satisfaire un utilisateur le plus souvent inconnu du système qui ignore la structure des informations contenues dans la base. On peut parler d'accès à un élément parmi plusieurs, à une "molécule" parmi un ensemble de molécules de structure semblable mais de contenu informationnel différent.

Nous placerons notre étude dans la deuxième catégorie de système c'est-à-dire les systèmes de recherche d'informations de type documentaire.

Définissons tout d'abord ce que représente un système documentaire.

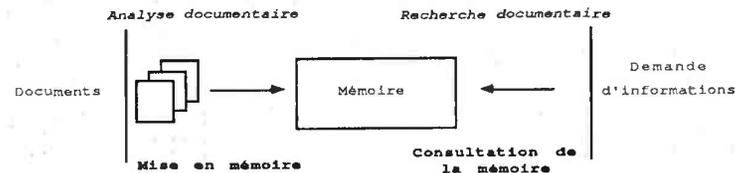
Chapitre 1	
Notion de base en recherche d'informations classique.....	7
I Introduction.....	7
II Problématique	8
III Notion de pertinence.....	8
III.1 La pertinence globale d'un système	9
III.2 La pertinence utilisateur.....	11
III.3 La pertinence système.....	12
IV Les fonctions d'un système de recherche d'informations	12
IV.1 L'indexation.....	13
IV.1.1 Cadre général	13
IV.1.2 Les modèles de documents.....	14
IV.2 L'interrogation	15
IV.2.1 Généralités	15
IV.2.2 La formulation booléenne	16
V Conclusion.....	17

Chapitre 1

Notion de base en recherche d'informations classique

I Introduction

La philosophie des systèmes documentaires peut être représentée par la figure suivante extraite de [Cro 64].



D'un côté, il y a les documents qui constituent les éléments de base et qui feront l'objet des futures recherches (utilisation du système). De l'autre côté, il y a les demandes d'informations que le système va tenter de satisfaire en proposant des documents en sa possession. Entre ces deux côtés, se trouve le système qui doit avoir, d'une part, une représentation des documents qu'il possède et, d'autre part, un moyen d'y accéder. L'analyse documentaire est l'opération qui transforme les documents sous une représentation qui facilitera plus tard la consultation. La recherche documentaire ou recherche d'informations est alors l'opération qui permet la consultation des données, c'est-à-dire des documents transformés, afin d'en extraire les documents conformes à une certaine demande.

Nous pouvons remarquer de suite que la représentation interne des documents dans le système constitue l'élément directeur de l'efficacité d'un tel système. En effet, la précision dans la recherche dépend essentiellement de la précision de la représentation des documents, et paradoxalement plus la représentation est précise plus l'obtention de tous les documents pertinents à une demande ("rappel") devient difficile.

Afin de mieux comprendre ces caractéristiques, étudions la problématique d'un tel système.

II Problématique

La figure 1 représente la problématique d'un système de recherche d'information. Elle met en évidence les parallèles existants.

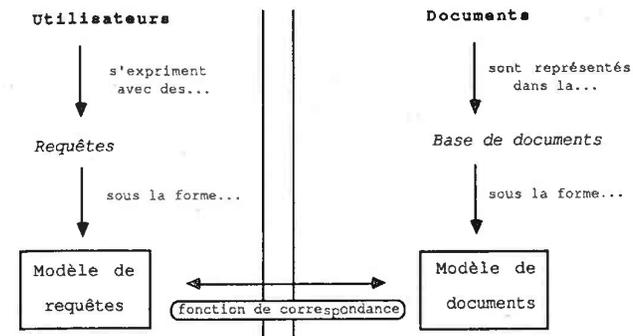


figure 1 : la problématique de la recherche d'information

L'utilisateur souhaite obtenir des documents ; il formule ses souhaits sous la forme d'une requête. Les documents enregistrés dans la base de documents sont représentés sous une autre forme qui n'est pas forcément identique à la forme des requêtes. La problématique est alors de trouver un modèle pour les requêtes, un modèle pour les documents et bien entendu une fonction de correspondance qui permettra de dire dans quelle mesure un document répond à une requête.

Cette fonction de correspondance ne doit pas être stricte car autrement il faudrait décrire entièrement un document dans la requête pour pouvoir l'obtenir en réponse. La requête contient des thèmes de recherche, la fonction de correspondance doit alors trouver les documents dont le contenu sémantique est une représentation des thèmes de recherche. Lorsqu'un document correspond à une requête via la fonction de correspondance, on dira que ce document est pertinent pour la requête.

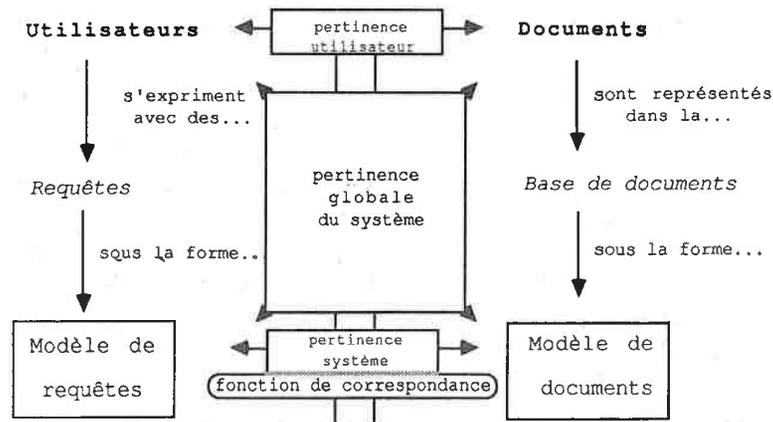
III Notion de pertinence

La notion de pertinence peut se placer à plusieurs niveaux :

- la pertinence système : lorsqu'un document est jugé pertinent pour une requête,

- la pertinence utilisateur : lorsque l'utilisateur émet un jugement sur les documents que le système lui propose,
- la pertinence globale d'un système : son évaluation permet de mesurer l'efficacité du système sur une ou plusieurs collections. Plusieurs mesures d'évaluation existent, elles doivent être effectuées sur des "bases-test" (ACM, INSPEC¹) permettant une comparaison des performances du système.

Ces différents niveaux de pertinence peuvent être visualisés sur notre schéma de la manière suivante :



Étudions séparément chacun de ces niveaux en commençant par celui qui ne pourra être évalué que dans un état avancé et même final d'un système.

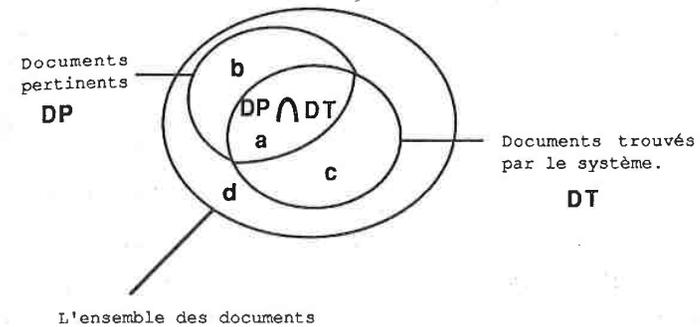
III.1 La pertinence globale d'un système

L'évaluation de la pertinence globale d'un système permet de mesurer ses performances. Elle comporte une série de mesures dont les deux principales sont [Cha 77]:

- le taux de rappel : mesure la capacité du système à retrouver le plus de documents pertinents à une requête,
- le taux de précision : mesure la capacité du système à ne retrouver que les documents pertinents à une requête.

Ces mesures peuvent être simplement expliquées par le graphique suivant :

¹ACM est une base contenant des articles en anglais des conférences ou revues de l'ACM, INSPEC est une base anglaise contenant des articles sur l'informatique, l'automatique, l'électricité ...).



Les documents pertinents DP sont les documents qu'un utilisateur jugera pertinents au regard des thèmes de sa recherche. Les documents trouvés DT représentent les documents pertinents pour le système en rapport avec l'interprétation des thèmes contenus dans la requête de l'utilisateur. L'intersection de ces deux ensembles DP, DT nous donne les documents pertinents effectivement trouvés par le système.

Soit les ensembles suivants :

- a : l'ensemble des documents pertinents retrouvés $DP \cap DT$,
- b : l'ensemble des documents pertinents non retrouvés (silence),
- c : l'ensemble des documents non pertinents retrouvés (bruit),
- d : l'ensemble des documents non pertinents non retrouvés.

C'est à partir des cardinaux de ces ensembles que le rappel et la précision sont calculés. Soit A, B, C et D les cardinaux respectifs de ces ensembles.

Le taux de rappel est égal à : $\frac{A}{A+B}$, c'est-à-dire à la proportion des documents pertinents effectivement trouvés par rapport au nombre total de documents pertinents. Le rappel se situe en général dans les systèmes classiques entre 0.6 et 0.8 [v Sly 86].

Le taux de précision est égal à : $\frac{A}{A+C}$, c'est-à-dire à la proportion des documents pertinents effectivement trouvés par rapport au nombre total de documents trouvés par le système. La précision des systèmes classiques se situe entre 0.2 et 0.8.

Il est clair que parmi ces deux mesures, c'est la précision qui est la plus facile à calculer puisque lorsque l'on propose à l'utilisateur des documents, celui-ci peut les classer en deux ensembles : les pertinents et les non pertinents. Par contre, la mesure du rappel est moins aisée : il

n'est pas facile de connaître l'ensemble des documents que l'utilisateur jugera pertinents puisque c'est ceux-ci que l'on recherche. Le seul moyen de trouver l'ensemble des documents pertinents est d'explorer le fonds :

- soit en examinant tous les documents un à un, ce qui prend du temps,
- soit en interrogeant le système avec des thèmes assez larges.

Deux mesures peuvent compléter les notions de rappel et de précision : le taux de bruit et le taux de silence.

Le taux de bruit représente la proportion de documents non pertinents trouvés (bruit) par rapport à l'ensemble des documents trouvés ($\frac{C}{A + C}$).

Le taux de silence mesure la proportion des documents pertinents qui n'ont pas été retrouvés (silence) par rapport à l'ensemble des documents pertinents ($\frac{B}{B + A}$).

L'utilité de ces mesures est de comparer les systèmes entre eux ainsi que les méthodes et modèles utilisés. Pour établir une bonne comparaison entre systèmes, il faut que les mesures soient effectuées sur un même fonds documentaire avec un jeu de requêtes identiques auxquelles sont associées les références des documents pertinents.

Le système idéal doit être un système où la précision et le rappel sont égaux à 1 (il y a ni bruit ni silence). Mais en général, l'amélioration d'un des deux critères se fait au détriment de l'autre. L'évaluation des performances est dépendante de la pertinence système, c'est-à-dire des modèles et de la fonction de correspondance utilisés. Mais avant de parler de la pertinence système, notion que nous développerons plus tard, parlons tout d'abord de la pertinence utilisateur et de son utilisation.

III.2 La pertinence utilisateur

La pertinence utilisateur, c'est-à-dire le fait de savoir pour chaque document trouvé s'il correspond aux besoins de l'utilisateur, est une information importante non seulement pour les mesures sur la pertinence globale du système mais aussi pour faire progresser les étapes d'une recherche.

La pertinence utilisateur et sa prise en compte dans les systèmes de recherche d'informations ont marqué une évolution dans les méthodes utilisées. En effet, permettre à l'utilisateur de donner son avis sur le résultat d'une recherche et de tenir compte de cet avis pour la recherche suivante ont permis aux systèmes d'améliorer leurs performances en augmentant le rappel dans les recherches [Sal 86].

La recherche d'informations n'est plus une série d'étapes de formulation de requêtes totalement gérées par l'utilisateur, mais tout un processus de formulation, choix et reformulation où le système et l'utilisateur tentent de trouver un accord sur les thèmes de recherche à utiliser. Dans ce type de système, le dialogue est omniprésent et la phase de visualisation des documents doit avoir une place importante afin de permettre à l'utilisateur d'effectuer ses choix avec facilité.

Nous reparlerons de la pertinence utilisateur car elle occupe dans notre système une place essentielle.

La dernière notion de pertinence qu'il nous reste à voir est celle qui permet de présenter des documents à l'utilisateur, c'est-à-dire la pertinence système.

III.3 La pertinence système

La pertinence système est la capacité d'un système à juger de la pertinence d'un document pour une requête donnée. Cette mesure de pertinence s'obtient grâce à la fonction de correspondance qui évalue un degré de similarité entre un document et une requête. Cette mesure de similarité est dépendante des modèles de requêtes et de documents utilisés. Pour permettre une mesure de similarité adaptée et cohérente, il ne faut pas que les modèles de requêtes et de documents soient trop éloignés.

Nous n'approfondirons pas plus dans ce paragraphe les notions de pertinence système et de mesure de similarité. Il faut auparavant introduire quelques exemples de modèles de documents et de requêtes et leur fonction de correspondance.

IV Les fonctions d'un système de recherche d'informations

La figure 2 reprend le schéma de la figure 1 représentant la problématique d'un système de recherche d'informations en mettant en évidence les deux grandes fonctions d'un tel système.

Les deux grandes fonctions mises en évidence sur la figure 2 sont :

- **L'indexation** : c'est l'étape la plus longue et la plus importante dans la mise en œuvre d'un système de recherche d'information. Cette étape consiste à mettre les documents sous la forme correspondant au modèle de document choisi afin de les enregistrer dans la base de documents.
- **L'interrogation** : elle représente la fonction principale d'un système de recherche d'information. Dans cette phase, l'utilisateur exprime ses besoins et le système met en correspondance ces besoins avec les documents de la base afin d'en extraire les plus pertinents.

Voyons plus profondément chacune de ces fonctions.

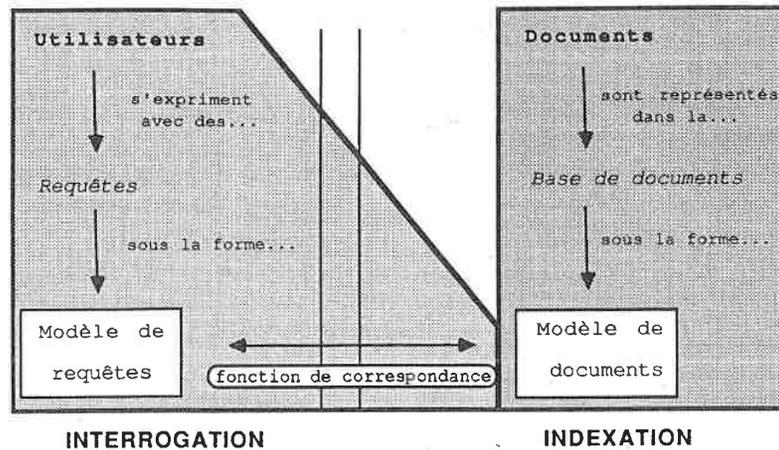


figure 2 : les fonctions d'un système de recherche d'informations

IV.1 L'indexation

L'indexation d'un document est l'étape fondamentale qui donne au document un statut conceptuel dans la base gérée par le système de recherche d'information.

IV.1.1 Cadre général

L'indexation est la mise sous une forme prédéfinie d'une représentation du contenu du document et de l'information qui lui est annexée (date, source, auteur ...). On distingue deux types de données résultant d'une indexation d'un document :

- les données informationnelles,
- les données de recherche.

Les données informationnelles représentent des informations que l'utilisateur peut consulter seulement lorsque le document les possédant a été sélectionné. Ces informations ne sont pas des critères de recherche possibles mais elles sont utiles pour mieux connaître un document.

Les données de recherche sont les informations qui vont permettre, via la langage d'interrogation, de retrouver le document parmi l'ensemble des documents. Ces informations jouent donc un rôle primordial pour la vie d'un document dans la base.

Quand un document est indexé, on obtient un second document que l'on peut appeler **document d'indexation** où les données, qu'elles soient informationnelles ou de recherche, sont regroupées en rubriques. Les rubriques permettent d'identifier la nature de l'information attachée au document ; on peut ainsi avoir des rubriques ayant pour nom "auteur(s)", "date de création" ou encore "description du contenu", ceci suivant le **modèle de document** choisi.

IV.1.2 Les modèles de documents

Dans une base de documents gérée par un système de recherche d'informations classique, un document est associé à un enregistrement contenant les informations du document d'indexation. Cet enregistrement est découpé en rubriques, le contenu de chaque rubrique pouvant être de types différents. Par exemple, le contenu d'une rubrique "résumé" peut être de type "texte", le contenu d'une rubrique "date" de type "date" et le contenu d'une rubrique "description du contenu" de type "liste de mots-clés". Les types "texte", "date", "liste de mots-clés" sont les types de données classiquement utilisés dans les systèmes opérationnels. Les opérations associées à ces types permettent d'effectuer les recherches des documents.

Le modèle de documents définit l'ensemble et le type des rubriques que contiendra le document d'indexation. En règle générale, le type des données de recherche est soit "date", soit "liste de mots-clés". Le type "texte" n'est pas utilisé comme type de donnée de recherche tout au moins dans une première étape de la recherche. Par contre, dans certains systèmes, la recherche dans un texte est utilisée pour sélectionner des documents dans un ensemble de documents provenant d'une première étape de recherche.

Les rubriques de type "liste de mots-clés" sont les rubriques qui vont représenter les données de recherche du document. Chaque concept présent dans le document est représenté par un terme, appelé descripteur. Le placement des termes dans les différentes rubriques est fonction du modèle de documents choisi. Les termes peuvent être placés, par exemple, les uns à la suite des autres séparés par un caractère prédéfini, sans donner de signification particulière à l'ordre des termes dans la liste.

D'autres formes d'indexations existent, celle que nous venons de présenter est la plus couramment utilisée car elle est facile à mettre en place.

Exemple de document d'indexation :

auteur (s) : Maniez J. (type liste de mots clés)
 titre : Les langages documentaires et classificatoires (type texte)
 mots-clés : systèmes documentaires ; langages classificatoires ; recherche d'informations ;
 (type liste de mots clés)
 date : 1987 (type date)
 résumé : Cet ouvrage est le premier livre de la langue française proposant une étude complète
 des classifications documentaires (type texte)

Dès qu'un document possède sa représentation interne dans la base, il devient alors candidat possible aux demandes formulées grâce à la fonction d'interrogation du système gérant cette base.

IV.2 L'interrogation**IV.2.1 Généralités**

L'interrogation d'une base de documents est la fonction principale des systèmes de recherche d'informations. Elle offre à l'utilisateur les moyens d'exprimer ses besoins sous la forme d'une requête en respectant une certaine syntaxe. Cette syntaxe qui définit le modèle de requête est naturellement conditionnée par le modèle de documents adopté lors de la phase d'indexation. Une requête est un ensemble de critères de recherche exprimant les caractéristiques des documents recherchés : *un ouvrage d'un certain auteur, publié à une certaine date, relatant certains faits*. La forme des critères dépend de la rubrique à laquelle s'adresse le critère. Ainsi, un critère relatif à une date peut avoir la forme suivante: *date ouvrage > 18/10/78*, exprimant le fait que l'utilisateur désire obtenir des ouvrages parus après le 18 octobre 1978.

Le résultat d'une interrogation est un ensemble de documents jugés pertinents par le système via la fonction de correspondance. La fonction de correspondance classiquement utilisée dans les systèmes opérationnels est basée sur un modèle de document et sur la présence des critères de recherche dans les documents.

La forme d'interrogation la plus couramment utilisée est celle utilisant des requêtes booléennes (modèle booléen).

Exemple de requête booléenne :

(auteur = Maniez) ET (mots-clés = Recherche d'informations) ET (date = >1/1/80)

Qui s'interprète : *Je désire les documents écrits par Maniez traitant de la Recherche d'informations qui sont parus après le 1er Janvier 1980.*

La formulation de requêtes booléennes est fondée sur la logique de Boole, la fonction de correspondance peut être interprétée de manière ensembliste.

IV.2.2 La formulation booléenne

La formulation booléenne consiste à combiner à l'aide d'opérateurs booléens un ensemble de critères afin d'obtenir une expression conceptuelle du besoin de l'utilisateur. Les principaux opérateurs booléens utilisés sont au nombre de trois : ET, OU et SAUF. Nous allons étudier la fonction de chaque opérateur en considérant que les documents de la base ne possèdent qu'une rubrique de type "liste de mots-clés". Un critère est alors un descripteur exprimant un thème de la recherche.

Le ET permet d'exprimer le souhait que deux descripteurs soient tous les deux présents dans le document d'indexation pour que celui-ci soit jugé pertinent. Cet opérateur peut être interprété de manière ensembliste par l'intersection de deux ensembles, chaque ensemble étant l'ensemble des documents indexés respectivement par le premier et le second des descripteurs.

Exemple :

besoin de l'utilisateur : *je veux des documents traitant de la formulation booléenne dans les systèmes de recherche d'informations,*

interprétation possible: FORMULATION BOOLEENNE ET SYSTEME DE RECHERCHE D'INFORMATIONS.

L'opérateur ET augmente la précision de la recherche tout en diminuant le rappel.

L'opérateur OU combine deux descripteurs dont la présence de l'un ou de l'autre ou des deux dans le document d'indexation suffit pour que ce document soit jugé pertinent. Cet opérateur peut être interprété par l'union des deux ensembles des documents indexés par ces descripteurs.

Exemple :

besoin de l'utilisateur : *Je veux des documents expliquant la formulation booléenne ou la formulation en langage naturel,*

interprétation possible : FORMULATION BOOLEENNE OU LANGAGE NATUREL.

Cet opérateur a l'effet inverse de l'opérateur ET puisque son utilisation permet d'augmenter le rappel, la précision devenant moins bonne.

Le dernier opérateur couramment utilisé est l'opérateur de négation d'un élément de la recherche, c'est-à-dire l'opérateur SAUF. Cet opérateur exprime le désir d'obtenir des documents traitant d'un sujet en exceptant certaines caractéristiques connues de ce sujet. L'interprétation ensembliste de cet opérateur est la complémentation d'un ensemble par rapport à un autre ensemble.

Exemple :

besoin de l'utilisateur : *Je désire des documents traitant de la formulation de requêtes, mais la formulation en langage naturel ne m'intéresse pas,*

interprétation possible : FORMULATION DE REQUETES SAUF LANGAGE NATUREL

Cet opérateur a le même effet sur une recherche que l'opérateur ET (augmente la précision tout en diminuant le rappel). En effet, si on considère l'ensemble $D(\bar{B})$ des documents n'ayant pas une certaine caractéristique B et un autre ensemble de documents $D(A)$ ayant une autre caractéristique A, le moyen d'obtenir en résultat d'une recherche, l'intersection de ces deux ensembles, est d'écrire la requête suivante : A ET \bar{B} . \bar{B} représentant le fait que les documents désirés n'ont pas le descripteur B dans leur document d'indexation. Cette requête s'écrit avec l'opérateur SAUF : A SAUF B. Ainsi, une formulation ayant un opérateur SAUF peut se réécrire avec un opérateur ET, mais l'effet du ET est, en général, atténué du fait de la supériorité de grandeur de l'ensemble $D(\bar{B})$ par rapport à l'ensemble $D(B)$.

Nous laisserons sous silence un opérateur qui est très peu utilisé : le OU exclusif, noté OUX . Le OUX se différencie du OU commun par le fait que l'on rejette les documents contenant à la fois les deux descripteurs reliés par cet opérateur. Cette expression peut, en fait, se réécrire avec les trois opérateurs que nous avons cités : A OUX B se réécrit (A OU B) SAUF (A ET B).

V Conclusion

Ce premier chapitre nous a permis de présenter dans un cadre général la problématique de la recherche d'informations en montrant comment, dans les systèmes classiques, un document est jugé pertinent pour une requête.

Pour cette première présentation, nous nous sommes volontairement placé dans un contexte assez simple afin de décrire les fonctions principales de tels systèmes, c'est-à-dire l'indexation et l'interrogation. Pour réaliser ces fonctions un modèle de document, un modèle de requête ainsi qu'une fonction de correspondance doivent être définis ; l'ensemble de ces informations constitue le modèle de correspondance ou encore le modèle de recherche d'informations.

Nous allons maintenant approfondir ces notions par une étude plus détaillée des composants et des modèles utilisés dans certains systèmes de recherche d'informations.

Chapitre 2	
Les principaux modèles et fonctions	20
I L'indexation	20
I.1 Les modèles d'indexation.....	20
I.1.1 Définition des données de recherche.....	20
I.1.2 Les langages d'indexation.....	21
I.1.2.1 Modes d'indexation.....	21
I.1.2.2 Les vocabulaires des langages libres	22
I.1.2.3 Les vocabulaires des langages contrôlés	23
I.1.3 Conclusion	23
I.2 Structure d'un thesaurus.....	24
I.2.1 Les unités lexicales.....	24
I.2.2 Les relations sémantiques	26
I.2.2.1 L'appartenance à un thème.....	26
I.2.2.2 L'équivalence sémantique.....	27
I.2.2.3 La relation de hiérarchie.....	28
I.2.2.4 La relation d'association	29
I.3 Conclusion	30
II Notion de modèles de correspondance	31
II.1 Généralités.....	31
II.2 Problématique.....	31
II.3 Le modèle vectoriel.....	32
II.4 Conclusion	34
III Autres formes d'interrogation	34
III.1 La formulation avec pondération.....	34
III.2 La formulation avec un thesaurus.....	35
IV Conclusion.....	36

Chapitre 2

Les principaux modèles et fonctions

I L'indexation

Nous avons vu dans le chapitre précédent que la phase d'indexation était une étape fondamentale, qui donnait au document son statut conceptuel dans la base du système de recherche d'information. La mise sous une forme prédéfinie du contenu du document est effectuée selon un ensemble de choix préalables : forme du document d'indexation (définition des rubriques et de leur types), choix de la forme d'indexation (automatique, manuelle), détermination des données de recherche (à partir d'une liste de termes prédéfinie, sans contrôle)...

Tous ces choix sont très liés au degré de précision que l'on souhaite attacher au document d'indexation. Ils sont nécessaires car ils définissent le modèle de l'indexation.

I.1 Les modèles d'indexation

Définir un modèle d'indexation, c'est choisir la structure du document d'indexation et surtout la forme du contenu des rubriques contenant les données de recherche, ainsi que les moyens de représenter ces données.

I.1.1 Définition des données de recherche

Les données de recherche sont les informations qui vont permettre de retrouver un document parmi l'ensemble des documents. Elles sont de deux types :

Il y a tout d'abord l'information signalétique du document. Ces informations sont connues et ne sont pas toujours présentes dans le document lui-même. Les informations appartenant à cette première catégorie ne prêtent pas à contestation. Elles décrivent le document en tant qu' "objet document" et non en tant qu' "objet véhiculant de l'information". Des rubriques comme "date de création", "auteur(s)", "source", et "titre" font partie de ce type d'informations.

La seconde catégorie d'informations qui permet de retrouver un document est l'information que véhicule ce document. Il faut alors trouver un moyen de décrire ce contenu, c'est-à-dire une représentation conceptuelle qui permette de retrouver un document par son contenu sémantique.

Un langage d'indexation est alors nécessaire pour que la représentation conceptuelle du contenu des documents soit la plus uniforme possible.

I.1.2 Les langages d'indexation

Choisir un langage d'indexation, c'est choisir d'une part, un mode d'indexation qui définit la syntaxe utilisée à l'intérieur des rubriques décrivant le contenu du document et d'autre part, un vocabulaire, c'est-à-dire un ensemble de termes pouvant être utilisés pour décrire le contenu.

G. van Slype dans son livre "Les langages d'indexations : conception, construction et utilisation dans les systèmes documentaires" [v Sly 86] définit deux grands types de langage d'indexation :

- les langages libres où le vocabulaire est construit "a posteriori", de manière automatique ou manuelle, à partir d'expressions en langage naturel du contenu des documents,
- les langages contrôlés, où le vocabulaire est construit "a priori", c'est-à-dire avant la phase de description du contenu des documents.

I.1.2.1 Modes d'indexation

Il existe deux grands modes d'indexation possibles :

- Indexation par juxtaposition : les termes sont placés les uns à la suite des autres, séparés par un caractère prédéfini. L'ordre des termes dans la liste n'a aucune signification. Ce modèle est le plus fréquent car très facile à mettre en place. Il correspond au modèle d'indexation appelé : **indexation binaire** : la présence d'un terme implique que le document est pertinent pour le concept exprimé par le terme ; sa non-présence implique la non-pertinence du document.

- Indexation par pondération : un document peut illustrer plus fortement certains concepts que d'autres. L'importance relative des concepts peut être mise en évidence par une certaine pondération sur les termes les représentant. Cette pondération peut être mise en place soit en découpant la rubrique en deux sous-rubriques : termes principaux, termes secondaires ; soit en affectant chaque terme d'un poids mesurant l'importance du concept dans le document. Un exemple d'utilisation de ce modèle d'indexation est donné dans ce chapitre.

Ces deux modes d'indexation peuvent être utilisés dans les deux grandes formes de langages d'indexation. Voyons, maintenant, quelles sont les différentes formes que peuvent revêtir les vocabulaires de ces langages.

I.1.2.2 Les vocabulaires des langages libres

Dans ce premier type de langage, on peut distinguer deux catégories de vocabulaire : les listes de mots-clés et les listes de descripteurs libres.

Une liste de mots-clés est constituée de mots simples unitermes reconnus automatiquement par le système comme ne relevant pas d'une liste prédéfinie de mots-vides (l'anti-dictionnaire*). On y retrouvera ainsi des mots de toutes catégories (noms communs, noms propres, adjectifs, verbes, ...), de toutes formes grammaticales (masculin, féminin, singulier, pluriel, ...). La liste de mots-clés a un volume très important et il y règne une très grande ambiguïté sémantique : dans ce genre de liste, un concept peut être représenté par plusieurs mots (synonymie), chaque mot peut représenter plusieurs concepts (polysémie) et un même mot peut avoir plusieurs formes infléchies.

L'avantage de ces listes est qu'elles ont un coût de construction quasi-nul et que l'actualisation des termes employés dans les documents est automatique. L'inconvénient des systèmes employant ce type de langage d'indexation apparaît lors de l'interrogation. En effet, l'utilisateur est confronté au problème de trouver des synonymes adéquats et il n'est pas du tout sûr d'obtenir des documents pertinents compte tenu de l'incidence importante des polysèmes et des formes infléchies des mots.

Les listes de descripteurs libres réduisent quelque peu les inconvénients des listes de mots-clés. En effet, il s'agit d'une première forme de contrôle du vocabulaire à utiliser lors de l'indexation des documents. Ici ce n'est pas l'ordinateur qui extrait les mots-clés des documents mais les documentalistes qui indexent les concepts avec des mots ou expressions sous une forme grammaticale normalisée. Les concepts nouveaux viennent s'ajouter à la liste des descripteurs libres. Les ambiguïtés sémantiques sont alors moins importantes qu'avec les listes de mots-clés du fait d'une certaine normalisation des expressions utilisées. Mais cependant, la synonymie et la polysémie du langage naturel demeurent puisque le langage n'est pas tout à fait contrôlé. L'avantage d'utiliser un tel langage est moindre, puisque l'interrogation est juste un peu moins ambiguë qu'avec les listes de mots-clés et que l'indexation est tout aussi coûteuse qu'avec un langage contrôlé, le parcours manuel de tous les documents étant nécessaire.

* L'anti-dictionnaire regroupe l'ensemble des mots qui ne sont pas des mots-clés (ex : le, la, les, et, ou, de, des)

I.1.2.3 Les vocabulaires des langages contrôlés

La normalisation de la forme des expressions des concepts à utiliser lors de l'indexation apporte une réduction des ambiguïtés sémantiques.

La première forme des vocabulaires des langages contrôlés est la liste d'autorité. Elle contient des descripteurs sous une forme normalisée qui seront les seules expressions des concepts à pouvoir être utilisées pour l'indexation des documents. De cette manière, les cas de synonymie et de polysémie deviennent plus rares, ce qui entraîne une meilleure approche de l'utilisateur à l'interrogation. Les inconvénients de ces langages sont d'une part le coût de construction de la liste, car il faut énumérer les concepts et choisir la meilleure représentation, et d'autre part le coût de l'indexation, celle-ci ne pouvant être que manuelle pour une bonne identification des concepts dans le document. Une réduction importante de la synonymie et de la polysémie n'est possible que si le vocabulaire possède une structure sémantique.

Le thesaurus est un vocabulaire des langages contrôlés qui possède cette structure. Il permet ainsi une indexation bien contrôlée et une interrogation très efficace [v Sly 86][Sal 86] par rapport aux autres langages. L'inconvénient : un coût de construction énorme (plusieurs dizaines de mois x hommes).

I.1.3 Conclusion

Après cette énumération des langages d'indexation, nous nous apercevons que plus nous souhaitons diminuer les ambiguïtés sémantiques possibles, plus le coût de mise en œuvre du langage est important. Il est clair qu'il ne faut pas négliger ce point, mais notre souci est d'offrir à l'utilisateur un système qui comprenne ses besoins. Cette compréhension n'est possible que si l'indexation est cohérente d'un point de vue sémantique. Il nous faut alors un langage d'indexation le moins ambigu possible. C'est pourquoi le choix d'utiliser un thesaurus nous semble raisonnable, et nous verrons plus tard qu'il est en fait le seul possible, compte tenu de la nature de nos documents : des images.

Etudions alors plus spécifiquement la structure sémantique des thesaurus.

I.2 Structure d'un thesaurus*

L'analyse documentaire est l'identification des concepts contenus dans un document et leur expression sous une forme lexicale. L'expression du sens d'un concept sous une forme lexicale à l'aide d'un thesaurus ne se fait pas comme avec un dictionnaire, où l'on consulte la définition d'une unité lexicale pour vérifier son sens, mais plutôt par une confrontation des termes à l'intérieur de "champs sémantiques". Un champ sémantique est un ensemble structuré d'unités lexicales. La structure des champs sémantiques lie les unités par des relations de signification qui permettent de préciser leur sens.

Exemple : Si l'on considère l'unité lexicale "voile" seule, on ne peut savoir sans aucune autre information si cette unité représente la voile d'un bateau ou le voile d'une femme. Par contre, si cette unité se trouve dans le champ sémantique "habillement", on l'assimilera aisément au concept de "pièce de toile, de dentelle ou de soie qui couvre le visage des femmes".

En résumé, la structure d'un thesaurus est constituée de deux éléments :

- des unités lexicales,
- des relations sémantiques entre unités lexicales.

I.2.1 Les unités lexicales

La principale forme des unités lexicales est le "descripteur". Il est représenté par un mot ou une expression et est associé à un document lors de l'indexation quand le concept, exprimé par le descripteur, est présent dans le contenu du document.

Une autre forme d'unité lexicale est le "non-descripteur". Un non-descripteur exprime un concept très voisin d'un descripteur, mais son utilisation pour indexer des documents n'apporterait pas grand chose si ce n'est une perte de rappel à l'interrogation par une augmentation de la synonymie ; il est donc présent dans le thesaurus, mais relié au descripteur de sens voisin par une relation d'équivalence. Un non-descripteur n'est jamais utilisé pour indexer un document, par contre son utilisation est possible à l'interrogation ; c'est grâce au descripteur de sens voisin que les documents pertinents seront sélectionnés.

* Cette partie reprend la description des thesaurus de G. van Slype dans [v Sly 86]. Un autre livre du même auteur peut être consulté [v Sly 79].

Exemples de non-descripteurs :

non-descripteur	descripteur synonyme
Averse	PLUIE
Professeur	ENSEIGNANT
Epigraphe	INSCRIPTION
hanap	RECIPIENT A BOIRE

Un thesaurus peut comporter quelques milliers de descripteurs. Pour manipuler une telle structure, il est plus commode de la diviser en sous-ensembles appelés "champs sémantiques" ou "micro-thesaurus", traitant chacun d'un domaine. Les intitulés de ces sous-ensembles constituent une troisième forme d'unité lexicale que l'on peut trouver dans les thesaurus, car ils jouent un rôle particulier, comme on le verra dans la suite..

Deux grands types de subdivision sont couramment utilisés :

- par "facettes" : les termes sont regroupés par catégories : phénomène, processus, matériaux, organisation, être vivant, équipement, propriété, discipline, etc...,
- par thèmes : les concepts sont regroupés par disciplines qui peuvent être enseignées ou faire partie d'une connaissance : biologie, chimie, architecture, habillement, etc ...

Une structure combinant les deux subdivisions n'est pas à éliminer, comme le montre le tableau extrait de [v Sly 86] :

Facettes \ Thèmes	Biologie	Construction mécanique	Transport
Phénomène	Digestion	Pesanteur	Retard
Processus	Manipulation génétique	Assemblage	Voyage
Organisation	Macromolécule	Chaîne de montage	Société de transport
Etre vivant	Bactérie	Mécanicien	Camionneur
Matériaux	Protéine	Tôle	Gasoil
Equipement	Microscope électronique	Robot de soudure	Locomotive
Propriété	Biomagnétisme	Flambage	Rapidité
Discipline	Cytologie	Construction assistée par ordinateur	Etude du trafic

En règle générale, c'est la subdivision par thèmes qui est la plus employée car elle correspond à un mode de raisonnement que nous avons adopté dès l'enfance. La subdivision par facettes est quelquefois utilisée à un deuxième niveau afin de préciser les thèmes.

Un thesaurus est donc constitué de sous-ensembles recouvrant des thèmes bien précis. Chaque sous-ensemble contient des unités lexicales de deux types, descripteur ou non-descripteur, reliées entre elles par des relations sémantiques.

I.2.2 Les relations sémantiques

Il existe dans les thesaurus classiques quatre grandes familles de relations sémantiques :

- l'appartenance à un thème, entre un descripteur et un intitulé de micro-thesaurus,
- équivalence sémantique entre non-descripteur et descripteur,
- les relations de hiérarchie entre descripteurs,
- l'association entre descripteurs.

I.2.2.1 L'appartenance à un thème

Cette relation permet la subdivision du thesaurus en micro-thesaurus. A chaque descripteur ou non-descripteur est attaché un nom de micro-thesaurus. Cette relation situe le terme dans un contexte sémantique, à l'intérieur duquel on peut rechercher d'autres descripteurs proches sémantiquement. Il existe des thesaurus acceptant le fait qu'un descripteur puisse appartenir à plusieurs thèmes (polyappartenance). Cette particularité entraîne un parcours du thesaurus différent et une indexation soit ambiguë soit plus fine (cf exemple).

Exemple : A l'intérieur d'un thesaurus on trouve :

VOILE
dom : HABILLEMENT
dom : TRANSPORT

"dom" matérialise la relation d'appartenance à un micro-thesaurus ou domaine, ici VOILE appartient à deux domaines.

Soit un document D traitant des voiles de bateaux,

D : VOILE

si D est indexé comme ci-dessus il y a une ambiguïté. Une solution pour enlever cette ambiguïté est de rajouter le domaine entre parenthèses pour tous les descripteurs ayant plusieurs domaines :

D : VOILE (TRANSPORT).

La polyappartenance est directement liée au problème de polyhiérarchie que nous traitons dans la suite (cf § I.2.2.3).

I.2.2.2 L'équivalence sémantique

L'équivalence sémantique est une association reliant un non-descripteur avec un descripteur. Le descripteur et le non-descripteur expriment le même concept. Un document relatant ce concept est indexé par le descripteur, on peut accéder au document avec le non-descripteur grâce à la relation d'équivalence.

Le non-descripteur est relié au descripteur par la relation "Employer". La relation réciproque est "Employé pour".

Exemple : Averse *Employer* PLUIE
PLUIE *Employer pour* Averse

Il existe deux grandes catégories d'équivalence sémantique :

- la synonymie forte et reconnue : quel que soit le thesaurus, l'équivalence est toujours vraie.

Exemple :

- * synonymie vraie : Concept et Notion,
- * variante orthographique : Clé et Clef,
- * sigle : CRIN et Centre de Recherche en Informatique de Nancy.

- La synonymie relative et conventionnelle : la relation d'équivalence entre un non-descripteur et un descripteur donnés va dépendre du thesaurus dans lequel on se trouve, et en particulier du domaine couvert par celui-ci.

Exemple :

- * terme relevant de la même famille : Voile et Voilette seront deux descripteurs différents dans un thesaurus spécialisé dans l'habillement, alors qu'ils seront reliés par une relation d'équivalence dans un thesaurus traitant de toute autre chose,
- * terme de niveaux de généralité différents : Félin, Panthère,
- * terme quasi-synonyme : Professeur, Enseignant.

En règle générale, à un non-descripteur correspond un seul descripteur. Mais, il se peut quelquefois qu'un non-descripteur soit associé à plusieurs descripteurs avec l'obligation de les utiliser ensemble (pluri-équivalence obligatoire) ou sans cette obligation (pluri-équivalence facultative).

Exemple :

Spectacle de télévision *Employer* RADIOTELEVISION et SPECTACLE

Entraineur *Employer* INSTRUCTEUR ou MENEUR.

La deuxième forme, la pluri-équivalence facultative est cependant très peu utilisée. Elle peut être évitée par d'autres moyens comme par exemple, la prise en compte du terme ENTRAI-NEUR comme descripteur.

1.2.2.3 La relation de hiérarchie

La relation de hiérarchie est une relation d'ordre partiel sur l'ensemble des descripteurs. Par cette relation, on peut dire qu'un descripteur A est supérieur à un descripteur B (A est un terme générique de B) et réciproquement que le descripteur B est inférieur au descripteur A (B est un terme spécifique de A). Si tous les descripteurs du thesaurus ne possèdent qu'un descripteur qui lui est supérieur, le thesaurus est alors mono-hiérarchique. Un thesaurus mono-hiérarchisé peut alors être représenté par une arborescence. Par contre, s'il existe un ou plusieurs descripteurs possédant plus d'un descripteur supérieur, le thesaurus est alors poly-hiérarchique.

Il y a deux types de polyhiérarchie :

- à l'intérieur d'un même micro-thesaurus :

TOMATE

dom : VEGETATION

TG : FRUIT

TG : LEGUME

La rubrique TG : indique le terme générique du descripteur.

- entre deux micro-thesaurus, ce qui entraîne l'appartenance du descripteur à plusieurs domaines (polyappartenance) :

OS

dom : ANATOMIE

TG : SQUELETTE

dom : MATERIAU

TG : MATERIAU D'ORIGINE ANIMAL

Deux grands types de relations hiérarchiques peuvent être distingués :

- la relation de "généricité/spécificité" qui relie deux termes A et B, où par exemple A est plus général que B et où B est une sorte de A :

BATIMENT et MAISON

- la relation de "partitivité" reliant deux termes A et B où B est une partie de A :

MAISON et MUR.

La distinction de ces deux relations de hiérarchie dans un même thesaurus complique à la fois sa construction et sa manipulation, puisqu'il faut gérer simultanément les deux structures hiérarchiques. C'est pourquoi dans de nombreux thesaurus ces deux relations sont confondues en une seule, ce qui simplifie la structure du thesaurus mais diminue de beaucoup sa précision sémantique.

Les relations de hiérarchie sont très utilisées, que ce soit à l'indexation des documents ou à l'interrogation.

Lors de l'indexation, les relations de hiérarchie sont très utiles pour choisir le descripteur désignant le plus spécifiquement le concept illustré par le document. Cette approche de l'indexation permet de réduire les bruits au cours des recherches. Si on compare le thesaurus à un arbre avec comme racine le terme le plus général et comme feuilles les termes les plus spécifiques, on parlera d'"indexation aux feuilles".

Dans quelques systèmes, toujours lors de l'indexation, l'ajout automatique de termes de niveaux hiérarchiques différents des descripteurs choisis est pratiqué afin d'augmenter le rappel lors des interrogations.

Réciproquement, lors de l'interrogation, le thesaurus et ses relations de hiérarchies sont utilisés de manières similaires : pour retrouver le ou les descripteurs désignant le plus spécifiquement le ou les concepts de la recherche ; et de même certains systèmes, n'ayant pas pratiqué lors de l'indexation l'ajout automatique de termes, se proposent d'enrichir automatiquement et par étapes la formulation de la requête par des termes supérieurs ou inférieurs hiérarchiquement aux termes déjà présents.

1.2.2.4 La relation d'association

La relation d'association est une relation sémantique reliant deux descripteurs susceptibles d'évoquer des idées proches. Cette relation est symétrique et elle ne peut exister qu'entre descripteurs non déjà liés par une relation hiérarchique.

Il existe de nombreux cas où cette relation peut être utilisée entre deux descripteurs. Ces cas sont ceux que les relations d'équivalence, de hiérarchie ou encore d'appartenance à un thème n'ont pas résolus.

Exemples :

causalité :

INTOLERANCE et RACISME,

objet d'une action, d'un processus, d'une discipline :

FONDS SOUS MARINS et OCEANOGRAPHIE,

localisation :

FETE et SALLE DES FETES,

concomitance :

EXORCISME et DIABLE,

instrumentation :

FUSIL et CHASSE,

matériaux constitutifs :

TUILE et TOIT,

similarité :

CAVERNE et GROTTTE,

antonymie :

HUMIDE et SEC, etc ...

On peut trouver, dans les thesauri existants, différentes appellations matérialisant cette relation :

- Terme Associé (TA),
- Voir Aussi (VA),
- Terme de Sens Proche (TSP).

L'apport de cette relation sémantique, que ce soit à l'indexation ou à l'interrogation, est considérable. Elle permet de se libérer des ornières sémantiques tracées par le découpage en micro-thesaurus. Elle permet d'atteindre, à partir d'un descripteur et par association d'idées, d'autres horizons sémantiques que ceux traités dans le champs sémantique du descripteur.

I.3 Conclusion

Le thesaurus, par rapport aux autres formes de vocabulaires existantes, permet de rendre l'indexation plus cohérente en réduisant les problèmes de polysémie et de synonymie.

Le modèle d'indexation choisi, il faut alors trouver un modèle de requête ainsi qu'un moyen de dire qu'un document correspond à une requête. On définit alors un modèle de correspondance.

II Notion de modèles de correspondance**II.1 Généralités**

Les modèles de correspondance sont à la base des systèmes d'informations. Ils permettent de définir la correspondance entre un document de la base et une requête de l'utilisateur. Pour ce faire, un modèle, c'est-à-dire une forme de requête et une forme de document, est défini ainsi que le moyen de les mettre en correspondance.

Il ne faut pas que le modèle des requêtes et le modèle des documents soient trop éloignés car alors la mise en correspondance peut être difficile ou arbitraire :

. En effet, si les documents ont une représentation du style :

(mot₁, relation₁, mot₂), ..., (mot_i, relation_i, mot_{i+1}), ..., (mot_n, relation_n, mot_{n+1}) où relation_i représente une relation quelconque entre deux mots,

. et si les requêtes ont une représentation moins complexes comme par exemple :

mot₁, mot₂, ..., mot_i, ..., mot_n,

. dans ce cas il y a de l'information au niveau du document "relation_i" qui est inutilisable compte tenu de la pauvreté de la requête, qui ne permet pas d'utiliser d'opérateur de relation.

Le modèle le plus classiquement utilisé et que nous avons décrit dans le premier chapitre, s'appelle : le modèle booléen. Ses particularités sont les suivantes :

- Indexation binaire (indexation par juxtaposition)
- Requête booléenne

Mais étudions plus précisément la problématique des modèles de correspondance.

II.2 Problématique

Prenons un modèle de correspondance simple :

- une requête est une liste de thèmes de recherche,
- un document est représenté dans la base par une liste de thèmes que son contenu recouvre,
- la mise en correspondance est définie de la manière suivante : on dira qu'un document correspond à une requête, si la liste des thèmes de recherche de la requête est contenue dans la liste des thèmes traités par le document.

Exemple :

Soit une requête R = recherche d'information, modèles et deux documents :

D1 : recherche d'informations, base textuelle, modèles

D2 : recherche d'images, modèles.

Ici seul le document D1 correspond à la requête R.

Nous venons de définir un modèle de correspondance. Mais il peut être utile d'en définir un autre, plus indulgent :

- on dira qu'un document correspond à une requête s'il contient au moins un des thèmes de recherche de la requête.

Exemple :

Avec ce nouveau modèle, D1 et D2 correspondent à la requête R.

En terme d'évaluation, on peut dire que ce nouveau modèle possède un meilleur rappel que le modèle initialement défini, mais par contre que sa précision risque d'être moins bonne. Cette nouvelle mise en correspondance va nous donner beaucoup plus de documents en réponse à une requête. Certains documents étant plus pertinents que d'autres, il est alors intéressant d'avoir une valeur attachée au document nous indiquant le degré de correspondance du document avec la requête. Nous pouvons ensuite utiliser cette valeur pour ordonner les documents afin de faciliter leur consultation par l'utilisateur.

On peut définir dans notre modèle le degré de correspondance de la manière suivante :

Soit LTR, la liste des thèmes de recherche de la requête R,

soit LTD, la liste des thèmes traités par le document D.

Le degré de correspondance dc de D avec la requête R peut se définir ainsi :

$$dc = |LTD \cap LTR|^*$$

Exemple :

En résultat d'interrogation de la base avec la requête R, on obtient :

1. D1 avec dc = 2,
2. D2 avec dc = 1.

On s'aperçoit qu'en modifiant uniquement la mise en correspondance, pour un modèle de requête et de document donné, les réponses à une même requête sont différentes. Voyons maintenant un exemple de modèle qui est utilisé dans le système SMART de Salton [Sal 83].

II.3 Le modèle vectoriel

Dans ce modèle, l'indexation des documents est automatique, elle utilise le mode d'indexation par pondération où les poids sont calculés automatiquement en parcourant le texte du document.

* |C| représente le cardinal de l'ensemble C.

Chaque document est représenté par une entité identifiée par un ensemble de propriétés (mots-clés ou termes). A chaque propriété est associé son poids, marquant l'importance du terme dans le document. Si on considère, pour un document donné, les poids des propriétés, on obtient un vecteur :

$$(1) \quad D_i = (a_{i1}, a_{i2}, \dots, a_{it})$$

où t est le nombre total des termes utilisés dans les représentations des documents de la base et où a_{ij} représente le poids ou l'importance du terme j attaché au document D_i . Un terme j qui n'est pas assigné au document D_i aura son poids égal à zéro.

Quand chaque enregistrement est représenté par un vecteur, une collection de n enregistrements et de t termes peut être représentée par une matrice C de dimension n,t :

$$C = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1t} \\ a_{21} & a_{22} & \dots & a_{2t} \\ \vdots & \vdots & \dots & \vdots \\ a_{i1} & a_{i2} & \dots & a_{it} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nt} \end{pmatrix}$$

où les lignes représentent les documents et les colonnes, les termes.

Dans ce modèle, les requêtes sont représentées de la même manière que les documents par un vecteur de poids. L'opération de recherche est effectuée par un calcul de similarité entre chaque document et la requête. La fonction calculant la similarité entre un document D_i et une requête R_i dépend normalement du nombre et des poids des attributs communs.

Une fonction classique de similarité est la fonction **cosinus** qui est égale à zéro, quand aucun terme n'est en commun et qui est égale à un, quand tous les termes correspondent parfaitement [Salt 83], [v Rij 79], [v Sly 86] :

Soit R_i une requête : $R_i = (r_{i1}, r_{i2}, \dots, r_{it})$

$$\text{Sim}(D_i, R_i) = \text{Cos}(D_i, R_i) = \frac{\sum_{k=1}^t a_{ik} * r_{ik}}{\sqrt{\sum_{k=1}^t a_{ik}^2 * \sum_{k=1}^t r_{ik}^2}}$$

Après la mise en correspondance (calcul de similarité pour chaque document), on obtient pour chaque document une valeur qui permet de classer les documents par ordre décroissant de similarité afin de les présenter à l'utilisateur.

L'avantage de ce modèle est qu'il est simple et qu'il peut être implanté facilement tout en répondant à beaucoup de besoins des systèmes de recherche d'informations. L'inconvénient est que ce modèle souffre d'un manque de sémantique, c'est-à-dire d'un manque de prise en compte de relations entre les termes, car dans ce modèle tous les termes ont la même importance et sont indépendants les uns des autres.

II.4 Conclusion

La notion de modèle de correspondance a permis de comparer et de nommer les méthodes utilisées pour faire correspondre un document et une requête. Le modèle booléen est le premier modèle qui a été utilisé et qui est utilisé encore. Le modèle vectoriel est un des premiers modèles qui a permis de formaliser simplement la mise en correspondance. Nous verrons dans le prochain chapitre, les tendances et les caractéristiques de modèles qui ont suivi le modèle vectoriel. Nous avons vu dans le chapitre précédent la forme classique d'interrogation (ie la formulation booléenne), mais l'apport des nouveaux modèles et l'intégration du thesaurus dans les systèmes ont permis d'élaborer d'autres formes d'interrogation.

III Autres formes d'interrogation

III.1 La formulation avec pondération

La formulation avec pondération a été créée pour palier à la rigidité de la formulation booléenne qui attribue à chaque terme une même importance. Il existe plusieurs modèles de requêtes avec pondération.

Ainsi, avec un modèle de requête de type vectoriel (cf § II.3), la pondération peut se trouver soit sur la requête, soit sur chaque terme de la requête, soit à la fois sur la requête et sur les termes. Le modèle de correspondance associé à ce type de requête peut être lui aussi de type vectoriel ; le résultat de la mise en correspondance des documents avec la requête est alors une liste de documents triée suivant leur coefficient de correspondance. La pondération sur la requête est alors un seuil de sélection dans cette liste. Les poids attachés aux termes sont des poids exprimant l'importance du terme relativement aux autres termes. Ces poids interviennent dans le calcul de correspondance entre les documents et la requête.

Exemple :

On supposera que les poids varient de 0 à 1.

FORMULATION BOOLEENNE (0,8) , LANGAGE NATUREL(0,4),
poids de la requête: 0,5.

Interprétation possible : *Je souhaite obtenir des documents traitant essentiellement de la formulation booléenne et un peu de la formulation en langage naturel.*

Dans ce genre de formulation, l'utilisateur doit interpréter ses besoins pour les traduire sous forme de poids attachés à des termes. Dans ce type de formulation, le sens des termes n'est pas pris en compte puisque les modèles de correspondance associés (de type vectoriel) ne prévoient pas l'utilisation d'une connaissance sémantique.

La formulation suivante est orientée plus vers l'utilisation du sens des termes puisqu'elle s'appuie sur un thesaurus.

III.2 La formulation avec un thesaurus

Le thesaurus représente une connaissance sémantique. L'utilisation de cette connaissance dans une formulation de requête permet de faire évoluer une recherche tout en gardant un sens compréhensible par un utilisateur novice.

L'accès aux relations sémantiques lors de la formulation de la requête peut être laissé à la charge de l'utilisateur qui peut ainsi choisir les termes les mieux adaptés à ses besoins. Une interface conviviale est alors souhaitable.

Dans certains systèmes comme STAIRS (logiciel développé par IBM en 1972) ou MISTRAL (logiciel développé par Bull en 1975) [Mis 75], l'ajout de termes liés sémantiquement à des termes de la requête est automatique lors d'un élargissement de la recherche. Dans le cas d'une formulation booléenne, les termes ajoutés sont connectés avec le terme d'origine par un opérateur OU, ce qui permet de respecter le sens des relations de proximité sémantique existant entre ces termes.

Exemple :

Soit la requête : BATEAU ET OISEAU MARIN,

l'ajout des termes spécifiques nous donne une nouvelle formulation de la requête :

(BATEAU OU BAC OU PETROLIER OU PAQUEBOT OU CARGO) ET (OISEAU MARIN OU GOELAND OU MOUETTE OU ALBATROS),

cette nouvelle requête aura un rappel supérieur à celui de la requête initiale.

L'utilisation d'un thesaurus à la formulation des requêtes permet de suivre une certaine cohérence sémantique entre les différentes formulations. Cette cohérence est encore plus forte lorsque les contenus des documents ont été décrits à l'aide de ce même thesaurus.

Les inconvénients majeurs sont bien entendu le coût de construction du thesaurus et l'aspect figé de sa structuration sémantique, qui ne correspond pas forcément à la représentation qu'a l'utilisateur du domaine couvert par la base.

Le gros avantage de ce type de formulation est que la suite des transactions entre l'utilisateur et la machine est compréhensible, puisqu'elle suit une certaine sémantique prédéfinie, ce qui rend le dialogue plus agréable. L'utilisateur montre alors un certain intérêt car il prend part à une recherche où il contrôle et comprend les évolutions et les résultats.

IV Conclusion

Cette partie a été consacrée à une étude un peu plus approfondie des principales fonctions d'un système de recherche d'informations. Nous avons introduit les notions de modèles d'indexation et de modèles de correspondance. Cette modélisation permet de caractériser plus précisément chaque système de recherche d'informations en indiquant quels sont les modèles qu'il utilise.

Nous avons exposé dans le détail la structuration d'un thesaurus, car nous pensons que ce type de structure sémantique est indispensable à un système qui a pour objectif d'effectuer des recherches d'informations cohérentes, c'est-à-dire compréhensibles par un utilisateur, et d'aider cet utilisateur dans ses démarches de formulation de requêtes. Il demeure que le coût de construction d'une telle structure est important et que certains systèmes de recherche d'informations de type textuel préfèrent utiliser une indexation automatique en langage libre et ensuite laisser l'utilisateur seul juge de la cohérence de ses formulations.

Tel n'est pas notre but, nous souhaitons construire un système qui aide l'utilisateur dans ses formulations et qui tente de comprendre ses besoins ; l'utilisation d'une structure sémantique est alors indispensable. Les modèles de correspondance que nous avons vus ne permettent pas d'intégrer une connaissance sémantique lors de la mise en correspondance des documents avec la requête. La seule intégration possible est lors de l'interrogation avec le modèle booléen qui permet de prendre en compte dans la formulation les relations sémantiques du thesaurus, mais cette prise en compte est le plus souvent laissée à la charge de l'utilisateur.

Voyons dans les tendances possibles, c'est-à-dire dans les nouveaux systèmes de recherche d'informations, quelles sont les nouveautés apportées pour la prise en compte de dépendances

entre les termes, sémantiques ou autres et pour l'amélioration des qualités (rappel et précision) d'une recherche.

Chapitre 3	
Les tendances possibles.....	38
I Le bouclage de pertinence.....	38
II Les modèles de correspondances.....	39
II.1 Les modèles probabilistes	39
II.2 Les modèles à base de logique booléenne	
II.2.1 Mise en correspondance	
II.2.2 Inconvénients des modèles booléens.....	44
II.2.3 Solutions possibles.....	44
II.3 Le langage naturel et les modèles linguistiques	46
II.3.1 L'interrogation en langage naturel.....	46
II.3.2 Le langage naturel et ses inconvénients.....	48
III Autres approches de l'interrogation	49
III.1 Avantages et inconvénients de la formulation booléenne	49
III.2 La pondération dans la formulation booléenne.....	51
III.3 Les solutions intermédiaires.....	52
III.4 Conclusion	54
IV Conclusion.....	55

Chapitre 3

Les tendances possibles

Dans ce chapitre, nous allons présenter un certain nombre de tendances qui représentent des choix possibles pour la conception d'un système de recherche d'informations. Ces tendances ont été pour la plupart expérimentées sur des systèmes mais elles font encore l'objet de travaux de recherche.

I Le bouclage de pertinence

La prise en compte de la pertinence de l'utilisateur pour faire progresser une recherche permet de considérer une recherche d'informations, non plus comme une série d'étapes de formulation de requêtes totalement gérées par l'utilisateur, mais comme un processus de formulation, choix et reformulation où le système et l'utilisateur tentent de trouver un accord sur les thèmes de recherche à utiliser.

Cette nouvelle approche de la recherche d'informations nous semble intéressante et indispensable à une meilleure communication entre l'homme et la machine. Les méthodes de **bouclage de pertinence*** [v Sly 86] [Sal 84] découlent directement de cette nouvelle approche. Ces méthodes tiennent compte du choix de l'utilisateur à une certaine étape pour lui proposer une nouvelle requête qui devient à son tour le point de départ d'une nouvelle étape de recherche (d'où le bouclage). L'intérêt de ces méthodes est de permettre à l'utilisateur d'effectuer ses recherches sous une forme évolutive. Cette forme de recherche s'inspire de celle du documentaliste : proposer des documents pour cerner les désirs du demandeur.

L'adaptation des modèles de correspondance pour la prise en compte d'un bouclage de pertinence n'est pas toujours évidente. Par exemple, pour le modèle booléen classique, même si l'analyse des choix de l'utilisateur donne un certain nombre de termes dit pertinents, leur intégration dans la requête initiale n'est pas évidente. Il faut alors trouver à quels endroits les termes pertinents doivent être insérés et trouver les opérateurs booléens adéquats.

* en anglais : Relevance feedback.

Par contre, dans le modèle vectoriel cette prise en compte est plus aisée. La modification de la requête après que l'utilisateur ait émis un avis sur les documents présentés, peut être modélisée facilement : il suffit d'augmenter le poids des termes pertinents et de décroître ceux des non pertinents.

Rocchio [Sal 84] a proposé pour ce modèle une méthode de bouclage de pertinence très simple :

Soit une requête initiale représentée par un vecteur Q_0 , un ensemble D de documents sélectionnés par cette requête.

Soit n_1 le nombre de documents jugés pertinents par l'utilisateur et n_2 le nombre de documents jugés non pertinents.

Si on considère le vecteur représentant la moyenne des poids des termes dans les documents pertinents $\frac{1}{n_1} \sum_{\text{per}} D_i$ et un même vecteur pour les documents non pertinents $\frac{1}{n_2} \sum_{\text{nonPer}} D_i$ où D_i représente le vecteur d'un document selon la formule (1) du chapitre 2, la nouvelle requête Q_1 s'obtient grâce à la formule vectorielle suivante :

$$Q_1 = Q_0 + \frac{1}{n_1} \sum_{\text{per}} D_i - \frac{1}{n_2} \sum_{\text{nonPer}} D_i$$

Cette formulation de la nouvelle requête fait apparaître de nouveaux termes provenant des documents pertinents et fait accroître ou décroître l'importance des termes déjà présents dans la requête initiale en fonction de leur importance dans les documents pertinents et non pertinents.

Ces méthodes de bouclage de pertinence améliorent en quelques étapes le rappel et la précision d'une recherche.

Une majorité des travaux sur la recherche d'informations a pour objet l'étude des modèles de correspondance. Voyons deux résultats de ces recherches qui nous semblent les plus importants et qui caractérisent bien les orientations prises.

II Les modèles de correspondances

II.1 Les modèles probabilistes

Les modèles probabilistes sont basés sur des fondements de la théorie de la décision où le but est de retrouver rapidement les documents qui ont une forte probabilité d'être pertinents pour une requête et qui ont en même temps une probabilité faible d'être non pertinents [Sal 85] [Fuh

86] [v Rij 79]. Les documents peuvent alors être ordonnés en sorte par ordre décroissant d'une fonction de classement $g(D)$ définie par :

$$g(D) = \frac{\Pr(\text{pert}/D)}{\Pr(\text{nonPert}/D)}$$

où $\Pr(\text{pert}/D)$ et $\Pr(\text{nonPert}/D)$ sont les probabilités de pertinence et de non pertinence d'un document D . La pertinence et la non pertinence formant un système complet d'événements, $\Pr(D)$, qui représente la probabilité d'obtenir aléatoirement D sans se soucier de la pertinence, peut s'exprimer de la manière suivante :

$$\Pr(D) = \Pr(D/\text{pert}) * \Pr(\text{pert}) + \Pr(D/\text{nonPert}) * \Pr(\text{nonPert})$$

où $\Pr(\text{pert})$ et $\Pr(\text{nonPert})$ sont les probabilités de pertinence et de non pertinence d'un document quelconque. $\Pr(D/\text{pert})$ et $\Pr(D/\text{nonPert})$ sont les probabilités de D d'apparaître respectivement dans l'ensemble des documents pertinents et dans l'ensemble des documents non pertinents.

En appliquant le théorème de Bayes, on obtient :

$$\Pr(\text{pert}/D) = \frac{\Pr(D/\text{pert}) * \Pr(\text{pert})}{\Pr(D)}$$

$$\Pr(\text{nonPert}/D) = \frac{\Pr(D/\text{nonPert}) * \Pr(\text{nonPert})}{\Pr(D)}$$

$g(D)$ devient :

$$g(D) = \frac{\Pr(D/\text{pert}) * \Pr(\text{pert})}{\Pr(D/\text{nonPert}) * \Pr(\text{nonPert})} \quad (2)$$

La fonction de classement $g(D)$ est inutilisable en absence d'information sur les probabilités d'apparition du document D dans les ensembles de documents pertinents et non pertinents. La manière la plus couramment utilisée pour estimer les probabilités est de ramener les propriétés de pertinence de D aux termes attachés à D . En prenant comme modèle de document le modèle (1) du chapitre 2, les probabilités peuvent s'écrire :

$$\Pr(D/\text{pert}) = \Pr(a_1, a_2, \dots, a_n / \text{pert}),$$

$$\Pr(D/\text{nonPert}) = \Pr(a_1, a_2, \dots, a_n / \text{nonPert}).$$

Ces deux formules sont faciles à évaluer si on suppose que les termes sont attachés aux documents indépendamment les uns des autres ; elles deviennent :

$$\Pr(D/\text{pert}) = \Pr(a_1/\text{pert}) * \Pr(a_2/\text{pert}) * \dots * \Pr(a_n/\text{pert}),$$

$$\Pr(D/\text{nonPert}) = \Pr(a_1/\text{nonPert}) * \Pr(a_2/\text{nonPert}) * \dots * \Pr(a_n/\text{nonPert}).$$

Quand on est en présence d'une indexation binaire qui s'exprime par le fait que a_i est égale à 1 quand le terme i est attaché au document et a_i est égale à 0 sinon, les probabilités peuvent s'écrire sous la forme [Sal 85] [Los 86]:

$$\Pr(D/\text{pert}) = \prod_{i=1}^t p_i^{a_i} * (1-p_i)^{1-a_i},$$

$$\Pr(D/\text{nonPert}) = \prod_{i=1}^t r_i^{a_i} * (1-r_i)^{1-a_i},$$

où $p_i = \Pr(a_i = 1 / \text{pert})$ et $r_i = \Pr(a_i = 1 / \text{nonPert})$.

En introduisant ces dernières formules dans l'expression (2) de $g(D)$ et en prenant le logarithme, on obtient une fonction de classement des documents de la forme suivante :

$$\text{Log } g(D) = \sum_{i=1}^t \left[a_i \text{Log } \frac{p_i}{r_i} + (1-a_i) \text{Log } \frac{1-p_i}{1-r_i} \right] + \text{Log } \frac{\Pr(\text{pert})}{\Pr(\text{nonPert})}.$$

Cette dernière formule de classement est inutilisable si on ne connaît pas pour chaque terme les probabilités p_i et r_i d'être pertinent ou non, et c'est bien là le point sensible de ces modèles [Sal 86].

Pour estimer ces probabilités, plusieurs méthodes peuvent être considérées :

- on peut supposer que les caractéristiques des occurrences des termes peuvent être décrites par une distribution de probabilités (Poisson, binomiale, ...) [Los 86].
- on peut aussi effectuer des statistiques sur un échantillon raisonnable de documents de la collection et répercuter les résultats sur la collection entière. Une solution consiste à utiliser un processus de bouclage de pertinence, qui produit des statistiques sur les occurrences des termes dans les documents pertinents et non pertinents. Les statistiques sont alors utilisées pour estimer p_i et r_i [Sal 85] [Los 86].

Le modèle probabiliste peut être adapté pour prendre en considération des dépendances entre les termes. Le modèle proposé par van Rijsbergen [v Rij 77] permet la représentation de dépendances entre termes par un arbre de dépendance dans lequel chaque terme dépend au plus d'un

autre terme. La racine n'est dépendante d'aucun terme et tout autre terme n'est dépendant que de son père. Les probabilités de pertinence et de non pertinence d'un document utilisent alors des probabilités conditionnelles représentant la dépendance entre deux termes. Les dépendances entre les termes sont des caractéristiques intrinsèques de la collection. Elles peuvent être estimées à l'aide d'un échantillon ou encore par des méthodes de bouclage de pertinence.

Le modèle probabiliste est très intéressant d'un point de vue théorique mais il demeure décevant en pratique, ceci pour plusieurs raisons [Sal 86] :

- les probabilités de pertinence et de non pertinence d'un document sont estimées et cette estimation est souvent imparfaite.
- des dépendances entre les termes existent. Nous avons vu que seulement certaines dépendances peuvent être prises en compte (dépendances entre deux termes) mais d'autres dépendances existent (dépendances entre n termes, $n > 2$) et leur intégration n'est pas évidente.
- une connaissance sémantique sur l'ensemble des termes, qui peut être considérée comme un ensemble de dépendances entre les termes, est totalement absente de ce modèle. Cette absence s'explique par la difficulté de construction d'une telle connaissance, construction qui ne peut être automatisée.

Les modèles que nous allons présenter par la suite, offrent la possibilité de prendre en considération des dépendances entre les termes, ces dépendances pouvant être sémantiques.

II.2 Les modèles à base de logique booléenne :

Les modèles à base de logique booléenne sont un peu plus raffinés que les modèles définis précédemment puisqu'ils permettent, lors de la formulation de la requête, d'identifier deux sortes de dépendance par l'utilisation des opérateurs booléens ou et et. L'opérateur ou va relier les termes synonymes ou quasi-synonymes, alors que l'opérateur et va relier l'ensemble des composants de la requête.

La formulation de requête booléenne, que nous avons largement illustrée au chapitre 1, offre un pouvoir d'expression supérieur à une formulation de type vectoriel. En particulier, les opérateurs rendent possible un élargissement de la requête avec l'opérateur ou ainsi qu'un rétrécissement avec l'opérateur et.

II.2.1 Mise en correspondance :

La mise en correspondance de ces modèles peut être formalisée de la manière suivante [Sal 85] :

Soit une requête booléenne R et un document D , la similarité $\text{Sim}(R,D)$ entre R et D peut être calculée en deux étapes :

- chaque terme t_j de la requête est remplacé par une fonction $F(D,t_j)$,
- l'évaluation des expressions booléennes s'effectue en combinant, à l'aide d'une table de vérité, les valeurs $F(D,t_j)$.

Exemple de table de vérité :

Clause	F
<u>x et y</u>	$x * y$
<u>x ou y</u>	$x + y - xy$
<u>non x</u>	$1 - x$

Exemple de mise en correspondance :

Soit une requête R ,

$$R = ((p \text{ et } u) \text{ ou } r) \text{ et } (\text{non } (s \text{ ou } t)),$$

et un document D ,

$$D = p, q, r.$$

1ère étape :

$$\text{Sim}(R,D) = ((F(D,p) \text{ et } F(D,u)) \text{ ou } F(D,r)) \text{ et } (\text{non } (F(D,s) \text{ ou } F(D,t))))$$

Pour mettre en oeuvre la deuxième étape, il est nécessaire de définir la fonction F . Pour cet exemple, nous allons prendre une définition très simple classiquement utilisée dans les systèmes de recherche à requêtes booléennes ayant une indexation binaire. Ainsi, on dira que $F(D,t)$ est égale à 1 si le terme t apparaît dans D , et F sera égale à 0 sinon.

Reprenons l'exemple en utilisant la fonction F , telle que nous l'avons définie, ainsi que la table de vérité citée plus haut.

Exemple :

2ième étape:

$$\text{Sim}(R,D) = ((1 \text{ et } 0) \text{ ou } 1) \text{ et } (\text{non } (0 \text{ ou } 0)),$$

$$\text{Sim}(R,D) = (0 \text{ ou } 1) \text{ et } 1 = 1 \text{ et } 1 = 1 * 1 = 1.$$

Ainsi, D est en correspondance avec la requête R puisque la fonction $\text{Sim}(D,R)$ est égale à 1. Si sa valeur avait été égale à 0, le document aurait été jugé non pertinent.

II.2.2 Inconvénients des modèles booléens

Les systèmes à formulation de requêtes booléennes sont très répandus parmi les systèmes de recherche opérationnels parce qu'un sous-ensemble de documents d'une base peut être retrouvé facilement si les termes adéquats sont utilisés. Cependant, l'utilisation de la logique booléenne, dans le cadre d'une recherche rétrospective, a été critiquée, ceci pour plusieurs raisons:

- d'une part parce que la formulation booléenne est difficile à manipuler par des novices,
- d'autre part parce que la quantité des documents obtenue en résultat d'une requête n'est pas facilement contrôlable,
- et enfin parce qu'il n'existe pas de critère pour ordonner les documents afin de les présenter à l'utilisateur en résultat d'une recherche.

De plus, l'effet de la recherche booléenne, si elle est mal utilisée, peut être mise en doute. En effet, en réponse à une formulation avec des ou, un document contenant un seul élément de la requête est aussi bon que celui qui les contient tous; de même, en réponse à une formulation avec des et, un document contenant tous les éléments sauf un est aussi mauvais qu'un document qui n'en contient aucun.

II.2.3 Solutions possibles

Par conséquent, un certain nombre d'approches proposent de relâcher les conditions assez strictes de la recherche booléenne où :

- seule une affectation binaire des termes est possible,
- et où un document est soit pertinent ($\text{Sim}(R,D) = 1$) soit non pertinent ($\text{Sim}(R,D) = 0$).

La première approche consiste à permettre d'attacher des poids aux termes d'un document. La fonction F est alors définie de la manière suivante :

$$F(D,t) = k, \quad k \in [0, 1].$$

Comme chaque terme possède ainsi un degré d'appartenance à l'ensemble des documents décrit par ce terme, l'approche des ensembles flous [Zad 65] [Tab 76] est utilisable pour le calcul de similarité entre un document et une requête [Mou 88] [Mou 89a].

Une nouvelle table de correspondance peut être utilisée :

Clause	Evaluation
$F(D, \text{et } s)$	$\text{Min}(F(D, t), F(D, s))$
$F(D, \text{ou } s)$	$\text{Max}(F(D, t), F(D, s))$
$F(D, \text{non } t)$	$1 - F(D, t)$

Après l'application de ces formules lors de la mise en correspondance, un ensemble de documents, chacun affecté d'un degré de correspondance, est obtenu. Cet ensemble peut alors être trié pour une meilleure consultation.

Cependant, l'application des formules montre que le degré de correspondance attaché aux documents reste dépendant de la valeur du poids d'un seul terme (terme au poids le plus grand pour le ou ; terme au plus petit poids pour le et). De plus, l'introduction de poids attachés aux termes de la requête a posé quelques problèmes et une méthode uniforme de calcul des degrés de correspondance n'a pas été facile à générer [Wal 79].

Une extension alors différente des systèmes conventionnels à requêtes booléennes a été suggérée [Sal 83b]. Cette extension autorise à la fois des poids sur les termes de la requête et sur les termes des documents. Elle permet aussi une relaxation de la portée des opérateurs booléens. En effet, les opérateurs peuvent être utilisés avec une force maximale. Dans ce cas le système est compatible avec les modèles conventionnels à base de logique booléenne. Cette force peut être progressivement relâchée en rendant l'interprétation du et et du ou de plus en plus similaire. A la limite, la distinction entre les deux opérateurs est perdue, le système est alors réduit à un système vectoriel où les requêtes n'ont plus d'opérateur.

Voici les formules de mise en correspondance dans le système étendu [Sal 83b][Sal 85][Sal 86] :

$$F(D, q_t \text{ et } q_s) = 1 - \left[\frac{q_t^p (1 - F(D,t))^p + q_s^p (1 - F(D,s))^p}{q_t^p + q_s^p} \right]^{\frac{1}{p}}$$

$$F(D, q_t \text{ ou } q_s) = \left[\frac{q_t^p F(D,t)^p + q_s^p F(D,s)^p}{q_t^p + q_s^p} \right]^{\frac{1}{p}}$$

$$F(D, \text{non } q_t) = 1 - F(D,t).$$

Les symboles t et s représentent deux termes, q_t et q_s sont leurs poids respectifs dans la requête. La variable p détermine la force d'interprétation de l'opérateur, p varie dans l'intervalle $[1, +\infty]$. $F(D, t)$ est la fonction que nous connaissons déjà, cette fonction donne le poids du terme t dans le document D .

A la vue de ces dernières formules, nous nous apercevons que la mise en correspondance d'un document avec une requête devient complexe, car de nombreux paramètres interviennent. Interpréter sémantiquement une requête, à un certain stade de sa mise en correspondance, par la valeur et la combinaison de ses paramètres n'est pas chose facile. Pourtant cette interprétation est nécessaire si le système veut expliquer à l'utilisateur les résultats qu'il obtient.

Les formes de mise en correspondance que nous venons de voir dans ce paragraphe ne tiennent pas compte d'une sémantique inhérente à la requête que l'utilisateur souhaite retrouver dans les documents. Ces formules de mise en correspondance ont avant tout pour souci (qui n'est pas des moindres) d'obtenir en réponse un bon rappel tout en conservant une précision correcte, ceci sans une compréhension du contenu de la requête. Cette compréhension, nous pensons qu'elle peut améliorer les résultats des recherches de documents tout en rendant la communication entre l'utilisateur et la machine meilleure. Pour rendre cette compréhension possible, il faut introduire plus de sémantique dans ces systèmes.

II.3 Le langage naturel et les modèles linguistiques

Les nouveaux modèles dits linguistiques utilisent des connaissances linguistiques pour extraire les concepts des documents textuels de manière automatique. En règle générale, dans ces systèmes, les requêtes sont formulées en langage naturel. L'extraction des thèmes de recherche est obtenue avec les mêmes traitements que ceux utilisés pour l'indexation des documents.

Voyons alors les problèmes que pose l'interrogation en langage naturel en sachant qu'ils sont quasi similaires à ceux posés par l'indexation des documents.

II.3.1 L'interrogation en langage naturel

L'interrogation en langage naturel permet à un utilisateur ne connaissant pas les caractéristiques d'un système (le vocabulaire utilisé, le modèle de correspondance) et plus généralement ne connaissant pas la problématique des systèmes de recherche d'informations, d'exprimer ses besoins sous la forme d'une phrase dont la correction du style et des termes à utiliser est laissée

à son seul jugement. Le système ensuite présente à l'utilisateur un ensemble de documents correspondant à l'interprétation de cette phrase sous la forme d'une requête.

Cette interprétation peut être de complexité variable allant d'une simple extraction de mots-clés à une certaine compréhension de la phrase.

L'extraction de mots-clés consiste à considérer la requête comme un document et à l'indexer en utilisant les méthodes d'indexation associées aux langages libres. Le résultat de l'indexation est alors une liste de mots-clés sous une forme normalisée ou non. La mise en correspondance avec les documents peut être effectuée en utilisant les formules des modèles vectoriels.

Cette interprétation de la phrase est la forme la plus simple que l'on puisse utiliser. Pour obtenir une interprétation plus précise, il faut prendre en considération d'autres caractéristiques qui peuvent être syntaxiques, morphologiques, ou sémantiques. On aboutit ainsi aux problèmes connus des chercheurs travaillant sur la compréhension du langage naturel et plus particulièrement sur la traduction automatique. Pour bien traduire une phrase d'une langue dans une autre langue, il faut comprendre cette phrase. Le problème est le même en recherche d'informations. Il faut comprendre la formulation des besoins en langage naturel afin de les traduire le plus précisément possible sous la forme d'une requête.

Dans les systèmes de traduction, plusieurs niveaux d'analyse de la phrase peuvent être distingués [Cou 86]. L'enchaînement chronologique de ces niveaux n'est pas toujours le même selon les méthodes utilisées ; il peut être le suivant [Bla 88] :

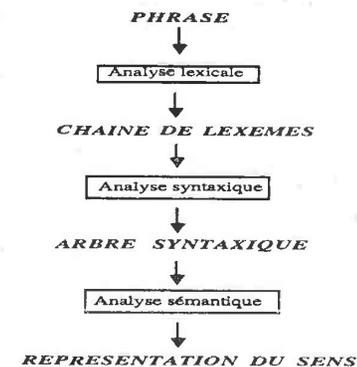
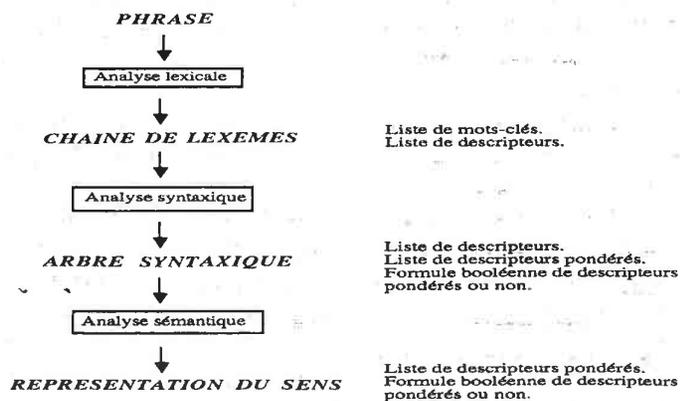


figure 3 : les différents niveaux d'analyse d'une phrase

Dans une phase d'interprétation d'une phrase en langage naturel, on peut retrouver tout ou une partie de ces niveaux d'analyse mis en évidence dans la figure 3. La présence de ces niveaux dépend tout d'abord des choix pris par les concepteurs du système mais surtout du pouvoir d'expression de la requête, c'est-à-dire du modèle de correspondance sous-jacent. Ainsi, en reprenant cette figure, on peut associer à chaque niveau le type de la requête, choisi parmi ceux que nous avons déjà décrits, pouvant représenter avec plus ou moins de fidélité le résultat de l'analyse :



A la vue de cette dernière figure, on s'aperçoit que pour bien prendre en compte les besoins de l'utilisateur, il faut tout d'abord les comprendre, c'est-à-dire en avoir une représentation correcte pour pouvoir les exprimer par une requête. Les recherches en cours dans la traduction automatique montrent que pour bien comprendre une phrase, de nombreuses connaissances sont nécessaires, ces connaissances pouvant être syntaxiques mais aussi et surtout sémantiques. Une connaissance sémantique indépendante d'un domaine particulier demeure difficile à concevoir. Seule une connaissance basée sur un domaine bien cerné peut être pour le moment modélisée. Compte tenu de ces remarques et des modèles de correspondance existant, le sens d'une phrase ne pourra pas être correctement représenté par une simple requête. La requête n'est alors qu'une expression approchée du sens contenu dans la phrase de l'utilisateur.

II.3.2 Le langage naturel et ses inconvénients

L'inconvénient principal de ce type de formulation est qu'il peut laisser croire à l'utilisateur que le système comprend réellement ses besoins et qu'il est en mesure de répondre à tous ces besoins. Nous pensons qu'un utilisateur ne peut jouir pleinement d'un système que si et

seulement si il en connaît les limites. Les limites des systèmes utilisant la formulation en langage naturel sont cachées dès le départ. Ces limites peuvent rester cachées pour un utilisateur non averti qui utilise le système sans finalité précise. Par contre, pour un utilisateur désireux d'obtenir des documents bien précis, ce masquage des limites va lui rendre sa recherche plus périlleuse car il essaiera de reformuler sans cesse ses besoins sans peut-être jamais trouver la bonne formulation.

La dénomination "langage naturel" de cette formulation, couramment utilisée et volontairement utilisée ici, est alors incorrecte. Nous lui préférons volontiers celle de "formulation libre" qui n'a elle aucune prétention. Nous n'approfondirons pas dans ce paragraphe l'utilisation des techniques de traitement et de compréhension du langage naturel pour l'interrogation. Quelques exemples d'utilisations pourront être trouvés dans [Jac 88], [Ant 88], [Deb 88], [Rau 88].

Nous ne croyons pas à l'avenir immédiat des systèmes ayant comme seule formulation une formulation libre. Nous pensons que la formulation des besoins doit s'adapter au niveau de connaissance qu'a l'utilisateur du système qu'il utilise. C'est pourquoi nous allons montrer par la suite d'autres formes de formulation se situant entre des formulations très précises comme la formulation booléenne et des formulations tous azimuts comme la formulation en "langage naturel".

III Autres approches de l'interrogation

III.1 Avantages et inconvénients de la formulation booléenne

Nous avons déjà relaté au § II.2 certains avantages et inconvénients de la recherche booléennes.

Un des avantages les plus forts est la capacité d'expression d'un tel langage de requêtes. Ce pouvoir vient du fait qu'une interprétation sémantique et simple d'une requête booléenne est possible. Le sens ainsi obtenu peut alors être manipulé aisément par un système. Cette interprétation peut être faite de manière ensembliste ou linguistique comme nous l'avons montré dans le chapitre 1.

Un autre avantage de cette formulation est que l'utilisateur peut, grâce aux influences des opérateurs sur la recherche, se définir des stratégies de recherche [Sal 83] [v Rij 79] [v Sly 86]. Ainsi, pour diminuer les réponses, on peut préciser la requête en introduisant des descripteurs connectés par des ET ou des SAUF. Inversement, pour obtenir plus de documents il suffit soit d'éliminer des connexions de type ET ou SAUF, soit d'ajouter des descripteurs avec

l'opérateur OU. Ces stratégies de recherches peuvent être guidées par l'utilisation d'un thésaurus comme dans la formulation de requête que nous avons décrite au chapitre 2.

Le dernier avantage de cette formulation est qu'elle est facilement réalisable à partir d'une indexation binaire. Il suffit d'associer une fois pour toutes à chaque descripteur les références des documents d'indexation dans lesquels ils apparaissent. Puis, à partir de cette structure appelée "fichier inverse", et à l'aide des opérations ensemblistes (intersection, union, complément) les réponses à une requête booléenne peuvent être évaluées [Sal 83][v Rij 79].

Les inconvénients de ce type de formulation sont les suivants:

- l'apprentissage et la manipulation des opérateurs booléens demeurent difficiles pour un novice qui, au moins dans un premier temps, ne pourra contrôler et utiliser toute la puissance d'un tel langage. En effet, la complexité d'une requête booléenne, qui a été formulée par un expert de la formulation booléenne, peut paraître incompréhensible pour un utilisateur non averti.

De plus, cette complexité n'est pas toujours justifiée et il est conseillé de limiter la forme d'une requête à un ensemble de groupes de descripteurs connectés par des ET ou des SAUF, l'opérateur utilisé dans les groupes de descripteurs étant le OU. Chaque groupe est un thème de recherche. Les descripteurs présents dans le groupe symbolisent les différentes représentations possibles du thème d'un point de vue paradigmatique¹ [Cro 64][Sal 86][v Sly 86]. La connexion de l'ensemble des groupes par des ET ou des SAUF exprime les relations syntagmatiques² que l'on souhaite avoir dans les documents recherchés.

Exemple :

(INTERROGATION OU FORMULATION DE REQUETES OU FORMULATION BOOLENNE OU INTERROGATION EN LANGAGE NATUREL) ET (INDEXATION BINAIRE OU INDEXATION PAR JUXTAPOSITION) SAUF INDEXATION PAR PONDERATION.

- L'interprétation de la conjonction de coordination "et", utilisée dans une phrase en langage naturel, ne se traduit pas systématiquement par l'opérateur booléen ET mais parfois par un OU.

Exemple :

besoin de l'utilisateur : *je recherche de l'information sur la formulation booléenne et sur l'interrogation en langage naturel,*

interprétation : FORMULATION BOOLENNE OU INTERROGATION EN LANGAGE NATUREL.

¹ Les relations paradigmatiques entre termes sont des relations qui existent entre ces termes quel que soit le contexte d'utilisation.

² Les relations syntagmatiques, par opposition aux relations paradigmatiques, sont des relations qui sont totalement dépendantes du contexte de leur utilisation.

- Le dernier inconvénient que nous mettrons en évidence est celui lié à la rigidité des opérateurs. Cet inconvénient a déjà été relaté au § II.2. Les solutions apportées (force d'interprétation attaché aux opérateurs) rendent l'interprétation sémantique de la requête très difficile car les relations paradigmatiques et syntagmatiques sont alors confondues.

III.2 La pondération dans la formulation booléenne

Avec des requêtes de type booléen, la pondération peut se trouver, comme nous l'avons montré au chapitre 2, sur les termes de la requête ou sur les opérateurs booléens. Une requête ayant tous ses opérateurs booléens avec un poids faible est alors équivalente à une requête de type vectoriel. La mise en correspondance d'une telle requête avec les documents fait appel à des formules assez complexes qui rendent une interprétation sémantique difficile.

Exemple :

On suppose que les poids associés aux opérateurs varient entre 0 et 100.

(INTERROGATION(1) OU⁷⁵ FORMULATION DE REQUETES(1) OU⁷⁵ FORMULATION BOOLENNE(0,8) OU⁵⁰ INTERROGATION EN LANGAGE NATUREL(0,4)) ET⁵⁰ (INDEXATION BINAIRE(1) OU⁹⁹ INDEXATION PAR JUXTAPOSITION(1)) SAUF⁹⁹ INDEXATION PAR PONDERATION(1).

interprétation : sans considérer les poids on peut aisément interpréter cette requête "l'utilisateur désire des documents traitant de l'interrogation en général dans le cadre de systèmes ayant une indexation binaire par juxtaposition et non une indexation par pondération". Si maintenant on désire ajouter dans cette interprétation les nuances apportées par les poids, celle-ci deviendra incompréhensible par un utilisateur qui ne connaît pas les techniques de mise en correspondance utilisées.

Les avantages apportés par la pondération dans la formulation des requêtes ne permettent pas d'améliorer la compréhension du sens de la requête. Un utilisateur novice ne manipule pas aisément les coefficients quand il méconnaît leur utilisation par le système c'est-à-dire l'influence qu'ils ont dans l'interprétation de ses besoins. Il est clair que l'accès d'un tel système de formulation nécessite un effort d'adaptation de la part de l'utilisateur. Cet effort ne sera récompensé qu'au bout d'un certain temps qui ne doit pas être supérieur au temps que l'utilisateur souhaite sacrifier pour sa recherche d'informations (temps de patience).

L'utilisation des poids est orientée système avant tout, car il est plus facile pour un ordinateur de manipuler des chiffres et des formules plutôt qu'une masse de symboles véhiculant un sens. Ces poids permettent une déformation progressive de la requête tout en contrôlant le rappel et la précision de celle-ci. L'inconvénient est qu'au fur et à mesure des transformations

l'interprétation du sens de la requête et de la suite des transformations devient complexe et même impossible.

Les systèmes utilisant ce type de requêtes doivent offrir à l'utilisateur une interface où la notion de poids est inexistante. L'affectation des poids est alors entièrement contrôlée par le système qui doit les utiliser pour présenter des documents pertinents.

III.3 Les solutions intermédiaires

Les solutions intermédiaires entre la formulation totalement libre et les formulations dans des langages de requête fixés sont celles qui ressemblent à l'interrogation que nous qualifierons de type "annuaire électronique" (le 11 sur minitel). Cette forme d'interrogation est orientée grand public et grâce à l'introduction de plus en plus grande des minitels dans les foyers, elle a maintenant fait ses preuves malgré sans doute quelques petits inconvénients que chacun a pu rencontrer. L'intérêt de ce type de formulation est qu'elle est abordable par n'importe quelle personne même par celles qui n'ont jamais été en contact avec un système informatique.

L'écran de formulation est découpé en rubriques indiquant tout de suite les différentes possibilités de recherche. De cette manière on s'aperçoit très vite que l'accès à un abonné par son numéro est impossible. Le remplissage des rubriques est laissé à la charge de l'utilisateur qui a alors toute liberté d'expression, nous appellerons ce type de formulation "la formulation libre par facettes" où chaque facette représente un caractère de la recherche, ce caractère faisant partie des données de recherche des documents. La formulation libre par facettes permet une recherche guidée et plus compréhensible pour un utilisateur non initié, tout en offrant à l'utilisateur plus averti de nombreuses possibilités de recherche multi-critères.

Il est clair que la phase suivante du système est de prendre la demande de l'utilisateur par facettes et de la traduire dans un langage de requête plus sophistiqué afin d'accéder aux documents. Cette traduction est dans le cadre de l'annuaire électronique assez facile puisque chaque facette contient au maximum un seul élément. Dès que cet élément est reconnu par le système, il peut être pris en compte pour la recherche. Voyons un exemple de traduction possible d'une demande de l'annuaire électronique dans une formulation booléenne mettant en évidence quelques stratégies de recherche possibles, ceci en supposant que les abonnés soient indexés par leur nom, prénom, rue, localité, département.

Exemple :

Demande d'abonné:

NOM : DUBREUIL

RUBRIQUE :

LOCALITE : TINQUEUX

facultatif

DEPARTEMENT: 51

PRENOM : ARTHUR

ADRESSE :

Première traduction possible :

(nom = DUBREUIL) ET (localité = TINQUEUX) ET (département = MARNE) ET (prénom = ARTHUR)

Deuxième traduction possible si la première n'obtient pas de résultat et si l'utilisateur n'a pas choisi les variantes orthographiques proposées par le système :

(nom = DUBREUIL) ET (localité = TINQUEUX) ET (département = MARNE)

Troisième traduction possible si la deuxième n'obtient pas de résultat :

(nom = DUBREUIL) ET (localité = TINQUEUX)

Si on n'obtient toujours pas de résultat, le dialogue avec l'utilisateur s'installe pour modifier sa demande avec par exemple la possibilité de rechercher l'abonné dans les localités voisines.

La traduction d'une demande d'un abonné bien précis est assez facile puisqu'il suffit d'utiliser l'opérateur booléen ET pour connecter les éléments de la recherche. Cette facilité n'existe plus lorsque par exemple l'utilisateur recherche un club de sport de combat près de chez lui.

Exemple:

NOM :

RUBRIQUE : CLUB DE SPORT DE COMBAT

LOCALITE : TINQUEUX, REIMS

facultatif

DEPARTEMENT: 51

PRENOM :

ADRESSE :

La traduction qui correspondrait le mieux à cette demande est :

[(rubrique = CLUB DE KARATE) OU (rubrique = CLUB DE JUDO) OU (rubrique = CLUB DE BOXE) OU (rubrique = CLUB DE KENDO) OU ...] ET [(localité = REIMS) OU (localité = TINQUEUX)] ET (département = MARNE)

Avec cet exemple nous atteignons les limites d'interrogation de l'annuaire électronique car il est incapable de répondre à ce type de demande : d'une part la notion de "sport de combat" est inconnue du système et d'autre part il est interdit d'indiquer plusieurs localités.

Les limites indiquées ici ne sont pas sans solutions, le problème est que les solutions apportées par l'annuaire électronique ne conviennent par pour ce type de demande que nous pouvons qualifier de "demande large". Des amorces de solutions sont apportées dans [Cle 88] où il est montré comment peut se résoudre l'interrogation des pages jaunes de l'annuaire par le "langage naturel", ce qui pose encore plus de problèmes que ceux mis en évidence ici.

Il n'en reste pas moins que nous considérons la formulation libre par facettes comme un bon intermédiaire entre la formulation dite en "langage naturel" et celle plus précise comme par exemple la formulation booléenne.

De nombreux systèmes proposent des formulations intermédiaires soit plus proches de la formulation libre soit plus proches d'une formulation précise. Par exemple, le système MenUSE [Pol 88] propose une forme d'interrogation utilisant une navigation dans un thesaurus, via les relations sémantiques, au cours de laquelle l'utilisateur sélectionne les critères de sa recherche. La navigation est rapide puisque le thesaurus est affiché dans une fenêtre où l'utilisateur sélectionne des concepts à l'aide d'une souris. Les concepts sont affichés avec certaines caractéristiques renseignant l'utilisateur sur la base descriptive comme le nombre de documents où le concept apparaît ainsi que le nombre de sous-concepts qui le précisent. L'utilisateur au départ ne fait que sélectionner des concepts puis au fur et à mesure que sa recherche se précise il peut connecter ces concepts par des opérateurs booléens.

III.4 Conclusion

A travers cet éventail des possibilités d'interrogation des systèmes de recherche d'informations nous avons voulu montrer les différents niveaux de précision d'une formulation des besoins. Une formulation très précise des besoins nécessite à l'utilisateur d'apprendre à manipuler un langage de requête à la syntaxe pas toujours facile ; cette forme d'interrogation est donc réservée à des spécialistes.

Une formulation en "langage naturel" rend l'utilisateur trop confiant dans un premier temps puis mécontent par la suite car il se rend compte qu'il éprouve des difficultés à faire comprendre au système ce qu'il désire réellement. Des solutions intermédiaires, de type annuaire électronique, sont alors préférables.

IV Conclusion

Nous avons évoqué dans ce chapitre quelles étaient les tendances actuelles ainsi que leurs apports dans les systèmes de recherche d'informations.

Ainsi nous avons vu que la prise en compte de la pertinence utilisateur dans un processus de recherche pour la reformulation de requête améliore les qualités d'une recherche mais que la reformulation n'est pas toujours évidente.

En présentant des modèles qui formalisent la mise en correspondance, nous avons remarqué qu'ils ne prennent pas en compte de connaissance sémantique sur les termes.

L'interrogation en "langage naturel" des systèmes utilisant un modèle linguistique n'est pas, à ce jour, le meilleur outil pour effectuer des recherches d'informations. Un outil qui a fait ses preuves est la formulation booléenne qui a beaucoup d'avantages mais qui a aussi de nombreux inconvénients que nous avons énumérés. L'avantage le plus important à notre avis est son pouvoir d'expression du sens d'un besoin d'un utilisateur ; cette expression peut être déformée par des stratégies de recherche qui peuvent s'appuyer sur une connaissance sémantique associée au système (ex: thesaurus). L'expression booléenne du sens d'un besoin peut être obtenue par le système à partir d'une première expression du besoin de l'utilisateur, non pas en langage naturel car là on devra se confronter à des problèmes souvent insolubles de la traduction automatique, mais sous une forme intermédiaire que l'on peut qualifier de type "annuaire électronique" c'est-à-dire guidée par le système. Cette expression booléenne obtenue, on peut dire alors que le système a une certaine compréhension des besoins de l'utilisateur.

L'intelligence artificielle apporte des outils, soit de représentation de la connaissance, soit de raisonnement sur cette connaissance, permettant de manipuler de l'information en ayant une certaine compréhension du sens qu'elle véhicule. Il est alors intéressant d'étudier quels sont les apports possibles de l'IA, pour l'amélioration de la compréhension des besoins de l'utilisateur dans la recherche d'informations.

Chapitre 4	
Apport de l'Intelligence Artificielle dans la recherche d'information	57
I Introduction.....	57
II L'indexation	58
III L'interrogation.....	59
III.1 Les systèmes experts.....	60
III.2 Les systèmes à base de connaissances.....	60
III.2.1 Le système IOTA.....	62
III.2.2 Le système I ³ R.....	64
III.3 L'apprentissage et ses applications en recherche d'informations.....	67
III.3.1 L'apprentissage à court terme.....	67
III.3.2 L'apprentissage à long terme.....	69
IV Conclusion.....	70

Chapitre 4

Apport de l'Intelligence Artificielle dans la recherche d'information

I Introduction

L'intelligence artificielle est un domaine de recherche vaste car il comprend tout un ensemble de problèmes tels que : la reconnaissance des formes, la résolution de problème, la représentation de la connaissance, le raisonnement, la réalisation de systèmes experts, l'apprentissage et beaucoup d'autres. Parmi les différentes manières d'aborder ces problèmes, on peut distinguer deux grands courants :

- le premier courant regroupe les personnes pensant que l'ordinateur pourra un jour raisonner comme l'homme et avoir une intelligence à son image. Ce courant fut très populaire au début des années 70 [Mis 77] et a été beaucoup critiqué [Dre 79]. Il redevient populaire aujourd'hui. Nous ne sommes pas en accord avec ce courant car nous pensons qu'il existe des différences entre le traitement de l'information effectué par un ordinateur et la manière dont raisonne et mémorise l'être humain. Cette approche pourra peut-être dans un avenir plus ou moins proche porter ses fruits si les expérimentations sont faites sur une nouvelle forme d'architecture de machine, fondée sur une étude sérieuse des problèmes de connexion, qui tente de reproduire les connexions existant entre les neurones de notre cerveau ; nous voulons parler bien sûr de ce qui est communément appelé "les réseaux neuronnux". Ainsi l'espoir existe de pouvoir, sur cette machine possédant une capacité de traitement de l'information de même nature que celle du cerveau humain, reproduire tous les mécanismes de raisonnement et de compréhension qui rendent l'homme "intelligent" ; alors, cette perspective accomplie, on pourra sans doute dire qu'une machine possède une intelligence à l'image de celle de l'homme.

- Le second courant est moins prétentieux, il considère l'IA comme branche de l'informatique à part entière. Les personnes appartenant à ce courant ne cherchent pas à calquer l'intelligence humaine, elles cherchent principalement à utiliser les capacités de traitements des machines actuelles, les modèles de données qui en découlent pour réaliser des systèmes informatiques ayant une utilité et une raison d'être. Le souci de ces personnes est de rendre les systèmes plus efficaces en leur permettant d'intégrer de nouvelles données appelées "connaissances" et en leur donnant les moyens de les exploiter pour que le service rendu par le système à l'homme soit le meilleur possible.

Ces deux approches de l'IA se rejoignent au niveau des traitements car les modèles des données choisies et les algorithmes de traitements mis en œuvre dans chacun des courants sont similaires. Ceci est normal car le dénominateur commun aux deux courants est la prise en compte du traitement du symbole, en tant qu'information élémentaire et non en tant que suite de caractères, dans les traitements informatiques. A chaque symbole est alors attaché un sens qu'il faut manipuler et confronter aux autres sens des autres symboles. Des représentations du sens et des moyens de l'exploiter ont alors été utilisées, des techniques nouvelles d'exploitation de données sont alors apparues [Far 87].

Nous allons donc nous intéresser dans ce chapitre à l'apport de ces nouvelles techniques dans la recherche d'information. Etudions cet apport en considérant les fonctions principales d'un système de recherche d'informations.

II L'indexation

Les techniques du traitement et de la compréhension du langage naturel ont modifié les conceptions classiques de l'indexation automatique de documents dits textuels. La notion de recherche d'information en texte plein* est alors apparue [Chi 86][Deb 88]. Le principe étant que l'indexation s'effectue sur tout le document, et non sur un résumé de celui-ci, en utilisant des connaissances de linguistiques qui viennent compléter les critères statistiques habituels [Sal 83][v Rij 79].

Un autre apport de l'IA à l'indexation provient de la reconnaissance des formes. En effet, le but de ce domaine est d'extraire des objets analysés certaines caractéristiques exploitées ensuite pour identifier l'objet. L'approche de l'indexation est similaire puisque le résultat de l'indexation est un ensemble de caractéristiques utilisées pour la recherche. Certaines méthodes d'obtention et d'exploitation de caractéristiques utilisées en reconnaissance des formes ont été d'une contribution importante pour la résolution de problèmes d'indexation ou de classification automatique de documents [Smi 80].

Nous n'étudierons pas plus en détail les problèmes d'indexation automatique de texte, car ils demeurent assez éloignés des problèmes d'indexation d'une image. L'indexation automatique d'image utilisant des techniques de reconnaissance des formes est limitée à des domaines très précis. En effet, pour indexer une image, c'est-à-dire pour identifier les objets présents, le système doit avoir un modèle de tous les objets qu'il peut rencontrer. Cette identification n'est pas toujours facile ; ces systèmes ne sont utilisés que pour des besoins précis comme par

* en anglais : full text retrieval (peut-être plus expressif qu'en français !).

exemple des besoins militaires. L'indexation manuelle est alors la seule méthode d'indexation à pouvoir être utilisée pour des images n'ayant pas de finalité précise.

Ce type d'indexation ne possède pas encore de système d'aide utilisant des techniques de l'intelligence artificielle. Une connaissance a priori sur les méthodes d'indexation et une connaissance sur le contenu de la base d'images à indexer peuvent être exploitées par un système à base de connaissance aidant les documentalistes dans cette tâche. La réalisation d'un tel système est à l'étude dans notre équipe, une première proposition a été faite dans [And 88]. Nous étudierons plus en détail dans la deuxième partie les problèmes relatifs à l'indexation d'images ainsi que les solutions que nous envisageons.

Le dernier apport de l'IA pour l'indexation concerne la structure du document d'indexation. Cette structure peut revêtir toutes les formes connues de représentation de la connaissance allant de l'ensemble de prédicats à la représentation sous forme de frames ou encore de réseaux sémantiques. Le problème qui se pose alors est celui mis en évidence au chapitre 2 concernant la mise en correspondance d'une requête avec un document ayant une structure assez complexe ; la notion de distance entre une requête et un document devient alors plus difficile à modéliser. Un autre inconvénient dans l'utilisation d'une structure de document plus sophistiquée réside dans le fait que le temps pris par la mise en correspondance sur un grand nombre de documents devient important. C'est pour ces raisons que les structures classiques des documents d'indexation demeurent dans les systèmes opérationnels.

Voyons maintenant ce qu'ont apporté les techniques d'IA pour l'interrogation des systèmes de recherche d'informations.

III L'interrogation

L'apport de l'IA dans la fonction d'interrogation des systèmes de recherche d'informations revêt plusieurs formes :

L'une d'elle provient de l'utilisation des techniques proposées par les chercheurs travaillant sur la compréhension et la traduction automatique dans les systèmes proposant une interrogation en langage naturel. Le principe général de ces formes d'interrogation est de considérer la requête comme un document, et de l'indexer de la même manière que l'ont été les documents (cf chapitre 3, § II.3). Ce sont en général les modes d'interrogation utilisés dans les systèmes de recherche plein texte [Chi 86][Deb 88].

Un autre apport de l'IA est l'approche système expert ; celle-ci a permis de reconsidérer la forme classique des fonctions d'interrogation des SRI (Système de Recherche d'Informations).

III.1 Les systèmes experts

L'apparition des systèmes dit "experts" a offert beaucoup d'idées d'extensions possibles aux systèmes existants [Bas 86][Bas 87][Vic 85][Bar 88][Chi 87][Cre 85]. La principale idée, qui respecte la notion de système expert, est que l'expertise d'un documentaliste puisse être représentée dans un système. Cette expertise, étant le plus souvent représentée sous la forme de règles de production s'appuie sur un ensemble de connaissances que possède le documentaliste. Les connaissances utilisées par un documentaliste effectuant une recherche de documents sont de trois sortes : sa connaissance des techniques de la recherche d'informations (formulation booléenne, parcours du thesaurus, ...), sa connaissance du fonds documentaire qu'il gère (contenu de certains documents, leur mise en forme, les domaines couverts par le fonds, la mémoire qu'il a gardée de certaines interrogations ...) et pour terminer sa culture personnelle (histoire, géographie, politique, scientifique ...). Les deux premières formes de connaissances peuvent être obtenues grâce à un long dialogue avec l'utilisateur, elles peuvent être ensuite matérialisées et utilisées de manière heuristique dans un système. Mais l'expertise demeure incomplète si la troisième forme de connaissance est inutilisée, et c'est souvent le cas car la masse de connaissance qui devrait être représentée est énorme et difficile à matérialiser, surtout lorsque les documents couvrent des domaines divers et variés [Sal 85]. De plus l'inconvénient de ces systèmes est qu'il faut trouver pour chaque application un expert différent car la connaissance du fonds est elle aussi primordiale dans le déroulement de l'expertise.

Il en ressort que l'approche purement "système expert" classique est mal adaptée aux problèmes posés par la recherche d'information : l'expertise d'un documentaliste étant à notre avis impossible à représenter dans sa totalité, donc impossible à utiliser. C'est pourquoi on assiste en ce moment à l'apparition d'approches différentes où la notion de "système expert" est remplacée par la notion de "systèmes à base de connaissances".

III.2 Les systèmes à base de connaissances

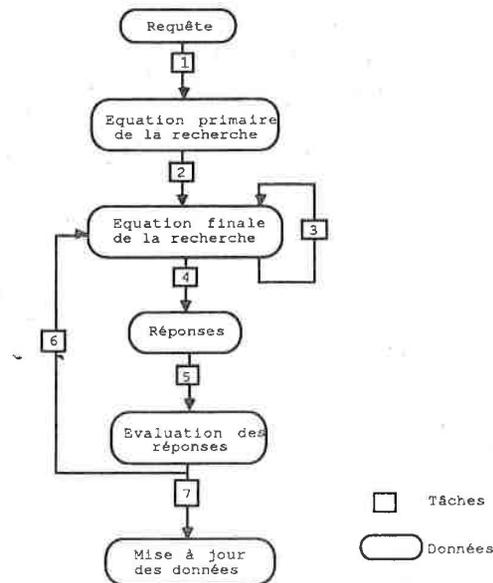
Un système à base de connaissances (SBC) est un système possédant une certaine masse de connaissances sur laquelle il est capable de "raisonner", c'est-à-dire d'effectuer des traitements logiques le plus souvent heuristiques. La différence avec un système expert est que ces traitements ne sont pas forcément ceux indiqués par un expert. Il n'en reste pas moins qu'un système expert est un système à base de connaissances.

Nous dirons dans un premier temps qu'un SBC doit apporter plus de souplesse au SRI (Système de Recherche d'Informations) : les connaissances présentes dans le système doivent permettre à celui-ci de réagir plus aisément et plus précisément aux problèmes que rencontre l'utilisateur. De plus, un SBC permet une gestion souple des connaissances ; elles peuvent être améliorées progressivement au fur et à mesure des expérimentations, ceci grâce à la séparation "connaissance" / "raisonnement". De nombreuses connaissances sont alors nécessaires et l'aide d'un documentaliste demeure souhaitable pour la collecte, la représentation et les utilisations cohérentes d'une partie de ces connaissances. Comme nous l'avons dit précédemment, un système ne sera jamais en possession de toute la connaissance nécessaire pour mener à bien une recherche à la manière d'un documentaliste. Par contre un système peut mettre en évidence des connaissances que les documentalistes ne connaissent généralement pas ou qu'ils utilisent très mal. Ces connaissances sont celles utilisées par certains systèmes pour la mise en correspondance, comme par exemple la pondération d'une indexation automatique, la fréquence d'apparition d'un terme dans les documents ou encore la fréquence d'apparition de co-occurrence de termes. Ces types de connaissances regroupés avec des connaissances dites "expertes" ou du "domaine" dans une même base de connaissances offrent au système de recherche d'information à base de connaissance (SRIBC) des possibilités de traitement de l'information qui ne sont pas semblables à une expertise classique d'un documentaliste, mais plutôt proches d'une combinaison de traitements heuristiques. Ces traitements heuristiques doivent être mis en œuvre avec des experts de la recherche d'information ayant pris conscience des types de connaissances pouvant être utilisés dans de tels systèmes.

Ces différents types de connaissances peuvent être par exemple des connaissances sémantiques sur les termes de l'indexation, des connaissances sur le domaine couvert par la base de document, des connaissances sur les méthodes de recherche et surtout sur les stratégies de déformation de requêtes, des connaissances sur les utilisateurs potentiels de la base et sur les utilisations possibles des documents extraits de la base. Toutes ces connaissances, dont nous ne prétendons pas avoir fait une liste exhaustive, peuvent contribuer, si elles sont judicieusement utilisées, à améliorer le processus de recherche. Voyons plus précisément et à titre d'exemples deux systèmes, IOTA et I³R, qui utilisent toutes ou une partie de ces connaissances. Bien évidemment ce ne sont pas les seuls ; d'autres systèmes comme PLEXUS [Vic 85], DIALECT [Bas 86], CANSEARCH [Del 86] mettent en place des processus qui sont paramétrés et guidés par une partie de ces connaissances.

III.2.1 Le système IOTA

Le système IOTA [Chi 86][Chi 87] se propose d'accéder à une base de documents textuels par une formulation de requête en "langage naturel". Le processus d'interrogation peut être représenté par le schéma suivant extrait de [Chi 86] :



La tâche 1 consiste à analyser la requête en langage naturel afin d'obtenir une première équation de recherche en utilisant principalement les concepts introduits par l'utilisateur. L'utilisateur a la possibilité de modifier les concepts pris en compte par le système.

La tâche 2 fait la correspondance entre les concepts utilisés dans la requête et les termes ayant servi à l'indexation. A chaque concept est alors associée une expression booléenne de termes d'indexation, cette association étant obtenue essentiellement par l'exploitation de connaissances syntaxiques sur les termes.

La tâche 3 évalue deux paramètres qui vont conditionner la recherche. Ces paramètres sont : le niveau de dégradation de la requête et le type de l'utilisateur. Le niveau de dégradation estime dans quelle mesure la tâche 2 et la tâche de reformulation (6) ont dégradé l'équation primaire de recherche. Le type de l'utilisateur est une connaissance qui est déduite de la forme brute de la

requête (caractéristiques sur les termes employés) et du comportement de l'utilisateur (participation dans la recherche).

La tâche 4 met en correspondance les documents et l'équation finale de la recherche. Cette mise en correspondance est décrite dans [Chi 87], elle est fondée sur l'utilisation d'un modèle de correspondance à base de logique booléenne décrit au chapitre 3.

La tâche 5 évalue les résultats de la tâche 4 en essayant de déterminer s'ils sont conformes à l'attente de l'utilisateur, celle-ci étant déduite de la typologie définie dans la tâche 3. Trois diagnostics mettant en évidence des caractéristiques qualitatives et quantitatives peuvent être déduits lorsque les résultats sont jugés non satisfaisants :

- le nombre de réponses est trop important,
- le nombre de réponses est insuffisant,
- la qualité des réponses n'est pas bonne.

La tâche 6 reformule l'équation de recherche en tenant compte du diagnostic transmis par la tâche 5. La reformulation met en œuvre des stratégies de déformation : remplacement de concepts ou modification d'opérateurs logiques.

La tâche 7 consiste à enregistrer les différents traitements réalisés lors de l'interrogation afin de pouvoir mettre à jour les données du système.

Toutes ces tâches manipulent des connaissances :

- connaissance sur les stratégies de déformation et les techniques de recherche d'informations,
- connaissance du domaine (thesaurus),
- connaissance linguistique,
- connaissance intrinsèque à la base : les données de l'indexation,
- connaissance déduite par le système :
 - . le type de l'utilisateur,
 - . la dégradation de la requête.

Les traitements utilisant ces connaissances sont représentés sous la forme de règles de production. Prenons comme exemple la déduction du type de l'utilisateur.

Le type de l'utilisateur est défini dans IOTA [Chi 87] par la définition de deux caractéristiques de l'utilisateur : d'une part sa connaissance du domaine couvert par le corpus et d'autre part sa connaissance du système. A partir de ces deux caractéristiques un usager est placé dans une des trois catégories : spécialiste, moyen, débutant.

L'évaluation de la première caractéristique s'effectue en mesurant la spécificité des termes utilisés par l'utilisateur :

exemple :

Si l'utilisateur manipule des termes assez spécifiques alors c'est un spécialiste du domaine.

La deuxième caractéristique est définie en fonction de la participation de l'utilisateur à l'exploitation de sa requête. A chaque caractéristique est associée une valeur parmi {spécialiste, moyen, débutant}, la déduction du type de l'utilisateur est obtenue grâce à la table de vérité suivante :

caractéristique 1	caractéristique 2	type
débutant	débutant	débutant
débutant	moyen	débutant
débutant	spécialiste	moyen
moyen	débutant	moyen
moyen	moyen	moyen
moyen	spécialiste	moyen
spécialiste	débutant	moyen
spécialiste	moyen	spécialiste
spécialiste	spécialiste	spécialiste

La connaissance sur le type de l'utilisateur est ensuite utilisée dans les traitements en partie droite des règles de production, par exemple, dans la tâche "évaluation des réponses et formulation de diagnostics" on pourra trouver des règles de la forme :

Si l'usager est un spécialiste et le nombre de réponses est faible et il y a des réponses qui ont été jugées bonnes par le système (mesure de pertinence) alors le résultat est correct.

Cette règle suppose que pour un spécialiste la qualité des réponses est plus importante que la quantité, cette règle exprime une connaissance que l'on peut appeler "experte".

Nous terminerons cette présentation en indiquant que le prototype IOTA est écrit en LELISP avec un document de 3000 mots d'un manuel technique du CNET. Les résultats de la version finale indiquent un bon rappel de 0,68 et une bonne précision de 0,77 (cf cha. 1, § III.1) mais il faut préciser que le corpus utilisé est très petit relativement aux systèmes classiques de recherche d'informations.

III.2.2 Le système I³R

Le système I³R (Intelligent Interface for Information Retrieval) [Cro 86] [Cro 88] fondé sur une architecture "multi-experte" se propose de fournir à l'utilisateur une assistance constante et

appropriée à chaque étape de sa recherche. Comme dans le système IOTA, l'interrogation en I³R s'effectue en trois grandes étapes :

- construction d'une représentation de l'information recherchée (formulation d'une requête),
- mise en correspondance de la requête avec les documents,
- évaluation des réponses et reformulation de la requête si nécessaire.

A chaque étape le système fournit toute une variété d'outils aidant l'utilisateur dans sa tâche. Les différents outils, associés chacun à une source de connaissances, sont organisés à la manière d'un système à base de "blackboard"¹[Lâa 88]. Il existe ainsi plusieurs systèmes à base de connaissances qui communiquent entre eux par l'intermédiaire du blackboard. Le déroulement des étapes de recherche et le choix de l'outil utile à un moment donné sont gérés par un SBC spécial appelé "directeur"².

Les différentes sources de connaissances mises en œuvre sont :

- Le constructeur du modèle de l'utilisateur : il collecte des informations sur l'utilisateur pour déterminer la forme d'interaction la mieux adaptée, le but général de la session de recherche, mais aussi pour évaluer les centres d'intérêt de l'utilisateur. Cette connaissance intervient dans la définition du modèle de la requête.

- Le constructeur du modèle de la requête : il fabrique une représentation de l'information recherchée par l'utilisateur. Ensuite cette représentation est indexée à la manière d'un document afin d'obtenir un ensemble de termes pondérés, ces termes étant ceux utilisés dans l'indexation des documents. La détermination des documents pertinents est effectuée en utilisant un modèle de correspondance probabiliste (cf cha. 3 § II.1).

- Le gestionnaire de la connaissance : il utilise la connaissance du domaine extraite du modèle de l'utilisateur et la base de connaissance générale du système pour en déduire les concepts sous-jacents à la demande initiale de l'utilisateur. Ces concepts sont alors évalués par l'utilisateur puis insérés dans le modèle de la requête.

- Le contrôleur de la recherche : il détermine les stratégies à appliquer pour rechercher les documents pertinents relativement à un modèle de requête.

- Le générateur de parcours : il fournit à l'utilisateur les moyens de trouver des documents en naviguant dans la base de connaissance. La connaissance acquise pendant un parcours peut être utilisée pour modifier le modèle de la requête ainsi que les stratégies à utiliser.

¹ : une architecture de blackboard est composée d'une base de donnée divisée en plusieurs niveaux d'abstraction propres à l'application, de différentes sources de connaissances qui doivent coopérer, seules ces sources ont accès au blackboard, et d'un contrôleur qui contrôle le déroulement des opérations.

² : en anglais "scheduler".

Remarque : Cette notion de parcours évoque sans nul doute la notion d'hypertexte dont nous reparlerons dans une prochaine partie.

Tous ces composants manipulent de la connaissance sur laquelle ils raisonnent, ils communiquent leur réflexion aux autres composants via le blackboard. Les connaissances manipulées sont de différentes sortes :

- le modèle de l'utilisateur,
- le modèle de la requête,
- les différentes stratégies de recherche,
- la connaissance sur le domaine.

Explicitons à titre d'exemple la connaissance du domaine telle qu'elle a été présentée dans [Cro 86].

La connaissance du domaine est un ensemble de concepts représentés sous la forme de frames ayant trois parties :

1. le nom du concept,
2. des informations sur la manière de reconnaître le concept,
3. les relations avec les autres concepts.

La deuxième partie est un ensemble de règles permettant de reconnaître la présence du concept. Ces règles sont de la forme :

Si p alors concept (c) .

p est une propriété mis en évidence dans un texte qui peut être la requête de l'utilisateur, ou un document à indexer, **concept** est le concept déduit, c est un coefficient de certitude ou de plausibilité, variant entre 0 et 1, associé à la déduction faite par cette règle.

Exemple :

Si adjacent (information, recherche) alors recherche d'informations 0,9.

Le détail de la forme de raisonnement combinant les plausibilités des règles est décrit dans [Cro 88], nous n'aborderons pas ce problème ici.

La troisième partie explicite les relations du concept avec les autres concepts, les relations utilisées sont celles qui ont été présentées en partie au chapitre 1 :

- synonymie,
- l'association entre termes de sens proche (Cross-reference),
- les relations de hiérarchies :
 - . partie de,
 - . sorte de .

Cette connaissance est principalement utilisée dans le composant "gestionnaire de la connaissance" mais elle conditionne grâce à la communication inter-composants les autres sources de connaissances.

Cette architecture de système a été testée sur une collection dite test (CACM collection d'articles des conférences ACM) avec des résultats jugés satisfaisants [Cro 88].

L'orientation future du système est la prise en compte d'un système traitant le langage naturel comme un nouveau composant de l'architecture "multi-expert" pour permettre au système de comprendre plus précisément les concepts exprimés par l'utilisateur. Cette introduction nécessitera une harmonisation entre les traitements statistiques actuellement existant lors de l'indexation de la requête ou des documents et les techniques de représentation des concepts utilisées classiquement en compréhension du langage naturel [Cou 87].

Après ces exemples de systèmes manipulant des connaissances pour améliorer le processus de recherche, voyons maintenant une des dernières applications de l'IA dans la recherche d'information. Cette application est la mise en œuvre de différentes formes d'apprentissage dans les systèmes de recherche d'information.

III.3 L'apprentissage et ses applications en recherche d'informations

Deux types d'apprentissage peuvent être mis en œuvre dans les systèmes de recherche d'informations : l'apprentissage à court terme et l'apprentissage à long terme [Smi 80]. L'apprentissage à court terme consiste à donner au système la possibilité d'améliorer l'étape courante d'une recherche en tenant compte des étapes précédentes. L'apprentissage à long terme, comme son nom l'indique, ne se place pas sur la même échelle du temps ; il permet au système d'améliorer son processus de recherche grâce à une prise en compte des sessions précédentes d'interrogation.

III.3.1 L'apprentissage à court terme

L'apprentissage à court terme, grâce à la prise en compte du choix de l'utilisateur parmi un ensemble de documents présentés par le système, permet la modification de la demande afin qu'elle corresponde le mieux possible aux documents jugés pertinents par l'utilisateur. Cette approche permettant une recherche progressive a fait l'objet de nombreuses études ; car la notion de bouclage de pertinence, fondée sur le calcul pour chaque terme d'un poids appelé poids de pertinence [Sal 83][v Rij 79], en est un exemple.

La prise en compte du jugement de l'utilisateur à différents stades de la recherche améliore la qualité de la recherche. Plusieurs systèmes proposent cette utilisation sous différentes formes. Le système THOMAS [Odd 77] offre une forme de dialogue où la participation de l'utilisateur est grande ; peut alors parler d'apprentissage avec professeur, où l'élève est le système qui doit

apprendre à formuler ce que l'utilisateur désire. L'utilisateur au départ introduit quelques termes caractérisant l'objet général de sa recherche qui peut être mal défini au départ. Le système ensuite propose à l'utilisateur le document semblant le mieux correspondre à sa demande, le document est présenté par son titre, ses auteurs ainsi que les descripteurs ayant servi à son indexation. L'utilisateur est alors invité à indiquer la pertinence des informations présentées, il peut dans un même temps introduire de nouvelles notions. Le système reconstruit alors une nouvelle expression des besoins de l'utilisateur et repropose le document qui semble être le plus proche de cette expression. Ainsi progressivement le système va construire une expression de plus en plus précise des besoins de l'utilisateur, sans que celui-ci ait eu à l'exprimer clairement dans un langage de requête quelconque.

L'approche de Wong & Ziarko dans [Won 86a] est quelque peu différente : ils proposent un système qui tente d'apprendre le besoin de l'utilisateur en l'exprimant sous la forme d'un ensemble de règles r_i :

r_i : Si $p_i(D)$ alors pertinence(D) coef,

où p_i est une propriété vérifiée par le document D. Cette propriété s'exprime par la conjonction de caractéristiques indiquant la présence ou la non présence d'un terme dans la description du document. La **pertinence(D)** est la conclusion de la règle indiquant si le document est pertinent, non pertinent ou ni l'un ni l'autre. Le **coef** est une mesure de certitude associée à la déduction faite par la règle.

Un premier jeu de règles est obtenu en proposant à l'utilisateur un ensemble de documents sur lesquels, pour chaque document, l'utilisateur indique si il est pertinent ou non. A partir de ce jugement le système construit une partition de l'ensemble des documents induite par la présence ou la non présence de termes dans la description des documents. Le système obtient ainsi un certain nombre de classes de documents, chacune décrite par une propriété p_i . Pour chaque classe il détermine si elle caractérise ou non le besoin de l'utilisateur en fonction du nombre de documents pertinents et non pertinents présents dans ces classes, une règle est ainsi obtenue pour chaque classe. La construction des règles utilise un algorithme appelé PLA (probabilistic learning algorithm) décrit dans [Won 86b].

L'application des règles sur l'ensemble des documents fournit un nouvel ensemble de documents ordonné relativement à un poids de pertinence associé à chaque document, et calculé en fonction des coefficients de certitude attachés aux règles. L'utilisateur indique de nouveau pour chaque document sa pertinence par rapport à ses besoins, de nouvelles règles sont alors construites par la modification des anciennes ou l'adjonction de nouvelles. Le processus continue jusqu'à ce que l'utilisateur soit satisfait ou que le système s'aperçoive qu'il ne peut le satisfaire. Ce processus de recherche est défini plus précisément dans un algorithme appelé AIR (Adaptive Information Retrieval) [Won 86a]. L'inconvénient majeur de ce système demeure

dans la complexité des algorithmes mis en œuvre. Cette complexité, si elle est importante, peut augmenter de beaucoup les performances de recherche ; par manque de résultat d'expérimentation nous ne pouvons rien indiquer à ce sujet.

III.3.2 L'apprentissage à long terme

Un apprentissage à long terme a pour objectif d'obtenir un système ayant la capacité d'améliorer son comportement grâce à l'étude et à l'exploitation des recherches antérieures. Cet apprentissage nécessite la mémorisation pour chaque session d'interrogation des traitements effectués ainsi qu'une appréciation sur le déroulement de la recherche fournie par l'utilisateur.

L'exploitation de ces données peut avoir différentes influences sur le système. Elle peut permettre la mise à jour des connaissances utilisées dans le système, comme par exemple l'évolution de la structure d'un thesaurus, la modification des utilisations des stratégies de recherche ou encore la définition de nouvelles stratégies de recherche adaptées à certaines catégories d'utilisateurs. Cette exploitation nécessite une mémorisation, sous une forme synthétique et réutilisable, des traitements effectués. Ainsi, le système pourrait être capable de retrouver une recherche qu'il a correctement menée par le passé, analogue à celle qu'un utilisateur est en train d'effectuer, afin d'aider cet utilisateur dans sa démarche. Cette approche peut être réalisée en définissant une sorte de "meta-SRI" sur le SRI lui-même. Dans ce meta-système, les informations à retrouver seraient les différentes recherches déjà réalisées, et la requête serait la recherche en cours ; mais encore faut-il déterminer les caractéristiques d'une recherche, ainsi que le modèle de correspondance associé.

La modification ou la construction de connaissance nécessite un contrôle extérieur ; un apprentissage automatique peut rendre un système inutilisable car la connaissance apprise peut s'avérer incompréhensible par l'homme. Un exemple de connaissance pouvant être apprise a posteriori est la connaissance des utilisateurs potentiels d'un système, car cette connaissance est souvent inconnue à l'origine. Il faut alors mettre en œuvre dans le système tout un ensemble de critères pouvant caractériser un utilisateur, puis par apprentissage en déduire une typologie des utilisateurs. A l'aide de cette typologie le système pourra adapter son comportement et mettre en œuvre les bonnes stratégies de recherche en fonction de l'utilisateur, cette adaptation pouvant être dans le meilleur des cas apprise grâce aux utilisations antérieures du système.

La construction d'un tel système n'est pas encore pour maintenant car l'apprentissage à un tel niveau nécessite une structuration du système très complexe. Pour le moment quelques systèmes comme IOTA (cf § III.2.1) proposent la mémorisation de quelques informations caractérisant une recherche dans le souci d'améliorer par la suite le processus de recherche. Connaissant les difficultés et les coûts de construction des bases de connaissances, et la diversité des stratégies de recherches pouvant être utilisées à un moment donné de la recherche, le système

idéal serait un système qui posséderait initialement les connaissances minimales pour effectuer des recherches et qui en fonction de la base de documents et au fur et à mesure des utilisations s'adapterait pour offrir aux utilisateurs potentiels le service attendu. Mais j'ai bien peur que cette vision du système idéal reste utopique et que les systèmes proposés dans l'avenir en demeurent assez éloignés. Car la connaissance initiale nécessaire à un tel système n'est pas encore représentée. Pour le moment seulement des solutions ponctuelles peuvent être apportées comme par exemple l'apprentissage de la typologie des utilisateurs.

IV Conclusion

Ce chapitre a tenté de montrer quel était l'apport de l'Intelligence Artificielle dans la conception des systèmes de recherche d'informations. Cette étude montre que le plus grand apport concerne la fonction interrogation de ces systèmes. En effet l'approche Système Expert puis l'approche système à base de connaissance ont permis au SRI d'intégrer et de gérer de nouvelles connaissances afin de rendre le processus d'interrogation à la fois plus performant et plus agréable. L'étude de deux systèmes IOTA et I³R a montré quelques exemples d'utilisation de connaissances adaptées à la recherche d'information.

Un autre apport important de l'IA concerne l'utilisation de techniques d'apprentissage symbolique. Deux formes d'apprentissage ont été mis en évidence :

- l'apprentissage à court terme où le système tente d'apprendre et d'exprimer les besoins de l'utilisateur,
- l'apprentissage à long terme où le système doit améliorer ses stratégies de recherche au fur et à mesure des utilisations.

La première forme d'apprentissage est illustrée très largement dans la quatrième partie de cette thèse. La deuxième forme constitue un problème beaucoup plus complexe et nous verrons qu'elle représente une de nos perspectives à long terme (cf quatrième partie, chapitre 3, § VI).

Conclusion

Cette première partie a été consacrée à l'étude de l'environnement d'un système de recherche d'informations car la recherche d'images est une recherche d'informations. Cette étude a été conçue avec une complexité croissante afin de permettre aux non initiés d'aborder la problématique de la recherche d'informations d'une manière évolutive.

Le premier chapitre a présenté dans un cadre général cette problématique en introduisant les différentes sortes de pertinences existantes ainsi que les principales fonctions de ces systèmes, à savoir l'indexation et l'interrogation.

Les principaux modèles et fonctions utilisés dans les SRI ont été présentés dans le chapitre 2. Les notions de modèles d'indexation et de modèles de correspondance ont été introduites ; ces notions permettent de caractériser plus précisément les SRI. La structure des thesaurus a été présentée dans le détail car, nous le verrons dans les prochaines parties, seule l'utilisation d'une connaissance sémantique de type thesaurus peut nous permettre de proposer un processus de recherche cohérent. L'utilisation du thesaurus conjointement à une formulation booléenne offre un pouvoir d'expression important mais qui n'est pas facilement contrôlable par un utilisateur non initié.

L'étude des tendances actuelles nous a permis de voir le bouclage de pertinence qui peut être facilement formalisé dans certains modèles de correspondance et plus difficilement dans d'autres, comme par exemple dans les modèles à base de logique booléenne. Cette étude a montré que, mis à part les modèles linguistiques qui doivent encore faire leurs preuves, les nouvelles approches sur les modèles de correspondance ne permettent pas la prise en compte de connaissance sémantique et que la formulation booléenne permettait un contrôle du sens donné à une requête.

L'apport de l'IA dans les SRI a été présenté dans le chapitre 4. Cet apport concerne surtout la fonction d'interrogation où la gestion de nouvelles connaissances offre une amélioration des processus de recherche. Cette étude a mis en évidence que l'apprentissage a un rôle important à jouer que ce soit à l'échelle de la vie du système ou à l'échelle d'une session d'interrogation.

Toutes ces études concernent la recherche d'information de type texte, car l'arrivée de l'image en tant que donnée d'un système informatique est encore récente. Or l'image est une information différente du texte écrit, sa prise en compte en tant qu'information à rechercher demande une

certaine adaptation des méthodes utilisées dans les SRI existants. En effet, certaines méthodes qui sont efficaces pour rechercher du texte ne le sont plus lorsque ce sont des images que l'on recherche et au contraire certaines méthodes prennent toute leur importance grâce aux spécificités de l'image. Cette prise en compte de l'image passe avant tout par une étude de ses caractéristiques et de leurs influences sur un processus de recherche, c'est cette étude qui est menée dans la deuxième partie.

Deuxième partie

Une recherche d'informations particulière : la recherche d'images

Introduction

Nous avons décrit dans la partie précédente les caractéristiques d'un système de recherche d'informations en considérant que l'information recherchée est représentée dans la base par un document la décrivant. Il nous faut maintenant prendre en compte la spécificité du type de document que notre système devra retrouver lors de ses recherches : l'image.

Le premier chapitre de cette partie s'intéressera à l'image en tant que support d'informations en présentant le problème assez complexe de l'analyse de l'image, et celui du rapport entre les mots et les images. Puis après l'étude de deux méthodes d'analyse de l'image, nous présenterons la recherche d'images à travers son évolution et dans son contexte actuel. Cette présentation se terminera par une description de quelques systèmes existants.

Le deuxième chapitre est consacré au projet EXPRIM, nous le présenterons d'une manière générale puis plus précisément à travers une étude critique des idées émises et des réalisations effectuées avant notre arrivée dans l'équipe.

Le dernier chapitre présente notre première étude et notre première réalisation du projet EXPRIM que nous avons effectués dans le cadre d'un projet ESPRIT. Même si cette étude n'a pas donné les résultats escomptés, elle a été déterminante dans l'évolutions de nos idées.

Chapitre 1	
Cadre général.....	75
I Les différences apportées par l'image	75
I.1 Le sens et l'image.....	75
I.2 Les images et les mots	77
I.2.1 Perception d'un mot.....	77
I.2.2 Le pouvoir d'imagerie	78
I.2.3 Application à la recherche d'images.....	79
I.3 L'analyse de l'image	80
I.3.1 Système descriptif des représentations.....	80
I.3.2 L'analyse d'une photographie	83
I.3.2.1 La morphologie d'une photographie.....	83
I.3.2.2 Le contenu informatif.....	84
I.3.2.3 La connotation.....	86
I.3.3 Conclusion	87
I.4 La recherche en elle-même.....	87
I.4.1 Historique	87
I.4.2 Les supports optiques.....	88
I.4.3 La visualisation	90
I.4.4 Conclusion	91
II La recherche d'images et ses applications.....	91
II.1 Les utilisateurs et les utilisations	91
II.1.1 Profil des demandeurs d'images.....	92
II.1.2 Les types de demandes	93
II.2 Les systèmes existants	94
II.2.1 Les systèmes liés à une application.....	94
II.2.2 Les systèmes construits pour la recherche d'images.....	96
III Conclusion.....	98

Chapitre 1

Cadre général

Le système que nous voulons construire est un système de recherche d'informations où l'information recherchée est l'image. Les caractéristiques générales d'un tel système sont les mêmes que celles décrites précédemment. Mais à l'origine les systèmes de recherche d'informations ont été conçus pour rechercher des documents de type textuels, et ils ne répondent pas forcément aux besoins des utilisateurs recherchant des images. C'est pourquoi nous allons tenter de montrer quelles sont les différences apportées par l'image par rapport à un document textuel, et comment ces différences interviennent dans un processus de recherche.

I Les différences apportées par l'image

La grande différence entre une image et un document textuel est que le sens véhiculé par celle-ci n'est pas mis à plat dans un langage connu de tous, mais représenté par des icônes dont l'agencement, à l'intérieur d'une surface souvent rectangulaire, fait transparaître un sens qui ne peut être perçu de la même manière par tous.

I.1 Le sens et l'image

Ainsi l'image véhicule un sens qu'il n'est pas facile de représenter. De nombreuses études ont été faites sur ce sujet mettant en évidence la difficulté d'identifier à un moment donné le message que transmet une image à un individu [Gau 81] [Por 76]. En effet, le sens d'une image peut varier selon la personne qui la regarde et selon le contexte dans lequel elle est placée. En effet, lorsque qu'un individu "lit" une image, il identifie certains composants en faisant passer l'image à travers un filtre que représente sa culture. Cette identification peut être différente dans un contexte différent ; de nombreux exemples montrant une image hors de son contexte ou associée à une légende influençant l'interprétation mettent en évidence cette caractéristique [Dur 79].

Le sens que véhicule une image dépend alors de nombreuses caractéristiques :

- les éléments présents dans cette image,
- l'agencement de ces éléments dans l'image,

- le contexte de l'image,
- la culture de la personne qui regarde l'image,
- le contexte qui fait que la personne regarde cette image ...

Cette liste n'est sûrement pas exhaustive mais elle montre la réelle difficulté que rencontre toute personne désirant décrire une image c'est à dire représenter son sens. Parmi cette liste on peut mettre en évidence certaines caractéristiques qui sont assez intrinsèques à l'image et qui peuvent, dans une certaine mesure, être communes à toutes interprétations : les éléments présents, leur agencement et le contexte de l'image. En effet ces caractéristiques, à un niveau de connaissance identique, peuvent être reconnues de tous.

Tous les éléments concrets d'une image peuvent être identifiés et énumérés ainsi que leur agencement ; le problème est que cette liste peut être très longue et qu'elle ne sera jamais exhaustive (cf § I.3.2).

De plus pour déterminer les agencements il faut utiliser un langage connu de tous. La description de ces agencements peut se faire globalement par rapport à l'image, mais aussi plus précisément entre les différents constituants de l'image.

Le contexte de l'image peut être décrit précisément si l'on connaît l'origine de l'image, ce pour quoi elle existe. Il peut être identifié par une date, un lieu, une époque, un événement, une histoire

Le dernier problème que l'on peut soulever relativement à l'identification du sens d'une image, est celui de son expression : quel langage utiliser ?

L'image possédant son propre langage, on pourrait s'arrêter là ; mais lorsque l'on veut communiquer le sens d'une image à une autre personne, on utilise la parole ou le texte écrit. Quand on veut qu'un système informatique puisse utiliser l'image en tant qu'information, il faut qu'auparavant cette image ait été décrite. Les systèmes informatiques ne peuvent manipuler l'image en tant qu'information élémentaire. Ainsi le document recherché n'est pas l'image mais sa description qui est un ensemble de mots agencés selon un modèle de document choisi.

Etudions alors quelles sont les relations existantes entre les images et les mots. Nous verrons par la suite comment décrire une image et quelles sont les informations à retenir pour rendre la recherche d'images possible.

I.2 Les images et les mots

Nous souhaitons montrer dans ce paragraphe les relations entre les mots et les images par l'intermédiaire d'une étude de M. Denis [Den 79] sur les images mentales. En effet, Denis montre que la représentation imagée est très utile pour l'apprentissage des mots que nous effectuons tout au long de notre existence. Grâce à cette étude, nous avons essayé de comprendre dans quelle mesure un mot permet l'évocation d'une image et inversement, comment une image fait référence à des mots. Nous pensons qu'il est intéressant de présenter succinctement les idées principales de cette étude, afin de montrer quelles peuvent être les démarches de notre esprit lorsque nous souhaitons associer des mots à une image.

L'image est une fenêtre sur le monde ; elle représente un support d'informations extrêmement conçus ("un bon croquis vaut mieux qu'un long discours" Napoléon 1^{er}). L'homme perçoit des millions d'images par jour, qu'il interprète et codifie dans sa mémoire. Il peut aussi se représenter mentalement ces images et les transformer à volonté. L'image est donc aussi une représentation mentale qui permet à toute personne de construire et de diriger ses pensées.

I.2.1 Perception d'un mot

Un mot peut être perçu de différentes manières, soit en le lisant, soit en l'entendant prononcer, ou encore en l'associant à un dessin qui l'illustre. Lorsqu'un mot est assimilé, il est enregistré dans notre mémoire avec une certaine représentation qui va dépendre de la manière dont il a été perçu et de la forme concrète qu'il peut avoir. En effet Denis met en évidence deux formes de représentations mentales, l'une étant verbale et l'autre imagée; un mot lu aura plus facilement une représentation mentale verbale et, par opposition, un mot identifié avec un dessin aura une représentation imagée.

La représentation imagée d'un mot sera facilitée si ce mot est concret, en effet lors de la perception d'un tel mot l'individu met en valeur des **traits sémantiques figuratifs** qui caractériseront l'image qu'il se construit.

Exemple :

Pour le mot cheval, les traits sémantiques sont : quatre pattes, crinière, hennit, taille, ami de l'homme, vitesse.

Ces traits sémantiques peuvent être physiques, fonctionnels ou évaluatifs. De plus il existe une hiérarchie entre ces traits lors de la perception : une personne perçoit et se représente en premier lieu les caractères physiques (quatre pattes) puis fonctionnels (ami de l'homme, vitesse) et avec un peu plus d'attention les traits évaluatifs (taille). Puis lors de l'évocation du mot, la

représentation imagée associée comportera une sélection des traits sémantiques en fonction du contexte de l'évocation (ex : lors de l'évocation de "cheval de course" les traits sémantiques présents seront "quatre pattes", "vitesse" et "crinière", on omettra le fait qu'un cheval "hennit").

Au contraire, un mot abstrait est pauvre en traits sémantiques figuratifs, la représentation mentale se fait soit par des symboles qui sont des acquis culturels (ex : justice -> balance) soit par l'évocation de situation ou expression physique illustrant le concept (ex : chagrin -> personnage pleurant).

I.2.2 Le pouvoir d'imagerie

Ainsi Denis attache au mot un pouvoir d'imagerie qui représente la facilité qu'a un individu d'évoquer une image lorsqu'il perçoit ou se rappelle un mot. Le pouvoir d'imagerie des mots concrets est alors supérieur à celui des mots abstraits, car ils sont plus riches en traits sémantiques figuratifs donc plus facilement représentables. Un concept abstrait a tout intérêt à avoir une représentation verbale car l'association d'une image influence la compréhension du concept, en limitant son sens.

Dans les mots concrets, on peut distinguer deux catégories de mots à pouvoir d'imagerie différent : les mots généraux et les mots spécifiques. La représentation d'un concept général se réalise par une analyse en traits sémantiques figuratifs, avec un recours systématique à des exemples l'illustrant. Ces exemples définissent des prototypes du concept.

Exemple :

Pour un oiseau, les traits sémantiques sont : ailes, bec, plumes, vole ... et un prototype qui illustre dans notre esprit le concept d'oiseau est l'image d'une hirondelle.

Deux sortes (non disjointes) de mots généraux peuvent être caractérisées : les mots généraux homogènes dont les prototypes partagent les mêmes traits sémantiques (ex : oiseau) et les mots généraux hétérogènes dont les prototypes ont des traits figuratifs différents (ex : meuble, avec comme prototypes : chaise, armoire, table). Les mots généraux hétérogènes ont un pouvoir d'imagerie inférieur aux mots généraux homogènes car l'absence de traits sémantiques communs ne facilite pas la représentation imagée du concept.

Denis met alors en évidence une échelle de mesure des pouvoirs d'imageries des mots :
Mots concrets spécifiques (ex : hirondelle)

Mots concrets généraux homogènes : (ex: oiseau)

Mots concrets généraux hétérogènes : (ex : meuble)

Mots abstraits (ex : pollution).

On peut remarquer que des termes ayant un fort pouvoir évocateur peuvent se retrouver dans plusieurs catégories et seul le contexte permettra de les situer dans l'échelle (ex: rouge, carré).

1.2.3 Application à la recherche d'images

Les résultats de l'étude de Denis nous poussent à croire que le pouvoir d'imagerie des mots a un rôle à jouer dans un système de recherche d'images, où le seul moyen de retrouver automatiquement l'image pour son contenu est de l'associer à une description constituée de mots.

Lors de la description de l'image, il nous semble important d'utiliser des mots à fort pouvoir d'imagerie. Car on peut imaginer que, lors de la visualisation des images, l'utilisateur perçoit plus aisément ces mots, leur représentation imagée étant immédiate. L'utilisation d'un thesaurus pour décrire les images peut alors aider le documentaliste à choisir les bons termes, c'est à dire choisir des termes spécifiques qui auront un pouvoir d'imagerie relativement fort. En effet, la hiérarchie de généralité, présente dans un thesaurus, reflète assez bien l'échelle de mesure de Denis sur le pouvoir d'imagerie des mots. Les termes les plus spécifiques d'un thesaurus ont de fortes chances d'avoir un pouvoir d'imagerie assez important.

L'utilisation de termes à faible pouvoir d'imagerie n'est pas à rejeter, mais leur présence dans une description introduit de l'ambiguïté, car l'image mentale relative à ce terme dépend de la personne et de son expérience. La forme prise par le terme dans l'image décrite n'est pas forcément similaire à la représentation imagée qu'a l'utilisateur.

A ce stade de notre étude nous préconisons l'utilisation d'un thesaurus pour décrire les images. En effet, le thesaurus apporte de la sémantique dans les descriptions, nécessaire à une recherche compréhensible de l'utilisateur, et il permet de choisir des termes qui auront une forte chance d'être évoqués par l'utilisateur lors de la visualisation des images. Voyons maintenant quelles sont les méthodes existantes pour décrire correctement une image.

1.3 L'analyse de l'image

L'image peut avoir plusieurs formes ; elle peut être une photographie sur un événement, un lieu, un objet ou encore un dessin, une représentation ancienne ou moderne, un tableau ... La démarche à suivre pour analyser une image ne sera pas la même suivant la forme de celle-ci.

Ainsi, nous pouvons donner en exemple deux méthodes d'analyse, de "lecture" de l'image. La première est spécialisée dans la description de représentations anciennes, elle utilise un thesaurus et une syntaxe de description. La deuxième est une méthode d'analyse pouvant s'appliquer à toutes formes de photographies traitant d'un thème assez général, le mode de description utilisé est de type "mots-clés".

1.3.1 Système descriptif des représentations

Le système descriptif des représentations [Gar 84], appelé aussi "Thesaurus iconographique" car principalement fondé sur l'utilisation d'un thesaurus assez conséquent, a été entrepris dès 1976 sur commande du Ministère de la Culture. Il permet de décrire des représentations anciennes illustrant des scènes de la vie, des événements et des personnes. Il a été aussi utilisé pour décrire des tableaux, des dessins et des sculptures faisant partie des collections de nombreux musées de France.

L'analyse d'une représentation s'effectue en répondant à un certain nombre de questions permettant d'énumérer les caractéristiques de l'œuvre [Gar 84]:

Sur la signification principale :

Quelle est la signification principale de l'image ?

Comment est-elle exprimée ?

Quels sont les personnages ? Quelles sont leurs caractéristiques ? Connaît-on leurs noms ?

Quels éléments concourent à l'expression de la signification principale ?

Quelles sont les relations importantes, signifiantes et signifiées ?

Quel est le lieu ? Peut-on l'identifier géographiquement ? Convient-il de le décrire ?

La compréhension de l'image nécessite-t-elle une référence à un contexte historique ? Pose-t-elle des problèmes de datation ?

Sur les significations secondaires :

L'image contient-elle d'autres éléments que ceux qui concernent la signification principale ?

Quels sont-ils ? Est-il utile de les noter ?

Dès que la personne qui analyse a répondu à toutes ces questions, il lui faut rédiger la description, c'est-à-dire le document qui permettra par la suite de retrouver l'image. La rédaction d'une description s'effectue suivant le modèle de document du système MISTRAL (cf figure 4), à l'aide de phrases de la forme suivante :

descripteur₀(descripteur₁,descripteur₂, ...,descripteur_n)

Le descripteur₀ représente la dénomination d'une signification de l'œuvre qui est ensuite décrite par les descripteurs mis entre parenthèses. La description comporte autant de phrases que de significations présentes dans l'image.

Le choix des descripteurs s'effectue grâce au thesaurus qui contient trois grands domaines :

- description de la représentation,
- sujet particulier de la représentation,
- source écrite de la représentation.

Le premier domaine est le plus important, il est très hiérarchisé et il permet de décrire le contenu du document. C'est dans ce domaine que la personne qui analyse devra chercher, pour chaque élément d'une signification, le descripteur adéquat et le plus précis, c'est-à-dire le plus spécifique possible.

Le système descriptif des représentations apporte une réflexion pertinente sur l'architecture d'un thesaurus et sur l'analyse d'un document par couches successives. Cette démarche peut, sans aucun doute, être appliquée dans d'autres domaines que celui de l'analyse des représentations.



213



215



216



214



217

- 213 - Scène (paganisme , groupe , prière , idole , intérieur) - idole (ronde bosse . colonne) -
- 214 - Scène (groupe , jeunes gens , divertissement , luth , profanation , eucharistie , prodige , punition : noyade) - pont (cours d'eau) - Scène (prêtre , portage , eucharistie , enfant de chœur , objet du culte , lanterne , cloche , bannière) -
- 215 - Scène (pénitent : nudité , mortification , fouet) -
- 216 - Scène (groupe , diable , écriture , supplice , colonne , lien) -
- 217 - Scène (satirique , groupe , chanoine , discussion , jeune homme , âne : vêtement religieux) -

figure 4 : exemple de représentations et de descriptions associées

I.3.2 L'analyse d'une photographie

La forme d'analyse que nous allons présenter dans ce paragraphe est celle proposée par G. Blery dans [Ble 81] et qui est celle aussi enseignée par M. Cluzeau-Ciry dans son cours sur l'analyse de l'image pour les étudiants en science de l'information de l'UT de Dijon.

Cette forme d'analyse s'apparente à une indexation par mot-clés, un mot-clé exprimant une caractéristique de l'image. Il est important de noter que ce modèle d'indexation ne permet pas de prendre en considération toutes les caractéristiques de l'image. En effet, avec ce modèle on ne peut exprimer qu'un objet est en relation avec un autre objet (un vase est sur une table) ou encore qu'un objet possède une caractéristique (un foulard rouge). Mais un tel niveau de précision dans un modèle rend la mise en correspondance entre requête et documents à la fois très complexe et peu performante (cf première partie, chapitre 2, § II). C'est pourquoi ce modèle d'indexation demeure celui qui est le plus utilisé dans les systèmes de recherche d'images.

L'analyse d'une photographie nécessite de prendre en considération plusieurs de ses aspects indispensables à toute recherche rétrospective. Car une photographie peut être intéressante pour son contenu, mais aussi pour sa qualité de lumière, pour son approche du sujet dans l'espace, pour son côté subjectif ...

La description de l'image, résultat de l'analyse, doit contenir chacun de ces éléments afin que la photographie puisse être retrouvée grâce à ses caractéristiques.

Nous pouvons retenir trois grandes catégories de caractéristiques d'une photographie :

- sa morphologie,
- son contenu informatif,
- son côté subjectif.

I.3.2.1 La morphologie d'une photographie

Pour décrire la morphologie d'une photographie, il faut se demander si une ou plusieurs techniques de prise de vue ont été utilisées pour donner un certain caractère à cette photographie.

La morphologie peut se définir par les éléments suivants :

- l'approche du sujet dans la photographie,
- le temps de pose,
- l'optique,
- la lumière.

La définition de l'approche du sujet dans la photographie n'est pas évidente. Un immeuble peut être vu en plan d'ensemble dans une ville qui est vue en plan moyen. Le définition d'une

approche se construit en fonction d'un élément de la photographie et non pour l'ensemble de la photographie. En se définissant un sujet principal de la photographie, on peut alors qualifier l'approche du sujet dans la photographie par les descripteurs suivants : plan d'ensemble, plan moyen, plan américain, gros plan, plongée, contre-plongée ...

Le temps de pose peut affecter le résultat d'une photographie. Par exemple, pour faire une photographie floue on peut augmenter le temps de pose.

L'optique, c'est-à-dire les objectifs utilisés, influe sur le rendu de la photographie. Par exemple l'utilisation d'un grand angle donne des images photographiques que l'œil humain ne peut percevoir en globalité dans la réalité.

Enfin, les caractéristiques de la lumière utilisée lors de la prise de vue sont des éléments qui peuvent être à l'origine d'un certain effet dans la photographie. Une lumière peut se définir par sa source (lumière naturelle, lumière artificielle), sa direction (contre-jour, éclairage naturel), son intensité (vif, diffus), sa qualité (contrasté, surexposition, ...).

L'analyse morphologique d'une image peut être assez rigoureuse si les descripteurs qui qualifient les caractéristiques sont définis clairement. L'analyse du contenu informatif ne peut être aussi exhaustif.

I.3.2.2 Le contenu informatif

Décrire le contenu informatif d'une image n'est pas une démarche naturelle. Il faut transformer ou plutôt "résumer" une image en mots. Effectuer cette opération surtout réductrice nécessite de se confronter à plusieurs problèmes :

- Une connaissance sur le sujet de la photographie est nécessaire. Une légende accompagnant la photographie est alors indispensable, elle doit apporter des renseignements sur la date, le lieu, le nom des personnes ou objets représentés. La légende décrit le contexte de l'image.

- Le choix des termes décrivant une image peut être facilité par l'utilisation d'un thesaurus, mais encore faut-il que celui-ci soit assez exhaustif et qu'il possède des termes suffisamment concrets pour permettre une description précise.

- Comment ne pas aller trop loin dans la description. On ne peut retenir tous les détails d'une image, il faut faire des choix. Plusieurs critères de choix sont alors possibles :

- . Un premier critère peut reposer sur la lisibilité dans l'image du détail à décrire.
- . Un second critère peut consister à se demander si le détail a son importance relative à la connaissance du type de la photothèque et de ses utilisateurs. Un détail

d'architecture peut avoir de l'importance dans une photothèque d'art et pas dans une photothèque d'actualité.

Un dernier critère nécessite de se demander si l'importance du critère n'est pas diminuée par le contexte des utilisations futures de l'image. Imaginons un personnage célèbre assis sur un fauteuil de style Louis XV. La photographie a peu de chance, dans une photothèque d'actualité, d'être utilisée pour le style du fauteuil, son principal intérêt réside dans la présence du personnage célèbre.

- Comment être sûr de ne rien oublier. La personne qui décrit les images doit s'imaginer quels seront les futurs utilisateurs. Elle doit définir toute la portée culturelle de l'image.

Afin de décrire le contenu informatif d'une photographie, il est indispensable d'avoir une méthode rigoureuse. L'exploitation de cette rigueur permet à l'interrogation d'être plus efficace. Comme dans le système descriptif des représentations, la personne qui décrit doit se poser un certain nombre de questions pour obtenir une description assez exhaustive de l'image. Ces questions peuvent être les suivantes : Où ? Quand ? Qui ? Quoi ? Comment ?

La question Où? permet de situer l'image dans son contexte géographique. L'analyste indiquera alors un nom de lieu (ville, pays, rue ...) s'il présente un intérêt pour la photographie.

La situation du document dans l'espace temps est déterminée en répondant à la question Quand? La date de la prise de vue, l'époque des objets photographiés sont les informations qui peuvent être données en réponse à cette question.

La question Qui? a pour but d'identifier les personnages présents dans la photographie en indiquant leur nom ou le rôle qu'ils ont dans la photographie (artisan, pêcheur, chanteur ...).

L'identification des objets, œuvres d'art ou monuments présents dans la photographie, ainsi que la description du décors dans lequel se trouvent ces objets, ont lieu en répondant à la question Quoi?

La question Comment? permet de définir l'action présente dans la photographie (manifestation, construction, inondation ...) ainsi que le contexte de cette action.

En répondant à ces cinq questions, le contenu informatif d'une image est décrit dans toutes ses dimensions. La précision de la description est relative au vocabulaire utilisé, au modèle de document choisi et à la personne qui décrit. On peut remarquer aussi que la personne qui décrit

sera constamment confrontée à la difficulté de trouver le ou les mot-clés qui expriment le mieux le concept présent dans la photographie.

Par exemple, pour une photographie d'un tableau du 19^{ième} siècle représentant la prise de la Bastille, l'analyste aura envie d'indiquer au minimum ces informations: tableau, prise de la Bastille, 19^{ième}. Ici ce posent quelques problèmes : prise de la Bastille est un événement, 19^{ième} exprime un intervalle de temps ; ces deux notions n'ont pas alors le statut de mots-clés dans un vocabulaire contrôlé et ne peuvent être présents dans un champ de type liste de mots-clés. La prise en compte de ces notions avec le modèle d'indexation par mots-clés nécessite d'avoir un modèle de document ayant des rubriques suffisamment précises pour englober ces notions. Dans notre cas, le document d'indexation devra posséder au moins les trois rubriques suivantes : "contenu de la photographie" à valeur dans un vocabulaire contrôlé (un thesaurus), "datation de l'objet photographié" de type date, "précision sur l'objet photographié" à valeur dans une liste de descripteurs libres contenant des noms propres. Avec ce modèle de document notre photographie pourrait avoir comme description :

contenu de la photographie : tableau, révolution ;
datation de l'objet photographié : 19^{ième} ;
précision sur l'objet photographié : prise de la Bastille.

La dernière caractéristique d'une image que l'on peut prendre en compte est son côté subjectif ce qu'on appelle la connotation de l'image.

I.3.2.3 La connotation

Décrire le côté subjectif d'une image est sûrement la tâche de l'analyste la plus difficile à accomplir. L'impression subjective, que peut lui donner une image, ne sera pas forcément celle qu'aura un individu quelconque, ou même celle qu'a voulu communiquer l'auteur de la photographie. Une image ayant un fort caractère subjectif, tel que la tristesse ou la joie, ne peut pas toujours se retrouver grâce à la description de son contenu informatif. Il existe des milliers et des milliers de représentations possibles de la tristesse ou de la joie. Une description du côté subjectif de ces images peut aider ce type de recherche basée sur le connoté des photographies. Cette description ne pas doit être trop rigoureuse ; elle doit permettre d'offrir à l'utilisateur, en résultat d'une recherche, une étendue assez importante des représentations concrètes d'une notion subjective, ceci ayant comme objectif d'aider l'utilisateur dans sa recherche en lui donnant des idées de représentations possibles.

Le choix des termes adéquats à l'aspect subjectif d'une photographie n'est pas aisé. De nombreux termes abstraits existent avec des nuances parfois très subtiles (harmonieux, beau, esthétique).

tique, ..) . La définition du vocabulaire est alors importante et l'avis de plusieurs personnes est souhaitable.

G. Bléry [Blé 81] a effectué une étude sur un échantillon de personnes à qui elle a demandé de formuler, à l'aide de mots abstraits, les impressions que leur suggéraient un certain nombre d'images. Cette expérience révèle que, sur 275 termes employés, seulement 25 termes n'ont pas pu être intégrés aux notions subjectives, regroupées par couples d'opposition, qu'elle propose :

(abstrait, sensuel); (actif, passif); (ancien, moderne); (apaisant, stimulant); (artificiel, naturel); (beau, laid); (sérieux, frivole); (chaud, frais); (coloré, terne); (gai, triste); (comique, tragique); (décontracté, angoissé); (érotique, froid); (ordonné, discordant).

Ce résultat montre que la diversité du vocabulaire, exprimant des sentiments ou des impressions, peut être résumée par un sous-ensemble de termes recouvrant, à quelques petites nuances près, les mêmes notions abstraites.

I.3.3 Conclusion

A travers cette étude sur l'analyse de l'image, il apparaît important d'avoir une méthode qui permette de ne rien oublier et d'obtenir une certaine homogénéité dans les descriptions. Chaque caractéristique de l'image est importante, mais chacune doit être décrite plus ou moins précisément, en fonction du type de la photothèque et de ses utilisateurs. Ce niveau de précision peut être déterminé grâce à l'utilisation d'un thésaurus, en choisissant les termes les plus précis possibles. L'utilisation du thésaurus apporte une cohérence sémantique entre les différentes descriptions, en réduisant le nombre de cas de synonymie et de polysémie, sources de bruit et de silence dans les recherches. Le thésaurus est donc un outil indispensable à la construction de descriptions d'images.

Voyons maintenant quelles sont les conséquences des propriétés de l'image dans un processus de recherche.

I.4 La recherche en elle-même

I.4.1 Historique

Le premier classement qui a été beaucoup utilisé, qui l'est encore dans certaines photothèques ayant peu de moyens, est le système des boîtes à chaussures, où les photographies sont classées par thèmes ou par date. L'arrivée de l'informatique a changé totalement la conception de la recherche d'images. De nouveaux systèmes ont permis d'accéder à des descriptions d'images

par des recherches multi-critères. L'analyse de l'image est alors devenue plus précise et plus conséquente. Mais à ce stade, la recherche d'images se différencie peu d'une recherche de documents classiques (livre, article ...), car l'accès à l'image ne s'effectue que tout à la fin du processus de recherche.

L'arrivée des supports optiques a apporté un grand potentiel supplémentaire à la recherche d'images en offrant les moyens d'accéder directement, en résultat d'une recherche, aux images elles mêmes et non plus seulement à leur description. De plus, l'image est un document qui possède la particularité d'être manipulé et "lu" facilement grâce à sa visualisation directe (sur un écran) et globale.

I.4.2 Les supports optiques

La lecture optique par laser a donné le jour à des nouveaux types de supports de l'information: les supports optiques [Pel 88] [Pla 88]. Ces nouveaux supports possèdent des qualités qui dépassent celles des supports classiques de type magnétique :

- inaltérabilité (lecture sans contact mécanique),
- très grande capacité de stockage,
- accès rapide à l'information,
- restitution très précise de l'information.

Si l'on en croit le succès du Compact Disc, les mémoires optiques sont vouées à un avenir brillant. Il existe à ce jour de nombreux types de supports optiques, qui se différencient essentiellement par leur diamètre, leur capacité, le type de l'information qu'ils contiennent et leur mode de stockage (numérique, analogique).

Ainsi, avant l'invention du Compact Disc, appelé aussi CD-DA (Compact Disc Digital Audio), le vidéodisque existait déjà. En effet, dès la fin des années 1960, des chercheurs tentaient de trouver un support pouvant stocker des images. C'est vers les années 1975 que sont apparus les premiers vidéolecteurs et vidéodisques.

Les années 80 ont permis au standard LaserVision de Philips, Sony et MCA, de s'imposer en Europe et aux Etats-Unis.

Les vidéodisques LaserVision existent en deux formats : 30 et 20 cm de diamètre. Les images sont codées en modulation de fréquence et stockées de manière analogique. Le pressage de l'information sur un disque plastique rendu réfléchissant par le dépôt d'une pellicule de métal a lieu en usine. Les vidéodisques existent en deux modes d'exploitation, l'un permettant l'arrêt sur l'image (CAV, Constante Angular Velocity, vidéodisque interactif) qui peut contenir jusqu'à 54000 images par face, ce qui correspond à 36 minutes de film vidéo (25 images par secondes),

l'autre n'autorisant pas l'arrêt sur l'image (CLV, Constante Linear Velocity) pouvant contenir jusqu'à 60 minutes de vidéo par face.

Le vidéodisque s'est très vite imposé comme le support adapté pour les bases d'images [Ima 88] car il offre une capacité de stockage importante et un accès à l'image très rapide. L'accès à l'image s'effectue à l'aide d'un numéro en quelques secondes, soit à partir d'une télécommande, soit par l'intermédiaire d'un ordinateur.

Depuis, de nombreux autres supports sont venus compléter la gamme des supports optiques, et principalement :

- Le CD-ROM ; il possède le même diamètre que le CD-DA (12 cm), les informations qu'ils contiennent sont sous une forme alphanumérique et peuvent être exploitées par un ordinateur. Sa capacité varie entre 540 et 600 méga-Octets¹. Son utilisation est orientée pour le moment vers le stockage de banques de données encyclopédiques et bientôt celui des banques d'images numérisées.

- Le CD-V (Compact Disc Video) est un support orienté grand public, il reprend le standard LaserVision pour l'image, le son associé à l'image est numérisé. Ce support est utilisé pour l'édition de films ou de clips vidéo.

- Le DON WORM (Disque Optique Numérique) permet d'enregistrer une seule fois des informations numériques (images, son, textes) par l'utilisateur lui-même (Write Once Read Many). Il existe de nombreux formats de DON, car ce n'est pas encore un support standardisé. La capacité peut être donnée sur un exemple : un DON de 30 cm peut stocker jusqu'à un giga-octets. Le DON a un avenir assuré si un standard est défini rapidement.

- Le LV ROM est un vidéodisque reprenant le standard LaserVision pour l'image avec, à la place de la bande son stéréophonique, la possibilité de stocker de l'information numérique. On peut stocker jusqu'à 6 kilo octets par image, ce qui permet d'attacher, à chaque image, les informations qui lui sont propres. Ce support est un des plus intéressants pour les banques d'images quand des informations définitives peuvent être associées aux images (titre, auteur, date, légende...).

Les recherches actuelles s'orientent vers la conception de supports optiques enregistrables et effaçables. Ces qualités sont nécessaires à l'intégration de ces supports dans les systèmes informatiques avec d'autres fonctions que celles de l'archivage et de la consultation.

La recherche d'images prend toute son ampleur lorsque le résultat d'une recherche n'est plus une liste de descriptions mais un ensemble d'images que l'on peut visualiser. Les vidéodisques,

¹ Une image couleur ayant une définition de 256x256 points occupe environ 100 Ko.

puis le CD-ROM et le DON ont ouvert cette voie, ils ont permis d'intégrer dans un processus de recherche la **visualisation** d'images.

I.4.3 La visualisation

Les systèmes de recherche d'images de l'époque pré-vidéodisque proposaient, en résultat d'une recherche, un ensemble de descriptions que l'utilisateur devait lire pour se représenter mentalement les images afin de ne retenir que les plus pertinentes. Cette tâche est peu naturelle et demande de l'imagination. L'arrivée des supports optiques a offert aux systèmes de recherche d'images la possibilité de manipuler l'image en tant qu'information à part entière. La visualisation est devenue l'étape fondamentale du processus de recherche d'images.

Les fonctionnalités des systèmes de recherche d'images ne sont plus les mêmes. La visualisation d'image en résultat d'une recherche a modifié à la fois les descriptions des images mais aussi les stratégies de recherche, et a donné le jour à de nouveaux postes de travail appelés "mosaïques d'images" permettant de manipuler des images.

La lecture d'une image est très rapide (environ 1 à 2 secondes pour une lecture complète), beaucoup plus que celle d'une description. De plus la visualisation successive ou simultanée d'images est évocatrice et amène à la synthèse, ce qui est faux pour un ensemble de descriptions dont la visualisation amènerait plutôt à un phénomène de lassitude. En conséquence, le nombre de documents en réponse à une requête peut être nettement plus important que dans les systèmes traditionnels. Le bruit dans le résultat d'une recherche est acceptable et même souhaitable, compte tenu de la pauvreté des descriptions par rapport à la richesse des images.

Les descriptions d'images ne jouent plus le rôle de descriptif permettant de se représenter mentalement les images. Elles ont comme nouvelle tâche d'une part d'apporter les informations qui ne se trouvent pas sur les images et d'autre part de permettre aux images d'être sélectionnées dans un processus de recherche automatisé. Cela ne veut pas dire que les descriptions doivent être moins précises, car la précision dans la recherche demeure nécessaire, mais elle est moins importante que dans les systèmes traditionnels compte tenu du nombre de documents images que peut visualiser une personne en 1 minute (environ 30 à 60 images).

La lecture presque instantanée de l'image est ralentie par une visualisation séquentielle des images. La personne qui recherche des images a besoin de comparer les images entre elles, de les trier, d'effectuer des choix. C'est ainsi que le concept de mosaïque d'images introduit par H. Hudrisier [Hud 83][Hud 84] a donné le jour à de nouveaux systèmes ayant pour but de rempla-

cer la traditionnelle table lumineuse qui permettait de trier et de comparer un ensemble de diapositives. Une mosaïque d'images est essentiellement un écran contenant plusieurs cases (9, 12 ou 16), chacune pouvant afficher une image. Des fonctions de manipulation d'images permettent à l'utilisateur de trier et de classer ces images en construisant des paquets d'images. Nous aurons l'occasion dans la partie suivante de décrire plus précisément ces fonctionnalités car nous avons utilisé, dans notre première réalisation d'un prototype, l'imageur documentaire de la SEP (Société Européenne de Propulsion) qui fut la première mosaïque d'images réalisée utilisant les concepts de Hudrisier.

I.4.4 Conclusion

La recherche d'images jusqu'à l'apparition des supports optiques se déroulait comme une recherche textuelle. Une grande différence existait lors de l'analyse des résultats : dans une recherche textuelle, l'utilisateur devait lire un résumé ou le document en entier pour évaluer la pertinence du document ; par contre, dans une recherche d'images, l'utilisateur devait s'imaginer mentalement les images à partir de leur description pour vérifier leur pertinence, le choix sur les images réelles n'étant possible qu'avec un petit échantillon.

L'arrivée des supports optiques, et surtout celle du vidéodisque, a révolutionné cette dernière recherche en offrant la possibilité de vérifier rapidement la pertinence d'une image en la visualisant.

Nous avons évoqué les apports de la visualisation d'images dans un processus de recherche, étudions maintenant quels sont les besoins des utilisateurs recherchant des images ainsi que quelques exemples de systèmes proposant une recherche basée sur l'image.

II La recherche d'images et ses applications

Ce paragraphe propose tout d'abord une étude des utilisateurs d'une base d'images puis fait le point sur les systèmes permettant de manipuler ces images.

II.1 Les utilisateurs et les utilisations

Cette étude sur les utilisateurs et les utilisations d'une base d'images est celle de Muriel Cluzeau¹ [Clu 88] qui fait partie de l'équipe EXPRIM. Après avoir décrit le profil des demandeurs d'images, elle propose une certaine classification de leurs demandes.

¹Maître de conférence à l'Université de Dijon, IUT, département Carrières de l'information.

II.1.1 Profil des demandeurs d'images

Muriel Cluzeau met en évidence deux sortes de demandeurs d'images :

- les consommateurs,
- les utilisateurs.

Le consommateur est un non spécialiste, il ne connaît pas le domaine des images de la base qu'il consulte. Seule la curiosité l'entraîne dans la quête d'images. Ce type de demandeur a besoin d'être guidé par le système d'une manière assez souple afin qu'il puisse conserver toute sa liberté.

L'utilisateur d'une base d'images a un profil différent ; c'est un professionnel, il a de nombreuses connaissances sur les images et un certain savoir-faire. L'aide attendue par l'utilisateur est celle d'un spécialiste de la base qu'il veut consulter, qui l'assistera de manière efficace dans ses recherches. L'utilisateur est aussi un intermédiaire, il recherche des images pour un type d'utilisation qu'il connaît au départ. C'est ce type d'utilisation qu'il devra communiquer au spécialiste de la base et qui guidera leurs recherches.

Le type d'utilisation de l'image va guider l'utilisateur et le spécialiste dans leur recherche en fonction de deux critères déterminants :

- le medium, c'est-à-dire le moyen de communication qui va être utilisé pour diffuser l'image,
- le public visé.

La connaissance du médium permet de déterminer les qualités techniques que doit avoir l'image recherchée :

- journal : qualité moyenne, médiocre acceptée,
- livre de luxe : très bonne qualité, bon cadrage,
- affiche : bonne qualité, possibilités d'agrandissement, effets spéciaux éventuels.

Ces différentes formes de qualités doivent apparaître dans les descriptions des images, afin qu'une image puisse être retrouvée en fonction de sa qualité.

La connaissance du public visé peut permettre de faire converger des recherches ; par exemple une typologie des différentes publications et des sujets qu'elles traitent, peut suggérer l'ajout de nouveaux critères de recherche :

- l'Equipe : sport,
- l'Événement du jeudi : information, culture,
- Science & Vie : science, vulgarisation.

La connaissance du type d'utilisation permet une certaine présélection dans les images de la base, mais encore faut-il pouvoir satisfaire la demande dans sa totalité, on peut alors s'intéresser aux différentes caractéristiques des demandes rencontrées.

II.1.2 Les types de demandes

Les caractéristiques des demandes ont permis à Muriel Cluzeau de les classer en quatre types :

- les demandes de consultation,
- les demande précises,
- les demandes thématiques,
- les demandes connotatives.

Les demandes de consultation correspondent à des recherches de contact visuel le plus immédiat possible, suivi d'une demande d'information de type culturel. Une liberté totale doit être fournie à cette catégorie de demande, associée à un accès rapide à l'information annexée aux images.

Les demandes précises sont clairement définies par des critères qui ne peuvent pas être changés : le nom d'une personne, un lieu, un événement, une date. Les réponses à ces demandes doivent être rapides et en cas d'échec un dialogue pertinent avec l'utilisateur doit être instauré. Des sujets de recherche doivent être proposés, ainsi l'utilisateur en voyant d'autres images peut découvrir des images de remplacement. La visualisation des images a ici aussi un rôle fondamental.

Les demandes thématiques servent à rechercher des images pour illustrer un thème assez général ; le besoin en image n'est pas encore clairement défini. Le but de ce type de demande est d'obtenir un grand nombre d'images afin d'avoir de grande possibilités de choix. Ces choix vont permettre de définir plus clairement le besoin, suivant tous ses paramètres : espace, temps, sujets, objets, contexte. Ce genre de demandes est en général mal formulé au départ, il comprend beaucoup de "non dit" ; il demeure cependant un des plus fréquents.

Les demandes connotatives ne font pas référence au sujet de l'image mais à l'impression qui s'en dégage. Ce type de demande est difficile à satisfaire car répondre à ce besoin correspond à prévoir le sens qu'aura l'image quand elle sera "lue" par l'utilisateur final, c'est-à-dire par le lecteur d'une revue, le passant dans la rue regardant une affiche... Un dialogue et une importante visualisation d'images sont nécessaires pour déterminer quels sont les caractères

concrets qui doivent intervenir dans l'image recherchée. L'introduction de termes connotatifs dans les descriptions d'images peut permettre de retrouver un premier ensemble d'images qui servira à amorcer ce type de dialogue.

Voyons quels sont les processus de recherche que proposent quelques systèmes de recherche d'images existants.

II.2 Les systèmes existants

Nous ne décrivons pas dans ce paragraphe les systèmes documentaires proposant un module de pilotage de vidéodisque – l'ensemble de ces systèmes étant répertoriés dans un guide de La documentation française [Sin 87] –. Les systèmes que nous allons décrire sont ceux qui apportent une originalité dans la prise en compte de l'image en tant qu'information recherchée.

Parmi les systèmes existants on peut distinguer deux catégories :

- ceux qui ont été construits pour une application particulière,
- ceux ayant été conçus pour la recherche d'images en général.

II.2.1 Les systèmes liés à une application

Les systèmes ayant été conçus pour une application particulière proviennent d'un besoin de mettre à la disposition du grand public un ou plusieurs vidéodisques existants ; les trois applications suivantes ont eu ce besoin :

- le projet VIDERALP ou l'exploitation du vidéodisque d'images régionales en Rhône-Alpes,
- le vidéodisque de la Bibliothèque Sainte-Geneviève,
- les vidéodisques du Louvre 1, 2 et 3.

Le projet VIDERALP [Dur 87] est un bon exemple d'utilisation de logiciel documentaire muni d'un module de pilotage de vidéodisque pour la gestion d'une base d'images. Le logiciel documentaire utilisé est le logiciel TEXTO bien connu des documentalistes. L'utilisateur recherchant une image est guidé par différents menus lui permettant de formuler sa demande. Sa demande ainsi formulée, le système lui propose alors de visualiser une à une les images sélectionnées. L'originalité du travail réside dans l'analyse de l'image qui est assez poussée, elle contient plusieurs éléments :

- éléments localisateurs : n° de référence, lieu de conservation, n° sur le vidéodisque,
- éléments descriptifs :
 - la légende du document,

- un commentaire fait par l'analyste,
- le nom de l'auteur,
- le format,
- la technique (photo, dessin, gravure ...),
- le type du document (affiche, tract, carte postale,...),
- l'éditeur,
- la date du document,
- éléments d'analyse :
 - la date de l'événement ou de l'objet photographié,
 - le lieu,
 - le sujet : décrit à l'aide de 900 mots.

Ce travail d'analyse est assez conséquent et très coûteux en temps.

Le vidéodisque de la Bibliothèque Sainte Geneviève a été créé à partir de photographies de miniatures, afin de communiquer ce patrimoine à un public plus varié. Le vidéodisque contient l'ensemble des photographies mais aussi des planches de 9 photographies qui sont utilisées par le système comme des menus visuels permettant aux utilisateurs de parcourir la base en choisissant l'image la plus intéressante dans la planche proposée. Cette génération de parcours a été réalisée grâce à un fichier analytique des photographies qui existait bien avant le passage du vidéodisque. Cette application est originale par la forme des parcours qu'elle offre aux utilisateurs, l'inconvénient est que les origines de ces parcours, les planches d'images, sont figées sur le vidéodisque. Une amélioration envisagée est de construire ces planches d'une manière dynamique selon une analyse effectuée sur un ensemble d'images pouvant être le résultat d'une recherche d'un utilisateur. Une autre perspective originale de ce système est de proposer aux utilisateurs le moyen de construire eux-mêmes leur propre indexation sur les ensembles d'images qu'ils obtiennent en résultat d'une recherche. Cette indexation personnelle est très utile pour les utilisateurs assidus du système qui pourront, sur leur base ainsi constituée, formuler des requêtes avec leur propre vocabulaire.

L'application le Louvre 1, 2, et 3 [Hat 88] qui est commercialisée par la société Laser Communication Plus propose au grand public 3 vidéodisques sur les collections du musée du Louvre :

- Louvre 1 : peintures et dessins,
- Louvre 2 : sculptures et objets d'art,
- Louvre 3 : antiquités.

Ces trois vidéodisques représentent les premiers vidéodisques d'une collection plus importante, devant être éditée prochainement, sur une partie du patrimoine artistique et architectural de la France.

Associé à ces vidéodisques est fourni un logiciel de consultation et de recherche d'images écrit en Hypercard [Goo 88] sur Macintosh. Ce logiciel permet de parcourir les vidéodisques d'une manière très agréable grâce à la forme du dialogue offert par Hypercard. De nombreuses fonctions sont offertes à l'utilisateur :

- recherche par auteur,
- recherche par école : française, italienne, flamande...
- recherche par date,
- recherche par style : gothique, renaissance, classique,
- recherche libre : l'utilisateur entre ses critères (auteur, école, date, style),
- étude d'une œuvre en détail : une vue générale de l'œuvre s'affiche à l'écran et il suffit de désigner un détail pour qu'il s'affiche sur l'écran vidéo. Cette forme de consultation ne peut s'improviser, elle doit être prévue avant le passage des vidéodisques.
- déroulement d'une séquence vidéo associée à une œuvre, si cette séquence existe.
- construction d'une base personnelle d'images commentées.

Ce logiciel qui demeure très lié à l'utilisation des vidéodisques du Louvre semble proposer une approche intéressante de parcours d'une base d'images.

II.2.2 Les systèmes construits pour la recherche d'images

Deux systèmes vont être présentés dans ce paragraphe, chacun de ces deux systèmes apporte une originalité dans l'utilisation de l'image en tant que document recherché.

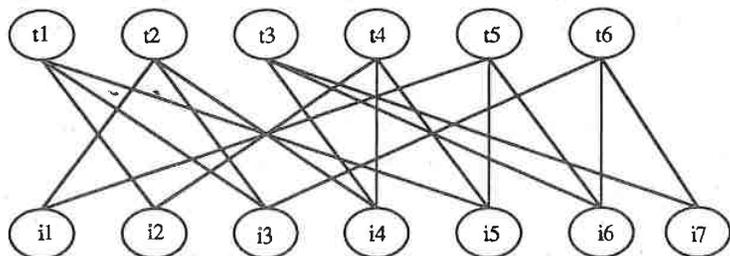
Le premier système, SIGMINI, qui a été développé par l'école des Mines de Paris [Mor 88][God 88] permet de gérer des descriptions de documents à structure hiérarchique. Ce système offre les moyens de définir des descriptions hiérarchiques et d'utiliser cette structure à l'interrogation grâce à l'utilisation d'opérateurs spécifiques. Des nouveaux opérateurs ont été ajoutés au modèle booléen classique, comme par exemple :

- C1 **PE** C2 : indiquant que le critère C1 doit être le père du critère C2 dans le document,
- C1 **FR** C2 : indiquant que le critère C2 doit être un des frères du critère C1.

Ce système, lorsqu'il est utilisé pour rechercher des images, permet une description des images plus précise. En effet, l'image peut être définie comme un ensemble d'éléments, chacun de ces éléments pouvant être décrit plus précisément et ceci de manière récursive. Ce système est utilisé par des archéologues, il leur permet de décrire des mosaïques grecques ayant comme caractéristique d'être constituées d'éléments emboîtés dans d'autres éléments. Cette approche

offre la prise en compte d'une typologie de l'image qu'est l'inclusion d'éléments, ceci constitue une première approche dans la recherche de modèle de documents dédié à l'image. Il n'en reste pas moins que la forme de cette interrogation n'est utilisable que par des spécialistes connaissant la base d'images et le système.

Le deuxième système, Brain, a été développé à l'IRISA de Rennes [Coc 88], son originalité se situe dans le modèle général utilisé pour mettre en correspondance une requête et un document. Ce modèle est basé sur une structure dite "neuronale" où chaque neurone est soit un document, soit un terme décrivant un ou plusieurs documents. L'ensemble du modèle peut être assimilé à un réseau de nœuds, un nœud étant équivalent à un neurone. Un lien relie un nœud "terme" à un nœud "image" lorsque le terme est un descripteur de la description de l'image. Un poids est attaché au lien, il représente l'importance dans l'image du concept représenté par le terme. Un réseau de la forme suivante est obtenu :



où les t_j représentent les nœuds "terme" et les i_j les nœuds "image".

Une requête est un ensemble de termes ayant chacun un poids. La recherche des images répondant à la requête se fait grâce à une propagation des poids dans tout le réseau en parcourant les liens des termes vers les images puis des images vers les termes (ce qui peut rendre pertinents de nouveaux termes) et ainsi de suite un certain nombre de fois ; cette propagation prend en compte les poids des termes de la requête, les poids des liens et affecte les nœuds images d'un poids ; les images ayant un poids important sont dites pertinentes pour la requête. Le résultat de la recherche est une liste d'images ordonnées en fonction de leur poids dans le réseau.

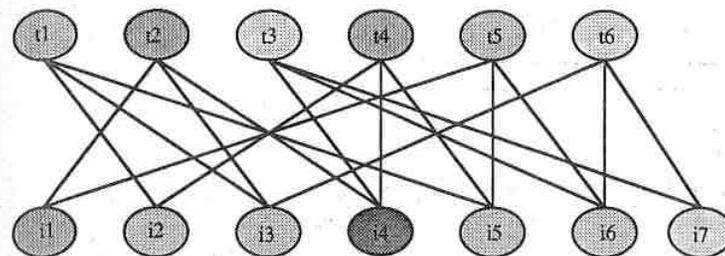
Exemple :

Soit la requête suivante :

- R = 1. t2,
- 2. t4,

les termes sont classés par ordre d'importance.

Le résultat de cette requête peut être matérialisé sur le réseau ci-dessus de la manière suivante :



après trois propagations (terme -> image, image -> terme, terme -> image), les nœuds les plus foncés représentent les nœuds ayant un poids fort. La liste des images les plus pertinentes sera :

1. i4,
2. i1,
3. i2, i3, i4,
4. i6,
5. i7.

Ce système a été conçu initialement pour la recherche d'images car les concepteurs pensent que peu de mots sont nécessaires pour décrire une image ; ce qui n'est pas vraiment le cas comme nous l'avons montré au § I.3, mise à part peut être pour les agences de presse qui indexent leurs images avec peu de termes compte tenu de la masse importante d'images qu'elles ont à gérer et du peu de temps qu'elles ont à y consacrer. Par contre, nous pensons que ce modèle de mise en correspondance peut être intéressant pour les systèmes de recherche textuelle où l'indexation est automatique. Toutefois cette démarche demeure originale puisqu'elle offre au système une connaissance implicite des relations sémantiques existant entre les termes à travers leur utilisation dans des descriptions d'images. Ces travaux ainsi que ceux de l'équipe EXPRIM qui ont inspirés cette approche, semblent se compléter et une collaboration entre les deux équipes est en cours d'étude.

III Conclusion

Ce premier chapitre a été consacré à l'étude de l'image en tant que support d'informations, et en tant que résultat d'une recherche rétrospective. Cette étude a mis en évidence que la prise en compte de l'image dans un système de recherche d'informations nécessite une analyse rigoureuse de l'image, car une image peut être recherchée pour de nombreuses raisons. Des méthodes d'analyse existent, elles ont pour but de mettre à plat toutes les informations qui sont liées à l'image. L'arrivée des supports optiques a offert à la recherche d'images sa raison d'être

car enfin le résultat d'une recherche pouvait être manipulé et surtout visualisé directement. Cet événement a permis aux systèmes de recherche d'images de proposer un dialogue plus interactif.

Le dialogue dans une recherche d'images est l'élément le plus important ; il doit être très interactif, et il doit permettre à l'utilisateur de voir beaucoup d'images pour lui donner les moyens de projeter rapidement son besoin sur des éléments concrets apparaissant dans les images qu'il visualise. L'étude des utilisateurs et des utilisations d'une base d'images a confirmé ce point.

Les systèmes que nous avons présentés dans le dernier paragraphe offrent chacun une forme de recherche d'images différente. Mais leur dialogue s'arrête à la visualisation des images, c'est l'utilisateur qui a la charge d'adapter sa requête en fonction des images qu'il a visualisées. Ces systèmes n'exploitent pas une des caractéristiques les plus importantes de l'image qui est sa lecture presque instantanée. Cette caractéristique peut permettre à l'utilisateur d'effectuer rapidement des choix lors de la phase de visualisation en choisissant ou en rejetant des images parmi celles proposées par le système. L'exploitation de ce choix doit sans aucun doute permettre au dialogue d'être encore plus interactif. C'est ce type de dialogue que se propose de réaliser le projet EXPRIM.

Chapitre 2	
L'approche EXPRIM	101
I Introduction.....	101
II Historique.....	101
III Les idées principales du projet EXPRIM.....	102
III.1. Le processus EXPRIM	102
III.2 L'architecture proposée	104
III.3 Les connaissances et expertises utilisées.....	105
III.3.1 L'expertise du domaine.....	105
III.3.2 L'expertise de base.....	108
III.4 Conclusion	109
IV Les études et réalisations	109
IV.1 Première réalisation.....	110
IV.1.1 Architecture du système et description des différents constituants	110
IV.1.2 Les connaissances et données utilisées et leur représentation.....	110
IV.1.3 Les fonctions réalisées.....	111
IV.1.4 Le jeu d'essai	112
IV.1.5 L'évaluation du système.....	112
IV.1.6 Conclusion	113
IV.2 Deuxième réalisation	114
IV.2.1 Architecture du système et description des différents constituants	114
IV.2.2 Les connaissances et données utilisées et leur représentation.....	114
IV.2.3. Les fonctions réalisées.....	115
IV.2.4 Le jeu d'essai	118
IV.2.5 L'évaluation du système.....	118
IV.2.6 Conclusion	118
V Conclusion.....	119

Chapitre 2

L'approche EXPRIM

I Introduction

Le projet EXPRIM, qui est à l'origine de l'équipe du même nom au CRIN, s'est fixé comme but d'aider le mieux possible et interactivement un utilisateur, même non spécialiste ni d'informatique ni éventuellement du domaine concerné, à trouver des images dans une base de données d'images couplée à une base de données descriptives ; le besoin en images de l'utilisateur peut être plus ou moins précis et plus ou moins bien connu au départ.

Ce chapitre a pour objectif de décrire l'origine et l'évolution de ce projet à travers les idées qui ont été érnises et les réalisations qui ont été faites.

II Historique

L'origine du projet est un besoin en recherche d'images de l'agence de presse SYGMA. Cette agence manipule tous les jours des centaines d'images qu'elle distribue partout en France. C'est pour cette recherche que H. Hudrisier a conçu la mosaïque d'images appelée "Imageur documentaire" avec une participation du CRIN pour la définition des fonctions de manipulation ; sa commercialisation été prise en charge par la SEP (Société Européenne de Propulsion). Puis la SEP, en collaboration avec le CRIN, a été chargée de réaliser un système d'interrogation utilisant cette mosaïque d'images. L'équipe EXPRIM, qui a alors trouvé son appellation, a proposé une architecture de système basée sur une approche système expert et sur un processus de recherche en trois phases que nous décrivons dans la suite.

La collaboration du CRIN avec la SEP s'est poursuivie dans le cadre du programme Européen ESPRIT. En effet, un projet Esprit¹ ayant pour but la réalisation d'un système intelligent de stockage et de recherche d'images destiné au grand public a été élaboré en réunissant Philips, la BBC-Entreprise, Logica (une société anglaise d'informatique) et la SEP. Ce projet comportait en réalité deux projets ayant comme point commun l'utilisation du vidéodisque LVRUM de Philips.

¹ Projet Esprit n° 901.

Le premier projet portait sur le projet national anglais DOMESDAY² devant permettre d'accéder, via un parcours cartographique à plusieurs échelles, à des photographies de n'importe quel endroit de l'Angleterre. Philips, la BBC et Logica étaient les participants à ce projet.

La SEP et son consortium, le Bureau Marcel van Dijk de Bruxelles et le CRIN, avaient pour projet l'étude d'une faisabilité technique et économique d'un système intelligent de recherche d'images utilisant le LVRUM de Philips et les idées du projet EXPRIM.

III Les idées principales du projet EXPRIM

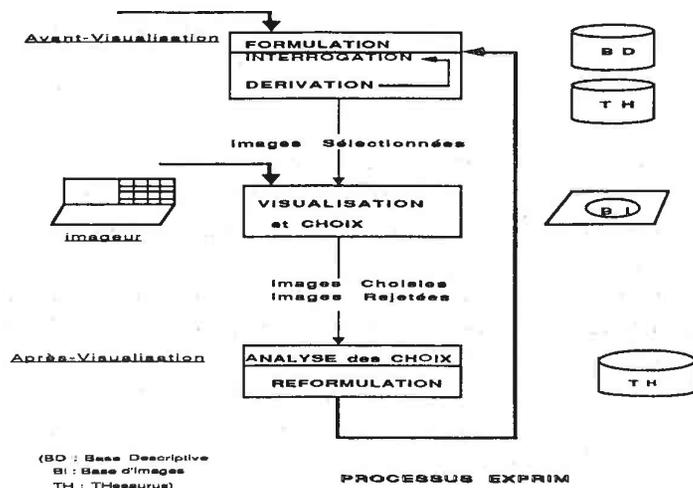
Le but de ce paragraphe est de montrer l'état du projet avant notre arrivée et avant le démarrage du projet Esprit. Cet état correspond à celui décrit avec plus de détails dans [Cre 85].

Le projet EXPRIM (système EXpert Pour la Recherche d'IMages), rappelons-le, porte sur un système qui a pour objet l'interrogation assistée d'une base d'images stockée sur vidéodisque, par l'intermédiaire d'une base de descriptions. Le système proposé a pour noyau un système expert qui procède au modelage progressif de la requête utilisateur, en s'appuyant sur différents types de connaissances et sur un processus de recherche en trois phases appelé le processus EXPRIM. Il est important de noter qu'à cette époque la jeunesse du projet nécessitait l'étude de nombreuses perpectives qui ont donné lieu à de nombreuses propositions, certaines ont été bonnes, d'autres moins, c'est ce que nous allons présenter maintenant.

III.1. Le processus EXPRIM

Le schéma de la figure 5 met en évidence les trois phases principales d'une étape de recherche dans le processus EXPRIM (qui itère sur cette étape) [Cre 85][Cre 86] :

² Domesday est le nom du cadastre établi par Guillaume le conquérant.



(BD : Base Descriptive
BI : Base d'images
TH : Thésaurus)

figure 5 : le processus EXPRESSIM

-la phase "Avant-Visualisation" qui est une phase d'interrogation souple (ou "approximative") de la base descriptive. "Souple" signifie qu'on cherche plus un voisinage qu'une identité entre requête et description d'image. Cette phase assure plusieurs fonctions : aide éventuelle à la formulation de requête sous forme d'une suite de mots, proposition de connecteurs logiques entre les mots, modification de certains descripteurs (inutiles ou trop sélectifs ...) en cheminant dans un thesaurus grâce à des règles, souvent heuristiques, etc...

-la phase "Visualisation" au cours de laquelle l'utilisateur observe les images issues de la phase précédente, en effectue des classifications et choisit celles qu'il désire retenir. Cette phase est tout à fait essentielle dans le processus. Dans de nombreux cas par exemple, la phase "Avant-Visualisation" initiale a pour but principal de fournir un ensemble important d'images à la phase de visualisation, et c'est dans celle-ci que l'utilisateur choisira effectivement les images qui l'intéressent et que lui-même (sur les images et les descriptions) et le système (sur les descriptions seulement) pourront acquérir par ce biais une meilleure connaissance des besoins réels.

-la phase "Après-visualisation" au cours de laquelle le système, et éventuellement aussi l'utilisateur, exploitent les résultats des phases précédentes pour mettre en évidence de nouveaux critères, grâce à certaines "régularités", ou certains traits marquants, découverts dans les choix. La suite de la phase "Après-visualisation" consiste à intégrer tous ces critères à l'ancienne re-

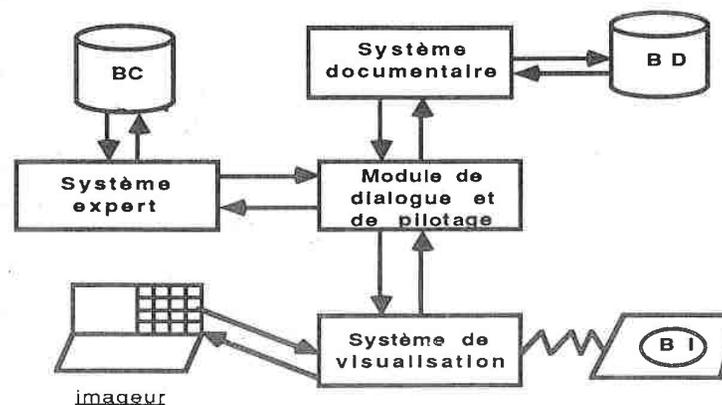
quête et à en déduire une nouvelle requête (reformulation), qui à son tour pourra jouer le rôle de requête initiale et être à l'origine d'une nouvelle étape.

Ce processus qui se veut très interactif a été proposé conjointement avec une architecture de système que nous allons présenter.

III.2 L'architecture proposée

La figure suivante met en évidence une architecture composée de quatre sous-systèmes :

- le système expert,
- le système documentaire,
- le système de visualisation,
- le système de dialogue et de pilotage.



Le système expert intervient dans les trois phases du processus. Il a pour tâche de manipuler les connaissances du système afin de contrôler le déroulement d'une recherche. Les connaissances utilisées par ce système seront détaillées dans le prochain paragraphe.

Le système documentaire assure la gestion et en particulier l'interrogation de la base de données textuelles contenant les descriptions des images. Ce système peut être des plus classiques, c'est-à-dire ayant les caractéristiques des systèmes de recherche d'informations décrits au chapitre 1 de la partie I (indexations binaire, formulation booléenne).

Le système de visualisation permet l'accès aux images qui sont stockées sur un support optique. Ce système peut être composé d'un simple lecteur connecté à un moniteur n'offrant qu'une visualisation séquentielle des images. Le système de visualisation proposé dans cette architecture est l'imageur documentaire développé par la SEP qui constitue un outil très puissant de visualisation interactive d'images. Il permet à l'utilisateur de comparer, de trier et de constituer les ensembles d'images nécessaires au système pour le bon déroulement du processus.

Le système de dialogue et de pilotage a, comme son nom l'indique, un double rôle : il doit communiquer avec l'utilisateur, puis en fonction de ce dialogue, il doit piloter les différents sous-systèmes en ordonnant les tâches à effectuer. Ce système est central dans l'architecture, il doit être en relation très étroite avec le système expert.

Nous ne décrivons pas plus en détail la communication entre ces sous-systèmes, un exemple de communication est donné dans [Cre 85], nous en donnerons un autre dans la prochaine partie qui présentera le système qui a été réalisé pour le projet Esprit.

Voyons plutôt quelles sont les types de connaissances nécessaires au système.

III.3 Les connaissances et expertises utilisées

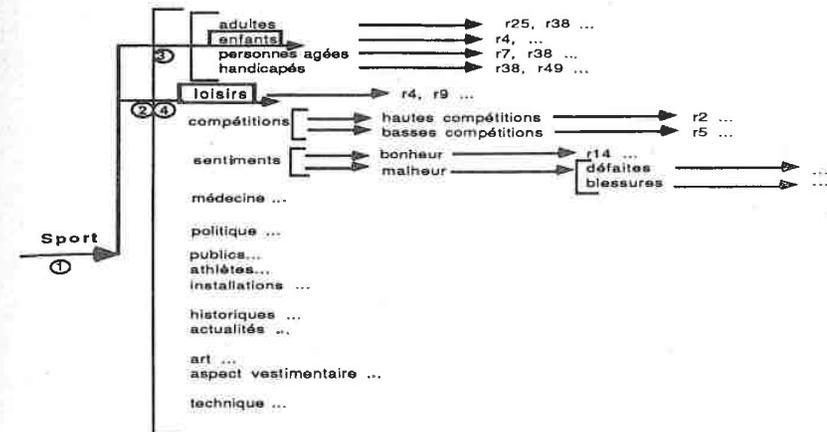
Le projet EXPRIM préconise pour la réalisation de la phase avant-visualisation la coopération de deux expertises :

- l'expertise du domaine,
- l'expertise de base.

III.3.1 L'expertise du domaine

L'expertise du domaine comprend tout ce qui est spécifique au domaine d'application, c'est-à-dire relatif au fonds iconographique. Ces connaissances regroupent à la fois le thesaurus qui a servi à décrire les images, mais aussi une certaine "expertise" du documentaliste qui a l'habitude de manipuler ce fonds. Dans [Cre 85], une structure logique basée sur un exemple tente de montrer comment une notion pouvant être subjective ou abstraite peut être interprétée et utilisée par un système.

Voici cet exemple :



Cette structure logique définit les différentes facettes possibles que peut revêtir la notion abstraite "sport". Ces facettes sont associées aux références d'images les illustrant (r4, r38 ...). Nous pouvons déjà remarquer que la construction de cette structure n'est pas aisée, car ici nous n'avons qu'une seule notion abstraite, et nous savons que la diversité des notions abstraites est grande et que ces notions peuvent se recouper. Ainsi, dans la notion de sport, il a été défini une facette médecine qui elle aussi est une notion abstraite, et qui doit avoir sans nul doute une facette sport etc... De plus cette structure doit cohabiter avec le thesaurus qui lui aussi demande un énorme investissement de construction. Nous émettons quelques réserves quant à la possibilité matérielle de construire une telle structure.

Mais voyons sur un exemple extrait de [Cre 85] comment cette structure devait être utilisée :

Soit la requête suivante :

"La revue "Modes et Travaux" recherche des photos d'enfants qui font de la course à pieds". Les mots-clés de la requête sont soulignés. Les chemins d'accès aux références sont indiqués dans le schéma logique couvrant la notion de "Sport".

Voici les raisonnements que le système est supposé effectuer :

- 1 : Le mot-clé "course à pied" est généralisé à "sport" grâce au thesaurus.
- 2 : Utilisation d'une autre connaissance dite "utilisateur" :
"Modes et Travaux" et "sport" ---> loisirs,
"enfant" et "sport" ---> loisirs.
- 3 : Accès direct grâce à la présence du terme "enfant" dans la requête.

4 : Renforcement de l'accès dû aux deux règles utilisées en 3 ayant la même conclusion.

Même si une telle structure existait dans le système, nous ne sommes pas tout à fait convaincus de son utilisation. Les différentes étapes du raisonnement associées à cet exemple, excepté l'étape 3, nous semblent quelque peu arbitraires, car trop spécifiques à l'exemple. Le point 1 s'appuie sur une utilisation du thesaurus tout à fait classique, cette généralisation permet d'accéder à la notion "sport". Mais un autre raisonnement utilisant des connaissances "utilisateur", semblables au point 2, peuvent conduire à des conclusions tout à fait différentes :

"Modes et Travaux" et "enfants" --> "mode enfantine"

"mode enfantine" généralisé à "mode".

Maintenant il se peut que dans le système il existe une structure logique associée à la notion "mode", et l'orientation de la recherche est alors tout à fait différente.

Ceci nous conduit au point 2, où l'utilisation de connaissances qui ne sont contenues ni dans le thesaurus ni dans les structures logiques, apportent de nouvelles informations. Nous nous posons alors quelques questions:

- Comment cette masse de connaissances exprimant une certaine culture générale a été obtenue ?

- Comment les règles exprimant cette connaissance s'enchaînent-elles entre elles ?

Et plus généralement nous nous demandons : comment l'utilisation du thesaurus, l'utilisation des connaissances dites "utilisateurs" et l'utilisation des structures logiques se coordonnent entre elles ?

Le thesaurus associé au fonds iconographique fait lui aussi partie de "l'expertise du domaine"; nous dirons plutôt que c'est une connaissance du domaine couvert par les photographies de la base. Dans la description du système présenté dans [Cre 85], le thesaurus est représenté par des règles de production de la forme :

sport d'équipe et violence --> rugby

Cette règle signifie que si l'on trouve à la fois "sport d'équipe" et "violence" dans la requête, il faut introduire le terme rugby.

Quelques remarques sur cette présentation sont nécessaires :

- ces formes de règles n'expriment pas réellement les informations présentes dans un thesaurus. Dans l'exemple on ne sait pas explicitement que le rugby est une sorte de sport d'équipe.

- ces formes de règles expriment une connaissance qui n'est pas présente dans un thesaurus et qui ne peut être déduite d'un thesaurus classique.

- la construction de telles règles demeure arbitraire et leur enchaînement reste à déterminer.

En conclusion, nous dirons que les seules connaissances du domaine réellement exploitables dans un tel système sont d'une part la base descriptive décrivant le fonds documentaire, d'autre part le thesaurus des termes utilisés dans les descriptions, représenté sous une autre forme que des règles de production. Ces deux connaissances peuvent avoir été construites avec des techniques tout à fait classiques (cf première partie, chapitre 2). L'utilisation d'autres connaissances semble difficile si la manière d'obtenir cette connaissance n'est pas clairement expliquée et si son exploitation dans le raisonnement du système ne semble pas locale à un exemple.

Nous avons vu dans la précédente partie que certaines connaissances d'un documentaliste, comme par exemple sa culture et sa sensibilité, ne peuvent être représentées compte tenu de l'étendue et du caractère de celles-ci, ces exemples le confirment.

Il reste un deuxième type de connaissances que le système doit manipuler, appelé "expertise de base".

III.3.2 L'expertise de base

Le rôle de cette expertise est de diriger la recherche d'images. Elle doit utiliser l'expertise du domaine pour coordonner le parcours du fonds documentaire et les différentes déformations de la requête.

Cette expertise est celle qui est la plus proche de l'expertise d'un documentaliste utilisant un système documentaire classique pour effectuer des recherches d'images. En effet, ce documentaliste va alors déformer sans cesse sa requête en fonction des réponses que lui fournit le système, et peut-être en s'aidant d'un thesaurus.

Si l'expertise du domaine est exprimée sous la forme donnée précédemment, l'expertise de base ne sera pas facile à déterminer car elle aura pour tâche d'ordonner l'enchaînement des différentes règles exprimant la connaissance du domaine.

C'est pourquoi nous pensons que l'expression des connaissances sur le domaine doit être statique, c'est-à-dire qu'elle ne doit pas contenir une certaine forme de raisonnement. L'utilisation de cette connaissance, qui peut être exprimée sous la forme de règles de production, doit être définie dans l'expertise de base.

L'expertise de base est alors constituée de traitements heuristiques exprimant des stratégies de recherche et de remodelage de la requête en s'appuyant sur la connaissance du domaine. Ces

heuristiques peuvent être définis avec l'aide d'un documentaliste et exprimés sous forme de règles de production.

Exemple :

Si la requête apporte peu d'images et le type de la recherche est large alors élargir la requête.

L'élargissement de la requête peut alors être effectué en utilisant le thesaurus (généralisation, utilisation des termes de sens proche) ou en modifiant la structure de la requête. L'information peu d'images sera elle aussi définie de manière heuristique à partir du nombre d'images sélectionnées par la requête et peut-être aussi à partir d'informations sur l'utilisateur.

III.4 Conclusion

Les principales caractéristiques du projet EXPRIM qui sans aucun doute définissent son originalité sont :

- Le processus en trois phases qui se veut très interactif grâce à un "feedback"¹ rendu puissant par la perception presque instantanée des images lors de la phase de Visualisation.
- Les deux types de connaissances manipulées par le système :
 - . la connaissance du domaine incluant le thesaurus et la base descriptive,
 - . la connaissance de base contenant les stratégies de recherche, de reformulation et de déformation de la requête.
- Son architecture basée sur l'utilisation d'un moteur d'inférence communiquant avec un système documentaire, un système de visualisation et un système de dialogue.

Cette approche de la recherche d'images que propose le projet EXPRIM a été en partie mise en œuvre dans quelques réalisations que nous allons maintenant décrire succinctement.

IV Les études et réalisations

Deux réalisations ont été développées dans le cadre de projets de DEA. Nous allons décrire chacune de ces réalisations en explicitant les points suivants :

- Architecture du système et description des différents constituants.
- Les connaissances et données utilisées et leur représentations : thesaurus, base descriptive, requête ...
- Les fonctions réalisées.

¹ rétroaction : ici c'est la phase Après-Visualisation qui joue ce rôle.

- Le jeu d'essai.
- L'évaluation du système.

IV.1 Première réalisation

Cette première réalisation a fait l'objet d'un rapport de DEA [Ato 84]. Il est à noter que la proposition de représenter l'expertise du domaine par un ensemble de schémas logiques, représentant différentes notions abstraites que manipule un documentaliste et permettant grâce à un certain raisonnement d'accéder aux documents, n'a pas été retenue pour cette première réalisation.

IV.1.1 Architecture du système et description des différents constituants

En fait cette première maquette été réalisée entièrement grâce à un moteur d'inférence développé par la société BULL appelé BOUM. Ce moteur fonctionne en chaînage avant et en largeur d'abord : il part des faits connus, et essaye d'en déduire tous les nouveaux faits possibles par une application de règles, puis il repart des nouveaux faits et ainsi de suite jusqu'à ce qu'aucun nouveau fait ne soit déductible.

Tous les objets manipulés ont été décrits dans la base de connaissance du système expert.

IV.1.2 Les connaissances et données utilisées et leur représentation

Les connaissances utilisées dans cette maquette se résument au thesaurus et à la base descriptive. Ces deux constituants sont représentés dans la base de faits initiale du système de la manière suivante :

Les descriptions des images : (*n° de l'image*, liste de (*mots-clés*)).

Les mots-clés : (*mot*, nombre d'*occurrences*, domaine principal).

le nombre d'occurrences correspond aux nombres de descriptions où le mot apparaît, le domaine principal est le champ sémantique du thesaurus dans lequel se trouve ce mot.

Le thesaurus : il est représenté par ses liens sémantiques ; un lien entre deux termes est défini par un fait de la forme :

(*type du lien*, *mot-clé de départ*, *mot clé d'arrivée*, *coefficient d'affaiblissement*, *domaine de validité*)

Le type du lien sémantique peut être de trois formes : sorte de, partie de, et associé à. Les deux premiers spécialisent la relation classique "générique/spécifique" et une orientation du lien

est nécessaire. Le dernier lien correspond à l'équivalence sémantique ; une orientation n'est pas nécessaire puisqu'il est symétrique.

Le coefficient d'affaiblissement exprime la proximité sémantique entre les deux mots-clés, il varie entre 0 et 1. Un coefficient proche de 1 exprime une forte proximité donc un affaiblissement peu important.

Les stratégies mises en œuvre dans le système sont représentées sous la forme de règles de production ; elles utilisent les faits représentant le thesaurus et la base descriptive, mais aussi les faits associés à la requête utilisateur.

La requête utilisateur :

(*n° de requête, n° de version, fidélité, domaine principal, utilisateur, liste de (mot-clé, indicateur d'importance, pureté)*)

Le numéro de version progresse de une unité à chaque déformation que subit la requête.

La fidélité reflète le degré de déformation qu'a subi la requête depuis sa forme initiale (fidélité = 1). Elle diminue à chaque déformation en fonction du type de la transformation appliquée à la requête.

Le domaine principal représente un des champs sémantiques du thesaurus dans lequel la requête se situe le mieux.

L'indicateur d'importance est attribué par l'utilisateur relativement à son attachement envers le mot-clé.

La pureté reflète les déformations qu'a subies le mot dans la requête depuis sa forme initiale.

IV.1.3 Les fonctions réalisées

Seule une partie de la phase Avant-Visualisation du processus EXPRIM a été réalisée dans cette maquette. Une politique de remodelage de la requête a été définie seulement dans le cas de l'élargissement. Cette politique consiste à ne considérer, lors d'une déformation, qu'un seul terme à la fois, puis à le supprimer où à le remplacer par un autre terme.

Le choix du terme s'effectue de la manière suivante : pour chaque terme t de la requête, est calculé un coefficient d'affaiblissement $hS(t)$ que subirait la requête si t était supprimé, et le coefficient $hR(t)$ si t était remplacé. La modification à appliquer est celle qui a un coefficient maximum.

Le calcul des coefficients d'affaiblissement est tout à fait heuristique, il utilise les différents coefficients présents dans la requête, le thesaurus et les mots-clés.

Voyons comme exemple celui de la suppression d'un terme :

Calcul de $hS(t)$:

Les paramètres utilisés sont les suivants :

a : l'affaiblissement qu'a déjà subi le terme t (pureté),

p : l'importance qu'a donnée l'utilisateur au terme t ,

kdm : qui mesure l'importance que donne l'utilisateur au champ sémantique de t .

$$h(t) = (1 - c_a * a^{e_a}) * (1 - c_p * p^{e_p}) * (1 - c_k * kdm^{e_k})$$

où c_a, c_p, c_k sont des coefficients à choisir et e_a, e_p, e_k sont des exposants à choisir.

Ce calcul d'affaiblissement intervient lorsque la requête ne rapporte pas d'image. Nous allons expliquer quand une description d'image répond à une requête, c'est-à-dire comment la mise en correspondance entre une requête et les documents s'effectue.

La mise en correspondance :

Soit une requête R ayant une liste de mots-clés LMR et une description D ayant une liste de mots-clés LMD ; D est en correspondance avec R si LMR est inclus dans LMD.

Ce qui est équivalent au modèle booléen classique avec un seul opérateur : le ET.

IV.1.4 Le jeu d'essai

Le jeu d'essai de cette application a été très pauvre : quelques images de l'agence SYGMA associées à des descriptions effectuées par des documentalistes. Un thesaurus a été bâti à partir seulement des mots-clés utilisés dans les descriptions ; des pondérations ont été introduites sur les arcs.

IV.1.5 L'évaluation du système

Compte-tenu de la pauvreté du jeu d'essai une évaluation en terme de rappel et de précision n'est pas envisageable. Seule une évaluation intuitive sur le comportement du système est possible mais elle demeure peu informative quant à la validité de l'approche utilisée.

IV.1.6 Conclusion

Notre avis sur cette expérimentation est que de nombreux tests sont nécessaires pour déterminer les utilisations des différents paramètres qui régissent la recherche. En effet, de nombreux "ingrédients" existent tels que : le nombre d'occurrences d'un terme, le coefficient d'affaiblissement d'une relation du thesaurus, la fidélité d'une requête, l'indicateur d'importance d'un terme, la pureté d'un terme, $hS(t)$, $hR(t)$ et leurs coefficients et exposants ... Déterminer le "dosage" permettant de manipuler tous ces paramètres n'est pas une chose facile à faire, car pour cela il faut contrôler l'effet de chaque paramètre sur une recherche.

De plus la valeur de certains paramètres n'est pas toujours facile à déterminer. Prenons par exemple le coefficient d'affaiblissement d'une relation du thesaurus. Qui a la charge et la possibilité d'associer à une relation d'un thesaurus un tel coefficient ? Connaissant le coût de construction d'un thesaurus, ce n'est sûrement pas les documentalistes iconographiques qui peuvent prendre cette responsabilité, et même s'ils avaient tout leur temps quelle méthode doivent-ils utiliser pour affecter ce coefficient aux relations ?

exemple : Noah-sort-de sportif, coefficient 0.8

Comment interpréter ce coefficient ? Est-ce que cela veut dire que Noah est un sportif à 80% ?

Cette première réalisation du système EXPRIM ne nous semble pas viable, car trop paramétrée, donc avec une complexité à terme incontrôlable. Par contre, il n'est pas à rejeter que certains des paramètres proposés aient un rôle à jouer dans le processus de recherche, mais la détermination de ce rôle nécessite, à notre avis, une réalisation assez avancée du processus EXPRIM, dans sa totalité, avec un échantillon d'images relativement important.

Au cours de cette première réalisation, seulement l'élargissement de la requête a été réalisé et ceci sans prendre en compte tous les paramètres donnés dans les structures comme, par exemple: le nombre d'occurrences d'un terme et le coefficient d'affaiblissement d'une relation du thesaurus ; de plus le modèle de mise en correspondance est limité à la prise en compte du ET, ce qui réduit les stratégies d'élargissement à la suppression ou au remplacement d'un terme.

Beaucoup de fonctions du système EXPRIM demeurent non réalisées et même non étudiées, comme par exemple la reformulation de requête dans la phase Après-Visualisation et les stratégies de rétrécissement de la requête dans la phase Avant-Visualisation.

Cette étude n'a pas été poursuivie, ceci pour plusieurs raisons : l'étudiant de DEA chargée de cette réalisation est parti, et la société BULL a abandonné l'exploitation du produit BOOM.

Voyons alors quelles ont été les caractéristiques de la deuxième réalisation.

IV.2 Deuxième réalisation

Cette deuxième réalisation a été obtenue grâce à la collaboration d'un stagiaire de DEA [Rua 86], d'un stagiaire DESS [Maz 86] et d'un stagiaire de maîtrise, chacun ayant une partie bien précise du système à réaliser. Voyons quelles étaient les parties de ce système.

IV.2.1 Architecture du système et description des différents constituants

L'architecture du système correspond à celle de la figure 7 et a été mise en place sur le système Multics.

Les composants du système sont les suivants :

Le système expert a été réalisé en Prolog par le stagiaire de DEA avec le langage Foll-Prolog. Nous étudierons plus particulièrement son travail dans la suite de la description.

Le système documentaire gérant la base descriptive est le système TEXTO.

Le système de pilotage et de dialogue ainsi que le système de visualisation ont été développés en PL1 par les deux autres stagiaires. Le système de visualisation se compose d'un moniteur et d'un vidélecteur connectés à l'ordinateur.

IV.2.2 Les connaissances et données utilisées et leur représentation

Les connaissances du domaine se résument au thesaurus ayant la représentation suivante de type Prolog :

relation (mot1,mot2,coef).

"relation" est un lien sémantique du thesaurus, du type : sorte de (générique/spécifique) ou associé à (équivalence sémantique).

"coef" détermine la proximité sémantique entre les deux mots via le lien sémantique. Nous ne reviendrons pas sur l'existence quelque peu arbitraire de ce coefficient.

exemple : Sorte_de (Noah, joueur, 0.90).

Les descriptions des images sont gérées par le système documentaire TEXTO qui a comme modèle de correspondance le modèle booléen classique (formulation booléenne ET, OU, SAUF; indexation binaire).

Le modèle de description choisi est composé d'un ensemble de dimensions représentant chacune un caractère de l'image, chacune des dimensions pouvant être découpée en champs mono ou multi-valués.

exemple extrait de [Rua 86] :

DIMENSION I :

Domaine : tennis

DIMENSION II : Événement

Date : Juin 83,

Tournoi : Inter-France

Joueur 1: Noah

Joueur 2 : Wilander

Vainqueur : Noah

Score : 6-3, 6-2, 7-6

Niveau : Finale

DIMENSION III : Action

Joueur 1 : Revers, grimace, au-filet

Joueur 2 : fond-de-cours, fatigue

La requête possède deux formes. La première forme est construite par l'utilisateur, elle se compose d'une liste de termes séparés par des virgules. La deuxième forme est construite par le système à partir de la première, elle est représentée par une liste de critères connectés par des ET, un critère étant un champ d'une dimension des descriptions, associé à une liste de valeurs possibles connectées par des OU.

exemple illustrant la deuxième forme :

(domaine = tennis) et (joueur = Noah ou Leconte)

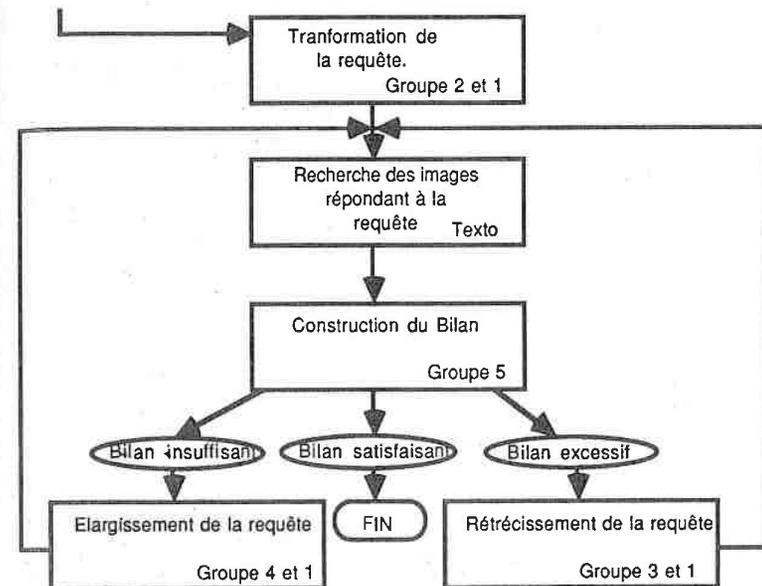
IV.2.3. Les fonctions réalisées

Dans cette deuxième réalisation, la phase Avant-Visualisation du processus EXPRIM a été réalisée avec plus de fonctions que dans la première réalisation.

En effet, cinq groupes de règles ont été définis, chaque groupe possédant une fonction bien précise :

- groupe 1 : règles permettant de calculer la proximité sémantique entre deux termes,
- groupe 2 : règles réalisant le passage de la première forme de requête à la deuxième forme,
- groupe 3 : règles réalisant le rétrécissement de la requête,
- groupe 4 : règles réalisant l'élargissement de la requête,
- groupe 5 : règles guidant la démarche du système.

Le cinquième groupe de règles régit le déroulement d'une étape de recherche de la manière suivante :



On peut remarquer que le groupe 1 contenant les règles de calcul de la proximité sémantique est utilisé à plusieurs endroits par d'autres groupes de règles. En effet, ce groupe de règles constitue l'élément central de la réalisation. Les règles de ce groupe tentent de déterminer si deux termes sont proches sémantiquement avec un certain coefficient, ceci en s'appuyant sur la structure du thesaurus, et surtout à l'aide des coefficients attachés aux relations sémantiques.

exemple de règle du groupe 1 :

Si R (motA, motB, coef 2) et coef2 >= coef1

alors Proche_sémantiquement (motA, motB, coef 1)

Le système Prolog fonctionnant en chaînage arrière, cette règle doit alors être interprétée de la manière suivante :

Pour que motA soit proche sémantiquement de motB avec un coefficient coef1, il faut que dans le thesaurus il existe une relation sémantique R reliant motA à motB avec un coefficient coef2, coef2 étant supérieur ou égal à coef1.

D'autres règles composant les relations sémantiques entre elles existent, elles permettent de dire qu'un terme est proche sémantiquement d'un autre terme s'il existe dans le thesaurus un chemin les reliant.

Les règles du groupe 2 ont pour fonction de construire la requête destinée au système documentaire. Elles cherchent d'une part à ventiler les termes de l'utilisateur dans les champs des facettes nécessaires à la sélection des images, puis d'autre part à étendre la liste des valeurs associées au champ avec des termes proches sémantiquement.

exemple :

Requête utilisateur : Joueurs français

Partie du thesaurus utilisé :

sorte_de (Noah, joueur, C2) ;

sorte_de (Noah, joueur français, C3) ;

sorte_de (Leconte, joueur, C4) ;

sorte_de (Leconte, joueur français, C5) ;

Le système s'aperçoit que Leconte et Noah sont des joueurs français et que ce sont aussi des joueurs. Comme joueur est un champ d'une dimension de l'image, la requête du système sera :

(joueur = Leconte ou Noah).

Les règles du troisième groupe ont pour fonction de rétrécir la portée de la requête si celle-ci apporte trop d'images. Ce rétrécissement s'effectue soit en ajoutant des critères en conjonction soit en supprimant des critères en disjonction.

Inversement les règles du groupe 4 élargissent la portée de la requête soit en supprimant des critères mis en conjonction ou en ajoutant des critères en disjonction.

Ces règles utilisent le calcul de proximité sémantique à travers les relations du thesaurus pour déterminer les critères à ajouter ou à supprimer. Nous ne donnerons pas d'exemple de ces règles

car la description complexe de leur fonctionnement dans [Rua 86] ne nous a pas permis de trouver un exemple simple et illustrateur.

IV.2.4 Le jeu d'essai

L'état final de la maquette comportait environ une quinzaine d'images extraites d'un vidéo-disque de l'agence SYGMA. Les photographies choisies avaient comme sujet principal le tennis, elles ont été décrites selon le modèle cité plus haut par des personnes de l'équipe EXPRIM c'est à dire par des non-spécialistes de la documentation iconographique.

Le thesaurus a été construit a posteriori à partir des descriptions des images et comportait une trentaine de termes.

IV.2.5 L'évaluation du système

Comme pour la première maquette, une évaluation en terme de rappel et de précision ne peut être significative compte-tenu de la pauvreté du jeu d'essai. Tout ce que l'on peut dire est que le système fonctionnait dans sa totalité avec peut-être un temps de transformation de la requête et de déformation un peu long par rapport aux autres fonctions du système.

IV.2.6 Conclusion

Cette réalisation a l'avantage d'avoir été effectuée en respectant l'architecture initialement prévue dans le projet. Elle a mis en évidence que la définition de la notion de proximité sémantique n'était pas facile à mettre en œuvre ; la solution réalisée ici avec Prolog étant déjà coûteuse en temps malgré la petite taille du thesaurus utilisé. On peut remarquer aussi que les descriptions des images ont un degré de précision trop important ; des informations relatives au contexte de l'image (le score du match et le vainqueur) et des informations relatives au contenu de l'image (le nom des joueurs) sont placées à un même niveau. Trop d'informations sont présentes dans les descriptions : comme par exemple le champ "joueur". Ce champ est inutile car, si dans le thesaurus on sait que Noah est un joueur, il y a de fortes chances que Noah apparaisse dans la description parce qu'il est en train de jouer au tennis dans l'image, et même s'il n'est pas en train de jouer et que l'utilisateur souhaite voir Noah en pleine action, il pourra rejeter rapidement cette image lors de la visualisation. Un tel niveau de précision n'est pas nécessaire compte tenu du choix rapide lors de la visualisation, de plus il a pour inconvénient de compliquer le passage de la première forme de requête utilisateur à la deuxième forme.

Cette expérimentation ne s'est pas poursuivie, car l'arrivée du projet ESPRIT et celle de nouveaux membres dans l'équipe ont orienté les recherches vers de nouvelles directions.

V Conclusion

Les idées émises jusqu'à notre arrivée ne concernaient que la phase "Avant-Visualisation". Beaucoup de ces idées n'ont jamais pu être concrétisées car la plupart d'entre elles étaient inadaptées aux problèmes posés par la recherche d'images ; la jeunesse du projet explique sans doute les propositions parfois hasardeuses qui ont été faites. D'autres idées surement plus raisonnables ont vu le jour dans deux réalisations. Une évolution des idées est remarquable dans la deuxième réalisation où une définition de la notion de proximité sémantique a été proposée. Ces résultats montrent que seule une expérimentation peut permettre aux idées d'évoluer tout en restant dans le domaine du réalisable.

L'arrivée du projet ESPRIT va donner les moyens à l'équipe EXPRIM de mener ces recherches sur une plus grande échelle.

Chapitre 3	
Première étude :le projet ESPRIT.....	121
I Introduction.....	121
II Spécification du système.....	122
II.1 Organisation générale.....	122
II.1.1 Le graphe d'enchaînement des processus.....	122
II.1.2 Description générale des processus.....	124
II.2 Présentation des deux principaux processus.....	125
II.2.1 L'évaluation d'une demande (phase "Avant-Visualisation").....	125
II.2.1.1 Position du problème.....	125
II.2.1.2 L'évaluation et la transformation de la requête utilisateur.....	126
II.2.2 La reformulation de la demande (phase "Après-Visualisation").....	128
II.2.2.1 Analyse des choix.....	129
II.2.2.2 Reformulation de la demande.....	132
III Réalisation.....	134
III.1 Introduction.....	134
III.2 L'application pilote.....	134
III.2.1 Les images et descriptions.....	135
III.2.2 Le thesaurus.....	137
III.3 L'expérimentation.....	138
III.3.1 Architecture du système.....	138
III.3.1.1 Présentation générale.....	138
III.3.1.2 Les choix des différents composants..	138
III.3.1.2.1 L'imageur documentaire.....	138
III.3.1.2.2 Le système documentaire.....	141
III.3.1.2.3 Le moteur d'inférence.....	141
III.3.1.2.4 Le module de dialogue.....	142
III.3.1.2.5 Le module de pilotage.....	142
III.3.2 Les parties réalisées.....	142
IV Conclusion.....	144

Chapitre 3

Première étude : le projet ESPRIT

I Introduction

Ce chapitre présente l'étude que nous avons effectuée dans le cadre du projet Esprit pour la réalisation d'un système "intelligent" de recherche d'images basé sur le processus EXPRIM et l'utilisation de l'imageur documentaire de la SEP. Elle représente à la fois notre première approche et la troisième étude du projet EXPRIM. Cette étude a été réalisée à travers une collaboration entre la SEP, et plus particulièrement la division traitement de l'image, le Bureau Marcel van Dijk de Bruxelles qui est un bureau d'ingénieur conseil ayant de sérieuses compétences en documentation automatique en la personne de Georges Van Slype et enfin l'équipe EXPRIM. Nous étions très impliqué dans ce projet car notre contrat CIFRE qui a été obtenu grâce à ce projet, nous a donné le statut de coordinateur entre les activités de l'équipe EXPRIM et celles de la SEP.

Les études menées à travers cette collaboration avaient pour finalité de déterminer la faisabilité à la fois technique et économique d'un système "intelligent" de recherche d'images utilisant l'imageur documentaire et plus particulièrement la technologie du vidéodisque LVROM (cf chapitre 1, § 1.4.2), et aussi les nouvelles techniques de l'intelligence artificielle. Notre participation dans ce projet a été exclusivement consacrée à l'étude technique pour laquelle l'ensemble des participants avaient décidé de réaliser un prototype intégrant les idées d'EXPRIM, l'imageur documentaire et une base d'images d'une importance suffisante pour que la faisabilité technique soit estimée dans les conditions réelles d'exploitation d'un tel système.

Il faut mettre en évidence, dès maintenant, les différences entre les objectifs de chacun des partenaires. Les partenaires industriels (SEP et BMvD) avaient pour objectif de réaliser un prototype proche d'un produit commercialisable, alors que pour l'équipe EXPRIM, ce projet apportait les moyens techniques et humains pour expérimenter enfin les idées émises jusqu'à maintenant. Les différences entre les objectifs visés ont orienté les choix vers des directions parfois opposées.

De nombreuses parties du prototype ont été réalisées, mais nous verrons dans ce chapitre les raisons qui ont conduit cette réalisation à l'échec. Cet échec a évidemment guidé nos choix pour la réalisation d'un deuxième prototype que nous avons mené seul et que nous présenterons dans la prochaine partie (chapitre 3).

II Spécification du système

La spécification générale du système a été construite sous la forme de deux vues, l'une interne et l'autre externe [Hal 86][Hal 88b] :

- la vue interne montre l'organisation du système vu de l'intérieur à travers une description générale des fonctionnalités des différents processus faisant partie de l'architecture globale du système et d'une description des principales structures impliquées dans les processus. Cette spécification propose une description détaillée des deux principaux processus réalisant en grande partie la phase "Avant-Visualisation" et la phase "Après-Visualisation". Nous en présenterons une partie dans ce paragraphe.

- la vue externe présente le dialogue à travers lequel un utilisateur peut mener sa recherche d'images. Cette vue comprend l'organisation générale du dialogue et les différentes transitions possibles entre les écrans en fonction des étapes de recherche.

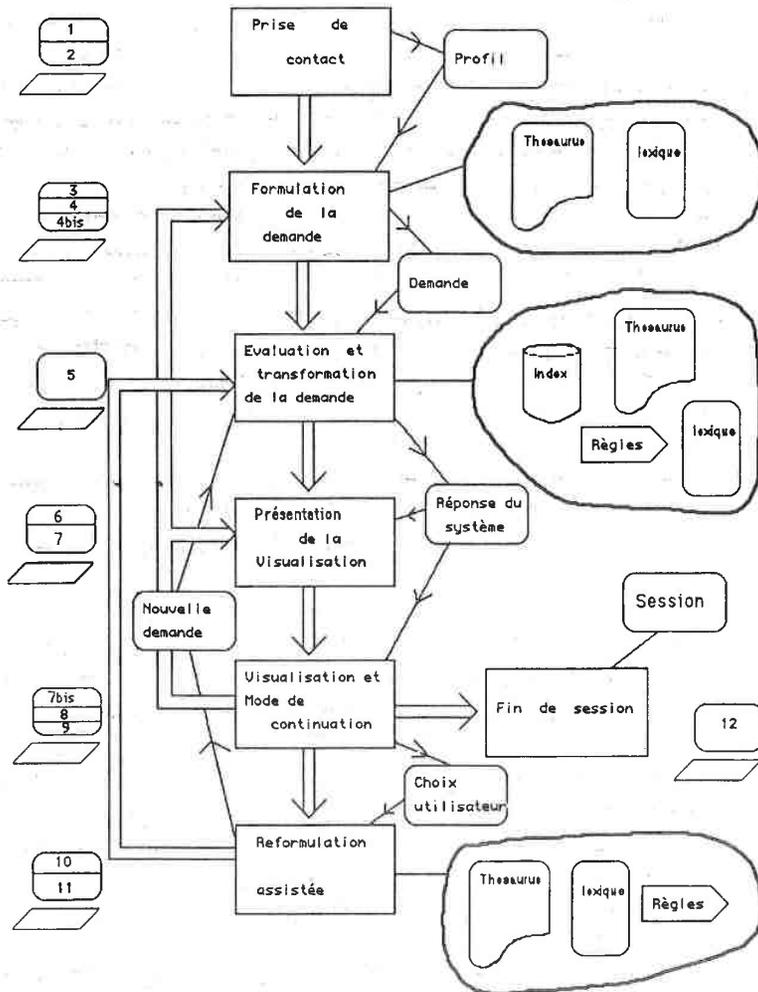
La description précise de ces deux vues est donnée en annexe (cf Annexe 4, volume séparé) ; nous ne donnons dans ce paragraphe que les principales fonctions du système à travers une présentation de l'organisation générale et de quelques exemples des traitements qui réalisent la phase "Avant-Visualisation" et la phase "Après-Visualisation".

II.1 Organisation générale

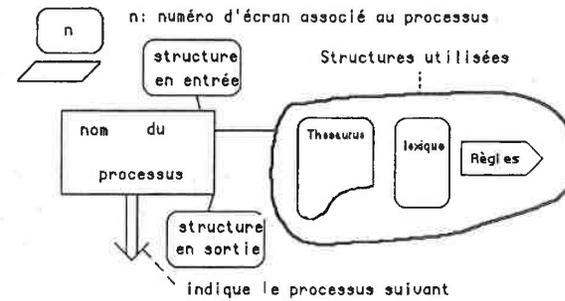
II.1.1 Le graphe d'enchaînement des processus

Le graphe d'enchaînement permet d'avoir une vue d'ensemble de l'architecture interne du système découpé en processus. Chaque processus est représenté avec ses structures de données en entrée, ses structures de données en sortie et les structures de données qu'il utilise. Un lien avec la vue externe est également présent par le rattachement des numéros des écrans au processus auquel ils sont associés.

Le graphe :



Légende:



II.1.2 Description générale des processus

Les processus sont présentés en indiquant pour chacun d'eux leur fonction :

Prise de contact: détermine le type de l'utilisateur, propose une aide pour l'utilisation du système et propose une information générale sur la collection.

Formulation de la demande: propose différents modes de recherche d'image en fonction du profil de l'utilisateur :

- recherche d'une image précise:
 - par un numéro: le système ne propose pas ce mode de recherche à un utilisateur ne connaissant pas le fonds d'images,
 - avec des critères :
 - . soit guidée par le système,
 - . soit en langage documentaire : ce mode de recherche n'est proposé qu'aux utilisateurs connaissant les techniques documentaires.
- feuilletage de la collection
- recherche sur un thème : mêmes propositions que pour une recherche d'image précise avec critères,
- la saisie de la demande, soit d'une manière assistée, soit en langage documentaire.

Evaluation et transformation de la demande: analyse et transforme la requête utilisateur en s'appuyant sur l'utilisation d'heuristiques et sur un dialogue avec l'utilisateur afin d'obtenir le plus rapidement possible des images. Les heuristiques sont représentées sous forme de règles afin de les rendre facilement modifiables et adaptables. (cf § II.2.1.2).

Présentation de la visualisation: présente le ou les ensembles d'images qui vont être manipulées lors de la visualisation.

Visualisation et mode de continuation: permet à l'utilisateur de choisir parmi les images proposées celles qui l'intéressent en utilisant les fonctionnalités de l'imageur

(Tri, classement, ...), propose un mode de continuation parmi: recherche d'images proches de son choix, reformulation assistée, fusion d'étapes, nouvelle formulation, fin de la recherche.

Reformulation assistée: propose à l'utilisateur une nouvelle demande composée de critères obtenus par une analyse des descriptions des images choisies ou rejetées par l'utilisateur ; reformule, à l'aide d'un dialogue, une nouvelle demande qui est transmise au processus **Évaluation et transformation de la demande.** Des heuristiques de proposition de critères et de reformulation sont utilisées et représentées sous forme de règles (cf § II.2.2).

Fin de session: évalue globalement la recherche en mesurant le degré de satisfaction de l'utilisateur; conserve l'ensemble des informations nécessaires à une évaluation externe de la recherche.

II.2 Présentation des deux principaux processus

Les principaux processus du système sont "l'évaluation de la demande" et "la reformulation assistée". Ces deux processus sont détaillés dans la suite d'une manière assez informelle car ces explications devaient pouvoir être comprises par tous les participants du projet.

II.2.1 L'évaluation d'une demande (phase "Avant-Visualisation")

La spécification de ce processus a fait l'objet d'une présentation dans le rapport de DEA de D. Leonard [Leo 86]. Ces descriptions prennent leurs exemples dans l'application pilote qui est décrite dans le paragraphe suivant (cf § III).

II.2.1.1 Position du problème

Les descriptions des images :

A chaque image stockée sur vidéodisque est associée une description textuelle gérée par un logiciel documentaire. Une description d'image est un ensemble prédéfini de champs auxquels sont associés des termes ou des valeurs.

Exemple de champs : "Date de prise de vue", "Auteur de la photographie", "Description principaux", "Descripteurs secondaires"...

Pour interroger la base, l'utilisateur décrit l'image qu'il recherche en étant assisté par le système. Cette description est réalisée à partir d'une grille d'interrogation constituée d'un ensemble de domaines sur lesquels l'utilisateur peut interroger la base d'images. Un domaine est une vue de plus générale de la notion de champ appartenant à la description de l'image. Par exemple, au domaine "Contenu de la photographie" peuvent être associés les champs

"Descripteurs secondaires" et "Précision sur l'objet de la photographie". Le résultat de cette étape de formulation est appelé "requête utilisateur".

II.2.1.2 L'évaluation et la transformation de la requête utilisateur

Ce processus doit analyser et transformer la requête utilisateur en s'appuyant sur des heuristiques et sur un dialogue avec l'utilisateur afin d'obtenir le plus rapidement possible des images.

Trois niveaux différents d'évaluation et de transformation ont été spécifiés :

niveau lexical :

Il a pour tâche d'identifier les termes utilisés dans la requête et, s'ils ne sont pas reconnus, de les modifier soit automatiquement, soit en dialoguant avec l'utilisateur. Les termes reconnus doivent être associés à un champ de la description de l'image en fonction du domaine de la requête utilisateur dans lequel ils figurent. Deux grandes étapes composent le niveau lexical :

- modification des termes non reconnus,
- passage des domaines aux champs : cette étape du niveau lexical a pour fonction d'associer à chaque terme utilisé au moment de l'interrogation un champ de la description des images. Nous appelons "requête intermédiaire" le résultat de cette transformation.

Niveau terminologique :

La fonction de ce niveau est de remplacer les non-descripteurs de la requête intermédiaire par les descripteurs équivalents.

Niveau sémantique :

Ce niveau soumet la requête intermédiaire à des heuristiques de compréhension en vue de la transformer en une requête documentaire adéquate aux indexations. Le résultat de cette transformation est une requête documentaire. A ce niveau, tous les termes présents dans la requête formulée par l'utilisateur ont été reconnus par le système, ou bien la décision de les ignorer a été prise.

Pour obtenir une requête documentaire booléenne, plusieurs types de coordinations logiques ont été déterminées :

Coordination logique entre les différents champs d'un même domaine :

Cette coordination doit être précisée application par application. D'une manière générale, il semble intéressant de coordonner les champs associés à un même domaine par des OU

Coordination logique entre champs de domaines différents :

Dans ce cas, il semble que la coordination logique entre les différents champs de la requête intermédiaire doit être un ET.

Coordination logique à l'intérieur d'un champ : un certain nombre d'heuristiques permettent d'effectuer cette étape, voici les deux heuristiques les plus importantes :

Heuristique 1 : Si deux descripteurs d'un champ de la requête intermédiaire ont même générique et que l'un n'est pas générique de l'autre, alors connecter logiquement ces deux descripteurs par une disjonction.

Exemple :

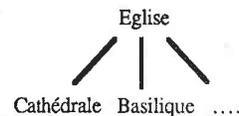
Requête utilisateur :

Contenu de la photographie : BASILIQUE, CATHEDRALE

Requête intermédiaire :

(De1 = BASILIQUE OU CATHEDRALE) OU (De2 = BASILIQUE OU CATHEDRALE)

Thesaurus :



Heuristique 2 : Les connecteurs logiques non encore instanciés par l'heuristique précédente sont des ET.

Exemple :

Requête utilisateur :

Contenu de la photographie : POUPEE, CHAPEAU

Requête intermédiaire :

(De1 = POUPEE ET CHAPEAU) OU (De2 = POUPEE ET CHAPEAU)

Des heuristiques modifiant la structure de la requête intermédiaire ont été définies afin de donner à la requête documentaire une certaine cohérence sémantique, voici les heuristiques les plus importantes :

Suppression de termes :

Heuristique 3 : Si dans un champ de la requête intermédiaire à valeurs dans le thesaurus, un descripteur est générique d'un autre, alors supprimer le terme le plus générique.

Exemple :

Requête utilisateur :

Contenu de la photographie : JOUET, POUPEE

Requête intermédiaire :

(De1 = POUPEE) ou (De2 = POUPEE)

Adjonction de termes :

Heuristique 4 : Si dans la requête utilisateur un terme est un générique, alors ajouter dans la requête documentaire la disjonction de tous ses spécifiques.

Exemple :

Requête utilisateur :

Contenu de la photographie : FONTAINE

Requête intermédiaire :

(De1 = FONTAINE OU FONTAINE ADOSSEE OU FONTAINE ISOLEE)

OU

(De2 = FONTAINE OU FONTAINE ADOSSEE OU FONTAINE ISOLEE)

Cette dernière heuristique nous paraît essentielle. En effet, dans l'exemple présenté, si l'on n'ajoute pas les termes spécifiques de fontaine dans la requête documentaire, l'utilisateur ne pourra visualiser que des images représentatives des fontaines contenues dans la base d'images et risquera de conclure qu'il n'y a pas d'image intéressante de fontaine.

Dès que l'évaluation et la transformation de la requête utilisateur est terminée, la requête intermédiaire prend alors le statut de requête documentaire afin de sélectionner les descriptions des images correspondant à la demande. L'utilisateur entre alors dans une phase de visualisation en manipulant les images à l'aide de l'imageur. Dès que cette phase est terminée, la phase "Après-visualisation" peut commencer en analysant les choix de l'utilisateur.

II.2.2 La reformulation de la demande (phase "Après-Visualisation")

Cette analyse constitue la première proposition de réalisation de la phase "Après-Visualisation" du processus EXPRIM. Elle comporte, sans aucun doute, de nombreux défauts de jeunesse, mais elle a l'avantage de montrer la difficulté de proposer une méthode cohérente permettant d'harmoniser une analyse de descriptions d'images avec une reformulation de demande.

Dans la phase de visualisation, l'utilisateur a choisi et rejeté des images et il a décidé de poursuivre sa recherche en laissant le système analyser ses choix afin de reformuler une nouvelle demande.

Deux formes de reformulation ont été étudiées :

- Avec dialogue
- Sans dialogue

Dans la première forme, le système va proposer une nouvelle formulation domaine par domaine qui comprendra des termes connus de l'utilisateur (présents dans son ancienne demande) et de nouveaux termes obtenus suite à l'analyse de ses choix.

L'utilisateur peut alors modifier la proposition en enlevant ou en ajoutant des termes. Cette formulation est le fruit d'une interaction entre le système et l'utilisateur. Cette interaction prend la forme d'un dialogue qui doit aboutir à une nouvelle demande.

Dans la deuxième forme de reformulation, l'utilisateur laisse toute liberté au système pour trouver des images proches de son choix. Les nouvelles images sont atteintes par une nouvelle demande que le système a construite de la même manière que dans la première forme de reformulation. Seul le dialogue est absent.

La reformulation peut être divisée en deux étapes :

- l'analyse des choix,
- la reformulation de la nouvelle demande,

II.2.2.1 Analyse des choix

Le premier but de cette analyse est de trouver des **propriétés intrinsèques** aux ensembles de descriptions d'images construits lors de l'étape de visualisation. Puis, parmi les propriétés trouvées, nous allons mettre en évidence celles qui caractérisent les ensembles entre eux. Nous appellerons **propriétés caractéristiques** ce type de propriétés. Les ensembles des descriptions manipulés sont : les descriptions des images choisies, les descriptions des images rejetées, les descriptions des images indifférentes : l'utilisateur n'a pas émis de choix sur ces images (cet ensemble de descriptions n'est pas analysé car son statut dans le choix de l'utilisateur n'est pas suffisamment discriminant).

La détermination des propriétés a été spécifiée en formulant un certain nombre de définitions, voyons à travers des exemples les principales définitions utilisées.

Recherche de propriétés intrinsèques :

Voyons comment une propriété de la forme $c = d$, où c est nom du champ et d un descripteur appartenant au thesaurus, peut être définie pour un champ à valeurs dans le thesaurus.

Def2 : Une propriété p d'un ensemble de descriptions D est de la forme $c = d$ si d est suffisamment présent à l'intérieur de c dans les descriptions de D .

Il faut alors définir ce que représente suffisamment présent pour un descripteur appartenant au thesaurus.

Def3 : Un descripteur d appartenant au thesaurus est suffisamment présent dans un champ c d'un ensemble de descriptions D si :

- son pourcentage d'apparition est supérieur à un pourcentage d'acceptation
- OU si
- ses spécifiques sont suffisamment présents.

Ici, suffisamment peut être aussi une comparaison entre un pourcentage d'apparition et un pourcentage d'acceptation, en considérant chaque occurrence des termes spécifiques comme une occurrence du terme générique.

Exemple :

Soit D un ensemble de descriptions :

$D = \{de_1, de_2, de_3, de_4\}$

$de_1 : c = \text{BAC, MER}$

$de_2 : c = \text{CARGO, MER}$

$de_3 : c = \text{CHALUTIER}$

$de_4 : c = \text{PORT, MER}$

Soit un pourcentage d'acceptation = 50 %, MER est suffisamment présent (pourcentage d'apparition = 75 %), BATEAU est suffisamment présent puisque ses spécifiques le sont (pourcentage d'apparition = 75 %).

Thesaurus :



D'où deux propriétés :

- $p1 : c = \text{MER}$
- $p2 : c = \text{BATEAU}$

Recherche de propriétés caractéristiques :

Appelons tout d'abord C l'ensemble des descriptions des images choisies et R l'ensemble des descriptions des images rejetées. Une propriété caractéristique d'un ensemble de description a été défini par la définition suivante :

Def6 : Une propriété de l'ensemble C (respectivement de R) est caractéristique de cet ensemble C, respectivement de R, si elle n'est pas une propriété de l'ensemble R (respectivement de C).

Cette définition est d'ordre général, elle doit être adaptée selon le type du champ présent dans la propriété. Par exemple dans un champ à valeurs dans le thesaurus, lorsqu'une propriété p est commune à C et R, elle peut être décomposée par l'intermédiaire du thesaurus en sous-propriétés. Il est alors intéressant de regarder dans les descriptions de C et de R s'il n'existe pas une discrimination entre les sous-propriétés.

Voyons comment cette discrimination a été définie :

Soit p une propriété commune à C et R. Sp l'ensemble des sous-propriétés de p.

Def7 : S'il existe une partition de Sp, $SpC = \{p_k \text{ appartenant à } C\}$ et $SpR = \{p_i \text{ appartenant à } R\}$ p_k différent de p_i , alors SpC (respectivement SpR) est un ensemble de propriétés caractéristiques de C (respectivement de R).

Exemple :

soit une propriété p commune à C et R :

p : c = BATEAU

Thesaurus :



$Sp = \{p_1, p_2, p_3, p_4, p_5\}$ avec

$p_1 : c = BAC$

$p_2 : c = BARQUE$

$p_3 : c = CANOE$

$p_4 : c = CARGO$

$p_5 : c = CHALUTIER$

$C = \{de_1, de_2\}$

$R = \{de_3, de_4, de_5\}$

$de_1 : c = BAC$

$de_2 : c = CARGO$

$de_3 : c = BARQUE$

$de_4 : c = CANOE$

$de_5 : c = CHALUTIER$

$spC = \{p_1, p_4\}$

$spR = \{p_2, p_3, p_5\}$

Nous obtenons à la fin de cette recherche un ensemble Pco contenant les propriétés communes à C et à R, et deux ensembles PcaC et PcaR contenant les propriétés caractéristiques de C et de R. Ces ensembles vont être utilisés pour reformuler la demande.

II.2.2 Reformulation de la demande

Un utilisateur ayant parcouru le chemin classique d'une étape de recherche (formulation, visualisation, reformulation), désire au moment de la reformulation que le système lui propose une nouvelle formulation proche de la version précédente et susceptible de lui procurer de nouvelles images.

Une demande D atteindra un ensemble d'images différents d'une demande D' si D n'est pas incluse sémantiquement dans D'. Quelques définitions sont alors nécessaires :

Définitions :

- une demande D est incluse sémantiquement dans une demande D' si la requête intermédiaire de D (RID) est incluse sémantiquement dans celle de D' (RID').

- une requête intermédiaire RID est incluse sémantiquement dans une requête intermédiaire RID', si quelque soit le contenu CO du champ C de RID, il est inclus sémantiquement dans le contenu CO' du même champ c de RID'.

L'inclusion sémantique de deux contenus de champ se définit en fonction du type du champ :

Champ à valeurs dans le thesaurus :

Un contenu CO est inclus sémantiquement dans un contenu CO' si chaque descripteur de CO est génériquement présent dans CO'.

génériquement présent : signifie quelque soit le descripteur d de CO, il existe dans CO' soit le descripteur d lui même, soit un générique de d.

Exemple :

RID : c = CARGO, MER, CHALUTIER

RID' : c = BATEAU, MER

Thesaurus :

Ici RID est incluse sémantiquement dans RID'.

Construction de la nouvelle demande :

La nouvelle demande va être obtenue grâce à l'application d'heuristiques modifiant la dernière requête intermédiaire à partir des ensembles Pco, Pca et PcaR définis précédemment. Ces heuristiques permettent la modification de la requête intermédiaire en ajoutant ou en supprimant des termes à l'intérieur des champs. Une dizaine d'heuristiques ont été définies (cf Annexe 4), voyons sur un exemple la forme générale de ces heuristiques.

Exemple :

Cette heuristique permet d'ajouter de nouveaux termes dans un champs prenant ses valeurs dans le thesaurus :

heuristique 7 : Si une propriété p , appartenant à PcaC, est de la forme $c = d$, c étant un champ à valeurs dans le thesaurus, d n'étant pas sémantiquement présent dans le champ c de la requête intermédiaire RID alors ajouter d dans le champ c de la requête RID.

sémantiquement présent : Un descripteur d est sémantiquement présent dans un champ c si d est présent dans c ou si au moins un descripteur lié à d dans le thesaurus est présent dans le champ c .

Exemple :

PcaC = { (c = BATEAU), (c = MER) }

RID : c = BATEAU

RID' : c = BATEAU, MER

La vue interne, telle qu'elle a été spécifiée ici, comporte de nombreux traitements heuristiques. La réalisation d'un prototype a été nécessaire pour la validation et la mise au point de ces traitements.

III Réalisation

III.1 Introduction

La réalisation du système s'est effectuée en deux étapes. La première a consisté à prospecter parmi les utilisateurs d'images pour trouver celui qui serait en possession d'une base d'images assez conséquente (1000 photos minimum), d'une base descriptive associée à ces images et enfin d'un thesaurus ayant été utilisé lors de la description des images. Les archives photographiques du Ministère de la Culture se sont avérées être en possession d'un vidéodisque de 4000 photographies représentatives de l'ensemble de leurs collections, d'un thesaurus en cours de construction pour décrire ces photographies, et enfin d'une envie très forte d'avoir une base descriptive associée à ces photographies.

La deuxième grande étape a été la réalisation même du système. Cette réalisation s'est effectuée au CRIN et nous étions chargé de coordonner et de diriger les différentes phases. Cette tâche difficile mais néanmoins intéressante nous a placé dans un certain contexte industriel où nous jouons le rôle du "chef de projet". En effet, plusieurs personnes ont participé à cette réalisation, nous étions chargé de la répartition et de la coordination des tâches à effectuer. Nous n'avons pas toutes les libertés nécessaires pour réaliser ce projet ; nous verrons dans ce paragraphe que le nombre important de contraintes et les choix d'architecture effectués ont placé cette réalisation dans une impasse.

III.2 L'application pilote

Pour que l'expérimentation des idées émises soit crédible, il a fallu trouver une application de taille suffisamment importante.

Notre choix s'est donc orienté vers les Archives photographiques du Ministère de la culture¹ qui possède près de 1 500 000 négatifs originaux de photographies noir et blanc de grands photographes de l'histoire photographique : Baldus, Le Secq, Normand, Atget, Nadar, Marville, Seeberger et beaucoup d'autres [Kor 87].

De nombreuses raisons ont motivé ce choix :

- un vidéodisque de 4000 images récemment pressé,
- un thesaurus en cours de construction,
- un besoin d'archivage de leurs photographies dans une base de données,

¹ Archives Photographiques Fort de Saint-Cyr 78182 St-Quentin en Yvelines Cedex.

- un ensemble de personnes prêtes à collaborer à un projet de recherche.

Les parties non achevées n'ont pas été limitatrices bien au contraire car cela a permis, à travers une collaboration fructueuse, de mettre en place une application assez proche de nos besoins tout en respectant les contraintes d'un archivage photographique.

III.2.1 Les images et descriptions

Les photographies disponibles sur le vidéodisque datent du début du siècle et une majorité d'entre elles sont l'œuvre d'Eugène Atget ; elles concernent Paris, ses monuments, ses petits métiers, ses quartiers D'autres auteurs sont présents comme Nadar et ses nombreux portraits de personnages célèbres, Seeberger, Mieusement, Lefèvre-Pontalis, Moreau, Gilleta...

Un échantillon de 1000 photographies a été choisi parmi les 4000 photographies du vidéodisque, la moitié des photographies choisies sont d'Atget et l'autre moitié est un mélange photographies d'auteurs différents traitant de sujets variés autres que le Vieux Paris.

Ces 1000 photographies ont été décrites par étapes successives par des documentalistes spécialistes des collections des Archives photographiques. Ces étapes ont été les suivantes :

- mise en forme d'une grille d'indexation associée à une notice d'utilisation,
- indexations de 50 photographies,
- analyse critique de l'indexation,
- indexation de 100 photographies,
- analyse critique de l'indexation,
- indexation des 1000 photographies,
- saisie informatique des 1000 descriptions.

La grille d'indexation ainsi qu'une notice de d'utilisation (cf annexe 1) ont été mises en forme par les documentalistes qui ont alors proposé un certain nombre de rubriques en indiquant pour chacune d'elles les caractéristiques suivantes :

- rubrique facultative (F) ou obligatoire (O),
- rubrique à valeur unique (U) ou à valeurs multiples (M),
- rubrique interrogeable (I) ou non (NI),
- type de la rubrique : date (D), texte (T), à valeur dans le thesaurus (TH), à valeur dans une liste (L).

Vingt six rubriques ont été définies, dont les principales sont :

- *la cote* (O, U, I, L) ; elle représente un numéro d'inventaire gérée par les documentalistes,

- *le titre ou légende* (F, NI, T) donné par le photographe lui-même,
- *auteur de la photographie* (F, M, I, L) ; c'est le nom du photographe,
- *objet de la photographie, descripteurs principaux* (O, M, I, TH) ; cette rubrique permet la description du sujet principal de la photographie à l'aide de quelques descripteurs.
- *objet de la photographie, descripteurs secondaire* (F, M, I, TH) ; description des autres sujets de la photographie ; ce découpage de la description du contenu de la photographie en deux rubriques a été demandée par les membres du projet ESPRIT afin d'attacher une certaine pondération aux descripteurs.
- *objet de la photographie, morphologie* (F, M, I, TH) ; description des caractéristiques morphologiques de la photographie.
- *objet de la photographie, connotation* (F, M, I, TH) ; description de l'aspect subjectif de la photographie ; cette rubrique a été demandée par les membres du projet ESPRIT afin de pouvoir traiter les demandes faisant appel à la subjectivité des photographies ;
- *précision sur l'objet de la photographie* (F, M, I, L) ; description des informations représentées sous la forme de noms propres (les noms d'édifices, les noms de personnes physiques, les noms de lieu, ...),
- *commentaire* (F, NI, T) ; c'est un commentaire donné par le documentaliste qui complète l'information mise dans les autres rubriques,
- *auteur de l'objet de la photographie* (F, M, I, L) ; cette rubrique contiendra des noms de peintre, de sculpteur, d'architecte,
- *datation de l'objet photographié* (F, I, D) ; date d'un tableau, d'un monument ...
- *lieu de prise de vue* (F, U, I, L) ; indique le lieu où la photo a été prise,
- *précision sur le lieu de prise de vue* (F, M, I, L) ; c'est un complément d'information lorsque le lieu de la prise de vue n'est pas complètement déterminé avec un seul terme (adresse complète, lieu-dit, ...),
- *datation de la prise de vue* (F, U, I, D) ;
- *numéro vidéodisque* (O, U, I, L) ; numéro de l'image sur le vidéodisque.

D'autres rubriques, souvent facultatives, existent, elles contiennent des données à caractère informationnelles qui ne feront pas l'objet des recherches futures, mais qui constituent un complément d'information important : propriété du cliché, appartenance du cliché à un ensemble, date d'entrée dans la collection, mode d'entrée dans la collection, la diffusion, mention obligatoire, bibliographie, historique du cliché, numéro attribué par le photographe.

Les différentes analyses critiques sont essentielles dans un tel processus d'indexation ; elles permettent une certaine harmonisation des contenus des descriptions (nombre de descripteurs à utiliser, précision des descripteurs, utilisation correcte des descripteurs ...) car les images sont

décrites par des personnes différentes qui n'ont pas toujours la même lecture d'une photographie.

Toute cette phase d'indexation prend en temps considérable. Sur ces 1000 descriptions d'images, 667 ont été saisies au CRIN et sont donc utilisables pour l'expérimentation, les 333 descriptions restantes n'ont pas été saisies pour des raisons de manque de budget.

III.2.2 Le thesaurus

Le thesaurus utilisé pour décrire les images a été construit par les documentalistes en étroite collaboration avec le Bureau Marcek van Dijk. Il a pour fonction de contenir tout le vocabulaire nécessaire à la description de l'ensemble des collections des Archives photographiques. Ceci explique le nombre assez important de champs sémantiques et de listes annexes qui ont été créés. En effet, il contient les champs sémantiques suivants (45 au total) : agriculture, alimentation, ameublement, anatomie, architecture, armement, artisanat, art graphique, chasse, commerce, communication, comportement, conflit arme, éducation, finance, géographie physique, habillement, hygiène, industrie, justice, loisir, mesure, mort, musique, ornementation, pêche, politique, religion, science, sculpture, société, spectacle, sport, transport, animal, art et période historique, affixe de position, analyse morphologique du document, état de l'objet photographié, connotation, matériau et technique, outillage, termes généraux, végétation. Viennent s'ajouter à cette liste de champs sémantiques, douze listes annexes, appelées aussi "listes ouvertes" car leur contenu augmente au fur et à mesure de l'indexation, contenant essentiellement des noms propres ; les intitulés de ces listes sont les suivants : groupe ethnique, nom de manifestation, ordre et décoration, personnage imaginaire, personnalité, personne morale, peuple de l'antiquité, site et édifice, sujet géographique, sujet religieux, titre de spectacle, titre d'œuvre d'art.

Le thesaurus, à ce jour, contient environ 6000 unités lexicales reliées entre elles par les relations sémantiques suivantes (cf première partie, chapitre 2, § I.2) :

- générique/spécifique,
- terme associé,
- employer/employé pour,
- appartenance à un champ sémantique.

Il faut noter que toutes les unités lexicales du thesaurus ne sont pas utilisées dans les descriptions, seulement un sous-ensemble (environ 2500) sont présentes dans les descriptions. Le thesaurus représente un outil d'indexation d'images indispensable, l'inconvénient majeur

réside dans son temps de construction qui est lui aussi important. L'ensemble du thesaurus a été saisi et contrôlé au CRIN.

Toutes les données étant disponibles, l'expérimentation du système peut donc avoir lieu.

III.3 L'expérimentation

La réalisation du système s'est effectuée au CRIN, nous étions chargé d'organiser les différentes étapes de travail et de répartir les tâches à effectuer au différentes personnes impliquées. Cette étude expérimentale a fait l'objet d'un rapport ESPRIT [Reb 88].

III.3.1 Architecture du système

L'architecture du système qui a été choisie pour réaliser le prototype, a été fortement influencée par l'utilisation de l'imageur documentaire comme système de visualisation.

III.3.1.1 Présentation générale

L'architecture générale du système est celle qui avait été proposée lors de la dernière réalisation du projet EXPRIM (cf chapitre 2, § III.2).

Cette architecture devait prendre place sur un IBM PC AT sous MSDOS avec les différents sous-systèmes suivants :

- le système de visualisation : l'imageur documentaire de la SEP,
 - le système à base de connaissance est réalisé avec le moteur d'inférence OPS5+,
 - le système documentaire : Micro-Questel de Télésystème,
 - le module de dialogue est de pilotage réalisé en C Microsoft,
- tous ces choix sont justifiés dans le paragraphe suivant.

III.3.1.2 Les choix des différents composants

Ces choix n'ont pas été le fruit du hasard, ils ont été principalement guidés par l'intégration de l'imageur dans le système.

III.3.1.2.1 L'imageur documentaire

L'imageur documentaire est un logiciel développé par la SEP dans le langage C Microsoft sur une configuration compatible PC. Cette configuration comprend un écran dit "Mosaïque",

permettant la visualisation simultanée de seize imagettes, un ordinateur compatible PC, un lecteur de vidéodisque LVROM et un écran couleur permettant la présentation d'une image plein écran ainsi que l'affichage du dialogue avec la machine.

Le logiciel "imageur" qui gère l'accès aux images et leur manipulation à l'aide de la mosaïque est interfacé avec un système documentaire permettant la gestion d'une base descriptive associée aux images stockées sur le vidéodisque LVROM. Le logiciel choisi par la SEP pour réaliser cette fonction est un système documentaire disponible sur PC développé par Télésystème : Micro-Questel. Il faut remarquer que ce logiciel était à cette époque écrit dans un Basic compilé, et que la SEP ne disposait pas des programmes sources ; tout ajout ou modification de fonction devait être clairement spécifié au concepteur du logiciel qui modifiait alors ces programmes et donnait une nouvelle version à la SEP qui la testait avec souvent d'agréables surprises. Cette coopération avec Télésystème qui était à notre avis un peu lourde provient du fait qu'à cette époque il n'existait pas de logiciel documentaire sur compatible PC pouvant être interfacé avec d'autres logiciels.

L'imageur documentaire possède de nombreuses fonctions permettant une visualisation d'un ensemble d'images efficace et agréable. Nous n'allons pas décrire l'ensemble des fonctions, mais seulement les plus importantes, afin de comprendre la philosophie du système (l'ensemble de toutes les fonctions est donné dans l'annexe 2).

Configuration de la mosaïque :

La mosaïque de l'imageur est composée de 16 cases permettant la visualisation simultanée de 16 imagettes. L'intérêt principal de l'imageur est de permettre la manipulation d'une image mais aussi la manipulation de listes d'images. L'écran mosaïque peut alors être configuré de plusieurs manières suivant la forme de manipulation de liste d'images que l'utilisateur choisit, deux formes de manipulation existent :

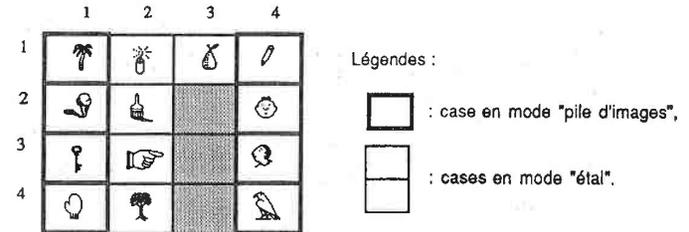
- l'étalage sur l'écran de la liste d'image : une ou plusieurs colonnes contiguës sont alors choisies par l'utilisateur ; les images de la liste à visualiser sont alors affichées les unes à la suite des autres dans les cases des colonnes désignées. Ce mode de visualisation est appelé le mode "étal"

- la pile d'images ; lorsqu'une colonne n'est pas choisie pour une visualisation en mode "étal", chaque case de cette colonne peut alors représenter une liste d'images avec seulement une image de la liste visible à l'écran. Ce mode de visualisation est appelé le mode "pile d'images", lorsqu'une colonne est en mode "pile d'images", l'utilisateur peut visualiser 4 listes différentes d'images dans cette colonne.

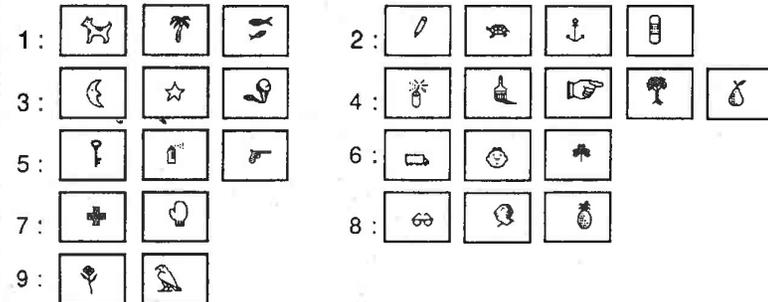
Exemple :

soit l'écran mosaïque avec la configuration suivante :

- colonne 2 et 3 en mode "étal",
- les cases des colonnes 1 et 4 en mode "pile d'images".



Soient les listes suivantes d'images :



Toutes les listes d'images (1 à 9) sont présentes dans l'écran mosaïque de la manière suivante :

- liste 1 : dans la case (1,1), seulement la deuxième image est visible,
- liste 2 : dans la case (1, 4), la première image est visible,
- liste 3 : dans la case (2, 1), la troisième image est affichée,
- liste 4 : dans les colonnes 2 et 3 qui sont en mode "étal" toutes les images sont visibles,
- liste 5 : dans la case (3, 1), la première image est visible,
- liste 6 : dans la case (2, 4), la deuxième image est visible,
- liste 7 : dans la case (4, 1), la dernière image est affichée,
- liste 8 : dans la case (3, 4), la deuxième image est à l'écran,
- liste 9 : dans la case (4, 4), la dernière image est affichée.

A l'aide de cet exemple, on s'aperçoit que l'imageur peut manipuler un nombre assez important d'images grâce à ce type de configuration et grâce aussi à de nombreuses fonctions associées telles que :

- parcourir une liste d'images d'avant en arrière,
- "zoomer" une image sur l'écran couleur,
- copier un image dans une autre case de la mosaïque (il y a duplication de l'image),
- déplacer une image d'une case vers une autre (il n'y a pas de duplication),
- détruire une image,
- nommer une liste d'images,
- afficher la légende de l'image (légende présente dans la description de l'image), etc ... (cf annexe 2).

Le logiciel documentaire interfacé avec l'imageur permet d'effectuer une recherche multicritère afin d'obtenir une ensemble d'images puis de les visualiser en utilisant la mosaïque. Le résultat de la recherche est alors une liste d'images que l'utilisateur associe à une ou plusieurs cases afin de pouvoir manipuler les images. De cette manière, un utilisateur, assez averti, peut mener plusieurs recherches de front, en gérant plusieurs listes d'images dans l'imageur.

L'intégration de l'imageur dans le prototype impose certains choix :

- le logiciel documentaire doit être Micro-questel car il est déjà interfacé avec l'imageur,
- le module de pilotage doit être écrit en C Microsoft, car une partie de ce module, écrit avec ce langage, existe déjà ; il gère la communication entre l'imageur et micro-questel.

III.3.1.2.2 Le système documentaire

Le système documentaire est donc micro-questel. La version utilisée est une version "prototype" car de nouvelles fonctions ont été ajoutées, sur demande de la SEP, afin de pouvoir communiquer avec l'imageur. Les principales caractéristiques de ce logiciel sont les suivantes :

- il est écrit en Basic sur PC avec MSDOS,
- les requêtes sont de type booléen,
- l'indexation est binaire par rubrique,
- il gère un dictionnaire de termes (lexique),
- il occupe 128Ko en mémoire centrale.

III.3.1.2.3 Le moteur d'inférence

Le besoin de mise au point incrémentale des traitements heuristiques définis dans la vue interne nous a conduit à choisir un outil adapté à cette forme de réalisation : un moteur d'inférence. Le moteur d'inférence OPS5+ nous a semblé correspondre à nos besoins, car d'une part il est disponible sur PC et il est écrit en C Lattice, et d'autre part des fonctions écrites

en C Lattice peuvent être appelées en partie droite des règles de production. Les principales caractéristiques de cet outil sont les suivantes :

- le moteur d'inférence est d'ordre 1,
- il fonctionne en chaînage avant,
- il possède deux stratégies pour choisir les règles à exécuter [ELH 87],
- il contient un algorithme rapide pour construire l'ensemble des règles exécutables à un moment donné (l'algorithme de RETE [ELH 87]),
- il occupe 320Ko en mémoire centrale.

III.3.1.2.4 Le module de dialogue

Ce module doit réaliser la forme du dialogue définie dans la vue externe. Plusieurs choix sont possibles pour réaliser le module de dialogue, il peut être :

- un sous-système autonome communiquant avec les autres sous-systèmes, le langage utilisé dans ce cas peut être le C Microsoft.
- intégré au système à base de connaissance, c'est à dire réalisé en OPS5+ à l'aide de fonctions écrites en C Lattice.

Nous verrons dans le chapitre suivant laquelle des deux solutions a été prise.

III.3.1.2.5 Le module de pilotage

Une partie de ce module existait déjà dans l'imageur documentaire, il était écrit en C Microsoft et il gérait la communication entre l'imageur et Micro-Questel. Le nouveau module de pilotage a du prendre en compte non pas 2 sous-systèmes, mais 4 sous-systèmes ; de nouveaux problèmes se sont alors posés :

- la gestion de la mémoire centrale,
- la gestion de l'écran,
- la communication entre logiciels écrits dans des langages différents.

III.3.2 Les parties réalisées

Ce travail d'organisation, fort intéressant, était nouveau pour nous et nous avons regretté de ne pas avoir pu mener la réalisation à terme. De nombreuses raisons ont conduit cette expérimentation à l'échec :

- la non disponibilité d'une bonne version de Micro-Questel, car nous ne sommes jamais arrivé à charger les indexations et le thesaurus dans le système malgré les versions successives utilisées,

- la taille mémoire gérée par le système MSDOS qui est insuffisante pour une telle architecture de système,
- la taille importante du système OPS5, qui était pourtant à l'époque le seul moteur d'inférence disponible sur le marché de si petite taille,
- la diversité des langages utilisés,
- la complexité du module de pilotage qui devait gérer la communication entre les différents sous-systèmes c'est-à-dire entre les différentes tâches du système.

Malgré ces différents problèmes certains modules ont pu être réalisés :

Le dialogue :

Une bonne partie du dialogue de la phase "Avant-Visualisation" a été réalisée en C. Le dialogue a nécessité la définition de plusieurs types de fenêtres et la gestion d'une souris à trois boutons. La réalisation de ce genre d'interface demande un temps considérable de mise au point, car la gestion de fenêtre demande la définition de nombreuses fonctions. La taille mémoire disponible étant insuffisante, nous étions contraint de réaliser cette gestion de fenêtres nous même car l'utilisation d'un produit existant, tel que MS-Window, était coûteuse en place mémoire.

Plusieurs modules de la vue interne ont été réalisés, tels que :

- la prise de contact,
- la formulation de la demande,
- deux premier niveaux de l'évaluation de la demande :
 - le niveau lexical,
 - le niveau terminologique.

Toutes ces réalisations ont été testées en simulant le système documentaire car une version correcte de ce système n'a jamais été disponible. Cette simulation n'a pas permis une expérimentation en grandeur nature de ces différents modules car le thesaurus n'a pas pu être exploité.

L'évaluation de la demande :

Le niveau sémantique de ce module a été réalisé en OPS5 afin d'expérimenter l'ensemble des heuristiques défini dans la vue interne. Cette expérimentation a fait l'objet d'un stage de DEA [Elh 87], elle a permis de déterminer un ordre d'enchaînement des règles heuristiques (cf chapitre 1, § III.3.1.2)

La requête documentaire ainsi obtenue était soumise manuellement à une version du système Micro-Questel qui ne pouvait gérer qu'une dizaine de descriptions d'images. Ces tests nous ont permis de vérifier la cohérence des heuristiques émises, mais ils ne nous ont pas permis de vérifier leur pertinence car ce type de vérification demande un échantillon plus important de descriptions.

Ces modules sont les seuls qui ont pu être expérimentés dans un environnement très peu adapté à ce type d'expérimentation. L'accumulation progressive des problèmes a réduit de plus en plus l'intérêt de l'expérience.

IV Conclusion

Cette première réalisation d'un système de recherche d'images n'est pas très encourageante, mais elle a permis de mettre en évidence des points délicats qui ne doivent plus être abordés de la même manière dans une prochaine réalisation. En effet, cet échec ne montre pas l'impossibilité de réaliser un tel système, il prouve seulement que l'environnement choisi et que l'architecture proposée n'étaient pas adaptés à la réalisation d'un prototype de recherche. Mais ces erreurs sont le fruit d'une incompréhension mutuelle entre les participants du projet, les uns souhaitant un produit final, les autres désirant "tester des idées". Les erreurs doivent être analysées pour que l'environnement de la prochaine réalisation soit plus adapté.

Les erreurs commises sont les suivantes :

- trop de composants différents dans le système,
- trop de problèmes de niveaux d'abstractions différents (assembleur pour le module de pilotage, règles de production dans OPS5),
- le système d'exploitation et la gestion de la mémoire doivent être adaptés à la réalisation d'un prototype de recherche (l'aspect mono-tâche de MS-DOS a été très contraignant)
- les objectifs à atteindre doivent être clairement définis au départ (réalisation d'un produit industriel, ou validation d'idées).

Mises à part toutes ces erreurs, il faut quand même indiquer que cette réalisation nous a permis de trouver une application pilote qui peut être réutilisée pour une autre expérimentation, d'expérimenter une forme de dialogue utilisant une souris et des fenêtres et enfin de valider la cohérence des heuristiques de la phase "Avant-Visualisation".

Cette première expérience sur la réalisation du processus EXPRIM nous a beaucoup apporté. En effet, elle nous a permis de nous confronter à la réalité industrielle et à ses contraintes, mais elle nous a donné aussi, par l'intermédiaire de nombreux contacts avec nos partenaires, une certaine connaissance sur tous les problèmes liés à la recherche d'informations et plus précisément à la recherche d'images.

La réalisation d'une version industrielle d'un tel système de recherche d'images n'est envisageable que lorsque l'ensemble des idées du projet EXPRIM auront été testées et validées sur un échantillon d'images suffisamment important.

Conclusion

Cette partie a été consacrée à la problématique d'une recherche d'informations où l'information recherchée est l'image.

Le premier chapitre s'est intéressé au problème lié- à l'utilisation de l'image dans les systèmes de recherche d'informations. Il met en évidence qu'une analyse de l'image est nécessaire et que cette analyse doit être menée en suivant une méthode qui doit permettre de ne rien oublier et d'obtenir une certaine homogénéité dans les descriptions. De plus, l'utilisation d'un thesaurus apporte une cohérence sémantique entre les différentes descriptions, en réduisant le nombre de cas de synonymie et de polysémie, sources de bruit et de silence dans les recherches. Le thesaurus est donc un outil indispensable à la construction de descriptions d'images. Ce chapitre a montré aussi que l'arrivée des supports optiques, et surtout celle du vidéodisque, a révolutionné la recherche d'images en offrant la possibilité à l'utilisateur de vérifier rapidement la pertinence d'une image en la visualisant. Grâce à cet événement un dialogue plus interactif a été proposé aux utilisateurs ; cette évolution est importante quand on sait combien la place du dialogue est importante dans les recherches d'images, comme l'a montré l'étude effectuée sur les utilisateurs et les utilisations d'une base d'images. Les systèmes qui ont été présentés dans le dernier paragraphe apportent chacun une certaine contribution dans l'évolution de la recherche d'images, mais leur dialogue s'arrête à la visualisation des images.

Or, une des principales caractéristiques du projet EXPRIM qui a été présenté dans le deuxième chapitre, réside dans la proposition d'un processus en trois phases qui se veut très interactif et qui permet à l'utilisateur de choisir ou de rejeter des images lors de la visualisation. L'utilisation des choix et la reformulation de requête de la phase "Après-Visualisation" permet au système de proposer un dialogue plus proche de l'utilisateur. Une étude des principales idées et des différentes réalisations a mis en évidence les difficultés à mettre en œuvre un tel projet. Seule quelques études sur la phase "Avant-Visualisation" ont été proposées, mais la phase "Après-Visualisation" reste non étudiée. A notre arrivée dans l'équipe, de nombreuses questions restaient sans réponses :

- le processus EXPRIM doit-il être analysé dans sa globalité ou phase par phase ?
- quelle forme doit-on donner aux requêtes, aux documents ?
- comment doit être utilisé le thesaurus dans les phases "Avant-Visualisation" et "Après-Visualisation" ?
- quelle forme doivent avoir les résultats de la phase "Après-Visualisation" ?
- comment doit être menée l'analyse des choix de l'utilisateur ?

Une partie de ces questions ont trouvé une réponse dans l'étude qui a été faite dans le cadre du projet ESPRIT. Cette première étude nous a permis de proposer un certain nombre de traitements heuristiques ayant pour fonction de réaliser les phases "Avant-Visualisation" et "Après-Visualisation". L'approche permettant de réaliser la phase "Après-Visualisation" a mis en évidence la difficulté d'extraire des choix de l'utilisateur un certain nombre d'éléments caractéristiques et de les intégrer dans la demande. Il faut remarquer que cette première proposition comporte une analyse séparée des deux ensembles : les descriptions des images choisies et les descriptions des images rejetées. Nous verrons dans la partie suivante une nouvelle approche proposant une analyse simultanée des deux ensembles qui simplifie de beaucoup les traitements et la compréhension globale de la méthode.

Beaucoup de temps et beaucoup de tests ont été nécessaires pour finalement conclure que les choix effectués ne correspondaient pas à ce type d'expérimentation qui se voulait être un prototype de recherche pour les membres de l'équipe EXPRIM, et un prototype industriel pour les autres participants du projet. L'analyse des erreurs commises a sans aucun doute guidé nos choix pour la prochaine expérimentation.

Cette première étude a proposé une première approche du processus EXPRIM, cette approche qui n'a été que partiellement testée a besoin d'un cadre sémantique plus général permettant une certaine compréhension des souhaits de l'utilisateur recherchant des images. Les techniques d'apprentissage symbolique semble offrir des éléments de réponse à ce besoin.

Troisième partie

Le processus EXPRIM vu comme un problème d'apprentissage

Introduction

Les recherches liées à la communication homme-machine sont un domaine d'application privilégié de l'Intelligence Artificielle (IA) [Pie 87]. Or l'interrogation d'une base de données, plus particulièrement d'une base documentaire, et encore plus lorsqu'il s'agit d'images, repose essentiellement sur un dialogue entre l'homme et un système [Cre 88a]. La notion d'apprentissage a alors un rôle important à jouer, et ceci à différentes échelles (cf partie 1, chapitre 4, § III.3). Cette partie s'intéresse à l'apprentissage utilisé comme procédé dans une étape de la démarche d'interrogation. Ce procédé est d'ailleurs dans la lignée de l'évolution du domaine de l'informatique documentaire, en intégrant des techniques d'autres domaines de recherche et en utilisant de nouvelles technologies [Sal 83b],[Opt 87],[Opt 89], [Sal 86].

Le processus EXPRIM vise, dans une certaine mesure, à essayer de comprendre le besoin en images d'un utilisateur. Cette compréhension est très proche d'un apprentissage à l'échelle d'une session d'interrogation. Ainsi, on peut considérer le processus EXPRIM comme un processus d'apprentissage où le concept à apprendre est le besoin de l'utilisateur et où les images choisies représentent des exemples du concept et les images rejetées des contre-exemples.

Afin de bien comprendre la problématique de cette forme d'apprentissage, qui doit prendre en compte le problème de la reformulation de requête, nous verrons donc, dans un premier chapitre, comment ce problème de reformulation est généralement traité dans les SKI. Puis, nous présenterons, toujours dans ce même chapitre, quelques notions de bases en apprentissage symbolique illustrées par plusieurs méthodes.

C'est dans le deuxième chapitre que nous exposerons notre solution, c'est-à-dire la mise en place d'un processus d'apprentissage sur les trois phases du processus EXPRIM.

L'expérimentation de la méthode dans le système RIVAGE ainsi que la présentation de ce système seront traitées dans le chapitre 3. Nous verrons dans ce chapitre comment cette méthode permet d'uniformiser le processus EXPRIM tout en réalisant les heuristiques qui ont été définies dans l'étude du projet ESPRIT. Ce chapitre propose aussi une évaluation du système en terme de rappel et de précision, qui montre que la méthode permet au système et à l'utilisateur de mener des recherches évolutives et cohérentes. Une présentation d'un certain nombre d'évolutions et de perspectives termine ce chapitre.

Le dernier chapitre expose plus en détail une première évolution en montrant comment une approche hypertexte peut être intégrée pour gérer l'interrogation dans le système RIVAGE.

Chapitre 1	
Position du problème.....	151
I Introduction.....	151
II La reformulation de requête, un problème de "Relevance feedback" .	152
III Notions générales sur l'apprentissage symbolique.....	154
III.1. Généralités.....	154
III.2 Quelques méthodes.....	156
III.2.1 Généralisation par points de vue et apprentissage de concepts, de Quinqueton & Sallantin	156
III.2.2 INDUCE de Michalski.....	158
III.2.3 AGAPE de Kodratoff.....	160
III.2.4 L'espace des versions de Mitchell.....	161
III.3 Conclusion	164
IV Conclusion.....	165

Chapitre 1

Position du problème

I Introduction

L'apprentissage symbolique développé dans le domaine de l'IA trouve de nombreuses applications dans d'autres domaines où la modélisation d'une expertise ou d'une connaissance initiale est difficile [Mic 83]. Dans un cadre général nous dirons qu'un système apprend lorsqu'il est capable à partir de ses connaissances initiales :

- de formuler de nouvelles hypothèses, de nouveaux raisonnements, ceci sans modifier son comportement, ni sa connaissance initiale.
- de faire évoluer son comportement, en modifiant ses propres connaissances et son propre raisonnement.

Dans le domaine de la recherche documentaire, l'apprentissage symbolique peut être appliqué à divers niveaux [Smi 80], soit pour le système lui-même (modification des structures de données, des connaissances et des stratégies de recherche), soit au niveau de la recherche en elle-même. La forme d'apprentissage que nous proposons se situe à l'échelle d'une recherche. Quoique nous pensons un jour mettre en place un apprentissage plus global à l'échelle de la vie du système (cf partie I, chapitre 4, § III.3.2), l'étude de cette première forme nous a semblé prioritaire.

Le problème à l'origine est la reformulation de requête, il se définit de la manière suivante : comment proposer à l'utilisateur une nouvelle formulation de sa requête en prenant en compte :

- sa formulation initiale,
- ses étapes précédentes de recherche,
- ses choix d'images.

Il est alors intéressant de voir comment les SRI, proposant une reformulation de requêtes à partir de choix effectués sur un ensemble de documents, analysent ces choix et comment ils obtiennent une nouvelle formulation de la requête. Cette approche communément appelée "Relevance feedback" a été déjà présentée dans la partie I (chapitre 3, § I), elle est approfondie dans le paragraphe suivant par l'étude de sa réalisation avec des modèles booléens. Le second

paragraphe de ce chapitre tentera de montrer quelles sont les caractéristiques de l'apprentissage que nous allons mettre en œuvre, à partir d'une étude de quelques méthodes existantes.

II La reformulation de requête, un problème de "Relevance feedback"

La phase "Après-Visualisation" du processus EXPRIM a pour tâche de reformuler une nouvelle requête à partir d'un ensemble de descriptions choisies, d'un ensemble de descriptions rejetées et d'une ancienne requête. Ce problème a déjà été abordé à propos de certains systèmes de recherche d'informations ; en effet l'approche du "Relevance feedback" introduit dès 1965 par Rocchio [Roc 65] montrait déjà qu'une nouvelle requête pouvait être construite en ajoutant à la requête initiale des termes extraits des documents pertinents et en éliminant des termes présents dans les documents non pertinents (cf Partie I, ch 3 § I).

Les différentes études [Dil 80] [Sal 83c] montrent que lorsque l'on se place dans un environnement de recherche à partir de requêtes booléennes, la reformulation n'est pas évidente. Il faut d'abord trouver des termes pertinents à faire intervenir dans la requête future ; il faut aussi effectuer une synthèse avec ces termes, et en particulier leur associer les opérateurs (et, ou, sauf) nécessaires à leur intégration dans la nouvelle requête.

Dillon et Salton ont construit plusieurs méthodes en suivant l'enchaînement ci-dessous [Dil 80][Sal 83c]:

- Formulation d'une première requête booléenne atteignant un certain nombre de documents
- L'ensemble des documents sélectionnés est ensuite séparé en deux:
 - . les documents pertinents
 - . les documents non pertinents
- Pour chaque terme présent dans les documents, un poids de pertinence lui est associé prenant en compte les occurrences du terme dans les documents pertinents et dans les documents non pertinents.
 - Grâce à des seuils prédéfinis ou définis dynamiquement, les termes sont rangés dans des classes en fonction de la valeur de leur poids.
 - Une nouvelle requête booléenne est alors construite en prenant la disjonction des termes dans chaque classe, d'abord un par un pour la classe des termes de poids élevé puis deux par deux consécutifs pour la classe des termes de poids inférieur puis trois par trois etc...

exemple:

Considérons trois classes de termes :

Soit S_i les termes de la 1ère classe (les poids les plus élevés), P_i les termes de la 2ème classe

La nouvelle requête aura la forme suivante :

(S_1 ou S_2 ou ... ou S_p) ou

(P_1 et T_2) ou (P_3 et P_4) ou ... ou (P_{p-1} et P_p) ou

(T_1 et T_2 et T_3) ou ... ou (T_{t-2} et T_{t-1} et T_t)

Les méthodes diffèrent entre elles essentiellement par le calcul du poids de pertinence attaché à un terme. Mais certaines prennent en considération le nombre de documents que la nouvelle requête est susceptible d'atteindre, ce paramètre leur permet de contrôler l'adjonction de nouveaux termes dans la requête [Sal 83c].

exemple de calcul du poids de pertinence d'un terme [Dil 80] :

Soit un terme t , N le nombre total de documents trouvés, P le nombre de documents pertinents, p_t le nombre de documents pertinents contenant t , n_t le nombre de documents trouvés contenant t , n_t le nombre total de documents contenant t .

$$P_{per,t} = \frac{p_t}{\min(P, n_t)} - \frac{n_t - p_t}{\min(N - P, n_t)}$$

Comme le dit l'auteur, cette formule montre que l'importance d'un terme sera d'autant plus grande que le terme apparaîtra plus souvent (en proportion) dans les documents pertinents, moins souvent dans les documents non pertinents, et dans l'ensemble de tous les documents.

Ces méthodes que nous venons succinctement de présenter ont été testées sur des bases textuelles avec des résultats satisfaisants [Sal 83c]. Sur ce type de base la compréhension des besoins d'un utilisateur est assez complexe car émettre un avis sur un document textuel (article, livre ...) à partir d'un résumé ou du document lui-même demeure difficile et long. Dans le cadre d'une base images où l'interaction avec l'image est immédiate et complète, la sélection est plus aisée et plus pertinente. Il est alors intéressant d'essayer de comprendre les besoins de l'utilisateur illustrés par ses choix. Ces besoins peuvent être au départ mal définis mais, petit à petit, la phase "Après-Visualisation" va aider l'utilisateur à affiner sa recherche en identifiant les éléments qui la caractérisent.

Cette recherche de compréhension des besoins ne peut pas s'effectuer seulement à partir d'un simple calcul de pertinence des termes présents dans les documents sélectionnés. Une approche plus sémantique où le sens des termes est pris en compte est nécessaire. Les techniques d'apprentissage symbolique développées en IA nous semblent bien adaptées à ce genre de problème. Elles permettent de mettre en évidence des concepts grâce à une certaine compréhension du système face aux situations qu'il rencontre. Nous pensons qu'une telle compréhension est nécessaire si l'on veut qu'une recherche ait une progression cohérente d'étape en étape.

Avant d'en montrer une application effective dans notre système, il nous faut tout d'abord expliquer la philosophie de ces méthodes.

III Notions générales sur l'apprentissage symbolique

Nous souhaitons donner dans ce paragraphe un aperçu des méthodes d'apprentissage développées en IA. Il est important de noter que ces méthodes sont toutes adaptées à des problèmes assez spécifiques et qu'il n'existe pas de méthode suffisamment générale pouvant s'adapter à tous les problèmes. C'est pourquoi, nous allons tout d'abord montrer quelles peuvent être les caractéristiques générales permettant une identification de ces méthodes. Ainsi nous essayerons, à travers l'analyse de certaines d'entre-elles, de déterminer les propriétés que notre méthode doit posséder.

III.1. Généralités

Il existe diverses formes d'apprentissages [Mic 83] que l'on peut classer selon plusieurs critères allant de la stratégie sous-jacente jusqu'au type du langage de représentation du concept appris. En se basant sur la stratégie d'apprentissage, Michalski, Carbonell et Mitchell mettent en évidence cinq formes d'apprentissage non réellement distinctes:

- . Apprentissage par coeur, où les informations sont enregistrées sans subir de transformation.
- . Apprentissage à partir d'instructions; l'acquisition de connaissance est dirigée par une source extérieure. Le système transforme les connaissances et les insère dans celles déjà existantes.
- . Apprentissage par analogie; le système transforme sa connaissance pour l'adapter à une nouvelle situation.

- . Apprentissage à partir d'observations et de découvertes, correspond à l'apprentissage le plus général sur un système allant jusqu'à agir sur son mode de fonctionnement.
- . Apprentissage à partir d'exemples, le système doit induire une description générale reconnaissant les exemples et réfutant les contre-exemples.

C'est cette dernière forme d'apprentissage qui nous paraît la plus proche de nos besoins. En effet, la description générale du concept à apprendre, c'est la requête à formuler; les descriptions des images choisies correspondent aux exemples et les descriptions des images rejetées aux contre-exemples. Cette idée n'est pas complètement nouvelle puisque Wong & Ziarko [Won 86a][Won 86b] ont proposé une approche basée sur un apprentissage à partir d'exemples, de construction de règles de recherche à partir de documents jugés pertinents ou non pertinents par un utilisateur (cf partie I, ch 4, § III.3.1).

Michalski & Diettrich dans leur comparaison de méthodes d'apprentissage à partir d'exemples [Mic 83] ont défini différents aspects permettant de caractériser ces méthodes:

- . La représentation: c'est-à-dire la définition du langage employé pour décrire les exemples et les concepts appris. De nombreuses représentations sont possibles allant du calcul des prédicats jusqu'aux réseaux sémantiques. La représentation doit à la fois définir la forme des exemples et des concepts ainsi que les opérateurs possibles (conjonction, disjonction, exception, quantificateurs ...).
- . Le type des descriptions générées: il correspond au type de la description obtenue à partir des exemples. La description peut être:
 - . Description caractéristique: elle permet de reconnaître une classe d'objets parmi d'autres objets.
 - . Description discriminante: description d'une classe d'objets construite avec juste les éléments qui permettent de différencier les objets des autres.
 - . Description taxonomique: description d'une classe d'objets divisée en sous-classes.
- . Les règles de généralisation: Ces règles permettent de rendre plus générale une description. Elles définissent les types de transformations possibles sur les exemples.
- . Induction constructive: Une méthode fait de l'induction constructive lorsqu'elle modifie l'espace des descriptions, par exemple en générant de nouveaux descripteurs.

- . La stratégie de contrôle: Deux stratégies peuvent être mises en évidence:

- . Dirigée par les données: les exemples sont traités un par un, ils sont mis en confrontation avec la généralisation courante. Celle-ci est modifiée lorsqu'elle ne reconnaît pas les exemples.
- . Dirigée par un modèle: le processus génère un ensemble ordonné de formes généralisées à partir des exemples (une forme généralisée est l'expression d'un concept, plus général que les exemples). Le processus s'arrête lorsqu'un certain nombre de critères sont vérifiés.

Voyons maintenant quelques méthodes en essayant de les caractériser à l'aide des propriétés citées ci-dessus.

III.2 Quelques méthodes

III.2.1 Généralisation par points de vue et apprentissage de concepts, de Quinqueton & Sallantin

La méthode proposée par Quinqueton & Sallantin [Qui 84] dans leur système CALM [Qui 87], a pour but d'expliquer un concept à partir d'exemples. Les exemples sont décrits avec des attributs ayant une valeur dans chacun des exemples. Un attribut est caractérisé par un ensemble de valeurs possibles. Leur approche consiste à construire des "points de vue" caractérisant les exemples. Un point de vue est en fait une forme généralisée expliquant la valeur d'un attribut (dans l'exemple l'attribut est "motif" et les valeurs sont : "présent ou non présent") en fonction des valeurs des autres descripteurs.

La construction des points de vue s'effectue en itérant sur trois opérations: expansion, sélection et compression:

- . L'expansion combine les descripteurs deux à deux avec un opérateur logique, ce qui donne un nouvel ensemble de "descripteurs étendus".
- . La sélection élimine les descripteurs qui ne caractérisent pas assez le concept à expliquer au sens d'une certaine mesure. Nous obtenons ainsi un ensemble de descripteurs étendus sélectionnés.
- . La compression classe les descripteurs étendus sélectionnés en fonction de leur "ressemblance" (analogue à celle de l'analyse de donnée) et définit pour chaque classe un "descripteur résumé". Ensuite un point de vue est défini pour chaque classe, il est exprimé par le descripteur résumé de la classe.

Exemple: (extrait de [Qui 84]).

Les **exemples** ci-dessous représentent des descriptions de drapeaux de pays:

Pays:	Description
VIETNAM	motif, central, étoile
ALBANIE	motif, central, étoile
CONGO	motif, haut gauche
URSS	motif, haut gauche
CAMBODGE	motif, central
MAROC	motif, central, étoile
CHINE	motif, haut gauche, étoile
TURQUIE	motif, central, étoile, croissant
TUNISIE	motif, central, étoile, croissant, rond
DANEMARK	croix
NORVEGE	croix
SUISSE	motif, central, croix

Le descripteur à expliquer est: **motif**

Après une expansion et une sélection on obtient les descripteurs étendus sélectionnés suivant:

Descripteur	N°
central	1
étoile	2
non croix	3
non haut gauche et central	4
non haut gauche et non croix	5
central et non rond	6
central et non croix	7
étoile et non croix	8
non croissant et non croix	9
non rond et non croix	10

Après une compression, les "points de vue" sont obtenus à partir des classes de descripteurs étendus sélectionnés. Ils ont la forme suivante: "descripteur résumé de la classe -> descripteur à expliquer", ce qui a pour signification: "Ce descripteur résumé et ceux de sa classe expliquent la présence dans les exemples du descripteur à expliquer".

Voici les points de vues:

non haut gauche et **non** croix -> **motif** avec la classe de descripteurs [5,7]

étoile -> **motif** avec [2,8]

non croissant et **non** croix -> **motif** avec [9,10]

central -> **motif** avec [1,4]

non croix -> **motif** avec [3]

central et **non** rond -> **motif** avec [6]

Cette méthode s'inscrit dans les méthodes "dirigées par un modèle". Puisque l'algorithme génère constamment des formes généralisées, il faut déterminer un critère d'arrêt qui peut être fondé sur la complexité des formes généralisées ou sur le nombre d'itérations. Les descriptions générées (formes généralisées) sont des descriptions caractéristiques puisque l'algorithme cherche à reconnaître les propriétés du concept, qu'elles soient discriminantes ou pas. Il n'existe pas réellement de règles de généralisation autres que la création des descriptions généralisées.

La méthode de Quinqueton & Sallantin est originale puisqu'elle repose sur une logique "majoritaire" [Qui 87] qui définit qu'une propriété est vérifiée sur l'ensemble des exemples si une majorité de ces exemples la possèdent. Cette approche nous semble intéressante dans le contexte d'un ensemble de descriptions d'images où l'imprécision est souvent présente. L'inconvénient de cette méthode est le manque d'une généralisation fondée sur une hiérarchie de descripteurs. Cette généralisation est pour nous essentielle compte tenu de l'importance du thesaurus des descripteurs utilisés dans les descriptions.

III.2.2 INDUCE de Michalski

Michalski & Larson [Lar 77] proposent une méthode d'apprentissage guidée par un modèle. Le langage utilisé pour décrire les exemples est le VL₂₁ qui est une extension de la logique des prédicats du premier ordre spécialisée pour l'induction.

Chaque exemple est représenté par une conjonction de sélecteurs, un sélecteur pouvant être une fonction ou un prédicat descripteur avec variables associé à une liste de valeurs que le prédicat peut prendre.

exemple: [Forme(v_1)=carré, rectangle]

RQ: Cette forme de représentation permet de représenter des disjonctions dans les exemples grâce à la liste des valeurs possibles associée au descripteur.

Le principe de la méthode est d'effectuer deux recherches de description, la deuxième affinant la première. La première description, générale, est construite à partir des descripteurs décrivant les objets par leurs relations (ex: descripteur de position : En-dessous(v_1, v_2)). La seconde description permet de détailler la première en introduisant les "descripteurs attributs" décrivant les propriétés des objets (ex: forme, taille,...).

L'algorithme construit à partir des exemples un ensemble d'hypothèses de formes généralisées. Ces formes généralisées sont ordonnées en fonction de plusieurs critères:

- . Le nombre d'exemples recouverts par la généralisation
- . Le nombre de sélecteurs contenus dans la généralisation
- etc...

Lorsque cet ensemble atteint une taille prédéfinie, la recherche s'arrête. Le système va alors compléter les hypothèses en les généralisant pour qu'elles décrivent le plus d'exemples possible et en les spécialisant pour qu'elles réfutent les contre-exemples.

Les règles de généralisation utilisées sont les suivantes:

- . Suppression de conjonction
- . Transformation des constantes en variables
- . Ajout de disjonction
- . Augmentation du domaine d'un intervalle
- . Grimper dans un arbre de généralisation (taxonomie de concept).

Ces règles sont applicables car le langage de représentation des exemples est assez riche puisqu'il comprend trois types de descripteurs:

- . Descripteur nominal à valeur dans un ensemble non ordonné
- . Descripteur linéaire à valeur dans un ensemble ordonné
- . Descripteur structuré à valeur dans un univers où il existe des relations de taxonomie entre les concepts.

Il existe une analogie entre les descripteurs manipulés dans INDUCE et ceux que nous utilisons dans les descriptions d'images, c'est pourquoi nous nous intéresserons plus tard à l'application des règles de généralisations de INDUCE dans la phase "Avant-Visualisation". Nous pouvons aussi remarquer que le modèle permettant de décrire les exemples est assez riche par rapport au modèle de documents classiquement utilisé en recherche d'informations. En effet,

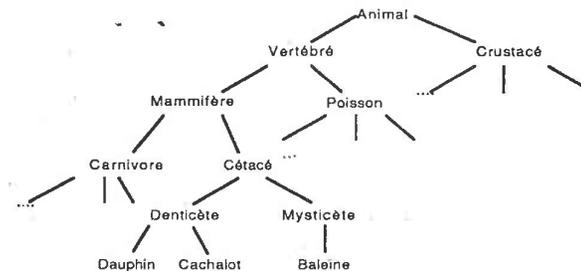
des relations entre objets présents dans les exemples peuvent être prises en compte ; ce qui augmente de beaucoup la complexité des algorithmes effectuant la recherche de caractéristiques. Les descriptions des images n'ayant pas une telle richesse, cette méthode ne peut être appliquée dans son ensemble. Une évolution des modèles de correspondance associés aux systèmes de recherche d'images, est nécessaire pour qu'une telle méthode soit applicable.

III.2.3 AGAPE de Kodratoff

Le système AGAPE proposé par Y. Kodratoff [Kod 85] met en oeuvre un algorithme "dirigé par les données" appelé "Appariement structurel" [Kod 86].

Les exemples à généraliser sont décrits avec des prédicats du premier ordre, chacun représentant une propriété de l'exemple. Le système possède des connaissances exprimées par des axiomes définis sur les prédicats et par des taxonomies de prédicats:

exemple de taxonomie de prédicats:



L'algorithme de généralisation comporte quatre étapes distinctes:

. L'appariement structurel où le système en plusieurs étapes, essaye de trouver des expressions plus générales des exemples, permettant de mettre en évidence des généralisations communes. La trace des transformations de la généralisation est gardée sous forme de liens attachés à l'expressions des exemples.

exemple:

E1: Baleine(Moby-Dick)

E2: Dauphin(Flipper)

après l'appariement structurel:

E1: Cétacé(u v) (x) [(u=Mysticète) (v=Baleine) (x=Moby-Dick)]

E2: Cétacé(u v) (x) [(u=Denticète) (v=Dauphin) (x=Flipper)]

Les liens sont exprimés entre crochets

. L'adjonction de liens: elle consiste à ajouter aux liens existants d'autres liens que l'appariement structurel n'aurait pas mis en évidence, tels que par exemple le lien de non égalité entre deux variables ($x \neq y$).

. L'élimination des différences où les prédicats et les liens non communs à tous les exemples sont supprimés. Cette étape constitue la partie inductive de l'algorithme puisqu'une forme généralisée est ainsi obtenue.

exemple:

G: Cétacé(u v) (x)

. La simplification où les liens d'égalité restants sont propagés dans l'expression et où les variables de prédicat non instanciées sont éliminées.

L'algorithme mis en oeuvre par Kodratoff a l'avantage d'être plus "rationnel" que les méthodes présentées précédemment puisqu'il procède à une analyse exhaustive des exemples, et qu'il n'applique les règles de généralisation que lorsque c'est nécessaire, ceci dans le but de créer une généralisation ou une description la plus caractéristique possible. Mais cette approche par les données nous semble inadaptée à nos problèmes puisque trop rigide et stricte, les connaissances contenues dans notre système n'étant pas assez rigoureuses et certaines. Une approche plus empirique nous paraît plus facilement adaptable.

III.2.4 L'espace des versions de Mitchell

La méthode que nous allons décrire maintenant a été utilisée par Mitchell dans son système LEX [Mit 83] réalisant la résolution d'intégration formelle. Ce système a été conçu pour apprendre les méthodes d'intégration à partir de ses propres expériences négatives et positives. Cet apprentissage peut être assimilé à un apprentissage à partir d'observations et de découvertes, mais il peut être adapté à l'acquisition de concept à partir d'exemples et de contre-exemples comme l'a montré Kodratoff dans [Kod 86].

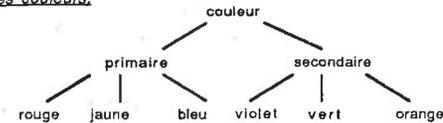
L'espace des versions se construit à partir d'un ensemble de taxonomies des descripteurs utilisés dans les exemples et contre-exemples. La finalité du processus est de définir un niveau de généralité d'un concept à l'intérieur des taxonomies. L'espace se dessine en faisant évoluer

des marqueurs sur les noeuds des arbres représentant les taxonomies. Voyons plus précisément sur un exemple extrait de [Kod 86] le déroulement de l'algorithme:

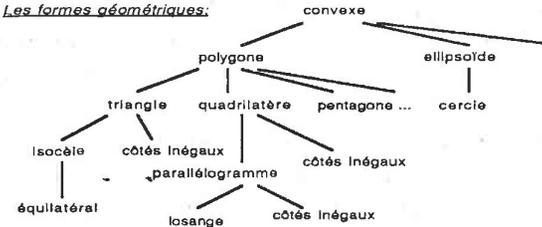
exemple:

Donnons-nous un ensemble de taxonomies:

les couleurs:



Les formes géométriques:



et des exemples:

E1: équilatéral & jaune

E2: losange & rouge

E3: triangle à côtés inégaux & bleu

E4: parallélogramme à côtés inégaux & rouge.

Au départ, on introduit les marqueurs de la manière suivante:

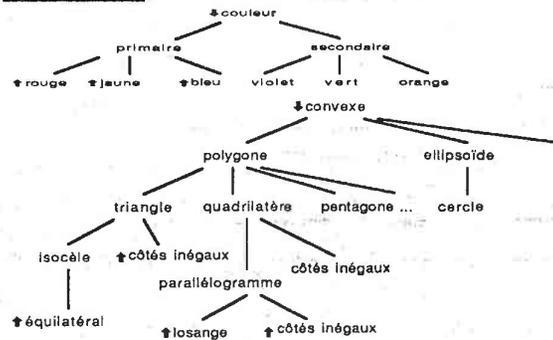
- . tous les descripteurs présents dans les exemples sont marqués d'une flèche dirigée vers le haut,
- . toutes les racines sont fléchées vers le bas.

Ce marquage a pour signification:

- . si un noeud est fléché vers le haut alors lui et tous ses fils sont des descripteurs du concept à apprendre,
- . si un noeud est fléché vers le bas alors lui et ses fils sont des descripteurs possibles du concept.

L'étude des exemples, des contre-exemples et l'application de règles de transformation va nous permettre de faire évoluer les flèches.

Situation initiale:



A ce stade on utilise les règles de transformation, l'une de généralisation et une de spécialisation à partir des contre-exemples.

Règle de généralisation:

Si tous les fils d'un descripteur sont fléchés vers le haut alors flécher le descripteur père vers le haut.

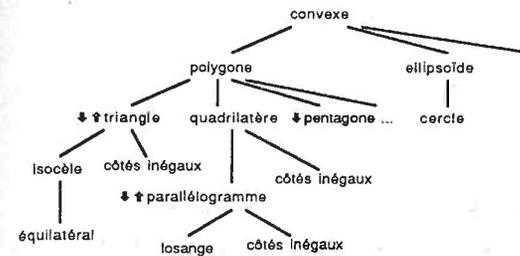
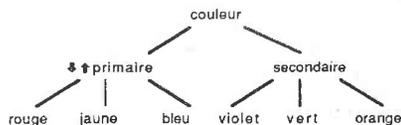
Règle de spécialisation:

Si un descripteur est présent dans un contre-exemple et qu'il est non présent dans les exemples alors descendre la flèche dirigée vers le bas du plus petit niveau possible permettant d'exclure le descripteur.

Soit les contre-exemples:

- CE1: violet & cercle
- CE2 : quadrilatère à cotés inégaux & vert

L'espace des versions après transformation sera:



Les noeuds fléchés à la fois vers le haut et vers le bas représentent les niveaux de généralité du concept appris dans les taxonomies. Lorsqu'il n'existe pas de noeuds de cette sorte, il faut évaluer le niveau de généralité en considérant les noeuds fléchés vers le bas les plus bas et les noeuds fléchés vers le haut les plus hauts. Ce niveau de généralité doit permettre la formulation du concept appris.

Dans l'exemple, la formulation du concept appris peut être la suivante :

(Triangle ou parallélogramme) & couleur primaire.

Ce concept reconnaît les exemples et réfute les contre-exemples.

Cette méthode est intéressante pour le cheminement qu'elle propose dans les taxonomies. Si nous considérons le thesaurus comme un ensemble de taxonomies de concepts, cette approche peut nous aider à évaluer un niveau de généralité de recherche dans le thesaurus. Il est bien évident que cette méthode ne peut s'appliquer globalement dans ses principes de bases, elle doit être adaptée compte tenu de la structure de notre thesaurus, des exemples et contre-exemples que représentent les descriptions d'images.

III.3 Conclusion

Cette étude succincte de la problématique de l'apprentissage symbolique et de quelques méthodes nous a permis de mettre en relief une potentialité de solutions à notre problème. Ainsi, un apprentissage à partir d'exemples, basé sur un ensemble de taxonomies, semble correspondre à notre besoin. L'approche de Mitchell paraît intéressante, mais elle doit être adaptée aux caractéristiques de notre système. Une approche dite "majoritaire", inspirée de Quinqueton & Sallentin, devrait permettre de combler l'imprécision présente dans nos exemples, c'est-à-dire les descriptions. Cette imprécision, mais aussi la quantité d'informations que notre système doit manipuler à un instant donné, rendent les approches d'INDUCE et d'AGAPE inadaptées.

Notre approche tentera donc de calquer le principe de la méthode de Mitchell tout en prenant en compte le concept de "majorité".

IV Conclusion

Le processus EXPRIM propose, dans la phase Après-Visualisation, une reformulation de requêtes. Cette reformulation de requêtes pourrait s'effectuer à l'aide de méthodes classiques de "Relevance feedback" où le sens des termes n'est pas pris en compte. Mais l'existence du thesaurus permet de voir le problème différemment ; une approche plus sémantique peut être considérée.

Le domaine de l'apprentissage symbolique nous offre des méthodes permettant d'évaluer le sens d'un concept à partir de structures sémantiques proches d'un thesaurus.

Après une étude de ces méthodes, il nous semble possible de proposer une approche permettant de considérer le processus EXPRIM comme un processus d'apprentissage. L'originalité de notre méthode, comme nous allons le voir dans la suite, est de placer l'approche classique du "Relevance feedback" dans un cadre sémantique inspiré des méthodes d'apprentissage symbolique.

Chapitre 2

La méthode proposée

I Introduction

L'assimilation du processus EXPRIM à un problème d'apprentissage offre un cadre sémantique cohérent pour l'ensemble du processus. Grâce à cette propriété, le processus EXPRIM peut être considéré dans sa totalité et non phase par phase comme dans les précédentes approches (cf partie 2, chapitre 2 & partie 3). Cette nouvelle approche permet d'avoir une représentation du besoin de l'utilisateur, un moyen de l'exprimer sous la forme d'un concept appris, et aussi des techniques de déformation et d'exploitation de ce concept quand celui-ci rapporte "peu d'images" ou trop "d'images".

Ce chapitre présente la méthode tout d'abord d'une manière intuitive qui, nous l'espérons, donne une compréhension globale du processus d'apprentissage. Ensuite, nous étudions le rôle des données principales impliquées dans la méthode. Enfin, dans le dernier paragraphe, la méthode est présentée d'une manière plus détaillée, ceci phase par phase, avec l'appui d'un exemple.

Avant de présenter la méthode, insistons sur le fait que nous souhaitons réaliser cet apprentissage avec une profonde interactivité entre l'utilisateur et le système car dans ce type de recherche ils sont complémentaires. En effet, l'utilisateur voit les images et le système "voit" leur moyen d'accès : les descriptions.

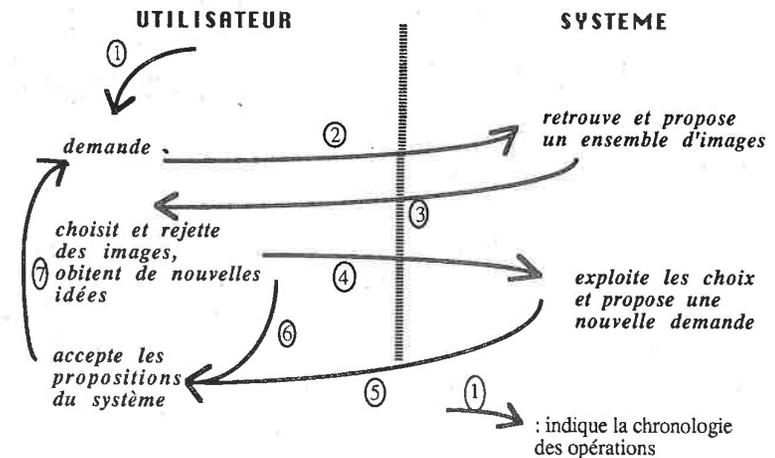
II Présentation intuitive de la méthode

L'apprentissage que nous proposons [Hal 88a][Hal 89a][Cre 89a][Cre 89b][Cre 89c] est un apprentissage à partir d'exemples où la demande correspond à une expression du concept à apprendre, et où les descriptions des images choisies correspondent aux exemples du concept et les descriptions des images rejetées aux contre-exemples. La méthode consiste à représenter dans le thesaurus l'expression de la demande par un **niveau d'expression** construit sous forme d'un marquage des termes. Le marquage d'un terme mesure sa capacité à exprimer la

demande. Cette représentation du niveau d'expression à l'intérieur du thesaurus est inspirée de l'Espace des Versions de Mitchell (cf chapitre 1, § III.2.4).

Le niveau d'expression du concept caractérisant la recherche est représenté dans le thesaurus en attachant à chaque terme un **poids d'expressivité**. Ce poids sera élevé si le terme exprime bien la demande. Si un terme possède un poids faible, cela signifie qu'il n'exprime pas correctement le concept de la demande et qu'il faut choisir parmi ses spécifiques ceux ou celui correspondant le mieux à l'expression de la demande.

Cette forme d'apprentissage permet d'offrir une très riche interactivité qui peut se représenter par le schéma suivant extrait de [Cre 89a] :



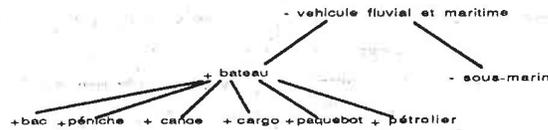
Cette interactivité est caractérisée par de nombreux échanges entre le système et l'utilisateur qui vont permettre à ce dernier de définir progressivement ses besoins

L'exemple suivant illustre l'échange n° 5 entre le système et l'utilisateur du schéma précédent.

Exemple :

Voici un morceau de thesaurus juste avant une reformulation, le niveau d'expression est représenté sur le dessin par :

- : si le terme n'exprime pas le concept, ce qui est révélé par un poids faible,
- + : si le terme exprime le concept: son poids est important :



Après un parcours du thesaurus nous allons obtenir :

Demande: contenu de la photo = bateau.

Ceci signifie que *bateau* est assez général, mais pas trop, pour exprimer les besoins de l'utilisateur.

Cet apprentissage est constitué essentiellement de deux grandes étapes :

- La première consiste à faire évoluer le niveau d'expression du concept de la demande en rapport avec la progression de la recherche.
- La deuxième étape consiste à utiliser le niveau d'expression exprimé dans le thesaurus pour bâtir une nouvelle expression du concept c'est-à-dire une nouvelle demande.

Mais indiquons quels sont la structure et le rôle des principales données impliquées : le thesaurus, les descriptions, la demande.

III Description des données impliquées dans la méthode

Les trois principales structures de donnée de notre système jouent chacune un rôle important dans la méthode. En effet, le thesaurus a le statut d'une connaissance sémantique, mais il est aussi le support permettant de représenter le concept à apprendre. Les descriptions des images jouent le rôle des exemples ; c'est leur structure qui définit la précision du concept que l'on peut apprendre. Enfin, la demande représente une formulation du concept qui est obtenue à partir du thesaurus.

III.1 Le thesaurus, la connaissance sémantique

Le vocabulaire est l'ensemble des descripteurs pouvant apparaître dans les descriptions. Sa composante principale est le thesaurus, qui constitue une importante connaissance sémantique.

Nous n'allons pas reprendre ici la présentation de la structure d'un thesaurus, ni celle de son rôle dans un SRI. Mais le cadre de cette méthode nous oblige à considérer le thesaurus sous un nouvel aspect : celui de support sémantique pour un apprentissage de concept.

III.1.1 Généralités sur les taxonomies

Un des supports habituels pour la représentation de concepts est la taxonomie. Une taxonomie, pour qu'elle puisse être utilisée lors d'un apprentissage automatique, doit posséder certaines propriétés dont l'une des plus importantes est sans nul doute :

- les fils d'un concept épuisent toutes les possibilités du concept.

Grâce à cette propriété, on peut dire qu'un concept est appris si toutes les possibilités de représentations de ce concept ont été présentes dans les exemples.

En se plaçant dans un univers bien circonscrit, les taxonomies peuvent facilement vérifier cette propriété. Par contre, dans un univers où les limites ne sont pas bien définies, cette propriété ne peut pas toujours être vérifiée.

En général, l'univers d'une recherche d'images est assez vaste ; un thesaurus ne peut en être une représentation de l'ensemble, il ne peut pas alors être considéré comme une taxonomie. Voyons comment on peut adapter la structure du thesaurus pour qu'il puisse avoir le statut de taxonomie.

III.1.2 Le thesaurus vu comme un ensemble de taxonomies

Le thesaurus est composé d'un ensemble de micro-thesaurus représentant différents domaines de connaissances (Architecture, Agriculture, Morphologie de la photographie, ...). Un micro-thesaurus est un ensemble de termes illustrant le domaine considéré. Les termes sont reliés entre eux par des relations sémantiques. La relation la plus importante est la relation "générique/spécifique" qui définit une hiérarchie entre les termes.

Exemple :

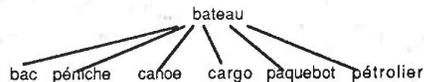
Cette relation va nous permettre de considérer le thesaurus comme un ensemble de taxonomies, une taxonomy étant un champ sémantique particulier.

Compte tenu de la propriété sur les taxonomies citée plus haut, pour que le thesaurus ait le statut de taxonomy, il faut que les spécifiques des termes génériques épuisent toutes les possibilités de représentation du concept exprimé par leur terme générique.

Mais l'utilisation classique du thesaurus lors de la phase d'indexation va à l'encontre de la vérification de cette propriété. En effet, malgré la méthode d'indexation, associée à l'utilisation d'un thesaurus, dite d'"indexation aux feuilles" c'est-à-dire "au plus spécifique", il existe des cas où la présence d'un concept ne peut être identifiée par une de ses représentations possibles, c'est-à-dire par un de ses spécifiques. Dans ce cas de figure c'est en général le terme générique qui est utilisé, donc le concept lui-même. Une ambiguïté est alors introduite, car dans ce cas le document est indexé avec un concept qu'il n'illustre pas entièrement.

Exemple :

Soit la taxonomy (ou partie du thesaurus) suivante :

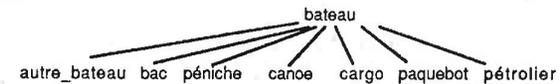


Si la photographie à décrire contient de nombreux bateaux (l'énumération de toutes les sortes présentes serait trop longue), ou si elle contient un bateau de si petite taille qu'on ne peut identifier son type, il est impossible, avec cette taxonomy, de décrire cette photographie sans introduire de l'ambiguïté. Avec la méthode classique une ambiguïté est introduite, car la personne décrivant la photographie utilisera, en règle générale, le terme "bateau".

Cette méthode d'indexation illustre bien le fait que le thesaurus ne peut être assimilé à une taxonomy car il existe des représentations de terme générique qui ne sont pas identifiées par ses spécifiques. Il est alors impossible d'apprendre avec cette structure, car on ne peut pas dire qu'un concept est appris lorsque toutes ces représentations ont été présentes dans les exemples.

Pour palier à ce problème, nous proposons de créer, pour tous les termes génériques n'ayant pas suffisamment de spécifiques pour épuiser toutes ses représentations, un nouveau terme spécifique dit "fourre-tout". Ce nouveau spécifique permettra d'identifier toutes les représentations du concept qui ne sont pas des représentations des spécifiques existants.

Exemple :



Ici le terme "autre_bateau" a pour charge d'identifier les représentations suivantes :

- beaucoup de bateaux,
- bateau non identifiables.

Cette modification de structure permet d'une part, d'assimiler le thesaurus comme un ensemble de taxonomies, et d'autre part elle permet aussi de réaliser réellement ce qu'on appelle l'"indexation aux feuilles".

Remarque : Pour être encore plus proche d'une représentation taxonomique nous avons limité la structure de notre thesaurus à une mono-hiérarchie (un terme ne peut avoir qu'un terme générique).

III.2 Les descriptions d'images, exemples du concept

Les descriptions des images constituent les exemples (ou contre-exemples). Le langage d'indexation utilisé pour décrire les images est simple ; une description est formée d'un ensemble de champs mono ou multivalués, le mode d'indexation de ces champs est une indexation par juxtaposition. On peut comme dans INDUCE (cf chapitre 2, § III.2.2) caractériser les champs (ou attributs) en fonction de la nature des descripteurs valant ces champs : les champs à valeur dans un ensemble non structuré, les champs à valeur dans un ensemble structuré, les champs à valeur dans un ensemble ordonné.

Ainsi, il existe des champs à valeurs dans une liste, des champs à valeurs dans un ensemble de champs sémantiques du thesaurus, et des champs de type date. Pour des raisons que nous développerons plus tard, nous avons ajouté une contrainte dans le modèle d'indexation : un champ sémantique du thesaurus ne peut être associé qu'à un et un seul champ de la description.

Exemple :

Soit la description d'une photographie :

Auteur de la photographie : ATGET JEAN-EUGENE.

Contenu de la photographie : FONTAINE ADOSSEE, PARC.

Morphologie de la photographie : PLAN D'ENSEMBLE.

Date de la photographie : 1900.

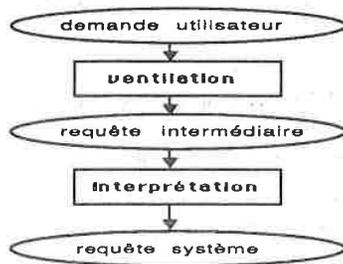
Dans cet exemple le champ "Auteur de la photographie" prend ses valeurs dans une liste de nom propre, le champ "Contenu de la photographie" dans un ensemble de champs sémantiques du thesaurus, le champ "Morphologie de la photographie" dans un champ sémantique appelé "Analyse morphologique du document" et enfin la date dans un ensemble ordonné.

Chaque image possède une et une seule description. L'accès aux descriptions s'effectue par l'intermédiaire d'une requête booléenne, combinaison des valeurs que l'on souhaite trouver dans les champs des descriptions des images recherchées. Les requêtes booléennes utilisées dans notre système sont générées lors de la phase "Avant-Visualisation", en interprétant les besoins de l'utilisateur, exprimés sous la forme d'une demande, et en utilisant le thesaurus. La mise en correspondance et l'interprétation de la demande feront l'objet d'un paragraphe particulier.

III.3 La demande, le concept à apprendre

La demande constitue le concept à apprendre ; tout au long des étapes de recherche, le système va aider l'utilisateur à formuler sa demande par raffinements successifs.

La demande possède trois formes différentes liées par des transformations :



-La première forme, appelée **demande utilisateur**, est orientée utilisateur : elle lui offre la possibilité d'exprimer ses besoins simplement en remplissant des "facettes" qui définissent les propriétés de l'(des) image(s) recherchée(s). Les facettes représentent des caractéristiques de l'image plus générales que les champs des descriptions. Elles permettent à l'utilisateur d'avoir une vue plus macroscopique des images. Cette forme de demande permet un dialogue avec l'utilisateur à un niveau de généralité proche de ses besoins ; de plus, elle le dispense d'utiliser des opérateurs booléens. A une facette peut être associé 1 ou plusieurs champs de la description.

- La **requête dite "intermédiaire"** est la deuxième forme de la demande, elle est obtenue à partir de la demande utilisateur en ventilant les termes valant les facettes dans les champs associés. Cette ventilation, que nous étudierons dans la suite, est effectuée grâce au champ sémantique des termes.

-La troisième forme de la demande est orientée système (**requête système**), c'est elle qui va permettre l'accès aux descriptions. Sa structure est plus complexe, elle est définie par une formule booléenne dont les atomes sont constitués d'un nom de champ (parmi ceux utilisés dans les descriptions) et d'une valeur. Cette requête est obtenue à partir de la requête intermédiaire en faisant une interprétation de son contenu essentiellement basée sur la structure sémantique du thesaurus.

Après cette description des principales données intervenant dans le processus d'apprentissage, voyons comment le niveau d'expression de la demande est représenté dans le thesaurus et comment il évolue globalement de phase en phase.

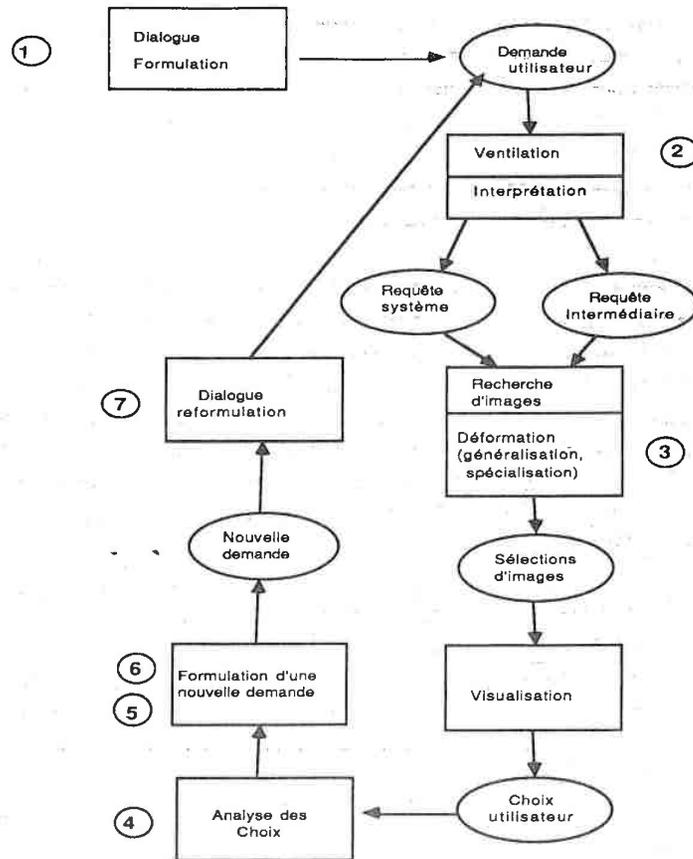
III.4 Niveau d'expression de la demande

III.4.1 Représentation

Le niveau d'expression de la demande est représenté dans le thesaurus en attachant à chaque terme un poids d'expressivité. Ce poids permet de mesurer la capacité d'un terme et de ses spécificités d'exprimer les besoins de l'utilisateur. La mesure de ce niveau d'expression, présentée dans la suite, prend en compte la définition des structures de données citées dans le paragraphe précédent (thesaurus, descriptions, demande).

III.4.2 Evolution globale

Le processus d'apprentissage que nous proposons, s'appuie sur le processus EXPRIM en trois phases. Le niveau d'expression de la demande évolue de phase en phase, cette évolution peut être résumée par la figure suivante :

**Légende:**

- 1 : Initialisation du niveau d'expression,
- 2 : Construction du premier niveau d'expression,
- 3 : Répercussion des transformations soumises à la requête intermédiaire,
- 4 : Calcul de pertinence des termes d'après les choix de l'utilisateur,
- 5 : Propagation du calcul de pertinence (étape 4) sur le niveau d'expression,
- 6 : Formulation d'une nouvelle demande à partir du niveau d'expression,
- 7 : Adaptation du niveau d'expression compte tenu des interactions avec l'utilisateur.

Le niveau d'expression évolue, les poids d'expressivité sont mis à jour à différentes étapes du processus de recherche ; nous proposons dans la suite un mode de calcul heuristique de cette

évolution. Nous illustrons l'évolution par un exemple unique que nous menons d'étape en étape. Etudions plus particulièrement son évolution pendant la phase "Avant-Visualisation".

IV Evolution du poids d'expressivité pendant la phase "Avant-Visualisation"

La valeur du poids d'expressivité est comprise entre -1 et 1. Un terme ayant un poids proche de -1 est mauvais pour exprimer le concept de la demande, par contre un terme de poids proche de 1 est nécessaire pour exprimer correctement la demande.

IV.1 Mise en place d'un premier niveau

Initialisation (1) :

Compte tenu de l'ignorance initiale des besoins de l'utilisateur, nous supposons que tous les concepts présents dans le thesaurus sont mauvais ; tous les poids d'expressivité sont mis à -1.

Après la ventilation de la demande (2) :

Chaque terme présent dans la requête intermédiaire a son poids d'expressivité mis à 0 ainsi que tous ses spécifiques de tous niveaux. Les poids sont alors propagés dans l'ensemble du thesaurus par la relation générique/spécifique.

Cette mise à zéro indique que les concepts introduits par l'utilisateur sont corrects mais non réellement appris, car seul le choix d'images peut valider un concept. En effet, les concepts utilisés par l'utilisateur n'ont pas forcément la même représentation dans la base d'images que dans l'esprit de l'utilisateur.

La répercussion aux termes génériques (non présents dans la demande) s'effectue ainsi: Soit t_g un terme générique, t_{si} ($i \in [1, n]$) ses spécifiques directs et $Pexp_i$ leur poids d'expressivité ; alors le poids d'expressivité $Pexp_{t_g}$ de t_g est défini par :

$$Pexp_{t_g} = F(Pexp_i)$$

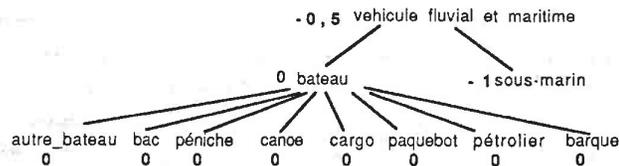
La fonction F doit permettre au poids $Pexp_{t_g}$ d'être représentatif des poids d'expressivité des spécifiques, une solution satisfaisante est de prendre la moyenne des poids ($Pexp_{t_g} = MOY(Pexp_i)$) comme l'illustre l'exemple.

Ainsi, les poids d'expressivité vont décroître en montant vers la racine du thesaurus.

L'exemple :

Demande : contenu de la photo = bateau.

La partie du thesaurus à proximité du terme *bateau*, après la mise à jour des poids d'expressivité, a la forme suivante :



Lors de la déformation (3) : la déformation intervient lorsque le résultat de la recherche d'images n'a pas été jugé satisfaisant. Les règles de déformation (généralisation, spécialisation, suppression, adjonction) peuvent agir soit sur la requête intermédiaire, soit sur la requête système. Mais pour expliquer quelles peuvent être les modifications du niveau associées à ces déformations, il nous faut voir plus en détail comment s'effectue la mise en correspondance entre la demande et les descriptions d'images.

IV.2 La mise en correspondance entre demande et descriptions d'images

Le problème de la mise en correspondance de notre processus se pose ainsi [Mou 89b] : comment à partir d'une demande utilisateur sélectionner un ensemble de descriptions d'images dites pertinentes pour cette demande.

Pour formaliser cette mise en correspondance nous allons décrire les structures impliquées en proposant une série de modèles. Cette formalisation n'a pas la prétention d'être des plus rigoureuses, elle nous est utile pour expliquer d'une manière assez précise comment la mise en correspondance s'effectue.

IV.2.1 Le modèle de mise en correspondance

Notre modèle de mise en correspondance est constitué d'un thesaurus, d'un modèle général de document utilisé pour les descriptions des images, et d'un modèle de demande : (TH, MDO, MDE).

1) Le thesaurus TH est constitué d'un ensemble de champs sémantiques contenant un ensemble d'unités organisé de façon arborescente par la relation générique/spécifique. Une unité connaît alors son père (le terme générique), ses fils (les termes spécifiques) et les unités ayant un terme de sens proche (les termes reliés par la relation "terme associé". Dans ce modèle

les relations du thesaurus utilisées sont ces deux relations (générique/spécifique, terme associé) car nous connaissons et contrôlons leur sémantique et leur effet dans le processus de recherche ; tout autre ajout de nouvelle relation dans le thesaurus nécessitera une étude approfondie sur la sémantique de cette relation et sur la portée de cette sémantique dans le processus de recherche.

TH = CHS avec CHS = { chs_m },

$chs_m = (\{ u_i \}, r_m)$ avec $u_i = (t_i, SPE_i, g_i, TA_i, Pexp_i)$ où t_i est le terme, SPE_i représente l'ensemble des unités des termes spécifiques du terme t_i : $SPE_i = \{ u_j \}$, g_i est l'unité du terme générique, TA_i contient les unités des termes associés du terme t_i , $TA_i = \{ u_k \}$, $Pexp_i$ représente le poids d'expressivité du terme t_i ; r_m est l'unité la plus générique du champ sémantique (racine de l'arbre).

2) Le modèle général de document MDO est constitué d'un ensemble de nom de champs C (exemple : contenu de la photographie, auteur de la photographie, morphologie de la photographie ...), d'un ensemble de couples liant les champs à leur domaine de valeurs CD. un domaine de valeurs est constitué d'un ensemble de champs sémantiques du thesaurus. Le modèle général de document contient aussi le modèle de document lui même MD qui est, comme nous l'avons déjà indiqué, un ensemble de couples associant un champ à une liste d'unités :

MDO = (C, CD, MD) avec

C = { c_i },

CD = { (c_j, Do_{c_j}) avec $c_j \in C$; $Do_{c_j} = \{ chs_j ; chs_i \in CHS \}$,

MD = { $(c_i, \{ u_j ; u_j \in ch_k \text{ et } ch_k \in Do_{c_i} \}) ; c_i \in C$ },

avec l'hypothèse suivante sur les domaines $\bigcap_j Do_j = \emptyset$. Cette hypothèse, qui est simplificatrice, est utile pour simplifier dans un premier temps la reformulation de la demande, car ainsi lorsque qu'un terme est sélectionné c'est grâce à son domaine que l'on saura dans quel champ des descriptions il a été pertinent. Sans cette hypothèse, un même terme peut alors apparaître dans plusieurs champs, ainsi lorsque qu'un terme est pertinent il faut savoir dans quel champ il a été pertinent pour pouvoir utiliser correctement cette information lors de la reformulation. Cette hypothèse est aussi utile, lors de la ventilation décrite dans la suite du paragraphe, pour connaître, sans ambiguïté possible, dans quel champ doit être placé un terme présent dans une facette de la demande utilisateur.

3) Le modèle de demande comprend trois modèles : le modèle de la demande utilisateur, le modèle de la requête intermédiaire et le modèle de la requête système. L'intérêt d'utiliser trois modèles pour représenter une demande est d'avoir trois niveaux de précision de la demande, et

cela permet aussi un découpage de l'analyse de la demande en plusieurs étapes (cf § III.3 et partie 3, chapitre 1; § III.2.2). Le premier modèle permet une certaine généralisation sur les champs des description et sur les concepts utilisées, le deuxième fait apparaître les champs des descriptions qui vont intervenir pour la mise en correspondance et garde l'aspect général sur les termes et le troisième modèle qui permet la sélection des documents contient l'expression précise de la demande, c'est-à-dire un ensemble de critères obtenu à partir des champs des descriptions et de l'interprétation des concepts présents dans ces champs :

$$\text{MDE} = (\text{MDU}, \text{MRI}, \text{MRS}).$$

Le modèle de la demande utilisateur MDU comprend un ensemble de facettes F, un ensemble de couples FC, associant les facettes aux champs du modèle de document auxquels elles sont liées et aussi le modèle de la demande MDe, qui est une liste de facettes auxquelles sont associées une liste d'unités :

$$\text{MDU} = (F, FC, \text{MDe}) \text{ avec}$$

$$F = \{fa_i\};$$

$$FC = \{(fa_i, CF_{fa_i}) \text{ avec } CF_{fa_i} = \{c_j; c_j \in C\} \text{ et } fa_i \in F\};$$

Remarque : les CF_{fa_i} ne sont pas forcément disjoints, une telle contrainte n'est pas nécessaire

car il n'existe pas d'ambiguïté que ce soit dans la mise en correspondance où lors de la reformulation si l'affectation des champs aux facettes est effectuée de manière cohérente.

$$\text{MDe} = \{(fa_i, \{u_j; u_j \in ch_k \text{ et } ch_k \in Do_{fa_i}\}) \text{ avec } Do_{fa_i} = \bigcup Do_{c_j}, c_j \in CF_{fa_i} \text{ et } fa_i \in F\}.$$

Le modèle de la requête intermédiaire MRI est constitué d'un ensemble de champs associé à une liste de termes :

$$\text{MRI} = \{(c_i, \{u_j; u_j \in ch_k \text{ et } ch_k \in Do_{c_i}\}) ; c_i \in C\}.$$

Le modèle de la requête système MRS correspond au modèle booléen classique de requête, c'est une formule booléenne conjonctive disjonctive de critères de recherche CR_j que l'on peut représenter par un ensemble d'éléments disjonctifs ; un élément disjonctif étant un ensemble de critères :

$$\text{MRS} = \{ELD_i\} \text{ avec } ELD_i = \{CR_j\} \text{ et } CR_j = (c, u) ; c \in C, u \in ch_k \text{ et } ch_k \in Do_c.$$

Exemple :

soit la requête système RS respectant le modèle MRS :

$$\text{RS} = \{ \{(c_1, u_2), (c_2, u_3), (c_2, u_1), (c_3, u_8)\}, \{(c_3, u_4), (c_4, u_6), (c_4, u_7)\} \}$$

cette requête peut se réécrire selon le formalisme classique des requêtes booléennes :

$$[(c_1 = u_2) \text{ OU } (c_2 = u_3) \text{ OU } (c_2 = u_1) \text{ OU } (c_3 = u_8)] \text{ ET } [(c_3 = u_4) \text{ OU } (c_4 = u_6) \text{ OU } (c_4 = u_7)]$$

ou encore :

$$[(c_1 = u_2) \text{ OU } (c_2 = u_3 \text{ OU } u_1) \text{ OU } (c_3 = u_8)] \text{ ET } [(c_3 = u_4) \text{ OU } (c_4 = u_6 \text{ OU } u_7)]$$

A l'aide de ce formalisme, nous allons pouvoir écrire, de manière algorithmique, les deux transformations liant la demande utilisateur, la requête intermédiaire et la requête système : la ventilation et l'interprétation.

IV.2.2 La ventilation et l'interprétation

La ventilation transforme une demande utilisateur en requête intermédiaire, elle transforme les facettes en champs en "ventilant" le contenu des facettes dans les champs.

1) La ventilation:

Soit une demande utilisateur D respectant le modèle MDU :

$$D = \{(fa_i, Lu_i)\}; Lu_i = \{u_i\},$$

soit une requête intermédiaire respectant le modèle MRI :

$$RI \leftarrow \{(c_i, \emptyset); c_i \in C\}$$

Pour tous les Lu_i de D faire

 Pour tous les u_i de Lu_i faire

 soit ch_k tel que $u_i \in ch_k$, soit c_j tel que $ch_k \in Do_{c_j}$,

Remarque : grâce à l'hypothèse sur les domaines des champs, on est sûr que c_j est unique.

 Soit lu_j tel que $(c_j, lu_j) \in RI$,

$lu_j \leftarrow lu_j \cup u_i$

 fin pour.

fin pour.

Exemple :

soit CHS = $\{ch_1, ch_2\}$, avec $ch_1 = \{u_1, u_2\}$ et $ch_2 = \{u_3, u_4, u_5\}$;

soit C = $\{c_1, c_2\}$ avec $Do_{c_1} = \{ch_1\}$ et $Do_{c_2} = \{ch_2\}$;

soit F = $\{fa_1\}$ avec $CF_{fa_1} = C$;

soit la demande utilisateur :

$$DU = \{(fa_1, \{u_1, u_3, u_4\})\}$$

après ventilation on obtiendra la requête intermédiaire RI suivante :

$$RI = \{(c_1, \{u_1\}), (c_2, \{u_3, u_4\})\}.$$

La ventilation terminée, l'interprétation entre en jeu en prenant comme données la requête intermédiaire, le type de la recherche et le thesaurus. Pour le moment l'expérimentation, qui sera décrite dans le chapitre suivant, a mis en évidence trois types de recherche différentes :

recherche large, recherche précise, recherche médiane. L'utilisateur, à la formulation de sa demande, a le choix entre la recherche précise et la recherche large. La recherche médiane se situe, comme son nom l'indique, entre les deux autres formes de recherche ; elle est choisie par le système lors de la phase de déformation si la sélection de documents avec la requête système n'est pas satisfaisante.

Trois interprétations ont alors été définies :

2) Interprétation d'une recherche large :

Soit une requête intermédiaire RI, $RI = \{ (c_i, lu_i) \}$

et une requête système RS, $RS = \emptyset$.

Pour tous les lu_i de RI faire

$ELD_i \leftarrow \emptyset$.

Pour tous les u_j de lu_i faire

$ELD_i \leftarrow ELD_i \cup (c_i, u_j)$.

fin pour.

$RS \leftarrow RS \cup ELD_i$.

fin pour.

RS <- extension (RS).

commentaire : cette interprétation revient à connecter l'ensemble des champs avec l'opérateur ET et leur contenu avec l'opérateur OU.

Exemple :

Soit la requête intermédiaire suivante :

$RI = \{ (c_1, \{ u_1, u_2, u_3, u_4 \}), (c_2, \{ u_5, u_6, u_7 \}) \}$,

la requête système obtenue par une interprétation d'une recherche large aura la forme :

$RS = \{ (c_1, u_1), (c_1, u_2), (c_1, u_3), (c_1, u_4), (c_2, u_5), (c_2, u_6), (c_2, u_7) \}$,

qui peut s'exprimer sous la forme booléenne suivante :

$[c_1 = (u_1 \text{ OU } u_2 \text{ OU } u_3 \text{ OU } u_4)] \text{ ET } [c_2 = (u_5 \text{ OU } u_6 \text{ OU } u_7)]$.

La fonction "extension" va modifier la requête système en étendant son sens grâce au thesaurus, cette fonction est commune à toutes les interprétations, nous l'étudierons par la suite plus en détail.

3) Interprétation d'une recherche précise :

Soit une requête intermédiaire RI, $RI = \{ (c_i, lu_i) \}$

et une requête système RS, $RS = \emptyset$.

Pour tous les lu_i de RI faire

Pour tous les u_j de lu_i faire

$RS \leftarrow RS \cup \{ (c_i, u_j) \}$.

fin pour.

fin pour.

RS <- extension (RS).

commentaire : cette interprétation revient à connecter l'ensemble des champs avec l'opérateur ET et ainsi que leur contenu.

Exemple :

Soit la requête intermédiaire suivante :

$RI = \{ (c_1, \{ u_1, u_2, u_3, u_4 \}), (c_2, \{ u_5, u_6, u_7 \}) \}$,

la requête système obtenue par une interprétation d'une recherche précise aura la forme :

$RS = \{ (c_1, u_1), (c_1, u_2), (c_1, u_3), (c_1, u_4), (c_2, u_5), (c_2, u_6), (c_2, u_7) \}$,

qui peut s'exprimer sous la forme booléenne suivante :

$[c_1 = (u_1 \text{ ET } u_2 \text{ ET } u_3 \text{ ET } u_4)] \text{ ET } [c_2 = (u_5 \text{ ET } u_6 \text{ ET } u_7)]$.

4) L'interprétation d'une recherche médiane :

Cette interprétation se situe entre les deux autres interprétations, son originalité réside dans le choix des opérateurs à mettre en œuvre à l'intérieur d'un même champ. En effet, dans la recherche large l'opérateur dans un même champ est le OU, et dans la recherche précise, le ET. Dans la recherche médiane, la connexion booléenne entre plusieurs unités dans un même champ va dépendre d'une "distance sémantique" évaluée grâce au thesaurus. Deux unités proches sémantiquement seront reliées par un OU, par contre deux unités éloignées sémantiquement seront connectées par un ET.

Avant de présenter l'algorithme de l'interprétation, faisons une remarque sur la notion de proximité sémantique entre deux unités. Si la symétrie ne fait aucun doute, il convient de se demander si cette notion est transitive.

Soit la fonction "ProcheSémantiquement" qui prend en paramètres deux unités du thesaurus et qui renvoie le booléen vrai quand les deux unités sont proches sémantiquement à l'intérieur du thesaurus :

ProcheSémantiquement (u_1, u_2) -> booléen.

Le cas de figure suivant peut se produire :

ProcheSémantiquement (u_1, u_2) -> vrai,

ProcheSémantiquement (u_2, u_3) -> vrai,

et ProcheSémantiquement (u_1, u_3) -> faux.

Ce résultat peut paraître normal pour certains et bizarre pour d'autres, mais dans notre approche, il est correct. En effet, il n'est pas intéressant de considérer la notion de proximité sémantique comme transitive, sinon la différence entre une recherche médiane et une recherche large ne serait pas assez importante.

Nous proposons l'algorithme suivant considérant la non transitivité de la proximité sémantique :

Soit une requête intermédiaire RI, $RI = \{ (c_i, lu_j) \}$

et une requête système RS, $RS = \emptyset$.

Pour tous les lu_j de RI faire

Pour tous les u_j de lu_j faire

ELD_j <- { (c_i, u_j) }.

UPS_{u_j} <- lesUnitésProchesSémantiquement (lu_j, u_j).

Pour tous les u_k de UPS_{u_j} faire

ELD_j <- ELD_j \cup { (c_i, u_k) }.

fin pour.

RS <- RS \cup ELD_j.

fin pour.

fin pour.

RS <- extension (RS).

Remarque : la fonction "lesUnitésProchesSémantiquement" ayant comme paramètres une liste d'unités lu_j et une unité u_j , recherche dans la liste lu_j , les unités proches sémantiquement de l'unité u_j ; le résultat est un ensemble d'unités.

Cette interprétation peut être suivit d'une étape de simplification de la requête système car certains éléments disjonctifs sont inclus dans d'autres éléments (cf exemple).

Exemple :

Soit la requête intermédiaire suivante :

$RI = \{ (c_1, \{ u_1, u_2, u_3, u_4 \}), (c_2, \{ u_5, u_6, u_7 \}) \}$,

et les évaluations suivantes à vrai de la fonction "ProcheSémantiquement" :

ProcheSémantiquement (u_1, u_4) -> vrai,

ProcheSémantiquement (u_1, u_2) -> vrai,

ProcheSémantiquement (u_2, u_3) -> vrai,

ProcheSémantiquement (u_5, u_6) -> vrai,

la requête système obtenue par une interprétation d'une recherche précise aura la forme :

$RS = \{ \{ (c_1, u_1), (c_1, u_2), (c_1, u_4) \}, \{ (c_1, u_1), (c_1, u_2), (c_1, u_3) \}, \{ (c_1, u_2), (c_1, u_3) \}, \{ (c_1, u_1), (c_1, u_4) \}, \{ (c_2, u_5), (c_2, u_6) \}, \{ (c_2, u_7) \} \}$,

qui peut s'exprimer sous la forme booléenne suivante :

$[c_1 = ((u_1 \text{ OU } u_2 \text{ OU } u_4) \text{ ET } (u_1 \text{ OU } u_2 \text{ OU } u_3) \text{ ET } (u_2 \text{ OU } u_3) \text{ ET } (u_1 \text{ OU } u_4))] \text{ ET } [c_2 = ((u_5 \text{ OU } u_6) \text{ ET } u_7)]$.

La fonction "ProcheSémantiquement" reste à définir, sa définition est heuristique. Nous verrons dans le prochain chapitre décrivant le prototype RIVAGE, la solution qui a été prise.

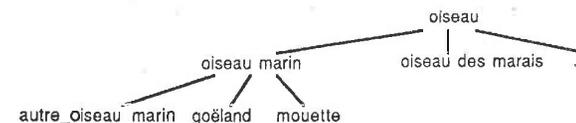
Cette présentation de la phase d'interprétation serait incomplète si la fonction "extension" utilisée dans chacune des trois interprétations n'était pas présentée.

5) L'extension :

L'extension est une fonction qui transforme la requête système en exprimant les concepts présents dans la requête, sous leur forme étendue. En effet, lorsqu'un utilisateur met un terme dans sa demande, le système doit rechercher dans la base toutes les formes possibles du concept représenté par le terme. Cette recherche est possible grâce au thesaurus qui a maintenant le statut de taxonomie. Voyons sur un exemple comment cette extension fonctionne.

Exemple :

Soit la partie du thesaurus suivante :



Si l'utilisateur met le terme "oiseau" dans sa demande, cela signifie qu'il veut voir toutes sortes d'images d'oiseaux. L'interprétation du concept sous la forme d'une requête booléenne doit être :

(goëland OU mouette OU autre_oiseau marin OU oiseau des marais OU ...)

L'extension va alors consister à remplacer toutes les unités qui ont des spécifiques par une disjonction de tous ses spécifiques étant des "feuilles" du thesaurus. Une feuille du thesaurus est une unité qui n'a pas de spécifiques et il est important de rappeler que la transformation du thesaurus en taxonomie a permis de réaliser une indexation des images en n'utilisant exclusivement que les feuilles des champs sémantiques. Cette étape est fondamentale dans notre processus. En effet si les concepts ne sont pas exprimés en extension dans la requête système, le système est incapable d'apprendre quelque chose à partir des choix de l'utilisateur car ceux-ci ne sont pas représentatifs des concepts introduits dans la demande puisque leur interprétation a été incomplète et que seulement une partie des images illustrant ces concepts a été vue par l'utilisateur.

Voici l'algorithme de cette fonction :

extension (RS) :

Pour tous les ELD_i de RS faire

Pour tous les (c_j, u_k) de ELD_i faire

$TSP_{u_k} \leftarrow \text{TousLesSpécifiques}(u_k)$.

Pour tous les u_j de TSP_{u_k} faire

$ELD_i \leftarrow ELD_i \cup \{(c_j, u_j)\}$.

fin pour.

$ELD_i \leftarrow ELD_i \ominus \{(c_j, u_k)\}$.

fin pour.

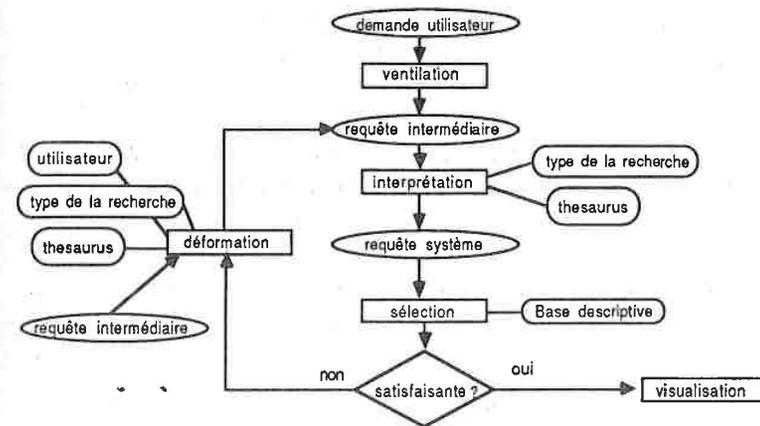
fin pour.

extension \leftarrow RS.

Commentaire : la fonction "TousLesSpécifiques" renvoie tous les spécifiques de l'unité en paramètres. Ce sont les spécifiques de tous les niveaux étant des feuilles. Cette fonction s'écrit de manière récursive ; nous n'allons pas la présenter ici, car elle ne présente que peu d'intérêt. L'opérateur \ominus symbolise la différence ensembliste.

La requête système, ainsi construite, peut alors jouer son rôle, c'est-à-dire sélectionner les descriptions d'images vérifiant le concept exprimé par la requête. Cette sélection correspond à la mise en correspondance classique entre une requête booléenne et un document ayant une indexation binaire (cf partie 1, chapitre 3, § II.2.1). Si le résultat de cette sélection n'est pas

satisfaisant, la phase de déformation intervient en mettant en œuvre quelques stratégies, qui sont naturellement heuristiques. Mais avant de voir quelles peuvent être ces déformations ainsi que leur influence sur le niveau d'expression, résumons toute cette étape de la mise en correspondance par un schéma :



Remarque : si l'on fait la correspondance avec les différentes étapes de la phase "Avant-Visualisation" présentées dans la partie précédente, nous dirons que la mise en correspondance commence ici à l'analyse sémantique. En effet, même si, dans ce schéma, on ne parle pas d'analyse lexicale et même morphologique, ces étapes existent toujours mais leur forme n'ayant subi de changement depuis leur définition dans le projet Esprit, nous n'avons pas jugé utile de les présenter dans ce schéma par souci de simplification.

IV.2.3 La déformation (3)

Comme l'indique le schéma précédent la déformation travaille sur la requête intermédiaire en utilisant le thesaurus et le type de la recherche, et en faisant intervenir l'utilisateur. Nous n'allons pas détailler ici l'ensemble des traitements pouvant être effectués lors de cette phase, ceci sera présenté dans le chapitre suivant (§ IV.2 chapitre 3). Mais donnons un aperçu général des possibilités de déformations de la requête intermédiaire.

Deux grands types de déformations existent relativement au résultat souhaité sur la recherche :

- déformations pour élargir la recherche,
- déformations pour rétrécir la recherche.

Dans les deux types de déformations, il existe deux niveaux possibles d'intervention sur la mise en correspondance :

- au niveau des termes présents dans la requête intermédiaire,
- au niveau de l'interprétation en changeant le type de la recherche.

Les déformations possibles sur les termes pour élargir la recherche peuvent être les suivantes :

- la généralisation (remplacer un terme par son générique),
- l'association d'idées (ajouter les termes associés d'un terme),
- l'ajout de terme par l'utilisateur.

Le changement du type de la recherche pour élargir la recherche dépend du type de la recherche précédente. Si la recherche est précise, alors le passage à une recherche médiane élargira la recherche. De même, si le type de la recherche est médiane, le passage à une recherche large augmentera le nombre d'images sélectionnées. Par contre si la recherche est de type large, le seul moyen d'élargir la recherche est d'intervenir sur les termes.

Réciproquement, les déformations possibles sur les termes, permettant de réduire l'étendue d'une recherche, sont :

- la spécialisation (remplacer un terme par un ou plusieurs spécifiques),
- la suppression de terme par l'utilisateur.

Le changement du type de recherche peut s'effectuer selon le schéma suivant :

recherche large -> recherche médiane,
recherche médiane -> recherche précise.

Dans le cas où la recherche est précise, seule l'intervention sur les termes peut permettre de rétrécir cette recherche.

Remarque : les règles classiques de déformations de requêtes consistant à introduire des ET pour rétrécir la recherche, ou à introduire des OU pour l'élargir existent dans ces stratégies grâce au changement de type de recherche lors de l'interprétation. De cette manière ces modifications de la requête sont alors justifiées par le calcul d'une distance entre des termes à l'intérieur du thesaurus et non d'une manière arbitraire.

Ces opérations peuvent être utilisées lors des déformations à l'intérieur de stratégies heuristiques de déformation. Cette partie de la mise en correspondance est en cours d'expérimentation. Quelques stratégies, mises en œuvre dans le prototype RIVAGE, seront présentées dans le chapitre suivant.

Seules les suppressions et les ajouts de termes nécessitent la modification du niveau d'expression dans le thesaurus. En effet, lorsqu'un terme est ajouté dans la requête intermédiaire, il devient candidat à la formulation du besoin de l'utilisateur, son poids d'expressivité est alors mis à zéro ainsi que ceux de tous ses spécifiques de tous les niveaux. La propagation aux termes génériques se fera avec la fonction F utilisée dans les autres propagations. Quand un terme est supprimé cela signifie qu'il n'exprime plus le besoin de l'utilisateur, son poids est mis à -1, ainsi que ceux de tous ses spécifiques de tous les niveaux, avec propagation aux termes génériques.

Dés que le résultat de la sélection des descriptions d'images est "satisfaisant" (nous verrons comment définir "satisfaisant" dans le chapitre suivant), la phase de visualisation peut commencer.

V Evolution du niveau d'expression pendant la phase de "Visualisation"

La phase de visualisation est primordiale dans le processus d'apprentissage puisqu'elle fournit les exemples et les contre-exemples du concept que le système doit apprendre. Il est alors important que l'utilisateur puisse choisir des images, mais aussi en rejeter. Le rejet d'images n'est pas une démarche habituelle dans les systèmes de recherche d'images, mais pour ce processus il est plus que nécessaire, car, nous le verrons par la suite, il permet d'éliminer certaines parties du thesaurus qui ne correspondent pas au besoin de l'utilisateur.

Dans cette phase, il faut que l'utilisateur puisse effectuer ses choix rapidement et sans contrainte. En effet, si l'utilisateur doit appuyer sur trois touches en même temps (^xc) pour rejeter une image, nous pouvons être sûr, qu'au bout d'un certain temps, il n'existera plus d'images rejetées dans le choix de l'utilisateur, ce qui appauvrirait les capacités du processus d'apprentissage.

Le résultat d'une phase de visualisation est constitué de trois ensembles d'images :

- les images choisies,
- les images rejetées,
- les images neutres.

Les images neutres sont soit des images que l'utilisateur n'a pas visualisées, soit des images sur lesquelles il n'a pu émettre un avis. Elles n'ont pas un grand intérêt, et surtout elles n'ont pas le statut d'images rejetées, car elles peuvent, au bout d'un certain nombre d'étapes, devenir

pertinentes. En effet, le besoin d'un utilisateur d'un tel système (système de recherche progressive d'images) évolue progressivement d'étape en étape, et au fur et à mesure des visualisations. L'utilisateur, pendant une phase de visualisation, doit avoir le comportement suivant : choisir les images qui correspondent, avec un fort degré de certitude, à son besoin, rejeter celles qui n'ont rien à voir avec son besoin, puis lorsqu'une image provoque une certaine hésitation dans le choix, il vaut mieux qu'elle ne soit ni choisie, ni rejetée c'est-à-dire neutre.

Quoique cette phase soit importante pour la mise à jour du niveau d'expression, celui-ci reste le même pendant tout le déroulement de la visualisation. Les ensembles d'images construits, la phase "Après-Visualisation" va alors exploiter ces résultats, c'est-à-dire analyser les exemples et les contre-exemples du concept que le processus doit apprendre. La phase "Après-visualisation" va permettre de faire évoluer le niveau d'expression en rapport avec les choix de l'utilisateur effectués dans la phase "Visualisation".

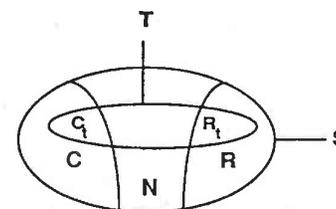
VI Evolution du niveau d'expression dans la phase "Après-visualisation"

Pour faire évoluer le niveau d'expression en fonction du choix de l'utilisateur, nous calculons pour chaque terme présent dans les descriptions des images sélectionnées un **poids de pertinence** calculé d'une manière similaire aux mesures de pertinence de Salton et Dillon (cf chapitre 1, § II). Les poids d'expressivité sont ensuite mis à jour en prenant en compte les poids de pertinence afin d'obtenir un nouveau niveau d'expression plus proche des besoins de l'utilisateur. Là encore, toutes les formules que nous proposons sont heuristiques, elles correspondent à celle utilisée dans l'expérimentation, elles nous donnent pour le moment entière satisfaction.

VI.1 La pertinence d'un terme (4)

La pertinence d'un terme est calculée en analysant les champs des descriptions et en associant à chaque terme présent un poids de pertinence P_{pert} . Ce poids peut se calculer d'une manière heuristique avec la formule suivante qui sera mise à l'épreuve lors de l'expérimentation :

Soit t un terme, S l'ensemble des descriptions des images sélectionnées, C l'ensemble des descriptions des images choisies par l'utilisateur, NC son cardinal, R l'ensemble des descriptions des images rejetées, NR son cardinal, N l'ensemble des descriptions des images neutres et T les descriptions des images contenant le terme t :



Appelons $C_t = T \cap C$, $R_t = T \cap R$ et NC_t , NR_t leur cardinal respectif.

Le poids de pertinence P_{pert_t} est alors obtenu de la manière suivante :

$$P_{pert_t} = \frac{NC_t}{\text{Max}(NC, 1)} - \frac{NR_t}{\text{Max}(NR, 1)}$$

Cette formule a été obtenue en supposant que lorsqu'un terme est présent dans une majorité des descriptions des images choisies, cela signifie qu'il caractérise bien les images choisies, et par opposition lorsqu'un terme est présent dans beaucoup de descriptions des images rejetées cela signifie qu'il caractérise bien les d'images rejetées. La mesure des proportions des documents choisis ou rejetés contenant un terme t par rapport à l'ensemble de tous les documents choisis ou rejetés, nous informe sur la pertinence du terme t dans chacun des deux ensembles choisis et rejeté. La différence de ces deux proportions nous informera alors sur la pertinence globale du terme dans le choix de l'utilisateur. De cette manière, la valeur de P_{pert} est comprise entre -1 et 1. Si la valeur se rapproche de 1, on dira que le terme caractérise bien les images choisies par l'utilisateur ; de même si la valeur est proche de -1, on dira que le terme caractérise bien les rejets. L'obtention de cette formule a été fortement inspirée par le concept de "majorité" du système de Quinqueton et Sallantin exprimant le fait qu'un concept est appris s'il est présent dans une majorité des exemples.

Remarque : Avec cette formule un terme qui est fortement présent dans les descriptions des images rejetées est considéré comme non pertinent ; il ne fera donc plus partie de l'expression du besoin de l'utilisateur. D'autres utilisations de ce type de termes pourraient être effectuées si le thesaurus possédait plus de relations sémantiques comme par exemple une relation "opposition" indiquant que deux termes ont un sens opposé (relation fréquente dans le champ sémantique de la connotation). Avec cette relation, la fréquence importante d'apparition d'un terme dans les rejets pourrait rendre automatiquement le terme de sens opposé pertinent.

L'exemple : la demande de l'utilisateur "contenu de la photo = bateau" a obtenu après interprétation (sans déformation) 9 images que l'utilisateur a visualisées. Il a émis un avis sur 8 images, une image est restée "neutre".

Voici les descriptions des images qu'il a choisies :

- description1: contenu : paquebot, goeland, ocean
- description2: contenu : pétrolier, mouette, mer
- description3: contenu : cargo, autre_oiseau marin,

les descriptions des images qu'il a rejetées :

- description4: contenu : autre_bateau, port
- description5: contenu : canoe, ruisseau
- description6: contenu : péniche, arbre
- description7: contenu : bac, rive
- description8: contenu : cargo, port.

et voici la description de l'image "neutre" :

- description9: contenu : oiseau des marais, étang, barque.

Après une analyse des descriptions, voici les poids de pertinence obtenus pour les termes présents :

paquebot	0,33	barque	0	port	-0,40
pétrolier	0,33	goeland	0,33	rive	-0,20
cargo	0,13	mouette	0,33	ruisseau	-0,20
autre_bateau	-0,20	autre_oiseau marin	0,33	étang	0
canoe	-0,20	oiseau des marais	0	arbre	-0,20
bac	-0,20	ocean	0,33		
péniche	-0,20	mer	0,33		

VI.2 Mise à jour du niveau d'expression (5)

Le choix de l'utilisateur ayant une influence sur le niveau d'expression, nous allons utiliser les poids de pertinence, calculés précédemment, pour mettre à jour les poids d'expressivité. Cette prise en compte du choix de l'utilisateur pour l'évolution du niveau d'expression repose sur l'hypothèse considérant que le besoin de l'utilisateur évolue d'étape en étape. Ainsi il se peut très bien qu'à une étape donnée un concept fasse partie de l'expression du besoin de l'utilisateur puis qu'à l'étape suivante ce même concept ne fasse plus partie de l'expression du besoin appris.

La répercussion des poids de pertinence dans le thesaurus va s'effectuer des termes les plus spécifiques (les feuilles) vers le plus générique (la racine) de la manière suivante :

soit u une unité du thesaurus, $u = (t, SPE, g, TA, Pexp_t)$ et $Ppert_t$ le poids de pertinence de t si le terme a été trouvé dans les choix de l'utilisateur.

Si $SPE = \emptyset$ (u est une feuille)

alors $Pexp_t = Ppert_t$

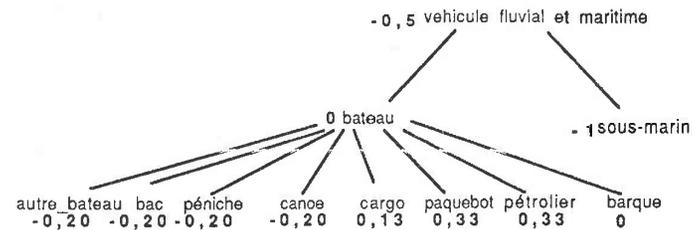
sinon Soit $SPE = \{u_i\}$ avec $u_i = (t_i, SPE_i, g_i, TA_i, Pexp_i)$,

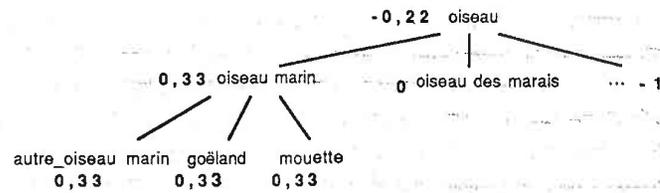
$Pexp_t = F(Pexp_i)$

La fonction F est la même fonction que celle utilisée pour la propagation des poids d'expressivité après la formulation de la demande (cf § IV.1)

L'exemple :

Considérons les parties du thesaurus représentatives du choix de l'utilisateur après la mise à jour des poids d'expressivité effectuée avec la formule (1) ; F étant une moyenne :





Remarque : Si nous appelons TDS l'ensemble de tous les termes présents dans les descriptions sélectionnées (choisies, rejetées et neutres) et TRS l'ensemble de tous les termes présents dans la requête système qui a sélectionné l'ensemble des descriptions. TRS n'est pas toujours inclus dans TDS, donc il existe des termes qui ont gardé un poids d'expressivité à 0 provenant de la phase "Avant-visualisation". Ce poids est normal car ce terme n'était présent ni dans les exemples, ni dans les contre-exemples. Ces termes ont le même statut que des termes n'ayant été trouvés que dans les descriptions d'images neutres (dans l'exemple : étang, oiseau des marais et barque). Cette remarque vient compléter notre argumentation sur la mise à zéro effectuée après la formulation de la demande.

Le niveau d'expression vient d'être mis à jour dans le thesaurus, en tenant compte de la présence des termes dans les descriptions des images choisies, rejetées ou neutres. Il faut maintenant parcourir le thesaurus pour en extraire les termes définissant le concept appris par l'analyse des descriptions. Le concept ainsi formulé, l'utilisateur doit donner son avis afin de recommencer une nouvelle étape de recherche.

VII Formulation du concept appris, construction d'une nouvelle demande

La construction d'une nouvelle demande nécessite d'abord le parcours de l'ensemble du thesaurus, champ sémantique par champ sémantique, pour en extraire les termes caractérisant le concept. Le concept est alors exprimé en ventilant les termes caractéristiques dans les facettes de la "demande utilisateur" auxquelles ils sont rattachés.

VII.1 Extraction des termes exprimant le concept (6)

Le concept de la demande est représenté dans le thesaurus sous la forme d'un niveau d'expression. Un parcours de l'ensemble du thesaurus est nécessaire pour en extraire les termes exprimant le mieux le concept. Un parcours de chaque champ sémantique, en partant de la racine vers les feuilles, est nécessaire car dans ce type de structure lorsqu'un concept est appris

on suppose que toutes ses représentations possibles (ses fils) le sont aussi et par contre lorsqu'un concept n'est pas appris il faut voir parmi ses fils quelles sont les représentations ou les sous-concepts qui ont été appris.

Le parcours, qui est inspiré de la construction de concept à partir d'un ensemble de taxonomies utilisé dans l'espace des version de Mitchell, va alors s'effectuer de la manière suivante : pour chaque noeud de l'arbre on va comparer son poids d'expressivité à un seuil ; si le poids est supérieur, on sélectionne le terme et le parcours s'arrête pour cette partie de l'arbre, on considère ainsi que le concept représenté par le terme est appris ; si le poids est inférieur, on descend dans l'arbre et on applique le même traitement aux noeuds de ce niveau, en considérant que le concept représenté par le terme n'est pas appris et qu'il est possible que des sous-concepts (des termes spécifiques) aient été appris.

Un ensemble de termes caractéristiques est ainsi obtenu. Pour exprimer correctement le concept sous la forme d'une demande, il faut alors ventiler les termes dans les facettes auxquelles ils sont rattachés (cf § III.3).

Voici l'algorithme de cette formulation :

Soit TH le thesaurus $TH = \{ chs_j \}$,

soit un seuil α fixé (nous verrons comment par la suite),

soit une demande utilisateur respectant le modèle MDU, $Du \leftarrow \{ (fa_i, \emptyset) \}$

Pour tous les chs_j de TH faire

$US_j \leftarrow \text{UnitésSélectionnées}(r_j, \alpha)$; US_j contient toutes les unités sémantiques du champ

Soit $c \in C$, tel que $chs_j \in Do_c$; ; sémantique courant ayant leur poids supérieur au seuil α .

soit $fa \in F$, tel que $c \in CF_{fa}$;

soit U tel que $(fa, U) \in Du$.

$U \leftarrow U \cup US_j$; ajout des unités sémantiques ainsi sélectionnées dans le

fin pour. ; contenu de la facette de la demande associé prenant ses

; valeurs dans le champ sémantique courant

Remarque : la fonction "UnitésSélectionnées" parcourt l'arbre ayant comme racine l'unité en paramètre et sélectionne les unités filles ayant un poids d'expressivité supérieur au seuil donné en paramètre ; le résultat est alors un ensemble d'unités :

UnitésSelectionnées(r_i, α):

$r_i = (t_i, SPE_i, g_i, TA_i, Pexp_i)$,

Si $Pexp_i \geq \alpha$

alors UnitésSelectionnées $\leftarrow \{ r_i \}$

sinon

Si $SPE_i \neq \emptyset$ alors

Pour tous les u_j de SPE_i faire

UnitésSelectionnées \leftarrow UnitésSelectionnées \cup UnitésSelectionnées (u_j, α)

fin pour.

fin si.

fin si.

L'exemple : Après un parcours du thesaurus avec un seuil α de 0,33 nous obtenons, en ce qui concerne la partie du thesaurus que nous avons montrée, la nouvelle demande suivante :

Demande:

contenu de la photo = paquebot, pétrolier, oiseau marin.

Remarque importante : On peut remarquer que la méthode permet de proposer à l'utilisateur de nouveaux critères de recherche, comme les oiseaux marins, qui étaient complètement absents dans sa première demande.

Le choix du seuil α est heuristique, sa valeur sera comprise entre 0 et 1 et plus précisément entre $\text{Max}(Pexp_i)$ et 0. Mais nous verrons dans le chapitre suivant les options qui ont été prises et celles qui peuvent l'être suite à l'expérimentation.

Suite à ce parcours du thesaurus, une nouvelle demande utilisateur Du, représentant le concept appris, est obtenue ; elle est alors présentée à l'utilisateur afin qu'il l'adapte aux mieux à ses besoins.

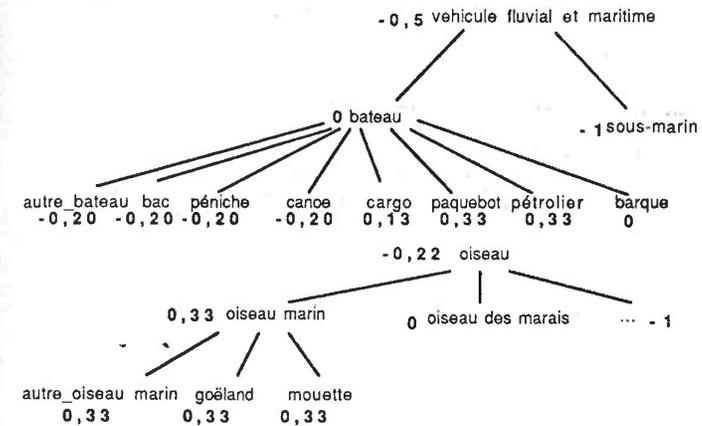
VII.2 Dialogue pour une reformulation (7)

La demande ainsi construite est proposée à l'utilisateur qui peut, s'il en ressent le besoin, la modifier. Un dialogue s'instaure où plusieurs actions sont offertes à l'utilisateur : ajout de terme dans une facette, suppression de terme dans une facette, voir les autres propositions du système, ...

Le système peut montrer d'autres propositions à l'utilisateur en faisant diminuer le seuil α d'un certain pas et en reparcourant le thesaurus avec ce nouveau seuil. De cette manière, l'utilisateur peut obtenir soit des concepts plus généraux, soit de nouveaux concepts.

L'exemple :

Si nous reprenons le morceau de thesaurus avec le niveau d'expression suivant :



Avec un seuil $\alpha = 0,33$, on obtenait la demande suivante :

contenu de la photo = paquebot, pétrolier, oiseau marin.

Avec $\alpha = 0,10$, on obtient :

contenu de la photo = cargo, paquebot, pétrolier, oiseau marin.

Et enfin avec $\alpha = 0$, on obtient :

contenu de la photo = bateau, oiseau marin, oiseau des marais.

Influence sur le niveau d'expression :

La suppression de termes par l'utilisateur est une action qui modifie le concept appris. Cette action a une influence sur les poids d'expressivité des termes. La répercussion de cette action sur les poids est la même que lors des déformations (cf § IV.2.3) : mise à -1 avec répercussion aux génériques. L'ajout de termes occasionne une mise à jour du niveau, mais celle-ci est prise en compte dans la mise à zéro globale effectuée après chaque ventilation de la demande utilisateur dans la requête intermédiaire (étape 2 de l'évolution globale).

Dès que l'utilisateur a terminé ses actions et qu'il juge sa demande satisfaisante, il laisse le système ventiler et interpréter celle-ci en appliquant les transformations nécessaires afin d'obtenir une *requête système*, puis des images ...

Le processus se poursuit jusqu'à ce que l'utilisateur et le système soient arrivés à exprimer clairement le concept exprimant les besoins de l'utilisateur et à obtenir les ou l'image le matérialisant.

VIII Conclusion

Nous venons de montrer que le processus EXPRIM de recherche d'images peut être assimilé à un processus d'apprentissage où le système tente d'apprendre et de formuler, tout au long des étapes de recherche, les besoins de l'utilisateur. L'apprentissage que nous avons mis en œuvre consiste à expliciter, pour chaque demande, un niveau d'expression, à faire évoluer par apprentissage ce niveau d'expression et à déduire, de l'état final d'une phase d'évolution, une nouvelle demande. Cette assimilation nous a permis d'offrir un cadre cohérent pour la mise en œuvre du processus de recherche d'images EXPRIM.

La réalisation de cette méthode, sur un échantillon d'images relativement important, va nous permettre de mettre au point plusieurs parties que nous avons définies comme "heuristiques", mais aussi de mettre à l'épreuve nos propositions. Ainsi seule une expérimentation "grandeur nature" peut nous aider à mettre au point les parties suivantes de la méthode :

- l'interprétation de la demande,
- la déformation de la demande,
- l'analyse des choix de l'utilisateur,
- la propagation des poids dans le thesaurus (la fonction F),
- le choix du seuil pour la reformulation.

Tous ces points vont être analysés dans le chapitre suivant où nous présentons le prototype RIVAGE qui fut le champ d'expérimentation de cette méthode.

Chapitre 3

Expérimentation de la méthode le prototype RIVAGE	199
I Introduction	199
II Présentation du prototype	200
II.1 Architecture	200
II.2 Le dialogue et le déroulement d'une étape de recherche.....	202
II.2.1 Le déroulement d'une étape	202
II.2.2 L'interface utilisateur en Smalltalk	203
II.2.3 Les fenêtres utilisées dans RIVAGE	205
II.2.3.1 La formulation de la demande	205
II.2.3.2 Visualisation et choix d'images	207
II.2.3.3 Les autres fenêtres	208
III Présentation des objets	211
III.1 Formalisme utilisé	211
III.2 Les objets manipulés	212
III.2.1 La session	213
III.2.2 Le thesaurus	216
III.2.3 La base descriptive	218
III.2.4 Les facettes	219
IV Description des fonctions principales	220
IV.1 La ventilation et l'interprétation de la demande	220
IV.1.1 Quelques remarques	220
IV.1.2 La fonction ProcheSémantiquement	222
IV.2 La déformation de la demande	225
IV.3 La Visualisation	230
IV.4 Analyse des choix	
calcul des poids de pertinence	232
IV.5 La propagation des poids	233
IV.6 La reformulation	235
IV.6.1 Le choix des seuils	235
IV.6.2 Parcours d'un champ sémantique en langage objet	
.....	238
V Evaluation du système	241
V.1 Avertissement	
V.2 La démarche choisie	241
V.3 Résultats de l'évaluation	242
VI Evolution & perspectives	247
VI.1 Les évolutions à court terme	247
VI.2 Les évolutions à moyen terme	249
VI.3 Les évolutions à long terme, perspectives du système	
RIVAGE	249
VI.3.1 Apprentissage à long terme	250
VI.3.2 Formalisation du processus EXPRIM	251
VI.3.3 Explicabilité associée au processus	251
VII Conclusion	252

Chapitre 3

Expérimentation de la méthode : le prototype RIVAGE

I Introduction

Ce chapitre décrit le prototype RIVAGE (Recherche Interactive et progressiVe d'imAGES) qui a été réalisé dans un environnement objet à l'aide du langage Smalltalk 80, la description de l'ensemble des objets que nous avons définis est donnée au § III à l'aide d'un formalisme inspiré des modèles sémantiques et des modèles objets de représentation de donnée.

Ce prototype qui a permis de tester nos idées, représente la première réalisation du processus EXPRIM dans sa totalité (Avant-Visualisation, Visualisation, Après-Visualisation). L'approche "apprentissage" est implantée entièrement selon la description qui en a été faite dans le chapitre précédent.

L'expérience de la première expérimentation, dans le cadre du projet ESPRIT, n'ayant pas été très concluante, nous avons orienté nos choix de matériel et de logiciel vers d'autres horizons. Cela a été possible, car les objectifs respectifs des deux réalisations ne sont pas identiques. Le but de la réalisation que nous allons présenter est de valider une idée, alors que celui de la réalisation précédente était une étude de faisabilité technique d'un produit. La différence est essentielle, car dans la réalisation actuelle, l'étude de faisabilité technique ne sera possible que si l'idée a été validée. Pour effectuer cette validation, les contraintes matérielles et logicielles ne doivent pas être trop importantes. C'est ainsi que nous avons choisi de développer RIVAGE sur un mini-ordinateur, avec un seul langage, où seules les fonctions nécessaires à la validation ont été écrites.

La base d'images du Ministère de la Culture, qui avait été choisie comme application pilote du projet ESPRIT (cf partie 3, chapitre 2, § I.1), a servi de champ d'expérimentation. Elle possède toutes les qualités nécessaires pour remplir ce rôle : nombre d'images important, thesaurus associé, vidéodisque existant.

Cette expérimentation a été réalisée avec l'aide d'une stagiaire de l'IUT du département carrière de l'information, Sabine Humbert. Elle a effectué de nombreuses recherches d'images qui nous

ont permis de mettre au point les différents composants du système.

L'ensemble de ce chapitre décrit le prototype RIVAGE dans sa première version, c'est-à-dire avec un dialogue de base permettant de réaliser le processus EXPRIM. Cette présentation débute par l'architecture générale du système, puis se poursuit par une "vue externe" des fonctions du système, écran par écran ou plutôt fenêtre par fenêtre. Nous rentrerons plus dans le détail de cette réalisation dans le troisième paragraphe, où la structure des principaux objets sera donnée. Le quatrième paragraphe est consacré à la présentation des différentes solutions choisies pour la réalisation des parties dites "heuristiques" de la méthode ; ces parties ont été mises en évidence dans le chapitre précédent.

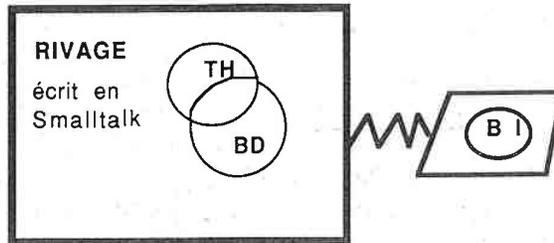
Toute cette présentation serait incomplète si une évaluation du système, mesurant la pertinence globale du système, comme nous l'avons définie au premier chapitre de la partie 1, n'était pas présentée. Cette évaluation fait l'objet du cinquième paragraphe. Les perspectives d'une telle réalisation sont détaillées dans le dernier paragraphe, l'une d'entre elles, déjà partiellement réalisée dans RIVAGE, est approfondie dans le dernier chapitre de cette partie.

II Présentation du prototype

Le prototype RIVAGE est la concrétisation du processus EXPRIM vu comme un problème d'apprentissage. Il représente aussi notre première réalisation dans un langage objet. En effet, RIVAGE est écrit en Smalltalk 80 [Gol 83] qui est un langage objet permettant une programmation structurée et modulaire. Il permet aussi de manipuler des objets à structure complexe, de décrire leur comportement grâce aux méthodes. Ces structures ainsi que leur comportement peuvent être modifiés au fur et à mesure que la réalisation évolue. Smalltalk 80 offre un environnement de programmation agréable et performant ; il constitue un outil de prototypage idéal, car il permet de concevoir des applications d'une manière incrémentale ; cette propriété étant fondamentale pour la réalisation de prototype de recherche.

II.1 Architecture

L'architecture du prototype RIVAGE est simple (cf figure 6), elle ne reprend pas les composants des architectures habituellement proposées pour la réalisation du projet EXPRIM. En effet, il n'existe plus de système documentaire, ni de système de visualisation, ni de système de pilotage, ni de moteur d'inférence. Il y a un langage et un ensemble de fonctions réalisées dans ce langage.



TH : THesaurus,
 BD : Base Descriptive,
 BI : Base D'Images.

figure 6 : architecture de RIVAGE

Les avantages de cette architecture sont nombreux [Léo 89a] :

- plus de problème de communication entre différents logiciels, (C'est ce point qui a bloqué la réalisation du projet ESPRIT)
- un seul langage à manipuler,
- seules les fonctions utiles sont réalisées :
 dans les anciennes architectures, seulement une partie des fonctions des différents logiciels étaient utilisées dans la réalisation du prototype. De plus la rigidité de certaines fonctions orientait la conception vers des choix limitant les champs d'expérimentation.
- plus de duplication de l'information :
 en effet, dans la dernière réalisation le thesaurus et la base descriptive étaient gérés par le système documentaire ; ainsi lorsque le SBC souhaitait raisonner sur une partie du thesaurus, il devait d'une part interroger le système documentaire pour accéder à la partie désirée, puis transformer l'information dans son formalisme pour pouvoir l'utiliser.

Il est important de noter qu'en plus de l'environnement de programmation performant, Smalltalk 80 offre la possibilité de définir des interfaces avec l'utilisateur utilisant le multi-fenêtrage, la manipulation de souris et l'utilisation d'outils de dialogue tel que :

- les menus déroulants,
- les boutons associés à une fonction,
- les fonctions d'édition (copier, couper, coller),
- les fonctions associées à une fenêtre (déplacer, changer de taille, transformer en étiquette) ...

C'est à l'aide de ces fonctions que nous avons défini le dialogue avec l'utilisateur que nous allons présenter maintenant.

II.2 Le dialogue et le déroulement d'une étape de recherche

Le dialogue, que nous allons présenter dans ce paragraphe, est un dialogue de base permettant de réaliser le processus EXPRIM dans sa totalité. Il est sûrement incomplet et de nombreuses aides pourrait y être ajoutées, mais il nous a permis de valider le processus proposé.

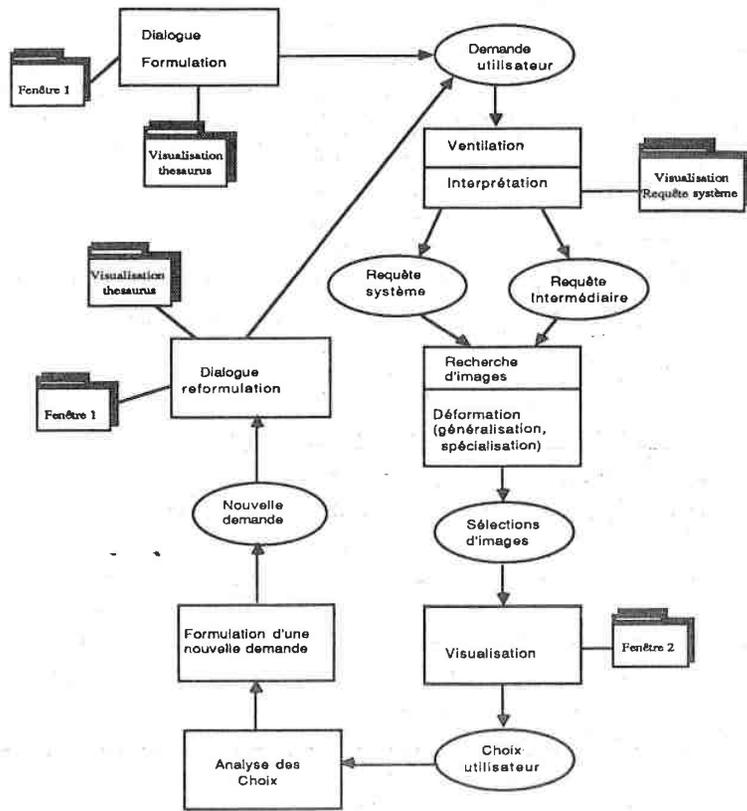
II.2.1 Le déroulement d'une étape

Le déroulement d'une étape est matérialisé par une figure dans § III.4.1 du chapitre précédent. Reprenons cette figure en y ajoutant les différentes interactions avec l'utilisateur représentées par des fenêtres. Ces fenêtres définissent le dialogue avec l'utilisateur. Seules deux fenêtres sont nécessaires au déroulement d'une étape :

- la **fenêtre 1** qui permet à l'utilisateur de formuler sa demande,
- la **fenêtre 2** qui permet à l'utilisateur de visualiser les images et d'effectuer son choix.

Les autres fenêtres ne sont affichées que sur ordre de l'utilisateur :

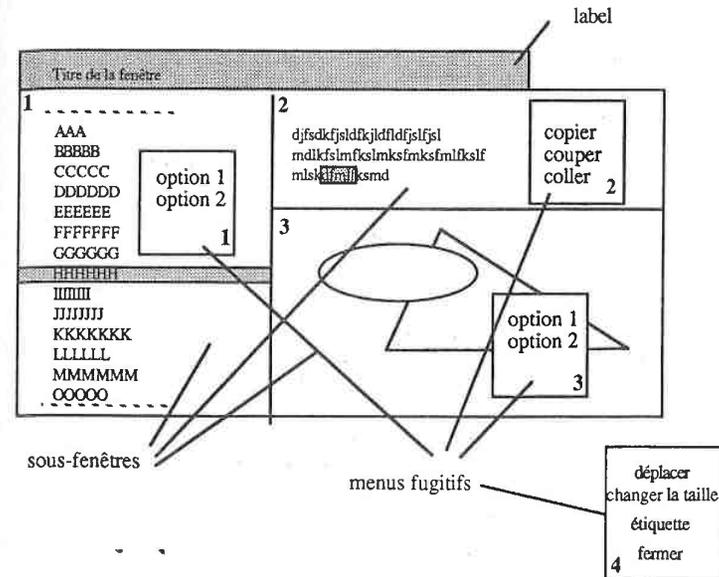
- la **visualisation du thesaurus** peut servir à l'utilisateur pour voir l'environnement sémantique d'un terme, pour choisir un terme à mettre dans sa demande, ou encore pour voir des images illustrant un concept donné (concept représenté par un terme du thesaurus).
- la **visualisation de la requête système** offre à l'utilisateur les moyens de contrôler l'interprétation de sa demande, et de la modifier si nécessaire.



Avant de voir plus en détails l'organisation des fenêtres, étudions la philosophie générale des outils de dialogue écrits en Smalltalk 80.

II.2.2 L'interface utilisateur en Smalltalk

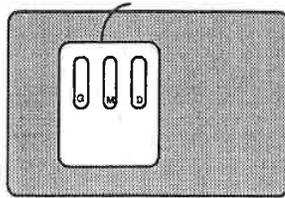
La figure suivante montre l'allure générale d'une fenêtre Smalltalk :



Une fenêtre est constituée d'un label, et de sous-fenêtres pouvant contenir des informations de types différents. Ainsi dans notre exemple la sous-fenêtre 1 contient une liste, la sous-fenêtre 2 du texte, et la sous-fenêtre 3 un graphique. A chaque sous-fenêtre peuvent être associées des actions sous la forme de menus déroulants, dits aussi "menus fugitifs", car ils n'apparaissent que si l'utilisateur a positionné le curseur sur la sous-fenêtre concernée et s'il a appuyé, en même temps, sur un certain bouton de la souris. Les menus fugitifs 1, 2 et 3 sont associés respectivement aux sous-fenêtres 1, 2 et 3. Le menu fugitif 4 concerne la fenêtre entière, il permet d'agrandir la fenêtre, de la déplacer, de la fermer ou encore de la réduire sous la forme d'une étiquette représentée par son label.

L'utilisateur manipule les fenêtres essentiellement à l'aide d'une souris à trois boutons, chaque bouton ayant des fonctions bien précises.

La figure suivante matérialise la souris optique à trois boutons utilisée dans RIVAGE. Les fonctions associées aux boutons sont les suivantes :



- le bouton gauche (G) permet d'interagir sur le contenu des sous-fenêtres :
 - sélectionner un élément dans une liste (sous-fenêtre 1),
 - sélectionner une partie de texte dans une sous-fenêtre texte (sous-fenêtre 2),
 - sélectionner un morceau d'un dessin (sous-fenêtre 3)...
- le bouton du milieu (M) fait apparaître le menu fugitif associé à la sous-fenêtre pointée par le curseur de la souris (menus 1, 2 et 3),
- le bouton de droite (D) fait apparaître le menu fugitif associé à la fenêtre entière (menu 4).

II.2.3 Les fenêtres utilisées dans RIVAGE

Les deux principales fenêtres, permettant d'effectuer une recherche d'images en suivant le processus EXPRIM, sont :

- la fenêtre n° 1 ou "fenêtre de formulation de la demande",
- la fenêtre n°2 ou "fenêtre de visualisation et choix".

II.2.3.1 La formulation de la demande

La figure suivante représente la fenêtre n° 1 :

Une Demande pour la base descriptive Ministère de la Culture; étape:3	
<p>1</p> <p>Morphologie de la photographie</p> <p>Auteur de la photographie</p> <p>Contenu de la photographie</p> <p>Connotation de la photographie</p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> consultation du thesaurus requête générée base résultat </div>	<p>2</p> <p>VALLEE</p> <p>MONTAGNE</p> <p>TORRENT</p> <div style="border: 1px solid black; padding: 5px; margin-top: 10px;"> consultation du thesaurus ajouter un terme enlever les termes sélectionnés garder les termes sélectionnés autres termes </div>
<p>3</p> <p>Recherche Large</p>	<p>4</p> <p>Recherche Précise</p> <p>Sélectionner le type de recherche désiré.</p>

La sous-fenêtre 1 visualise les facettes de la demande utilisateur ; quand une facette est sélectionnée, son contenu, c'est-à-dire une liste de termes, est affiché dans la sous-fenêtre 2. Les sous-fenêtres 3 et 4 représentent des boutons réagissant globalement comme un interrupteur: quand l'un est allumé (en inverse vidéo), l'autre est éteint. Ces boutons permettent à l'utilisateur de choisir le type de sa recherche ; sur la figure c'est la recherche large qui a été choisie.

Les menus fugitifs associés aux sous-fenêtres indiquent les options que peut choisir l'utilisateur.

Le menu de la sous-fenêtre 1 possède les options suivantes :

- "consultation du thesaurus" ; cette option permet à l'utilisateur de visualiser le thesaurus et de choisir les termes de sa demande en parcourant le thesaurus. Si lors de l'appel de cette option une facette est sélectionnée, seuls les champs sémantiques associés à la facette seront présentés. Cette nouvelle fenêtre est étudiée dans la suite.
- "requête générée" ; lorsque l'utilisateur choisit cette option, la formulation de sa demande est considérée comme terminée et le système lui présente son interprétation, c'est-à-dire la requête système.
- "base résultat" ; le choix de cette option termine la phase "Avant-Visualisation", le système interprète la demande et sélectionne les images pertinentes, puis propose à l'utilisateur de les visualiser grâce à la fenêtre 2.

Les options de la sous-fenêtre 2 sont :

- "consultation du thesaurus" ; cette option est similaire à celle, de même nom, de la sous-fenêtre 1, excepté que lorsqu'un terme est sélectionné le thesaurus s'affiche en visualisant l'environnement sémantique du terme.
- "ajouter un terme", "enlever les termes sélectionnés" et garder les termes sélectionnés" sont des options permettant de mettre à jour la demande.
- "autres termes" ; cette option n'est utilisable que lorsqu'une phase "Après-Visualisation" a eu lieu. Elle permet de proposer à l'utilisateur d'autres termes que ceux déjà proposés par la reformulation. Cette option correspond à l'utilisation du niveau d'expression présentée dans le § IV.5.2.

Si l'option "base résultat" a été choisie, la fenêtre 2 permettant la visualisation et le choix des images, apparaît à l'écran.

II.2.3.2 Visualisation et choix d'images

La fenêtre permettant de visualiser et de choisir les images se présente ainsi :

Base résultat extraite de la base descriptive du Ministère de la Culture ; étape: 3		
12: personnage sur...pont...: choisie	n°du vidéodisque connotation	VALLEE_D TORRENT PAYSAGE_D PERSONNAGE PONT_D PIERRE_D
5 : promeneurs...du xxeme siècle:	contenu	
239: paysage ... de Grèce: rejetée	morphologie commentaire précision sur l'objet... auteur titre	
<div style="border: 1px solid black; padding: 5px; width: fit-content;"> choix rejet annulation du choix </div>	<div style="border: 1px solid black; padding: 5px; width: fit-content;"> Fin de la Visualisation </div>	
1	2	3
	4	

Quatre sous-fenêtres sont nécessaires pour visualiser les informations attachées aux images. La sous-fenêtre 1 présente la liste des titres des descriptions pertinentes ; lorsqu'un titre est sélectionné dans la liste, l'image associée s'affiche sur le moniteur. L'utilisateur peut alors à l'aide du menu fugitif choisir ou rejeter cette image. Dans cet exemple, l'image 12 a été choisie, la 239 a été rejetée et la 5 est une image neutre. Les autres sous-fenêtres servent à visualiser le contenu des descriptions. Dans la sous-fenêtre 2 s'affichent les libellés des champs de la description ; quand un libellé est sélectionné, le contenu du champ s'affiche dans la sous-fenêtre 3 si c'est une liste de termes (contenu, morphologie, ...), ou dans la sous-fenêtre 4 si c'est du texte (titre, commentaire).

Remarque : dans la sous-fenêtre 3, on aperçoit des termes ayant la terminaison "_D", cette notation permet de considérer le thesaurus comme une taxonomie (cf chapitre 2, § III.1.2). Un terme muni de ce suffixe identifie des représentations du concept, le générique (terme sans suffixe) que ses spécifiques n'identifient pas. Avec cette notation, "autre_bateau" du § III.1.2 du chapitre 2 s'écrit "bateau_D".

Lorsque l'utilisateur a terminé ses choix, il sélectionne l'option "fin de la visualisation". Le système va alors analyser ses choix, mettre à jour le niveau d'expression dans le thesaurus, puis reconstruire une nouvelle demande, qu'il va proposer à l'utilisateur à l'aide de la fenêtre 1 décrite précédemment.

II.2.3.3 Les autres fenêtres

Deux autres fenêtres peuvent être affichées sur demande de l'utilisateur :

- la visualisation du thesaurus pour aider l'utilisateur à formuler sa demande,
- la visualisation de la requête système pour montrer à l'utilisateur l'interprétation par le système de sa demande.

La visualisation du thesaurus :

Consultation du thesaurus du Ministère de la Culture.			
ARRIERE PAYS COLLINE GLACIER GROTTE ILE	Alphabétique	Par thème	Le champ sémantique
	2	3	GEOGRAPHIE PHYSIQUE
MONTAGNE PLAINE PLATEAU ROCHER	Les spécifiques		Le générique
	COL DE MONTAGNE EPERON ROCHEUX PRECIPICE SOMMET VALLEE VOLCAN MONTAGNE-D		RELIEF TERRESTRE
<div style="border: 1px solid black; padding: 5px; width: fit-content;"> sélectionner plus spécifique plus général sens proche voir les images </div>			Les termes de sens proche
			GEOGRAPHIE MORPHOLOGIQUE COLLINE
			Les synonymes
			Chaîne de montagnes Mont
			Le nombre d'images
		150	
		Le poids d'expressivité	
		0.50	
1			

La fenêtre visualisant le thesaurus comporte de nombreuses sous-fenêtres permettant de montrer toutes les informations attachées à un terme. Le thesaurus peut être vu sous deux formes :

- alphabétique : la sous-fenêtre 1 contient alors la liste alphabétique des termes ; quand un terme est sélectionné dans cette liste, ses caractéristiques apparaissent dans les autres sous-fenêtres.
- par thème : cette forme de visualisation permet de parcourir le thesaurus en empruntant les liens sémantiques entre les termes. La figure précédente illustre ce type de visualisation, elle présente le résultat de l'option "consultation du thesaurus" associée à la sous-fenêtre 2 de la fenêtre "formulation de la demande" (cf § II.2.3.1). Avant l'appel de cette option le terme "MONTAGNE" était sélectionné dans la demande, la visualisation du thesaurus propose alors la partie du thesaurus où se situe ce terme. La sous-fenêtre 1 visualise le terme et ses frères, c'est-à-dire les termes spécifiques de son générique (RELIEF TERRESTRE). Le terme MONTAGNE

apparaît comme sélectionné, et ses caractéristiques sont affichées dans les autres sous-fenêtres :

- ses spécifiques,
- son générique,
- son champ sémantique,
- les termes de sens proche,
- les synonymes (non-descripteurs),
- le nombre d'images représentant le concept,
- son poids d'expressivité.

A l'aide des options du menu fugitif, l'utilisateur peut parcourir le thesaurus, en montant dans la hiérarchie ("plus général"), en descendant ("plus spécifique"), ou transversalement ("sens proche"). Il peut aussi sélectionner un terme ("sélectionner") pour l'ajouter dans sa demande, ou visualiser les images représentant le concept ("voir les images") pour se faire une idée sur les représentations imagées du concept dans la base d'images.

Le choix entre les deux formes de visualisation du thesaurus est possible grâce aux sous-fenêtres 2 et 3 matérialisant des boutons de choix.

La visualisation de la requête système :

Une requête système pour la base descriptive Ministère de la Culture; étape:3
20 images sélectionnées dont 15 nouvelles.

<p>MONTAGNE...150 images TORRENT.....30 images PLAN D'ENSEMBLE..300 ima</p> <p>ajouter enlever fin de formulation</p> <p>1</p>	<p>contenu</p> <p>ajouter enlever</p> <p>2</p>	<p>COL DE MONTAGNE EPERON ROCHEUX PRECIPICE SOMMET FJORD VALLEE_D EUPTION VOLCANIQUE VOLCAN_D MONTAGNE_D</p> <p>ajouter enlever</p> <p>3</p>
------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

La requête système, que montre la figure précédente, correspond à une interprétation en recherche intermédiaire de la demande suivante :

Contenu de la photographie : MONTAGNE, VALLEE, TORRENT.

Morphologie de la photographie : PLAN D'ENSEMBLE.

La sous-fenêtre 1 montre les différents éléments disjonctifs de la requête associés aux

nombre d'images qu'ils rapportent. Nous rappelons brièvement que dans notre système, une requête est constituée d'un ensemble d'éléments disjonctifs connectés par des ET, chaque élément étant un ensemble de critères connectés par des OU.

Dans notre cas, la requête est constituée de trois éléments nommés MONTAGNE, TORRENT et PLAN D'ENSEMBLE connectés par des ET (le nom des éléments est obtenu en prenant le nom du terme le plus générique des termes présents dans les critères). Lorsqu'un élément est sélectionné, ici "MONTAGNE", le ou les champs figurant dans les critères sont affichés dans la sous-fenêtre 2 ; en sélectionnant un de ces champs, les valeurs des critères où figure ce champ sont affichées dans la sous-fenêtre 3. Ainsi l'élément "MONTAGNE" est constitué des critères : contenu = COL DE MONTAGNE, contenu = EPERON ROCHEUX, contenu = PRECIPICE... Cette requête peut s'écrire sous la forme booléenne suivante :

contenu = (COL DE MONTAGNE OU EPERON ROCHEUX OU PRECIPICE OU ...) ET (contenu = TORRENT) ET (morphologie = PLAN D'ENSEMBLE).

A ce niveau du dialogue l'utilisateur, s'il le souhaite, peut modifier cette requête en ajoutant de nouveaux critères et de nouveaux éléments disjonctifs. Cette mise à jour nécessite de connaître les champs des descriptions, le thesaurus, l'effet des opérateurs booléens, ainsi que la signification des différentes colonnes de cette fenêtre. Ce type de formulation de requêtes est quelque peu similaire à ceux des systèmes documentaires que l'on trouve dans le commerce. C'est ce type de formulation qui nous servira, à titre de comparaison, lors de l'évaluation de notre système.

Lorsque l'utilisateur a visualisé ou modifié la requête système, il peut visualiser les images sélectionnées en choisissant l'option "fin de formulation".

Nous venons de voir le dialogue de base du prototype RIVAGE, qui nous a permis de mettre au point la méthode proposée, en réalisant les trois phases du processus EXPRIM. Ce dialogue a évolué, puisqu'une structure de type hypertexte a été ajoutée afin d'améliorer les conditions d'une recherche. Il nous a semblé important de présenter ce dialogue, afin de pouvoir montrer, par la suite, ce que peut apporter l'ajout d'une telle structure dans un système de recherche d'images, utilisant le processus EXPRIM. Cette présentation fera l'objet du dernier chapitre ; mais auparavant, étudions plus profondément ce prototype en examinant la structure des objets manipulés.

III Présentation des objets

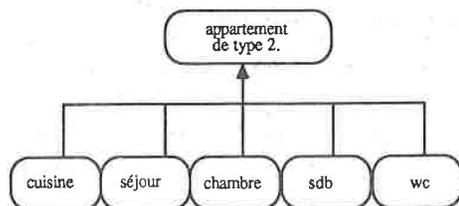
Ce paragraphe a pour but de présenter la structure des objets que manipule RIVAGE. L'utilisation d'un langage objet nous a permis de définir des objets à structure complexe. Pour décrire ces objets, il nous faut donc utiliser un formalisme adapté à la représentation de structures complexes. Le formalisme que nous allons utiliser se situe à la frontière entre les modèles dits "sémantiques" [Hul 87] et les modèles "objets" [Cau 89][Léo 89], car nous ne manipulerons que certains concepts d'abstraction permettant de définir la partie statique des objets, mais nous ne définirons pas leur comportement, c'est-à-dire la dynamique, car sur ce point les chercheurs ne sont pas toujours en accord sur la représentation à prendre.

III.1 Formalisme utilisé

Pour la représentation de la partie statique des objets nous allons utiliser une représentation graphique de quelques concepts d'abstraction.

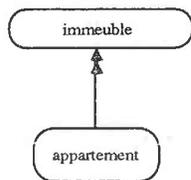
L'agrégation [Smi 77]: cette abstraction permet de considérer plusieurs objets de types différents comme un unique objet (objet agrégat). Ce qui peut se représenter graphiquement de la manière suivante :

Exemple :



Le regroupement [Cau 89]: permet de considérer plusieurs objets de même type comme un seul objet (objet collection).

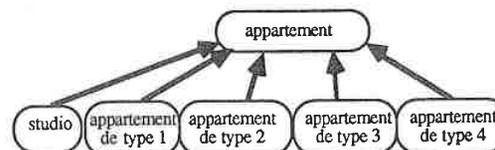
Exemple :



La spécialisation/généralisation [Smi 77]: cette abstraction permet de définir des types d'objets qui spécialisent un type d'objet appelé type générique. La notion de

spécialisation/généralisation s'apparente à la hiérarchie des classes et au mécanisme d'héritage des langages orientés objets. En effet, avec cette abstraction, un type d'objet spécifique hérite de la structure et du comportement de son type d'objet générique.

Exemple :

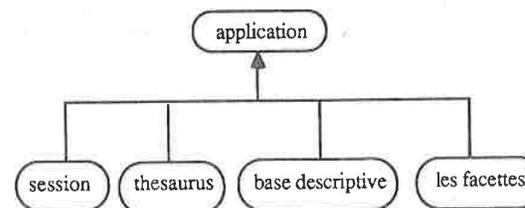


Remarque : dans les descriptions qui vont suivre, lorsque nous utilisons cette abstraction, les types objets qui spécialisent le type objets générique épuisent toutes les possibilités de représentations. Ainsi, avec l'exemple précédent, il n'existe pas d'autres types d'appartements que ceux qui spécialisent le type objet appartement.

Ces trois concepts d'abstraction vont nous permettre de décrire les structures des objets manipulés dans RIVAGE.

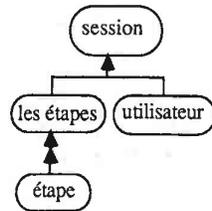
III.2 Les objets manipulés

L'objet principal dans RIVAGE est un objet agrégat qui décrit l'ensemble de l'application, cet objet s'appelle **application** :



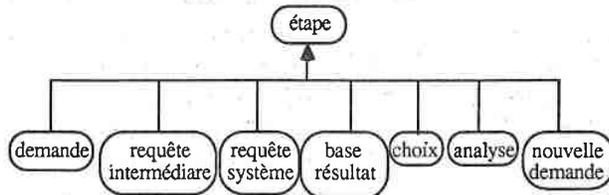
A partir du type d'objet **application** on peut atteindre n'importe quel type d'objet utilisé dans le système. Ainsi le type d'objet **session** décrit une session d'interrogation et le type d'objet **les facettes** décrit l'ensemble des facettes que l'on peut trouver dans une demande. Etudions plus précisément chaque constituant.

III.2.1 La session



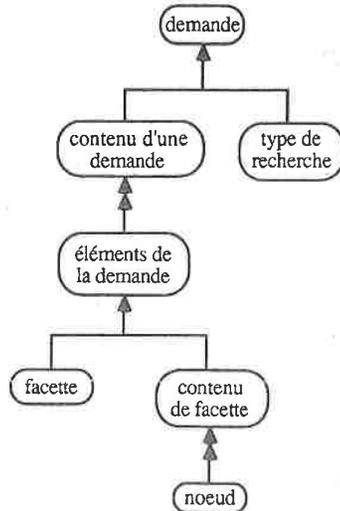
Une session d'interrogation est constituée d'étapes de recherche effectuées par un utilisateur.

L'étape :



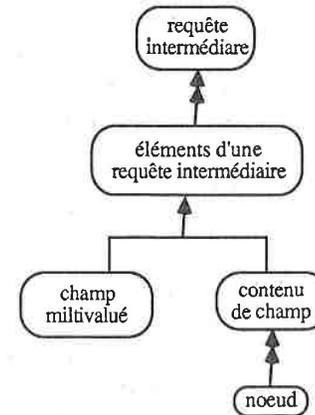
L'objet étape est sans nul doute un des objets les plus complexes du système, voyons successivement chacun de ses constituants :

La demande :



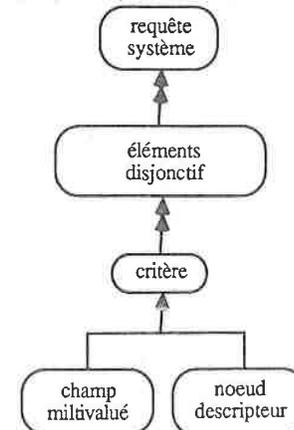
En considérant que les structures des objets **facette** et **noeud** sont décrites plus loin, l'objet **demande** est entièrement défini avec ce schéma. Le type de la recherche représente le paramètre qui intervient lors de l'interprétation de la demande (cf chapitre 2, § IV.2.2.2).

La requête intermédiaire :



La requête intermédiaire est le résultat de la ventilation du contenu des facettes de la demande dans les champs utilisés dans les descriptions d'images. Les objets **noeud** et **champ** sont décrits plus loin.

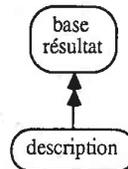
La requête système :



La requête est le résultat de l'interprétation de la requête intermédiaire, c'est une formule

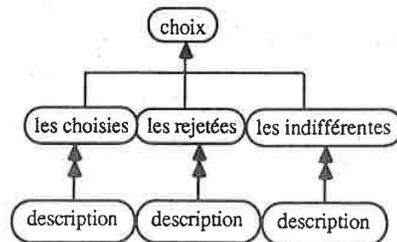
booléenne où les éléments disjonctifs sont connectés par des ET, et les critères à l'intérieur des éléments par des OU. L'objet **nœud descripteur** est le type du nœud utilisé pour décrire une image, sa structure est donnée plus loin.

La base résultat :



La base résultat est un ensemble de descriptions des images pertinentes pour le système, c'est-à-dire qui sont en correspondance avec la requête système.

Le choix :

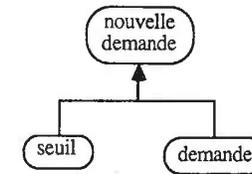


Le choix est constitué de trois ensembles de descriptions contenant respectivement les descriptions des images choisies, rejetées et indifférentes (neutres).

L'analyse :

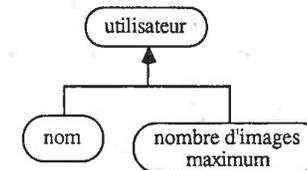
L'analyse contient à la fois les résultats de l'analyse en pertinence des choix de l'utilisateur, et le résultat de la propagation des poids de pertinences dans le thesaurus. Pour le moment le résultat de la propagation se résume à conserver le poids maximum des poids d'expressivité du thesaurus, ces poids vont servir à définir le seuil nécessaire à la reformulation de la demande. Nous verrons plus tard que pour mieux contrôler la reformulation, le résultat de la propagation doit être plus complexe.

La nouvelle demande :



La nouvelle demande est constituée du seuil, calculé à partir du résultat de la propagation, et de la demande, construite à partir du niveau d'expression dans le thesaurus et du seuil.

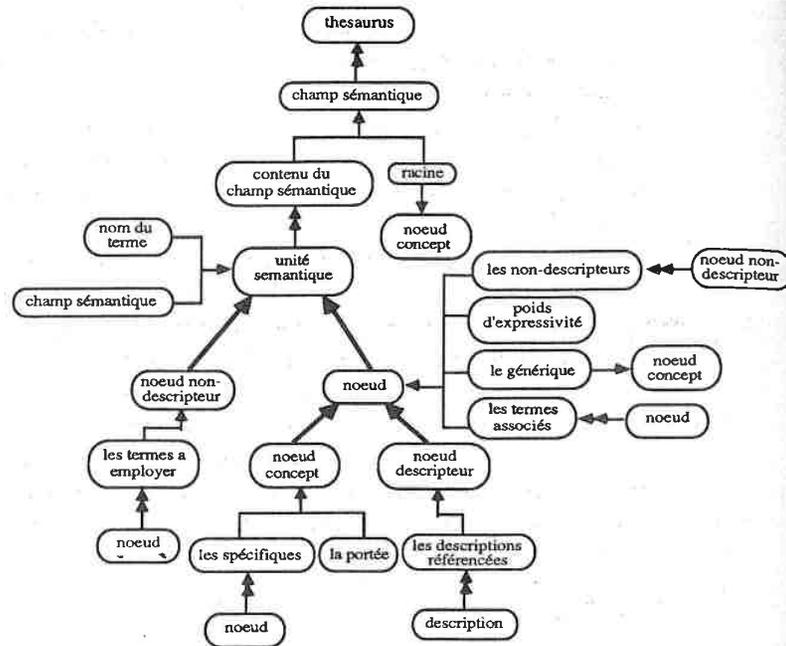
Pour terminer la description d'une session, il nous faut voir la structure de l'objet décrivant un utilisateur.



Cette structure possède, dans la version actuelle du prototype, deux composants : le nom de l'utilisateur et le nombre d'images maximum que l'utilisateur accepte de visualiser. Ce nombre permet de savoir à quel moment le système doit rétrécir la portée d'une recherche. Cette structure est, pour le moment, assez pauvre, nous souhaitons y ajouter, dans un avenir assez proche, d'autres informations (niveau de l'utilisateur, le type d'utilisation de l'image, ...) qui permettraient un meilleur contrôle de la recherche.

III.2.2 Le thesaurus

Voici la structure générale du thesaurus :



Dans cette structure complexe, l'abstraction de spécialisation/généralisation est utilisée pour représenter les différentes formes possibles de l'objet **unité sémantique**. En effet, l'unité sémantique peut avoir plusieurs aspects :

- noeud non-descripteur,
- noeud,

l'objet **noeud** étant lui aussi une généralisation des types d'objets suivants :

- noeud concept,
- noeud descripteur.

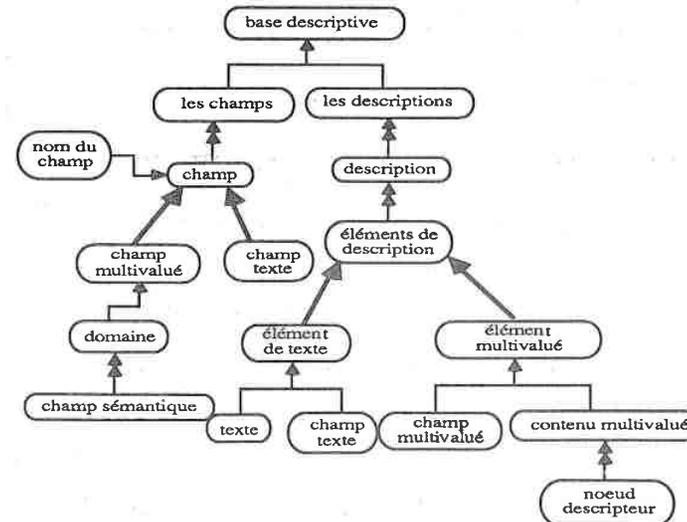
Le **noeud non-descripteur** permet de représenter les termes du thésaurus qui ont le statut de non-descripteur. Le **noeud** représente tous les autres termes du thésaurus. Pour transformer le thésaurus sous forme de taxonomie, nous avons fait la différence, parmi les termes qui ne sont pas des non-descripteurs, entre les termes qui expriment un concept (les termes génériques) et les termes qui sont une représentation du concept (les termes spécifiques). Cette transformation nous a permis de préciser une indexation exclusivement aux feuilles, c'est-à-dire aux termes n'ayant pas de spécifiques. C'est pour cette raison que le type objet **noeud** a été décomposé en deux sous-types : le **noeud concept**, et le **noeud descripteur** utilisé pour

décrire les images. Tous ces sous-types héritent des propriétés de leur type générique, ce qui permet une certaine "factorisation" de la structure. Par exemple, le noeud descripteur possède les propriétés suivantes : le nom du terme, le champ sémantique, les non-descripteurs, le poids d'expressivité, le générique, les termes associés et les descriptions référencées ; dans cette liste seule la dernière propriété n'est pas héritée.

La portée d'un noeud concept représente le nombre d'images qui possèdent une représentation du concept, c'est-à-dire le nombre d'images que l'on obtiendrait en réponse à une demande ne contenant que ce concept. L'objet **les descriptions référencées** contient l'ensemble des descriptions qui ont été décrites avec le terme du noeud descripteur auquel il est rattaché ; cet objet joue le rôle des fichiers inverses que l'on trouve dans les systèmes classiques de recherche d'informations.

Nous verrons, lors de la présentation de quelques traitements, comment cette structure est utilisée, mais auparavant étudions la structuration d'une base descriptive.

III.2.3 La base descriptive

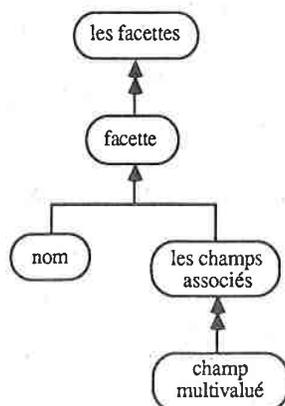


Une **base descriptive** est constituée d'une liste de champs et d'un ensemble de descriptions. Un champ peut avoir plusieurs formes suivant le type d'information qu'il contient.

Une description peut alors être décrite par un ensemble d'éléments de description. Un élément de description peut être de différentes natures suivant le type du champ qui figure dans cet élément. D'autres types de champ peuvent être considérés, les deux types définis ici sont les plus importants. En effet, les champs de type texte contiennent les données informationnelles annexées à l'image (titre, commentaire) et les champs dit "multivalués" des données de recherche.

Terminons la description des objets par la présentation de l'environnement de formulation de la demande, c'est-à-dire la description des facettes.

III.2.4 Les facettes



L'objet **les facettes** contient l'ensemble des facettes qui peuvent figurer dans une demande. A chaque facette est associée une liste de champs, appartenant à la liste de champs de la base descriptive ; c'est à l'intérieur de ces champs que seront ventilés les termes présents dans la facette.

La structure assez générale des objets offre au prototype RIVAGE la possibilité de s'adapter à toute application de recherche d'images ayant un thesaurus et un ensemble de descriptions construit avec l'aide du thesaurus. L'adaptation nécessite de déterminer les différentes facettes et de les relier aux champs des descriptions, et de donner au thesaurus le statut de taxonomie. Ces adaptations sont nécessaires à l'utilisation du processus EXPRIM, basé sur l'apprentissage des besoins pour l'interrogation.

Les objets que nous venons de décrire possèdent des structures complexes, l'utilisation du langage orienté objet Smalltalk 80 nous a permis d'implanter facilement ces structures.

La structure des objets définie, voyons comment les parties heuristiques ont été réalisées dans la version actuelle de RIVAGE, ainsi que les modifications de structure qu'elles ont entraînées.

IV Description des fonctions principales

Dans le chapitre 2, nous avons donné une modélisation plus ou moins complète formalisant l'univers de notre système. Cette modélisation nous a permis de présenter les algorithmes les plus importants de notre méthode, et aussi les parties nécessitant une expérimentation que nous avons appelées "parties heuristiques". Ces parties heuristiques interviennent dans les fonctions suivantes de la méthode :

- la ventilation et l'interprétation de la demande,
- la déformation de la demande,
- l'analyse des choix de l'utilisateur,
- la propagation des poids dans le thesaurus (la fonction F),
- le choix du seuil pour la reformulation.

Les choix pris pour réaliser ces parties heuristiques sont décrits, lorsque cela n'introduit pas trop de lourdeur, en se plaçant dans le modèle défini au chapitre 2. Un exemple, plus proche de la réalisation, utilisant les structures objets décrites dans le paragraphe précédent, est donné. Nous verrons à travers cet exemple comment, en attachant des méthodes aux objets, c'est-à-dire un comportement, une de ces solutions peut être réalisée.

IV.1 La ventilation et l'interprétation de la demande

IV.1.1 Quelques remarques

Lors de la ventilation de la demande dans la requête intermédiaire, on suppose que tous les termes présents dans la demande existent dans le thesaurus et que ce ne sont pas des non-descripteurs, c'est-à-dire qu'on suppose que les analyses lexicale et terminologique, définies dans la partie 2 (cf chapitre 3, § II.2.1.2), ont été effectuées.

Un problème subsiste toutefois lorsqu'il existe, dans la demande, un terme qui ne rapporte pas d'image. Ce terme ne peut être gardé car lors de l'interprétation de la requête intermédiaire, si ce terme est connecté avec un ET avec d'autres éléments de recherche, la portée de la requête système sera nulle. Plusieurs solutions peuvent être choisies :

- supprimer le terme ; c'est la solution la plus simple, mais elle ne nous satisfait pas, car en faisant cette suppression, on enlève de l'information attachée aux besoins de l'utilisateur.

- remplacer le terme par son terme générique ; cette solution est satisfaisante, c'est celle que nous avons choisie.

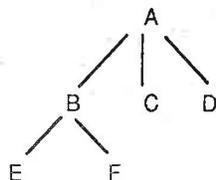
- modifier le thesaurus pour qu'il n'existe plus ce genre de problème, il suffit de donner aux termes qui ne rapportent pas d'images le statut de non-descripteur, et de les relier à leur terme générique par la relation de synonymie "employer / employé pour". Nous n'avons pas choisi cette dernière solution, car notre base d'images n'étant pas complète, nous n'avons pas voulu appauvrir la structure du thesaurus.

Remarque concernant les heuristiques définies dans le chapitre 3 de la partie 2 : la requête intermédiaire peut être interprétée de trois manières différentes selon le type de recherche. La dernière étape de cette interprétation est l'extension. Cette étape permet au concept présent dans la demande d'être exprimé sous une forme étendue : les concepts sont remplacés par une disjonction de tous les termes spécifiques "feuilles" du terme exprimant le concept. On peut remarquer que cette étape réalise l'heuristique n°4 définie dans le chapitre 3 de la partie 2 .

Une autre remarque peut être faite lorsque deux concepts, l'un étant générique de l'autre, sont présents dans une même demande. En effet, dans ce cas de figure l'heuristique n°3 définie dans le chapitre 3 de la partie 2 indiquait que dans le cas d'une recherche large il fallait supprimer le terme le plus spécifique, alors que dans le cas d'une recherche précise c'est le terme le plus générique qui devait être supprimé. Voyons, à travers un exemple, ce que donne ce cas de figure dans notre approche :

Exemple :

Soit la partie d'un thesaurus :



et la requête intermédiaire :

RI = A, B

après une interprétation sans extension dans le cadre d'une recherche précise, on obtiendra la requête système suivante :

RS = A ET B et après l'extension RS = (E OU F OU C OU D) ET (E OU F) ce qui est équivalent à (E OU F) c'est à dire à RI = B.

Par contre après une interprétation sans extension dans le cadre d'une recherche large, on obtiendra la requête système suivante :

RS = A OU B et après l'extension RS = (E OU F OU C OU D) OU (E OU F) ce qui est équivalent à (E OU F OU C OU D) c'est à dire RI = A.

Ainsi on s'aperçoit que l'heuristique n°3 donnée dans le chapitre 3 de la partie 2 est réalisée par l'interprétation.

Ces quelques remarques prouvent que la méthode proposée offre un cadre assez général puisque certains traitements heuristiques donnés dans le projet ESPRIT sont pris en compte dans la méthode.

Voyons maintenant l'interprétation dans le cas d'une recherche intermédiaire, et plus précisément comment définir la fonction ProcheSémaniquement utilisée dans cette interprétation.

IV.1.2 La fonction ProcheSémaniquement

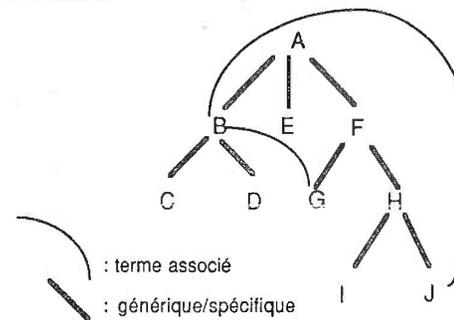
Cette fonction indique la proximité sémantique entre deux termes du thesaurus ; cette évaluation ne peut s'effectuer qu'à partir de la structure du thesaurus. Le thesaurus que nous utilisons possède deux grandes relations :

- générique/spécifique,
- terme associé.

Nous pouvons définir d'une manière simple que deux termes sont proches sémantiquement s'ils sont reliés par la relation "terme associé" ou s'ils ont même générique. Mais cette définition demeure trop restrictive.

Exemple :

Soit le morceau d'un thesaurus



Il ne fait aucun doute sur la proximité sémantique entre C et D, J et B ou encore E et F, mais que dire sur celle entre G et E, H et B, ou J et G? En effet, quand les liens sont directs cela ne pose aucun problème, mais il existe des cas où deux termes peuvent être proches sémantiquement alors qu'ils n'ont pas même générique direct, et qu'ils ne sont pas reliés directement par une relation "terme associé". Ainsi la notion de proche sémantiquement doit être relativisée et elle doit être définie de telle manière que le système puisse agir dessus.

Voici comment nous avons défini la notion de proximité sémantique entre deux termes :

ProcheSémantiquement(u_i, u_j):

Deux nouveaux paramètres doivent être introduits:

- nbLienGen, qui représente le nombre maximum de liens "générique/spécifique" que l'on peut parcourir pour rechercher un terme générique commun à deux termes,
- nbLienTA, qui représente le nombre maximum de liens "terme associé" que l'on peut parcourir pour déterminer si deux termes sont de sens proches.

La fonction devient :

ProcheSémantiquement($u_i, u_j, nbLienGen, nbLienTA$):

ProcheSémantiquement <- mêmeGénérique($u_i, u_j, nbLienGen$)

OU

sensProche($u_i, u_j, nbLienTA$)

La fonction mêmeGénérique détermine si les deux termes ont un générique commun en parcourant au plus nbLienGen liens. Si cette fonction a comme paramètres les deux mêmes unités, le résultat renvoyé sera "vrai".

Exemple :

Sur la figure, si nbLienGen = 2, on a :

- J et G qui sont proches sémantiquement, car ils ont un générique commun à au plus deux niveaux au-dessus, qui est F,
- I et E ne sont pas proches sémantiquement, car ils ont bien un générique commun A mais il se trouve à trois niveaux au-dessus de I.

La fonction sensProche regarde dans le thesaurus si deux termes sont séparés par au plus nbLienTA lien de type "terme associé" :

Exemple :

Sur la figure, si nbLienTA = 2, on a G et J qui sont proches sémantiquement.

Les deux paramètres, nbLienGen et nbLienTA, permettent un certain contrôle de la proximité sémantique, mais quelques remarques doivent être faites sur quelques cas particuliers :

Remarques :

- nbLienGen = 1 et nbLienTA = 1 : on obtient la définition simple de la proximité sémantique donnée au début du paragraphe,
- nbLienGen = 0 et nbLienTA = 0 : il n'existe pas de termes dans le thesaurus proches sémantiquement, l'interprétation correspond alors à une interprétation de recherche précise.
- nbLienTA = +∞ : si le thesaurus possède beaucoup de relations de type "terme associé", on peut arriver à dire que deux termes sont proches sémantiquement alors que leurs sens réels sont totalement différents.

Exemple :

mer ↔ port ↔ phare ↔ tour ↔ chateau.

Par contre si le thesaurus est pauvre en relation de type "terme associé", la solution de fixer une fois pour toute la valeur de nbLienTA à +∞ peut être prise.

- nbLienGen = +∞ : deux termes sont proches sémantiquement dès qu'ils appartiennent au même champ sémantique du thesaurus. Cette remarque engendre d'autres. En effet, on s'aperçoit que si un champ sémantique ne possède pas une structure hiérarchique très profonde, le paramètre nbLienGen ne peut varier que dans un intervalle de valeurs étroit. Réciproquement si un champ sémantique a une structure hiérarchique très profonde l'intervalle de variation de nbLienGen sera plus large. En raisonnant à l'extrême, on peut se demander : est-ce que le paramètre nbLienGen doit varier ? N'est-il pas figé à une certaine valeur ? Doit-il être global ou spécifique à un champ sémantique donné ? N'ayant pu donner de réponses à ces questions, nous avons choisi de faire varier ce paramètre d'une manière globale pour l'ensemble des champs sémantiques. Ce choix nous donne un moyen supplémentaire de déformation de la requête.

Cette fonction prend en compte certaines formes de proximité sémantique, mais d'autres formes de proximité auraient pu être considérées. Par exemple en manipulant les deux relations sémantiques comme une seule relation, ainsi le calcul de la proximité sémantique revient à évaluer la taille du chemin reliant deux termes. De cette manière, le terme J et C de la figure pourrait être proche sémantiquement en parcourant seulement 2 arcs. Mais comme la sémantique de ces deux relations est totalement différente, nous n'avons pas choisi cette solution qui à notre avis demande une étude assez sérieuse sur la manière de représenter et d'utiliser des sémantiques différentes dans un graphe. Toutes nouvelles solutions apportées auront sans nul doute un caractère heuristique, elles devront être expérimentées afin d'évaluer leur influence sur

le processus de recherche.

IV.2 La déformation de la demande

Nous tenons à dire que, comme pour le paragraphe précédent, de nombreuses solutions existent pour résoudre le problème de la déformation, nous n'avons pas étudié toutes ces solutions car cette étude pourrait faire l'objet d'une thèse complète ; nous avons choisi la solution qui nous semblait la mieux adaptée dans l'état actuel de nos réflexions.

La déformation intervient lorsque la sélection des documents par la requête système n'est pas satisfaisante. Une sélection n'est pas satisfaisante quand le nombre de documents est trop important ou insuffisant. La déformation agit sur la requête intermédiaire (cf chapitre 2, § IV.2.2.3) soit sur les termes, soit sur son interprétation.

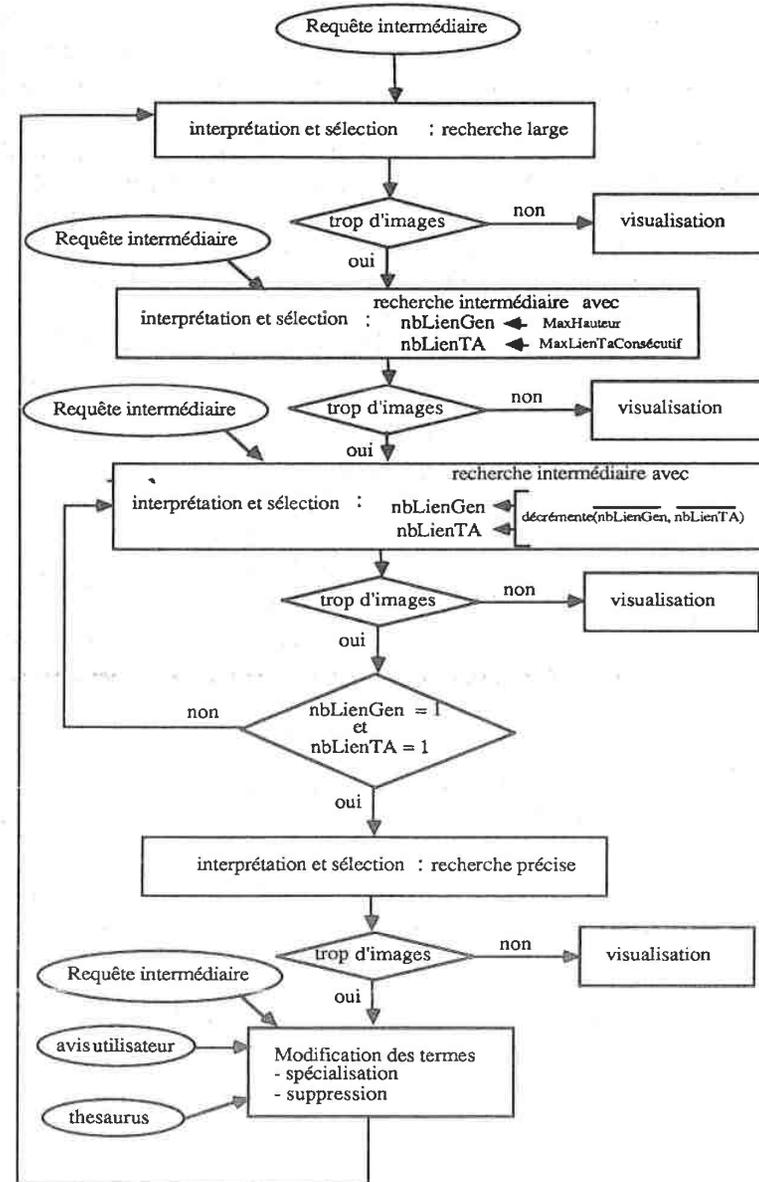
La stratégie de déformation que nous avons adoptée est progressive, elle intervient tout d'abord sur l'interprétation, puis en cas d'échec, sur les termes et lorsque cela est nécessaire l'utilisateur est sollicité. Nous avons choisi cette stratégie afin de proposer des images le plus rapidement possible à l'utilisateur sans que celui-ci ait à se poser de nombreuses questions sur le bon usage d'un terme avec un autre, ou d'un terme dans une facette. La recherche étant progressive, plus vite l'utilisateur fait un choix, plus vite le système l'analyse pour essayer de comprendre les besoins, et de les formuler sous la forme d'une demande qui doit donner à l'utilisateur des idées sur la manière d'exprimer ses besoins.

Nous avons mis au point quatre stratégies de déformation associées aux situations suivantes :

- le type de la recherche est "large" et la sélection a trop d'images,
- le type de la recherche est "précise" et la sélection n'a pas de nouvelles images,
- le type de la recherche est "large" et la sélection ne rapporte pas de nouvelles images,
- le type de la recherche est "précise" et la sélection comporte trop d'images.

Voyons chacune de ces stratégies sous la forme d'organigramme montrant la progression dans la phase de déformation :

Une recherche large avec une sélection ayant trop d'images :



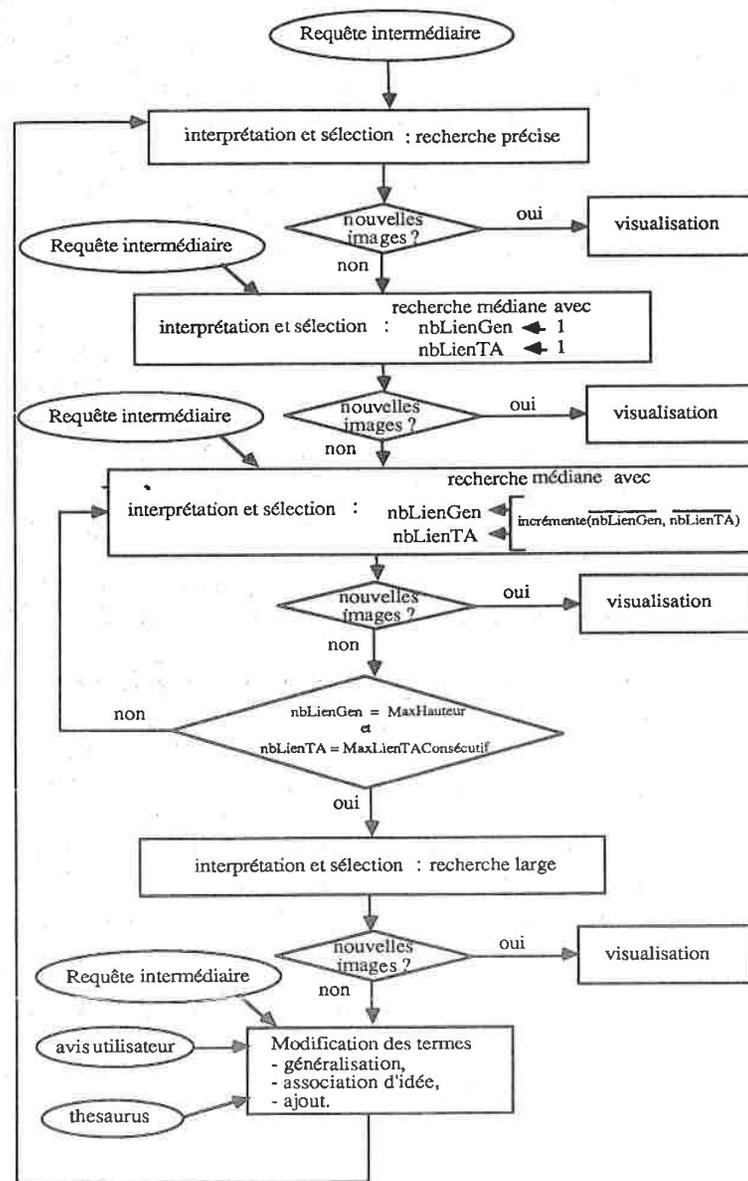
Une sélection contient trop d'images lorsque son nombre d'images est supérieur au nombre maximum d'images que l'utilisateur accepte de visualiser (cf § III.2.1).

Les constantes MaxHauteur et MaxLienTaConsécutif représentent respectivement la hauteur maximale des hiérarchies des champs sémantiques et le nombre maximum de lien "terme associé" que l'on peut parcourir, d'une manière consécutive, à travers le thesaurus.

La fonction "décrémente" prend comme paramètres les deux anciennes valeurs de nbLienGen et nbLienTA et renvoie un couple contenant leurs nouvelles valeurs. Cette fonction assure une régression cohérente et simultanée des valeurs successives de nbLienGen et de nbLienTA de leur valeur maximale jusqu'à 1. Son fonctionnement exact est du domaine de l'heuristique.

Lorsque la recherche précise apporte toujours trop d'images, il faut alors préciser ou supprimer certains concepts de la requête intermédiaire. L'avis de l'utilisateur est important, le système lui propose une liste de termes parmi lesquels il doit choisir ceux qui correspondent le mieux à ce qu'il recherche. Cette liste de termes est construite à partir du thesaurus en proposant les termes spécifiques des termes présents dans la demande.

Une recherche précise avec une sélection n'ayant pas de nouvelles images :

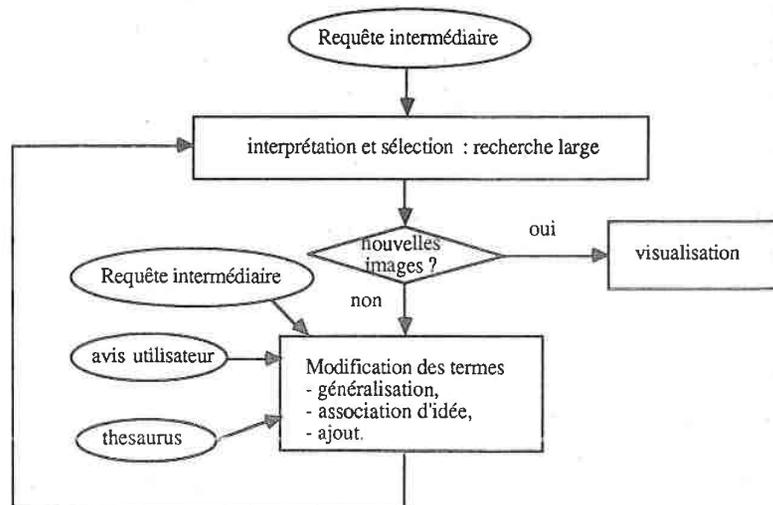


La condition permettant de dire qu'une sélection ne rapporte pas de nouvelles images est expliquée dans le paragraphe suivant, cette condition tient compte des ensembles d'images contruits dans les étapes précédentes.

La fonction "incrémente", à l'inverse de la fonction "décrémente", assure une progression cohérente des valeurs successives de nbLienGe et de nbLienTA de 1 jusqu'à leur valeur maximale.

Lorsque la recherche large ne rapporte pas de nouvelles images, il faut alors élargir la portée de la requête système en généralisant des termes ou en ajoutant des termes reliés par la relation "terme associé" à des termes déjà présents dans la demande. La généralisation peut être effectuée automatiquement en généralisant les termes qui ne rapportent pas beaucoup d'images, par contre les ajouts de termes de sens proche nécessitent le contrôle de l'utilisateur afin que la recherche ne deviennent pas "hors sujet". L'utilisateur choisit parmi un ensemble de termes que le système lui propose (termes de sens proche), il peut aussi choisir d'en ajouter de complètement nouveaux.

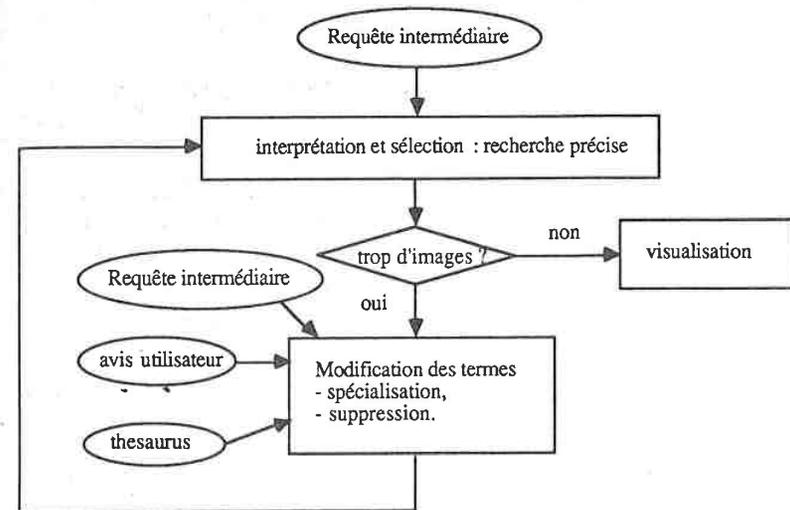
Une recherche large avec une sélection n'apportant pas de nouvelles images :



Les déformations dans ce cas ne peuvent s'effectuer que sur les termes de la requête

intermédiaire. Les modifications apportées à la requête intermédiaire sont les mêmes que dans le cas précédent.

Une recherche précise avec une sélection ayant trop d'images :



Comme dans le cas précédent, seule la modification de la requête intermédiaire peut modifier la portée de la requête système. Les types de modifications possibles sont semblables à ceux que nous avons énumérés pour la première stratégie de déformation.

IV.3 La Visualisation

La phase d'analyse des choix qui suit la phase de visualisation, a comme paramètres trois ensembles d'images : les images choisies, les images rejetées, les images indifférentes. Ces ensembles peuvent être construits à une étape donnée à partir exclusivement des documents sélectionnés par la requête système, ou bien ils peuvent tenir compte des choix effectués dans les étapes précédentes. Voyons plus précisément, pour le cas particulier de la deuxième étape donnée, comment ces ensembles peuvent être construits : à la première étape, trois sous-ensembles C_1 (les descriptions des images choisies), R_1 (les descriptions des images rejetées), I_1 (les descriptions des images indifférentes ou neutres) de l'ensemble des descriptions sélectionnées S_1 ont été obtenus ; l'ensemble S_1 ayant été construit à partir de la requête système

RS_1 . A la deuxième étape, une requête système RS_2 a sélectionné un ensemble S_2 de descriptions. A ce stade de la recherche un problème se pose : quel ensemble d'images doit-on proposer à l'utilisateur pour la visualisation ?

Mais avant de répondre à cette question faisons quelques remarques sur l'ensemble S_2 :

- $C_1 \cap S_2 = C_s$; si cet ensemble n'est pas vide, il faut se demander s'il est nécessaire de remontrer à l'utilisateur les images de cet ensemble, et si oui pourquoi ne pas présenter l'ensemble C_1 entièrement en indiquant que les images de cet ensemble ont déjà été choisies.

- $R_1 \cap S_2 = R_s$; de la même manière, si cet ensemble n'est pas vide, il convient de se demander s'il est astucieux de remontrer les images de cet ensemble à l'utilisateur alors qu'il les a rejetées à l'étape précédente.

- $I_1 \cap S_2 = I_s$; les images appartenant à cet ensemble doivent-elles encore être montrées à l'utilisateur ?

L'expérimentation a montré qu'il semblait intéressant pour l'utilisateur de voir constamment, lors de la phase de visualisation d'une étape donnée, toutes les images qu'il a déjà choisies dans les étapes précédentes (C_s), et il nous semble important aussi que les descriptions de ces images figurent dans l'ensemble d'images qui va servir à calculer les poids de pertinence de cette étape, afin que l'analyse des choix tienne compte des choix positifs des étapes précédentes. Pour les images rejetées (R_s), l'expérimentation nous a montré qu'elles ne devaient plus apparaître dans les images à visualiser, car une image rejetée a de fortes chances de le rester durant toute une session d'interrogation. Les images indifférentes (I_s) doivent faire partie des images à visualiser, car au fur et à mesure des étapes le besoin de l'utilisateur évolue et il se peut qu'à une certaine étape, il désire choisir ou rejeter une de ces images.

Résumons tout cela plus précisément :

Soit C_{n-1} l'ensemble de toutes les images que l'utilisateur a choisies dans les étapes précédentes, R_{n-1} l'ensemble de toutes les images qu'il a rejetées, I_{n-1} l'ensemble de toutes les images qui ont été sélectionnées et qui n'ont été ni choisies, ni rejetées, S_n l'ensemble des images sélectionnées par la requête système RS_n .

Alors l'ensemble V_n des images que l'utilisateur va visualiser à une étape n se calcule de la manière suivante :

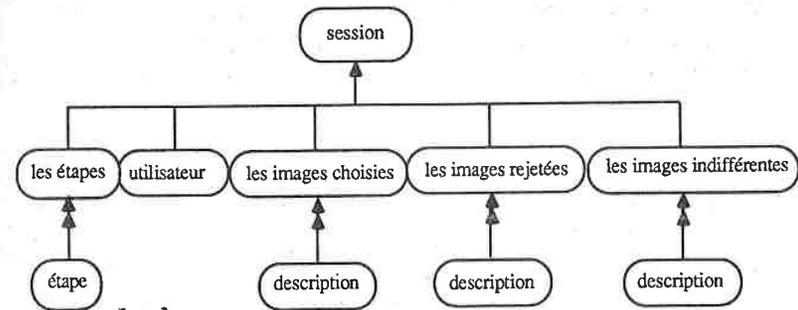
$$V_n = (S_n \cup C_{n-1}) \ominus R_{n-1}$$

A l'aide de ces ensembles, on peut déterminer, d'une manière plus formelle, à quelle

condition une sélection ne rapporte pas de nouvelles images ; cette condition provoque une déformation de la requête système (cf § précédent). On dira qu'une sélection S_n d'une étape n ne rapporte pas de nouvelles images quand :

$$S_n \subset (C_{n-1} \cup R_{n-1} \cup I_{n-1})$$

La prise en compte de ces choix demande une modification de structure de l'objet session ; la nouvelle structure doit comporter les ensembles C_n , R_n et I_n :



IV.4 Analyse des choix : calcul des poids de pertinence

L'analyse des choix telle qu'elle a été présentée dans le chapitre précédent comporte un calcul de poids de pertinence pour chaque terme présent dans les images sélectionnées. La formule proposée a donné satisfaction, car le poids de pertinence calculé à l'aide de cette formule offre un bon aperçu de l'importance du terme dans le choix de l'utilisateur.

Cette formule, calculant la pertinence, ne prend pas en compte la fréquence d'apparition du terme dans la base comme le proposent les méthodes classiques de calcul de pertinence présentées au § II du chapitre 1. Cela n'a pas semblé nécessaire, car ce que nous recherchons dans cette phase c'est de déterminer quelles sont les représentations des concepts qui caractérisent le choix de l'utilisateur. La formule heuristique donnée dans le chapitre précédent réalise correctement cette fonction. Nous verrons dans le prochain paragraphe que la fréquence des termes peut être utilisée lors de la propagation, si l'on suppose que l'importance d'un concept, dans l'expression du besoin de l'utilisateur, dépend du nombre de ses représentations dans la base d'images.

Le calcul des poids de pertinence va alors s'effectuer à partir des ensembles C_n ($C_{n-1} \cup C$) et R ; C et R étant les ensembles des images choisies et rejetées de l'étape n . Une fois ces poids

calculés, le niveau d'expression de la demande va être mis à jour en propageant ces poids dans le thesaurus. Cette propagation fait partie des propositions heuristiques de la méthode.

IV.5 La propagation des poids

La propagation des poids s'effectue en remplaçant le poids d'expressivité des termes présents dans les descriptions sélectionnées par le poids de pertinence calculé lors de l'analyse des choix. Cette façon de procéder se justifie par le fait que l'on suppose que le dernier choix de l'utilisateur remet en cause les choix effectués dans les étapes précédentes. L'expérimentation a montré que grâce à cette hypothèse, les propositions données par le système sont assez bien représentatives de l'évolution du besoin d'étape en étape.

Une des parties heuristiques de la propagation est la détermination de la fonction F calculant le poids d'expressivité P_{exp_t} d'un terme t à partir des poids d'expressivité P_{exp_i} de ses spécifiques.

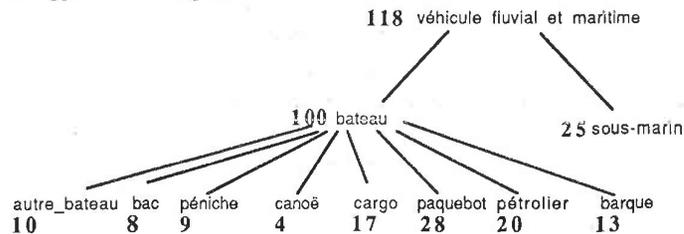
Dans les exemples que nous avons montrés ainsi qu'au début de notre expérimentation la fonction F était une moyenne :

$$F(P_{exp_t}) = \frac{\sum_{i=1}^n P_{exp_i}}{n}$$

Avec cette fonction tous les termes spécifiques ont une même importance. Mais si l'on considère le thesaurus comme une taxonomie, on dira qu'un concept est caractéristique d'un choix, si ce choix contient beaucoup de représentations de ce concept. Le poids d'expressivité du concept doit alors être une sorte de poids moyen d'une représentation du concept ; une représentation du concept est ici une image qui a été décrite par un des termes spécifiques du concept.

Exemple :

Soit un morceau de thesaurus où la portée des concepts, c'est-à-dire le nombre d'images qu'ils rapportent est indiquée :



Dans cet exemple, il y a 8 images de bac, 9 de péniche, 4 de canoë...et globalement il y a 100

représentations du concept de "bateau" sur un total de 100 images, car il y a des images qui possèdent plusieurs représentations. Le poids d'expressivité du terme bateau doit permettre d'évaluer l'intérêt que porte l'utilisateur au concept quelque soit sa représentation dans les images. En d'autres termes le poids d'expressivité de bateau doit être un poids moyen évaluant la pertinence, pour l'utilisateur, d'une représentation quelconque d'un bateau.

F peut alors être une moyenne pondérée des poids d'expressivité :

$$P_{exp_t} = F(P_{exp_i}) = \frac{\sum_{i=1}^n (P_{exp_i} * coef_i)}{\sum_{i=1}^n coef_i}$$

$$\text{avec } coef_i = portée_i * C, \text{ et } C = \frac{portée_t}{\sum_{i=1}^n portée_i}$$

ici C est un coefficient correcteur pour obtenir la propriété suivante : $\sum_{i=1}^n coef_i = portée_t$

et $portée_i$ est le nombre d'images que rapporte le terme spécifique i du concept représenté par le terme t.

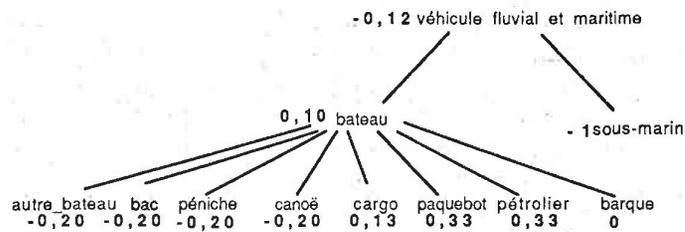
En développant la formule on obtient :

$$F(P_{exp_i}) = \frac{\sum_{i=1}^n (P_{exp_i} * portée_i * C)}{\sum_{i=1}^n (portée_i * C)}$$

$$\text{donc } F(P_{exp_i}) = \frac{\sum_{i=1}^n (P_{exp_i} * portée_i)}{\sum_{i=1}^n portée_i}$$

Exemple :

Reprenons l'exemple du chapitre 2, en utilisant la nouvelle formule F et les portées données dans l'exemple précédent. Le niveau d'expression calculé est le suivant :



Ainsi on s'aperçoit que le poids du terme bateau est plus important qu'auparavant (0,10 au lieu de -0,08), ce qui peut s'interpréter, très approximativement, de la sorte : "si le système montre à l'utilisateur des images quelconques de bateaux, il a une chance sur dix de lui montrer un bateau intéressant" ; de plus si l'utilisateur a déjà rejeté un certain nombre d'images de bateaux qui ne l'intéressaient pas, ce pourcentage de chance peut alors être plus important, car à une étape donnée le système ne propose plus les images rejetées dans les étapes précédentes.

Le niveau d'expression de la demande modifié, la demande peut alors être reformulée.

IV.6 La reformulation

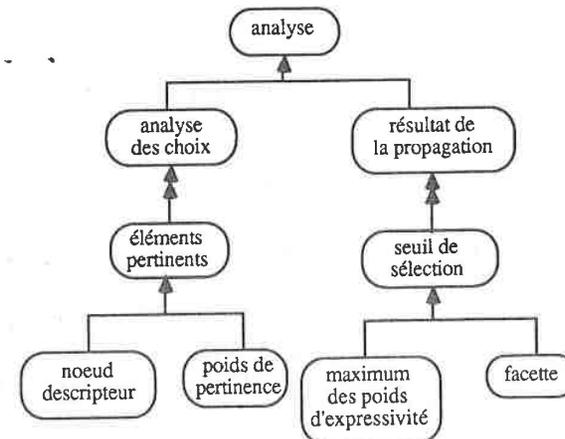
IV.6.1 Le choix des seuils

La reformulation de la demande nécessite un parcours du thesaurus pour sélectionner les concepts caractérisant les besoins de l'utilisateur. L'algorithme du parcours a été donné au chapitre précédent, un de ses paramètres est le seuil permettant de sélectionner les termes. La valeur de ce seuil doit être déterminée pour que les concepts les plus pertinents fassent partie de la demande. Les concepts les plus pertinents sont ceux qui ont un poids d'expressivité maximal, c'est pour cela que dans notre première expérimentation nous avons fixé le seuil de sélection au maximum des poids d'expressivité.

L'utilisateur voit dans sa nouvelle demande les termes les plus pertinents, voir d'autres termes signifie : sélectionner dans le thesaurus des termes avec un poids d'expressivité plus faible. Comme le seuil est à la valeur maximale des poids d'expressivité, il suffit de reparcourir le thesaurus avec un seuil plus faible. Ce nouveau seuil est obtenu en décrémentant l'ancien seuil d'un certain pas. Ce pas peut être fixé arbitrairement ou à partir d'une étude statistique des écarts des poids entre les termes spécifiques et leur terme générique. Dans notre réalisation ce pas a été choisi arbitrairement, car la propagation des poids dans le thesaurus étant déjà assez coûteuse en temps nous n'avons pas souhaité augmenter ce coût. Nous avons effectué une série de tests en

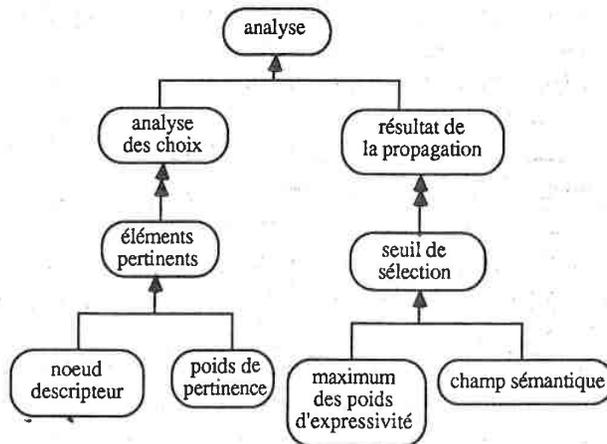
prenant plusieurs valeurs pour ce pas ; nous avons choisi une valeur suffisamment importante qui permet d'obtenir, dans la majorité des cas, d'autres termes à chaque nouveau parcours du thesaurus. En effet, si la valeur du pas est trop faible, la différence entre deux seuils successifs sera tellement peu importante que le parcours du thesaurus n'apportera pas de nouveaux termes ; cela est à éviter car un parcours du thesaurus prend un certain temps non négligeable.

Le dialogue de la formulation de la demande, présenté au § II.2.3.1, montre que l'utilisateur peut obtenir d'autres termes en choisissant l'option "autres termes" sur un menu déroulant. Le choix de cette option n'est possible que si l'utilisateur a sélectionné une facette. Ainsi, lorsque l'utilisateur choisit cette option, il faut reparcourir le thesaurus avec un nouveau seuil, seulement à l'intérieur des champs sémantiques associés à la facette. Il paraît alors nécessaire d'avoir un seuil de sélection par facette, calculé lors de la propagation, qui serait égal au maximum des poids d'expressivité des termes des champs sémantiques associés à la facette ; le résultat de la propagation est alors différent :

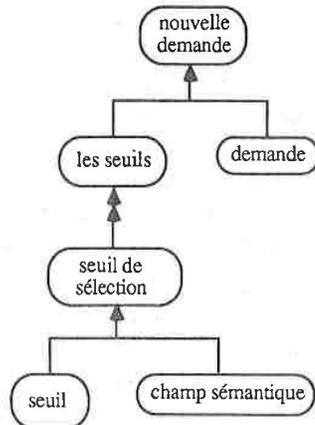


Avec ce nouveau type de reformulation, que l'on peut appeler "par facette", nous nous sommes aperçu que la qualité de la nouvelle demande était meilleure. En effet, la particularisation des seuils de sélection permet à chaque facette du concept d'avoir ses propres caractéristiques, alors qu'avec un seul seuil, les facettes sont supposées posséder les mêmes caractéristiques. En effet, pour une recherche donnée, il se peut qu'une facette, par exemple "la connotation", ne fasse pas partie de l'expression des besoins d'un utilisateur recherchant des représentations concrètes d'objets.

Dans le même ordre d'idée, il est possible que certains champs sémantiques du thesaurus soient moins importants pour l'expression du concept que d'autres champs. Ainsi, ce n'est pas un seuil par facette qu'il faut déterminer, mais bien un seuil par champ sémantique du thesaurus :



La structure de la nouvelle demande doit elle aussi être modifiée :



La reformulation s'effectue alors avec autant de seuils que de champs sémantiques ; un champ sémantique, ayant un seuil négatif, ne doit pas être parcouru, car cela signifie que le besoin de l'utilisateur ne se situe pas dans le domaine de ce champs.

IV.6.2 Parcours d'un champ sémantique en langage objet

Voyons à titre d'exemple comment est implantée une partie de la reformulation sous la forme d'une programmation objet, c'est-à-dire en attachant des méthodes aux types d'objets que nous avons définis précédemment. Cette partie consiste à parcourir un champ sémantique du thesaurus avec un certain seuil afin d'en extraire une liste de termes.

La méthode générale effectuant ce parcours est rattachée au type objet **champ sémantique** ; le résultat de cette méthode est une liste de **nœuds** à faire figurer dans la facette de la nouvelle demande qui possède ce champ sémantique dans son domaine de valeurs.

champ sémantique

extraire: unSeuil

| lesNœudsCaractéristiques |

; variable locale.

lesNœudsCaractéristiques <- racine *extraireConcepts*: unSeuil.

- ; la méthode "extraireConcepts" est envoyée à l'objet "racine", qui est un composant de ; l'objet champ sémantique, avec un paramètre "unSeuil" qui est aussi le paramètre de la ; méthode courante. Le résultat de l'appel est placé dans la variable locale.

!lesNœudsCaractéristiques

; le contenu de la variable locale est renvoyé comme résultat de l'appel de la méthode

Commentaires: la syntaxe employée dans cet exemple est proche de la syntaxe Smalltalk ; les nom des types objets sont écrits en caractères gras. Les commentaires sont placés après un point virgule. Le profil de la méthode est souligné, le nom de la méthode se termine par deux points car elle possède un paramètre qui se trouve juste après le nom. Dans le corps de la méthode les variables locales sont déclarées, au début, à l'intérieur des barres verticales. L'opérateur '<-' est une méthode spéciale réalisant l'affectation. Un appel de méthode dit aussi "envoi de message" est représenté par le nom de la méthode invoquée, écrit en italique (*extraireConcepts*), et le nom de l'objet receveur placé devant le nom de la méthode (racine). Un message est toujours envoyé à un objet, si la déclaration du type de l'objet (classe en smalltalk) connaît la méthode, elle est exécutée, sinon il faut parcourir la hiérarchie des types (types génériques) pour rechercher la méthode appelée. Si la méthode n'est pas trouvée un message d'erreur est généré.

L'objet "racine" est, si l'on regarde la structure d'un champ sémantique, un objet de type **nœud concept** ; la méthode "extraireConcepts" doit être une méthode de ce type ou d'un de ses types génériques. Mais comme une demande peut contenir des nœuds concepts ou des nœuds

descripteurs, c'est-à-dire des objets de type **nœud** en général, la méthode "extraireConcepts" doit être définie sur le type **nœud**.

nœud

extraireConcepts: unSeuil

| lesNœudsCaractéristiques |

lesNœudsCaractéristiques < **liste nouvelle**.

; la variable "lesNœudsCaractéristiques" devient un objet de type **liste**.

ponds d'expressivité > unSeuil

; on teste si le poids d'expressivité du nœud courant est supérieur au seuil

siVrai: [lesNœudsCaractéristiques *ajouter:* self]

; si oui, le nœud courant est ajouté dans la liste "lesNœudsCaractéristiques" et le parcours
; s'arrête.

siFaux: [self *estUnConcept*

; sinon, on teste si le nœud courant est un nœud concept

siVrai: [lesSpécifiques *pourChaque:*

- ; si c'est un nœud concept il possède des nœuds spécifiques, il faut alors continuer le
- ; parcours dans le thesaurus pour tous les spécifiques du nœud courant.

[| unNœud |

lesNœudsCaractéristiques *ajouterTous:* (unNœud *extraireConcepts:* unSeuil)]

; à chaque nœud spécifique du nœud courant est envoyé le message "extraireConcepts", ce
; qui consiste à appeler la méthode que nous sommes en train de décrire. Le résultat de cet
; envoi de message est donc une liste de nœuds ayant un poids supérieur au seuil. Cette
; liste est alors ajoutée en totalité dans la liste courante "lesNœudsCaractéristiques".

siFaux:[]

; si le nœud courant n'est pas un nœud concept, c'est alors un nœud descripteur sans
; spécifiques, le parcours s'arrête.

]; ; fin du *siFaux*.

↑lesNœudsCaractéristiques

; la liste de nœuds ainsi construite constitue le résultat de la méthode.

Commentaire: beaucoup de méthodes sont utilisées, nous n'allons pas toutes les détailler ; certaines font partie de l'environnement de programmation de base, elles sont attachées à des types préexistants. Par exemple, le type objet **liste** est un type existant qui possède les méthodes "nouvelle", "ajouter", "ajouterTous" et la méthode "pourChaque" qui prend comme paramètre un morceau de méthode qu'elle évalue pour chaque élément de la liste. La méthode ">" est définie sur les réels, elle compare le réel qui reçoit le message avec le réel en paramètre et elle renvoie un

objet de type **booléen** qui est soit **vrai** soit **faux**. La méthode "siVrai: siFaux:" est définie sur chacun des booléens **Vrai** ou **Faux**, elle prend comme paramètres deux morceaux de méthode ; le premier est exécuté si le booléen qui reçoit ce message est de type **vrai** ; si le booléen est de type **faux** c'est le deuxième qui est exécuté.

Si l'on considère que l'on possède une méthode "évaluer" qui permet d'exécuter un morceau de méthode la méthode "siVrai: siFaux:" peut s'écrire de la sorte :

Vrai

siVrai: unMorceau siFaux: unAutreMorceau

↑unMorceau *évaluer*.

Faux

siVrai: unMorceau siFaux: unAutreMorceau

↑unAutreMorceau *évaluer*.

La méthode "estUnConcept" doit être définie à la fois sur le type objet **nœud concept** et sur le type objet **nœud descripteur** :

nœud concept

estUnConcept

↑Vrai

nœud descripteur

estUnConcept

↑Faux

Le but de cette présentation est de montrer un aperçu de la philosophie de la programmation dite "objet". L'aspect condensé de la programmation est important, il faut peu de lignes de programme pour écrire les méthodes, ceci grâce à :

- la hiérarchie des types objets qui permet une certaine factorisation des traitements,
- l'aspect très modulaire des méthodes,
- la réutilisation des méthodes et des types déjà existants.

RIVAGE a été réalisé d'une manière incrémentale grâce à l'utilisation de Smalltalk 80. Dès que le système est arrivé à un état stable, c'est-à-dire dès qu'il a effectué correctement les traitements demandés, nous avons mis en œuvre une évaluation globale afin d'estimer la pertinence des résultats fournis.

V Evaluation du système

L'évaluation du système a été effectuée par Sabine Humbert, une stagiaire de l'IUT département Carrière de l'Information de Nancy II [Hum 89]. Cette évaluation a pour but de mesurer la pertinence globale du système en termes de rappel et de précision (cf partie 1, chapitre 1, § III.1), et de la comparer à celle d'un système classique de recherche documentaire, car c'est ce type de système qui est généralement utilisé pour rechercher les images.

V.1 Avertissement :

Pour effectuer de telles mesures, il faut se fixer un certain nombre de demandes assez générales, et rechercher soit manuellement, soit d'une manière très large, l'ensemble des documents de la base qui correspondent à chaque demande, c'est-à-dire l'ensemble des documents qui seront jugés pertinents par un utilisateur qui poserait la demande. La comparaison s'effectue sur une base test ; des demandes pré-formulées sont soumises au système et la confrontation du résultat obtenu avec le résultat escompté permet le calcul des mesures de pertinence. Cette méthode est classiquement utilisée pour comparer les systèmes de recherches d'informations textuelles. Les mesures demeurent assez fiables pour ce type d'informations, car il est assez facile de comparer une demande textuelle avec un document de même type et d'indiquer si le document traite des sujets présents dans la demande.

Mais lorsque l'information recherchée est l'image, la méthode devient moins fiable. En effet la comparaison d'une demande, formulée en langage naturel, avec une image fait appel à un processus de lecture qui n'est pas le même d'une personne à l'autre (cf partie 2, chapitre 1, § I). La pertinence d'un système de recherche d'images ne peut pas être mesurée avec certitude, elle peut être estimée pour un certain nombre d'utilisateurs. L'évaluation consistera plus à mesurer la capacité qu'a le système à orienter une recherche vers les images dites pertinentes pour un utilisateur, que sa capacité à retrouver des images pertinentes pour n'importe quel utilisateur. Les mesures de cette évaluation devront être prises à chaque étape du processus de recherche.

Voyons quelle a été notre démarche pour réaliser cette évaluation.

V.2 La démarche choisie

La démarche est composée de plusieurs phases qui ont été menées de front par Sabine Humbert. En effet, il fallait trouver un certain nombre d'utilisateurs potentiels de la base ayant chacun un besoin spécifique en images qu'il devait soumettre au système RIVAGE, et il fallait aussi une documentaliste, connaissant le fonds d'images, capable d'interpréter ces besoins pour

interroger la base d'images d'une manière classique, c'est-à-dire à l'aide de requêtes booléennes.

Huit utilisateurs, ne connaissant ni la base ni le fonctionnement du système RIVAGE, ont été trouvés dans l'équipe. Chacun a défini son besoin sous la forme d'une phrase en langage naturel. Puis chaque utilisateur, après une courte initiation au système RIVAGE, a soumis au système une première demande exprimant, à sa manière, son besoin. Puis il s'est laissé guider par le système en choisissant et rejetant les images et en acceptant ou rejetant les propositions du système. A chaque étape des différentes sessions d'interrogation, les références des images sélectionnées par le système et celles des images jugées pertinentes par les utilisateurs ont été conservées.

En parallèle, Sabine Humbert, connaissant parfaitement la base d'images, le thesaurus, la structure des documents, et la formulation en langage documentaire (ici la formulation booléenne), a construit, pour chaque utilisateur, une requête système correspondant à une interprétation de son besoin en images. La construction des requêtes s'est effectuée par étapes successives comme dans les systèmes documentaires classiques. Les références des images sélectionnées par ces requêtes ont été conservées.

La dernière phase de la démarche, qui a été la plus longue, a permis aux utilisateurs de visualiser un nombre assez important d'images (proche de l'ensemble total des images de la base) et de construire l'ensemble des images pertinentes relativement à leur besoin. Ce dernier ensemble doit contenir au minimum l'ensemble des images choisi lors de l'utilisation de RIVAGE.

Ces différentes phases ont été nécessaires pour le calcul des mesures de pertinence associées à l'interrogation d'une base d'images avec le système RIVAGE et aussi, à titre comparatif, associées à l'interrogation classique avec des requêtes booléennes.

V.3 Résultats de l'évaluation

La dernière phase de la démarche d'évaluation a permis de constituer l'ensemble IP des images pertinentes pour l'utilisateur.

Lors de l'interrogation avec le système RIVAGE, les ensembles suivants ont été constitués progressivement, d'étape en étape :

a : l'ensemble total des documents pertinents retrouvés,

c : l'ensemble total des documents non pertinents retrouvés (bruit).

Ces mêmes ensembles ont été constitués pour l'interrogation classique. C'est à partir du cardinal de ces ensembles que nous avons calculé les mesures de pertinence pour chaque interrogation de la base :

NIP : le nombre total d'images pertinentes,

A : le nombre de documents pertinents retrouvés à une étape,

C : le nombre de documents non pertinents retrouvés à une étape.

Le taux de rappel $R \left(\frac{A}{NIP} \right)$ et le taux de précision $P \left(\frac{A}{A+B} \right)$ ont été calculés à chaque étape, voici le tableau récapitulatif de l'interrogation classique :

N	NIP	Etape 1				Etape 2				Etape 3				R	P
		A	C	R	P	A	C	R	P	A	C	R	P		
1	17	2	0	0,12	1	11	8	0,64	0,57	14	11	0,82	0,56		
2	5	3	14	0,6	0,18	5	71	1	0,06						
3	17	0	0	0	0	1	1	0,05	0,5	5	12	0,29	0,29		
4	11	3	0	0,27	1	7	9	0,63	0,43	11	18	1	0,37		
5	25	-18	2	0,4	0,83	25	72	1	0,25						
6	29	14	16	0,48	0,46	16	41	0,55	0,28						
7	17	1	0	0,05	1	15	24	0,88	0,38	17	40	1	0,29		
8	25	0	0	0	0	21	9	0,84	0,7	23	12	0,92	0,65		
Moyenne				0,24	0,49			0,70	0,40			0,81	0,43	0,58	0,44

On peut remarquer que de bons taux de rappel ($> 0,6$) sont obtenus dans 6 recherches sur 8, cela signifie que Sabine Humbert connaît bien le vocabulaire et la base d'images et qu'elle domine parfaitement la formulation booléenne. Voyons sur l'exemple du test n° 1, les différentes étapes d'une telle recherche ; le détail des autres recherches est donné en annexe (cf annexe 3) :

Exemple :

Formulation du besoin en langage naturel :

Plan très large d'un fleuve avec des monuments sous un ciel brumeux.

17 images de la base répondent à ce besoin.

Interprétation sous la forme d'une requête booléenne :

étape 1 :

((contenu = FLEUVE OU RIVIERE) [1]

ET

(contenu = (MONUMENT OU ARC DE TRIOMPHE OU OBELISQUE OU STELE OU PYRAMIDE OU EDIFICE OU CHATEAU OU PALAIS OU MANOIR OU HOTEL PARTICULIER OU BAPTISTERE OU COUVENT OU ABBAYE OU PAGODE OU TEMPLE OU CHAPELLE ABSIDIALE OU CHAPELLE CASTRALE OU CHAPELLE RAYONNANTE OU CHAPELLE SYNODALE OU CATHEDRALE OU BASILIQUE OU AQUEDUC OU PONT_D OU VIADUC) [2]

ET

((contenu = BROUILLARD OU NUAGE) OU (morphologie = SOMBRE OU GRISAILLE)) [3]

ET

(morphologie = VUE PERSPECTIVE OU PLONGEE OU PLAN D'ENSEMBLE)) [4]

A l'aide de cette requête 2 images ont été sélectionnées et ce sont deux images pertinentes :
R = 0,12 ; P = 1.

étape 2 :

[1] ET [2] ET [4] : 19 images sont sélectionnées, 11 sont pertinentes et 8 ne le sont pas :
R = 0,64 ; P = 0,57.

étape 3 :

[1] ET [2] : 25 images sont sélectionnées, 14 sont pertinentes, 11 ne le sont pas :
R = 0,82 ; P = 0,56.

Quelques remarques doivent être faites :

- la taille d'une requête peut être très grande,
- cette formulation nécessite un gros investissement en temps pour trouver l'ensemble du vocabulaire et le moyen de l'intégrer dans la requête,
- une bonne connaissance des champs des descriptions et de leur influence dans la sélection des documents est importante ; l'exemple montre que le champ morphologie est utilisé pour réduire la taille de la sélection,
- ce type de formulation n'est pas orienté grand public.

Voyons maintenant les résultats obtenus avec le système RIVAGE. Pour chacun des tests nous avons demandé aux utilisateurs d'effectuer les deux types de recherche possibles : recherche large L et recherche précise P :

N	NIP	TR	Etape 1				Etape 2				Etape 3				Etape 4				R	P
			A	C	R	P	A	C	R	P	A	C	R	P	A	C	R	P		
1	17	L	16	20	0,94	0,44	16	21	0,94	0,43	17	8	1	0,68						
		P	3	0	0,17	1	11	4	0,64	0,73	12	1	0,7	0,92	16	13	0,94	0,55		
2	5	L	3	13	0,6	0,18	4	19	0,8	0,17	4	5	0,8	0,44	5	32	1	0,13		
		P	3	13	0,6	0,18	3	9	0,6	0,33										
3	17	L	2	5	0,11	0,28	5	8	0,29	0,38	12	0	0,7	0,15						
		P	2	5	0,11	0,28	5	8	0,29	0,38	12	0	0,7	0,15						
4	11	L	6	8	0,54	0,42	6	6	0,54	0,5	6	2	0,54	0,75	6	3	0,54	0,66		
		P	1	0	0,09	1	1	1	0,09	0,5	1	2	0,09	0,33	1	1	0,09	0,5		
5	25	L	12	5	0,48	0,7	13	1	0,52	0,92	14	0	0,56	1	19	19	0,76	0,5		
		P	12	5	0,48	0,7	14	0	0,56	1	14	1	0,56	0,93						
6	29	L	10	5	0,34	0,66	14	21	0,48	0,4	16	31	0,55	0,34	16	4	0,55	0,8		
		P	10	5	0,34	0,66	11	4	0,37	0,73	11	1	0,37	0,91	20	27	0,68	0,42		
7	17	L	10	18	0,58	0,35	11	20	0,64	0,35	11	25	0,64	0,3	11	33	0,64	0,25		
		P	0	6	0	0	1	1	0,06	0,5	3	6	0,18	0,33	4	5	0,24	0,44		
8	25	L	10	3	0,4	0,76	12	6	0,48	0,66	12	2	0,48	0,86	18	32	0,72	0,36		
		P	10	3	0,4	0,76	12	6	0,48	0,66	18	44	0,72	0,29	25	20	1	0,55		
Moyenne		L			0,5	0,47			0,59	0,48			0,66	0,56			0,7	0,45	0,61	0,49
		P			0,27	0,57			0,39	0,6			0,47	0,55			0,59	0,49	0,43	0,55

En moyenne générale sur l'ensemble des recherches, le taux de rappel de la recherche large est meilleur que celui des recherches avec formulation booléenne (0,61 contre 0,58). La différence n'est pas énorme mais cela montre que RIVAGE permet à un utilisateur du grand public de rivaliser avec un spécialiste de la formulation booléenne connaissant la base d'images.

Voyons quelles ont été les étapes de recherche du test n°1 :

Exemple :

Cette une recherche précise :

étape 1 :

Demande utilisateur :

Contenu de la photographie = FLEUVE, MONUMENT

Avec cette demande 3 images sont sélectionnées et ce sont des images pertinentes, l'utilisateur les choisit : R = 0,17 ; P = 1,

étape 2 :

Après l'analyse des choix de l'utilisateur le système propose comme nouvelle demande :

Contenu de la photographie : PONT_D, FLEUVE

Cette demande propose 15 images à l'utilisateur comprenant 12 nouvelles images plus les trois images choisies dans l'étape précédente. Parmi les 12 nouvelles images l'utilisateur en choisit 8, il en rejette 2 et il laisse les deux autres neutres : R = 0,64 ; P = 0,73,

étape 3 :

le système propose comme nouvelle demande :

Contenu de la photographie : QUAI, ARCHE

Cette demande propose 13 images à l'utilisateur comprenant 2 nouvelles images plus les 11 images choisies dans les étapes précédentes. Parmi les 2 nouvelles images l'utilisateur en choisit une, il en rejette une : R = 0,7 ; P = 0,92,

étape 4 :

le système propose comme nouvelle demande :

Contenu de la photographie : FLEUVE

Cette demande propose 29 images à l'utilisateur comprenant 17 nouvelles images plus les 12 images choisies dans les étapes précédentes. Parmi les 17 nouvelles images l'utilisateur en choisit 4, il en rejette 13 : R = 0,94 ; P = 0,55.

L'utilisateur étant satisfait de sa recherche, il décide de s'arrêter.

Dans cet exemple, on peut remarquer que l'utilisateur a mené la totalité de sa recherche en n'introduisant que 2 termes dans sa première demande puis grâce à ses choix lors de la phase de visualisation, le système l'a orienté vers des directions proches de son besoin mais formulées de manière différente. L'utilisateur n'a alors plus besoin de connaître le vocabulaire manipulé par le système, il peut le découvrir d'étape en étape grâce aux propositions du système.

Beaucoup de choses peuvent être dites sur cette évaluation, mais son intérêt principal réside dans le fait qu'elle nous a servi à montrer que le système RIVAGE permettait à un utilisateur connaissant le principe du système (les trois phases), mais ne connaissant ni le vocabulaire ni la base d'images, d'obtenir, seul, des images correspondant à son besoin d'une manière beaucoup plus aisée que celle utilisée par un documentaliste utilisant un système documentaire classique.

Grâce à cette évaluation, nous avons pu mettre en évidence les points suivants :

- un bon outil de visualisation permettrait une sélection plus importante d'images. La précision pourrait alors avoir une valeur moins importante, de l'ordre de 0,2, si l'on suppose qu'un utilisateur est prêt à visualiser 1 image pertinente toutes les 5 images,
- l'importance des champs dans la demande doit être prise en compte lors des déformations,
- certains termes dans la demande ont plus d'importance que d'autres,
- la structure du thesaurus ne correspond pas toujours à la représentation qu'a l'utilisateur des mêmes concepts,
- l'indexation des images est souvent incomplète.

La mise en évidence de ces différents points nous a permis d'entrevoir un certain nombre d'évolutions à ce système.

VI Evolution & perspectives

Les évolutions du système RIVAGE se placent à des échelles de temps différentes. En effet, on peut distinguer les évolutions qui peuvent être effectuées à très court terme, celles à moyen terme et enfin celles qui demandent un temps non négligeable.

VI.1 Les évolutions à court terme

A court terme les évolutions possibles sont celles qui vont permettre d'améliorer l'utilisation et les qualités du système RIVAGE sans changer fondamentalement la structure du système.

Une des plus urgentes est la réalisation d'un outil de visualisation des images plus performant, permettant à l'utilisateur de voir plusieurs images en même temps, de les comparer, de les trier. Cet outil devra être très proche, sur le plan des fonctionnalités, de l'Imageur Documentaire de la SEP, mais il doit être réalisé avec l'outil Smalltalk afin que son intégration dans le prototype soit maximale.

Une autre amélioration, qui a sans aucun doute une grande importance, est la prise en compte de critères de recherche prenant leur valeur dans un ensemble ordonné, comme par exemple la date. La prise en compte de ces types de critères à l'interrogation nécessite d'utiliser des techniques adaptées permettant une mise en correspondance souple entre requête et document. En effet, une date, que ce soit dans les documents ou dans les requêtes, peut avoir une valeur imprécise, incertaine, vague ou encore avec des limites mal définies ; exemple : 16 janvier 1928, vers le 15 septembre 1934, mars 1939, entre 1920 et 1925 Ce type d'information s'apparente aux informations floues [Zad 65] ou nuancées [Mou 88] et leur intégration dans le processus EXPRIM doit être étudiée sérieusement. Une partie de cette étude peut reprendre quelques résultats des travaux de Mouaddib qui a étudié la prise en compte d'informations nuancées dans le problème de l'identification d'espèces naturelles [Mou 89a]. Mais la plus grosse partie de l'étude concerne l'utilisation de ce type d'informations dans le processus d'apprentissage permettant la reformulation de requête. Une proposition de cette prise en compte avait été étudiée dans le projet ESPRIT (cf partie 3, § III.3.2.1), elle avait montré que l'apprentissage pouvait s'effectuer par une étude des intersections entre les différents domaines de valeurs des informations présentes dans les documents. Il s'est révélé que la construction de

ces intersections, en considérant les documents choisis et les documents rejetés, n'est pas immédiate.

L'étape de déformation de la demande peut être améliorée, car notre étude de cette partie du système a été succincte, puisque nous souhaitons, avant toute chose, mettre en place un processus d'apprentissage, pour donner au système la capacité de reformuler une nouvelle demande à partir des choix de l'utilisateur. De nouvelles stratégies peuvent être mises en place. En effet, l'expérimentation de RIVAGE a mis en évidence que les champs présents dans la requête ont une importance différente les uns des autres dans la sélection des documents qui est très liée à l'importance qu'ils ont dans l'expression du besoin de l'utilisateur. Cette importance peut être générale pour toutes les formes de recherche, mais elle peut être aussi spécifique à une utilisation du système. Il arrive souvent, lorsqu'une demande rapporte peu d'images, que certaines facettes se trouvent immédiatement vidées de leur contenu car leur importance dans l'expression du besoin est faible. Il apparaît alors intéressant d'attacher un poids aux facettes de la demande, relativement à l'importance qu'elles ont dans l'expression du besoin de l'utilisateur. Ces poids peuvent être attribués globalement pour toutes les recherches ou par chaque utilisateur lors de la formulation de la demande. Ainsi, lors de la déformation, les stratégies (généralisation, suppression) agiront tout d'abord sur les facettes ayant un poids faible, ce qui permettra à la demande de l'utilisateur d'être transformée tout en conservant son intérêt principal.

De la même manière, on peut remarquer que les termes, à l'intérieur d'une facette, ont une importance différente dans l'expression du besoin. Une pondération des termes peut être mise en place, elle donnerait au système, en phase de déformation, le pouvoir de choisir les termes qui peuvent être supprimés ou généralisés en premier sans que l'expression du besoin s'en trouve trop altérée. Il nous semble important que la prise en compte de ces pondérations n'ait pas un caractère contraignant pour l'utilisateur. Elles peuvent être déduites par le système à partir des agissements de l'utilisateur lors de la formulation : la première facette remplie a de fortes chances d'être la facette la plus importante, de même que le premier concept placé dans une facette a de fortes chances d'être le plus important pour cette facette ; ou elles peuvent être obtenues à partir d'un dialogue agréable et fondé sur l'utilisation de qualificatifs (très important, moins important, ...) plutôt que sur la manipulation de poids par l'utilisateur.

Ces améliorations peuvent être effectuées à court terme car elles ne modifient pas fondamentalement la philosophie de RIVAGE. Par contre, pour les améliorations suivantes, des études plus approfondies doivent être menées car leur intégration dans le système n'est pas immédiate.

VI.2 Les évolutions à moyen terme

Une évolution, qui est à notre avis indispensable à la mise en œuvre du système, est le développement d'outils d'aide à l'indexation et d'aide à la construction de thesaurus intégrés au système RIVAGE et utilisant le processus EXPRIM. Ces outils doivent permettre de construire de nouvelles applications qui seront gérées par le système RIVAGE, ils ont pour tâche de contrôler à tout moment la cohérence du thesaurus et la cohérence de l'indexation. Le processus EXPRIM peut être utilisé pour vérifier la cohérence de l'indexation en effectuant des recherches d'images pour comparer les descriptions entre elles, ou pour amorcer la construction de plusieurs descriptions à partir de l'analyse d'un choix d'images déjà décrites ; les images qui doivent être décrites ressemblent aux images choisies et sont totalement différentes des images rejetées, ainsi la phase "Après-Visualisation" peut donner une première forme aux futures descriptions, à partir d'une analyse des choix et d'une formulation du concept appris. Cette utilisation du processus EXPRIM pour l'indexation est en cours d'étude à travers une collaboration entre un expert du domaine de l'analyse de l'image fixe, Muriel Cluzeau, et quelques membres de l'équipe.

Une perspective à moyen terme qui a déjà fait l'objet d'une étude soutenue, est l'utilisation des concepts de l'hypertexte pour l'interrogation interactive et progressive d'une base d'images. Cette étude consiste à proposer une structure dite "hypertexte" [Con 87] pour le système RIVAGE. L'intérêt de cette structure est de permettre à un utilisateur de parcourir la base d'images selon ses besoins, mais aussi suivant les raisonnements, les associations d'idées que suscite sa recherche. Ainsi, l'utilisateur doit pouvoir formuler une première demande, puis s'orienter vers d'autres horizons plus ou moins éloignés de sa recherche soit par curiosité, soit par nécessité, et revenir à sa formulation, la modifier pour continuer le processus normal d'une recherche (visualisation, après-visualisation). Un tel parcours intégré à RIVAGE a été proposé par une étudiante de DEA [Ham 89], nous le présenterons avec plus de détails dans le chapitre suivant.

Voyons maintenant quelles sont les études à plus long terme qui peuvent être menées à partir de l'état actuel du système RIVAGE.

VI.3 Les évolutions à long terme, perspectives du système RIVAGE

Nos préoccupations à long terme concernent essentiellement les problèmes liés à la formalisation de la méthode d'apprentissage proposée, à l'intégration de capacités d'explicabilité

dans le processus de recherche, mais aussi au problème plus général d'apprentissage à long terme au niveau du système lui-même.

VI.3.1 Apprentissage à long terme

La problématique générale de la mise en œuvre d'un apprentissage à long terme sur un SRI a été présentée dans le chapitre 4 de la partie 1. En résumé, un tel apprentissage doit, à partir de l'analyse de ses recherches antérieures, permettre au système de :

- modifier son comportement, modifier ses stratégies de recherche,
- modifier ses connaissances (évolution du thesaurus, évolution des indexations),
- caractériser les utilisateurs potentiels du système,
- définir des processus de recherche adaptés.

La spécification d'un tel système nécessite une expérimentation importante car la définition d'un modèle d'apprentissage adapté n'existe pas encore.

Les démarches à effectuer pour une telle étude peuvent être les suivantes :

- Définition des différents constituants possibles d'un modèle de l'utilisateur recherchant des images.

- Détermination des éléments caractérisant une recherche :

Pour le moment nous distinguons deux grandes catégories de recherche : la recherche large et la recherche précise. A chacune de ces catégories nous avons associé des stratégies de recherche différentes. Le but de cette étape sera alors la détermination des éléments qui caractérisent ces types de recherche : nombre d'étapes, nombre d'images visualisées, les stratégies de déformation utilisées, le niveau de généralité des termes utilisés, les éléments de l'image qui ont permis la sélection de l'image (son contenu, sa morphologie, le lieu de prise de vue). Cette liste peut être importante et risque de grossir au fur et à mesure des expérimentations.

- Intégration du modèle de l'utilisateur dans la maquette RIVAGE :

Permettre à l'utilisateur d'indiquer le but de sa recherche d'images (illustration d'un article, recherche d'une photographie qu'il a aperçue dans un magazine, recherche de photographies sur un sujet qu'il ne connaît pas ...), et de se présenter au système (spécialiste du domaine couvert par les images, spécialiste de la recherche d'informations, novice ...). Certaines caractéristiques d'un utilisateur pourront être déduites automatiquement par le système en fonction de la manière dont l'utilisateur aborde le système.

- Mise en place de l'apprentissage à long terme :

Le but de cet apprentissage est de permettre au système de caractériser automatiquement des utilisateurs-types pouvant avoir différents types de besoins, et de leur associer un certain nombre de processus de recherche adaptés. L'identification des bons processus pour un type d'utilisateur ayant un type de besoin particulier peut s'effectuer de la manière suivante : le système propose à chaque type d'utilisateur une série de processus de recherche, à la fin d'un processus l'utilisateur juge le résultat de sa recherche (inadapté, bon, moyen, trop long), puis par apprentissage le système devra être capable de proposer, à un utilisateur donné, le processus de recherche permettant au système de satisfaire au mieux son besoin.

Exemple : Ne pas utiliser les stratégies d'une recherche précise pour un utilisateur qui ne connaît pas le fonds d'image. Un premier accès à la base nécessite d'avoir une vue assez générale sur les images de la base.

VI.3.2 Formalisation du processus EXPRIM

Nous souhaitons généraliser le processus d'apprentissage mis en œuvre dans RIVAGE, afin de définir un modèle de recherche progressive de documents fondé sur une connaissance sémantique.

Cette démarche nécessite de définir d'une manière générale la forme des documents, des requêtes, des connaissances nécessaires à l'utilisation du modèle. Les traitements effectués au niveau des différentes phases du processus sont pour le moment heuristiques mais nous avons grand espoir de pouvoir les rendre plus généraux grâce à une identification de leurs effets sur la recherche et par un contrôle de ces effets. Cette étape passe par une étude expérimentale importante et demande des champs d'expérimentations nombreux et variés.

La formalisation du modèle de recherche progressive ainsi obtenue, une méthodologie de réalisation de système utilisant cette forme de recherche sera alors plus facile à élaborer. La réalisation d'outils permettant la mise en œuvre de tels systèmes pourra alors être envisagée (prise en compte de la connaissance, description des informations ...).

IV.3.3 Explicabilité associée au processus

Une des principales perspectives que nous envisageons est l'amélioration du dialogue homme-machine lors du déroulement du processus de recherche. En effet un dialogue devient plus intéressant lorsque le système coopère avec l'utilisateur pour l'aider dans sa recherche. Cette coopération passe par une compréhension des besoins de l'utilisateur mais elle est fondée

aussi sur une capacité d'explication du système. Si le système comprend plus ou moins les envies de l'utilisateur il est nécessaire, dans une recherche coopérative, que l'utilisateur comprenne ce que fait le système. Cette qualité d'un système d'expliquer les traitements qu'il effectue ne consiste pas à définir une simple trace des modules appelés, mais nécessite la manipulation annexe d'une connaissance sur les traitements effectués. De plus la représentation de cette connaissance doit prendre en compte les différents niveaux d'explicabilité adaptés aux expériences des utilisateurs potentiels du système.

Toutes ces perspectives liées à l'évolution du système RIVAGE doivent faire l'objet d'études approfondies et d'expérimentations soutenues et diversifiées. De nouveaux domaines d'application sont en cours d'études, ils concernent la publicité, la médecine et surtout l'industrie métallurgique où une collaboration étroite avec le centre de recherche de Pont-à-Mousson SA et un thésard CNAM est en bonne voie pour aboutir à une réalisation concrète utilisant le système RIVAGE. Ce nouveau champ d'expérimentation permettra de mettre à l'épreuve le système RIVAGE dans de nouvelles conditions, il devra répondre à des besoins concrets d'utilisation d'une base d'images.

VII Conclusion

Le prototype RIVAGE a permis de valider l'approche proposée pour la réalisation du processus EXPRIM. Il a été réalisé d'une manière incrémentale grâce au langage Smalltalk 80 qui représente un outil efficace et performant de prototypage d'applications interactives. Avec cet outil la structure des objets complexes de l'application a pu être décrite et mise à l'épreuve aisément.

Grâce à cet environnement performant, une certaine abstraction est atteinte car les problèmes classiquement rencontrés dans le prototypage, comme la communication entre différents composants, la réalisation d'un dialogue agréable et la mise au point parfois longue de certains modules sont éliminés. De cette manière, certaines parties de la méthode ont pu être mises au point progressivement grâce à l'expérimentation et aux utilisations nombreuses et variées qu'une stagiaire en "Carrière de l'Information" a effectuées sur le système.

Une évaluation des performances a montré que RIVAGE, grâce à son processus de recherche progressive, peut satisfaire toute une variété de besoins ; le processus de recherche étant le plus souvent amorcé avec, au départ, une demande assez pauvre. Un bon rappel est en général obtenu. La précision, quant à elle, demeure correcte, malgré sa faible importance en recherche

d'images.

L'utilisation de RIVAGE dans de nouveaux champs d'application, et surtout dans le milieu industriel, permettra de mesurer sa réelle capacité à satisfaire des besoins concrets d'images.

RIVAGE est actuellement dans un état stable où il répond à une certaine forme de besoins, mais il demeure ouvert à toutes formes d'évolutions que nous souhaitons mettre en œuvre d'une manière progressive afin de contrôler à tout moment la complexité du système. Ces évolutions sont nombreuses et l'une d'entre elles a déjà vu le jour à travers la réalisation d'une structure hypertexte adaptée au processus d'interrogation progressif et interactif. C'est la présentation de cette approche qui fait l'objet du dernier chapitre de cette partie.

Chapitre 4	
Implantation d'une structure hypertexte dans le système RIVAGE.....	255
I Introduction.....	255
II Notion générale d'hypertexte.....	255
II.1 Historique.....	255
II.2 Les principales idées.....	256
II.3 Les solutions apportées.....	256
II.3.1 Le lien.....	257
II.3.2 Le nœud.....	258
II.4 Les problèmes liés à l'utilisation de système hypertexte.....	258
II.5 Les remèdes.....	259
III Définition du concept d'Hypermedia.....	260
IV RIVAGE et les besoins d'un utilisateur recherchant des images.....	260
V Réalisation d'une structure hypertexte dans le système RIVAGE.....	261
V.1 Le réseau de nœuds.....	261
V.2 Le nœud étape.....	263
V.3 Les aides associées au parcours.....	265
VI Conclusion.....	266

Chapitre 4

Implantation d'une structure hypertexte dans le système RIVAGE.

I Introduction

Une des perspectives que nous avons envisagées pour le système RIVAGE était d'améliorer le dialogue avec l'utilisateur afin de lui permettre de mener des recherches en suivant ses propres envies, c'est-à-dire en passant aisément du texte à l'image, d'une recherche à une autre ... Cette nouvelle conception du dialogue fait appel à des concepts liés à l'hypertexte, c'est pourquoi une étudiante de DEA, Catherine Hamon, s'est penchée sur ce problème et c'est sa solution que nous allons exposer dans ce chapitre [Ham 89].

Ce chapitre est découpé en quatre paragraphes : le premier et le deuxième présentent les différents concepts liés à l'hypertexte et plus généralement à l'hypermédia, le troisième met en évidence les besoins que peut avoir un utilisateur au cours d'une recherche, enfin le quatrième expose la solution proposée par Catherine Hamon, qui, par la définition d'une structure hypertexte associée à l'interrogation de RIVAGE, permet de prendre en compte une partie des besoins de l'utilisateur. Cette dernière présentation sera illustrée par quelques fenêtres qui ont été implantées dans le système RIVAGE.

II Notion générale d'hypertexte

II.1 Historique

Le concept d'hypertexte a été introduit dès 1945 par Bush [Bus 45] qui décrivait dans son article "As we may think" une machine permettant de parcourir en ligne un document d'une manière non séquentielle. En effet, cette manière de parcourir un document est souvent celle utilisée par l'être humain en quête d'une information. Quelques systèmes, permettant le parcours non linéaire d'un texte, ont été proposés à partir des années 60 [Con 87][Jul 88], mais c'est l'arrivée du multi-fenêtrage associé à l'utilisation d'une souris qui a remis le concept d'hypertexte au goût du jour. L'arrivée du logiciel HyperCard, offert par Apple à tous les acquéreurs de Macintosh, a renforcé l'intérêt de ces systèmes auprès d'un public plus vaste. Voyons alors plus en détails quelles sont les idées sous-jacentes à cette nouvelle vague de système.

II.2 Les principales idées

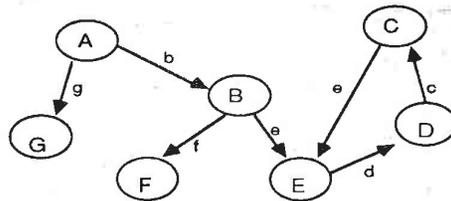
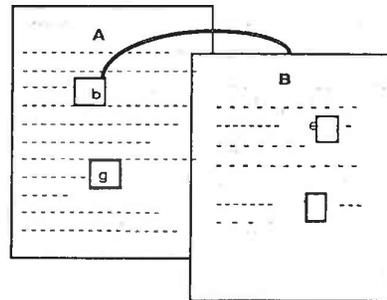
Les systèmes hypertextes doivent proposer aux utilisateurs le moyen de consulter un ou plusieurs documents en suivant un schéma de pensée qui peut être différent d'un utilisateur à l'autre. Les concepteurs de ces systèmes doivent répondre aux questions suivantes :

- comment prévoir l'ensemble des idées présentes dans un document qui seront susceptibles d'intéresser un utilisateur ?
- comment mettre en évidence dans un document un concept ou une idée importante ?
- comment relier les idées entre elles ?
- comment rendre l'utilisateur maître de ses décisions ?
- comment permettre à l'utilisateur certaines digressions dans son raisonnement sans qu'il perde son objectif initial ?
- comment permettre à l'utilisateur de faire une synthèse correcte des parcours qu'il a effectués à l'aide du système ?

II.3 Les solutions apportées

Les systèmes hypertextes possèdent essentiellement deux composants : le nœud et le lien. Le nœud est une représentation d'une idée ou d'un concept, les liens relient les nœuds entre eux; ils permettent d'associer les idées entre elles, donc plus globalement ils définissent des parcours possibles dans l'ensemble des idées que gère le système hypertexte. Cet ensemble de nœuds et de liens définit un réseau ou graphe associé à l'hypertexte appelé **hyperdocument**.

En règle générale les nœuds sont associés à des fenêtres sur l'écran et les liens sont matérialisés par des marques à l'intérieur des fenêtres, comme le montre la figure extraite de [Con 87] :



II.3.1 Le lien

Trois types de lien sont couramment utilisés pour relier les informations entre elles :

- le lien dit de "référence" ; c'est ce type de lien qu'illustre la figure précédente, il relie deux nœuds de manière explicite ; il est en général orienté.
- le lien organisé ; ce lien relie les informations d'une manière hiérarchique en introduisant les notions de nœuds "parents" et de nœuds "fils" qui définissent un arbre représentant alors un sous-graphe du réseau de nœuds associé à l'hypertexte.
- le lien par mot-clé ; ce lien matérialise une certaine forme d'indexation des nœuds. A partir d'un nœud plusieurs autres nœuds peuvent être atteints, ces nœuds ayant été indexés par le concept du nœud source. Ce lien est dit implicite car il n'est pas représenté dans le réseau de nœuds ; il pourrait être qualifié de lien "multi-référence".

Tous ces types de liens permettent de structurer l'information, ils permettent aussi à l'utilisateur de passer d'une idée à une autre en choisissant dans les fenêtres les marques représentant les liens.

II.3.2 Le nœud

Le nœud dans un système hypertexte représente une idée ou un concept. Ce découpage en nœuds permet de décrire tout un ensemble d'idées d'une manière modulaire et de créer des "associations d'idées" en les reliant par des liens. Un nœud d'une structure hypertexte prend toute son importance quand il possède une représentation concrète qui peut être affichée à l'écran et manipulée par l'utilisateur.

Trois types de nœuds sont souvent rencontrés dans les systèmes hypertexte :

- le nœud de type **note** réservé à la prise de note ; ce nœud permet à l'utilisateur de documenter son parcours dans le réseau de nœuds.
- le nœud de type **contrainte** ; ce nœud est un nœud de passage obligatoire dans le réseau hypertexte sans quoi les parcours seraient incohérents.
- le nœud de type **décision** qui permet à l'utilisateur de choisir la suite de son parcours parmi plusieurs alternatives.

Un quatrième type de nœud permet de regrouper en un seul nœud plusieurs nœuds : le nœud **composite**. Ce nœud permet une certaine abstraction dans le parcours, la consultation des nœuds regroupés ne se fera que sur demande de l'utilisateur.

Un système possédant une structure hypertexte en terme de nœuds et de liens, va offrir à l'utilisateur toute une panoplie de parcours qu'il pourra emprunter au gré de ses pensées. Mais cette forme d'interactivité pose un certain nombre de problèmes.

II.4 Les problèmes liés à l'utilisation de système hypertexte

Deux grands types de problèmes résultent de l'organisation complexe de l'information des systèmes hypertextes :

- le problème d'orientation,
- le problème de "saturation intellectuelle" (cognitive overhead).

Le problème d'orientation provient des nombreuses formes de parcours que l'utilisateur peut emprunter et qui finissent par le désorienter ; l'utilisateur n'arrive plus à se repérer dans le réseau, il n'arrive plus à accéder à des nœuds bien précis. L'hypertexte offre beaucoup de degrés de liberté qui permettent de se mouvoir aisément mais aussi de se perdre.

Le problème de "saturation intellectuelle" est provoqué chez un utilisateur par le fait qu'il

puisse mener plusieurs recherches différentes en même temps. En effet, dans ce genre de parcours l'utilisateur doit faire un effort important de concentration pour analyser et synthétiser la masse d'informations qu'il a obtenues en résultat de ses différentes recherches.

II.5 Les remèdes

Différentes solutions ont été proposées pour résoudre les problèmes liés à l'utilisation d'un système hypertexte.

Des browsers¹ graphiques permettent à l'utilisateur d'avoir une vue d'ensemble de son parcours dans le réseau hypertexte. Un problème se pose lorsque le réseau est trop complexe pour pouvoir être affiché à l'écran. Pour palier à ce problème, on peut utiliser des filtres qui sélectionnent dans le réseau les nœuds et liens les plus importants qui seront affichés.

Des mécanismes de recherche permettent à l'utilisateur de se situer dans le réseau. L'utilisateur indique les coordonnées de l'endroit où il se trouve (nom du nœud, type du nœud ..) et le système doit alors être capable de lui montrer la localisation du nœud où il se trouve par rapport à l'ensemble du réseau ; cette localisation peut être matérialisée sur le browser graphique.

Ces deux solutions permettent de limiter le problème d'orientation de la structure hypertexte. Le problème de "saturation intellectuelle" peut être réduit en proposant à l'utilisateur différents outils lui permettant d'effectuer une synthèse de ses recherches :

- la construction d'un ou plusieurs historiques des nœuds parcourus,
- la prise de notes associée à la construction de l'historique.

Cette construction d'historiques associée à une prise de notes offre le moyen à l'utilisateur de construire son réseau de nœuds personnalisé ; réseau qu'il peut reparer à volonté pour effectuer les synthèses nécessaires.

Nous venons de présenter l'environnement des systèmes hypertextes ; voyons maintenant comment un nouveau concept est venu quelque peu étendre le concept d'hypertexte, nous voulons parler du concept d'hypermedia.

¹ browsing signifie en français "feuilletter". Un browser est un outil permettant une visualisation d'un ensemble d'informations, mais il permet aussi de "feuilletter" cet ensemble.

III Définition du concept d'Hypermedia

La définition du concept d'hypermedia est un peu controversée. En effet, on peut considérer l'hypermédia seul comme étant l'ensemble des media réunis en un seul médium, ou on peut considérer les systèmes dits "hypermédia" comme pouvant être :

- un système hypertexte intégrant différents media autres que le texte (image, graphique, son, film ...),
- un système informatique manipulant plusieurs media.

Nous choisirons la deuxième définition des systèmes hypermedia car elle se veut plus générale et elle a l'avantage de contenir la première. De cette manière nous pouvons dire que RIVAGE est un système hypermedia car l'utilisateur manipule à la fois le texte et l'image. Ce choix n'est pas définitif, il pourra être changé lorsque le concept de système d'hypermedia sera clairement défini et entrera sans ambiguïté dans le monde des informaticiens.

IV RIVAGE et les besoins d'un utilisateur recherchant des images

L'état du système RIVAGE que nous avons décrit dans le chapitre précédent ne permet à l'utilisateur d'effectuer qu'une seule recherche à la fois, car cette réalisation n'avait pour but que de mettre en œuvre une méthode. Or les besoins et les envies d'un utilisateur recherchant des images peuvent être très variés et dépasser le cadre limité d'une recherche en plusieurs étapes.

Ainsi un utilisateur peut avoir envie de mener plusieurs recherches en même temps, de comparer les résultats d'une recherche avec ceux d'une autre ; cette comparaison pouvant aller jusqu'à remettre en cause les résultats d'une étape d'une recherche, en modifiant soit la demande, soit les choix lors de la visualisation. Il peut souhaiter aussi documenter ces différentes recherches afin de pouvoir en faire une synthèse à la fin de sa session d'interrogation.

Une recherche d'images peut être amorcée par d'autres manières que celle qui consiste à formuler une demande. En effet, l'utilisateur peut avoir envie de feuilleter la base d'images, de choisir les images qui l'intéressent le plus, puis de demander au système de lui trouver des images qui ressemblent à ces choix et de commencer ainsi un processus de recherche. Le processus peut aussi débiter grâce à un parcours dans le thesaurus en empruntant les relations sémantiques et où l'utilisateur demande à voir des images d'un concept qu'il a sélectionné.

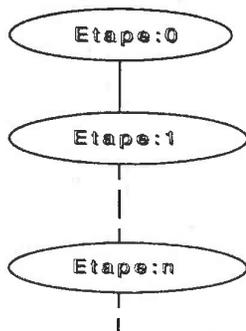
Ainsi un système d'interrogation d'une base d'images peut proposer toute une série de parcours possibles à l'intérieur la base, mais la mise en place de toutes ces fonctions nécessite la définition d'une structure permettant le contrôle et l'enchaînement des différentes actions de l'utilisateur. L'approche hypertexte semble apporter les concepts nécessaires à la réalisation d'une interface permettant de prendre en compte les différentes envies et besoins d'un utilisateur interrogeant une base d'images.

V Réalisation d'une structure hypertexte dans le système RIVAGE

La structure hypertexte proposée [Ham 89] [Hal 89b] repose sur la construction par l'utilisateur lors de son interrogation d'un réseau de nœuds de même type : le nœud **étape**. Ce nœud est à la fois un nœud composite et un nœud décision car il est composé de plusieurs nœuds et il permet à l'utilisateur de choisir entre plusieurs alternatives. Ce réseau de nœuds est construit dynamiquement lors d'une session d'interrogation. Voyons plus dans le détail quelles sont la structure du réseau et la structure des nœuds.

V.1 Le réseau de nœuds

Le parcours de l'utilisateur pendant une session d'interrogation dans l'ancienne version de RIVAGE peut se représenter de la manière suivante :

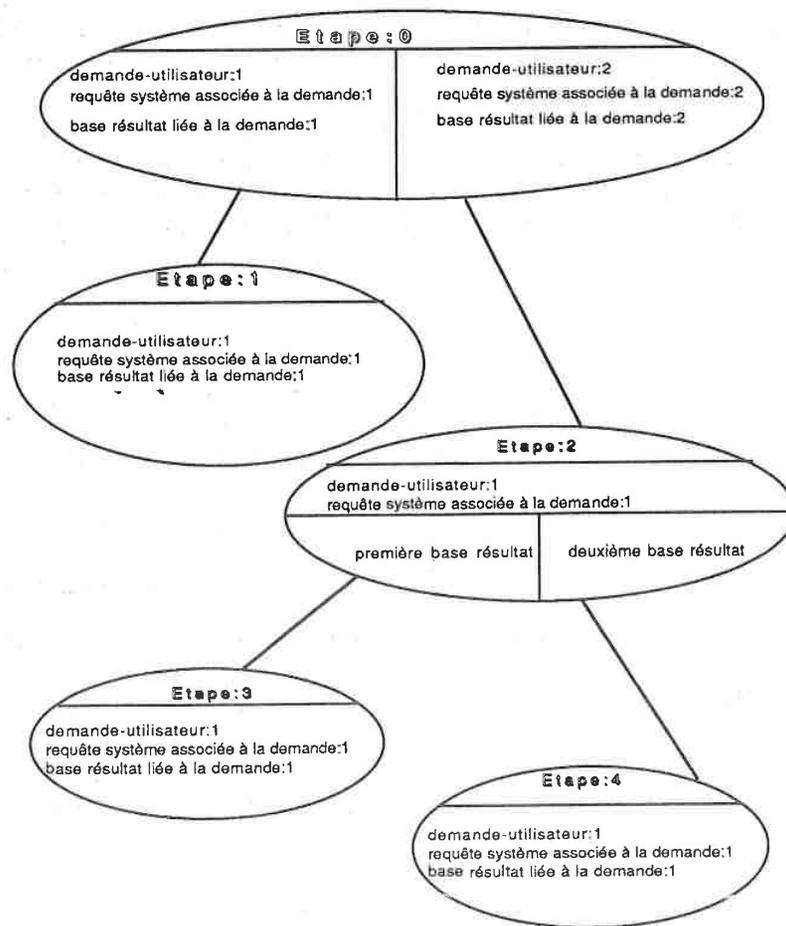


Le parcours est alors une liste d'étapes, une étape étant un nœud composite contenant un nœud demande, un nœud requête système, et un nœud base résultat, chacun de ces nœuds ayant une représentation à l'écran matérialisée par les fenêtres que nous avons décrites au chapitre 3, § II.2.

Mais si l'on veut permettre à l'utilisateur d'effectuer plusieurs recherches dans un même

temps, de les comparer, de les documenter, et de remettre en question certains choix, il faut modifier la structure du réseau de nœuds étapes de la session.

La structure proposée est la suivante :



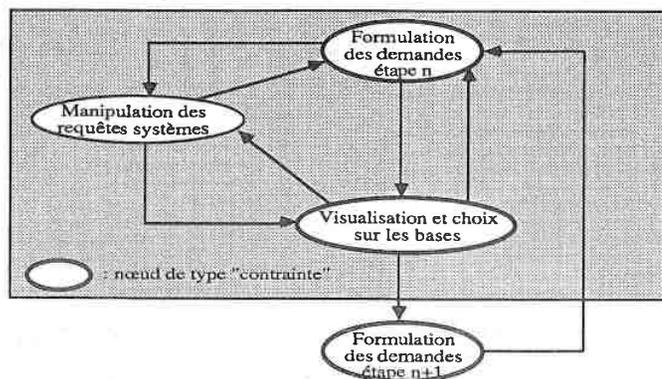
Lors de ce parcours l'utilisateur a effectué cinq étapes. A la première étape, ici étape 0, l'utilisateur a formulé deux demandes, chacune lui ayant fourni deux bases résultats différentes.

Après avoir effectué un choix sur la première base, il a demandé au système d'analyser ses choix, le système lui a donc proposé une nouvelle demande dans l'étape 1. Après avoir visualisé les images qu'a sélectionnées cette demande, l'utilisateur a décidé de revenir à l'étape précédente pour exploiter sa deuxième recherche. Le second choix qu'il a effectué sur la deuxième base de l'étape 0 lui a permis de voir une nouvelle demande, celle de l'étape 2 ; avec cette nouvelle demande, l'utilisateur a obtenu une base résultat sur laquelle il a effectué deux choix différents, le premier choix a permis de générer l'étape 3, et le deuxième l'étape 4.

Grâce à cette nouvelle forme de parcours, l'utilisateur peut mener plusieurs recherches dans un même session d'interrogation, effectuer des choix différents et remettre en cause des choix qu'il avait déjà effectués.

V.2 Le nœud étape

Le nœud étape est un nœud composite contenant un ensemble de nœuds permettant la définition du parcours dans une étape donnée. Le parcours dans une étape n est représenté par le réseau de nœuds suivant :



Le parcours dans une étape peut être matérialisé par trois nœuds :

- formulation des demandes,
- manipulation des requêtes systèmes,
- visualisation et choix sur les bases d'images.

Chacun de ces nœuds est représenté par une fenêtre à l'écran possédant un certain nombre d'options permettant à l'utilisateur de parcourir le réseau de nœuds qu'il a construit. Par

exemple le nœud "visualisation et choix sur les bases d'images" est associé à la fenêtre suivante:

Les bases formulées à partir de la base descriptive Ministère de la Culture; étape 0				
base : 1	240: Moutons broutant ... : rejetée	N° vidéodisque	MONTAGNE	Arrêt
	12: Personnage ... pont ... : choisie	contenu	PAYSAGE	Requête
base : 2	5: Promeneurs en file ... : choisie	connotation	PERSONNE	Demande
	239: Paysage brumeux : rejetée	morphologie	GLACIER	Etape suivante
		commentaire	EXCURSION	Nouveau choix
		précision ...		Annuler choix
		auteur		Parcours
		titre		Informations
			NOTES :	
				Fin de la visualisation

Les options permettant à l'utilisateur de parcourir le réseau de nœuds qu'il a construit pendant sa session d'interrogation sont matérialisé sur la fenêtre par des boutons. En sélectionnant les boutons avec la souris, les actions suivantes sont effectuées :

- Arrêt : la session s'arrête,
- Requête : l'utilisateur passe au nœud "Manipulation des requêtes systèmes",
- Demande : il passe au nœud "Formulation des demandes" de la même étape,
- Etape suivante : il passe au nœud "Formulation des demandes" de l'étape suivante, si elle existe, c'est-à-dire si l'utilisateur a déjà demandé au système d'analyser ces choix sur la base résultat sélectionnée à l'écran, ici la base n° 2,
- Nouveau choix : cette option permet à l'utilisateur d'effectuer nouveau choix d'images sur la base sélectionnée ; pour ce faire le système crée une nouvelle base semblable à celle sélectionnée, c'est sur cette base que l'utilisateur pourra effectuer son nouveau choix,
- Annulation choix : cette option annule le choix sur la base sélectionnée,
- Parcours : cette option sera détaillée par la suite, elle permet à l'utilisateur d'avoir une vue globale sur son parcours,
- Informations : cette option donne une vue plus précise de la base résultat sélectionnée,

- Fin de la visualisation : cette option indique au système qu'il doit passer en phase "Après-Visualisation", c'est-à-dire qu'il doit analyser les choix de l'utilisateur sur la base résultat sélectionnée puis proposer une nouvelle demande dans un nouveau nœud "Formulation des demandes" de l'étape suivante.

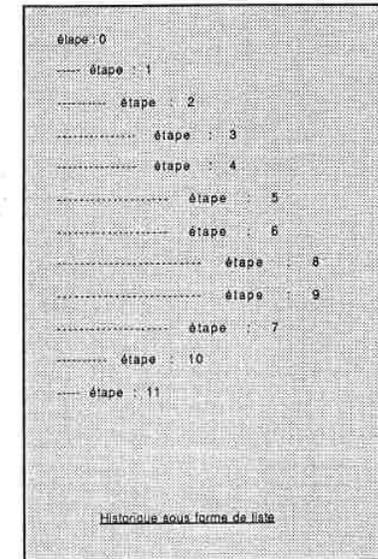
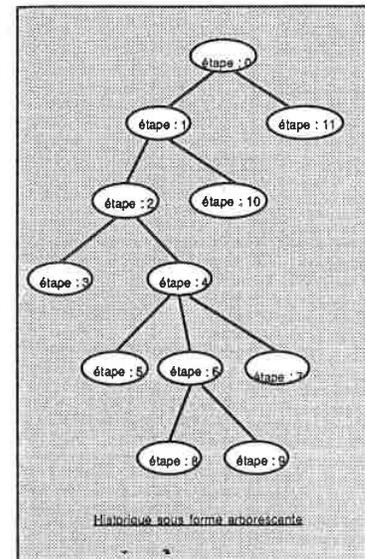
Connaissant les problèmes liés à l'utilisation d'un système hypertexte quelques aides ont été réalisées permettant à l'utilisateur d'éviter de se perdre et de documenter ses recherches.

V.3 Les aides associées au parcours

Une des principales aides qu'offre le système est la possibilité donnée à l'utilisateur de consulter à tout moment l'historique de son parcours. Cet historique offre à l'utilisateur une vue générale sur son parcours, il lui permet aussi de se repérer en lui indiquant l'endroit où il se trouve lors de la demande de consultation. Voici la fenêtre qui visualise cet historique à l'écran :

Visualisation du parcours	
Etape 0	Informations sur les demandes
.....Etape 1	
.....Etape 2	
.....Etape 3	
	Informations sur les bases
	Retour
Aller à l'étape sélectionnée	

Cet historique représente un nouveau nœud accessible à partir de chaque nœud du parcours. Il montre les différentes évolutions des recherches : le numéro des étapes représente l'ordre chronologique de l'enchaînement des étapes, l'indentation indique l'origine de chaque étape. Cette forme de représentation est équivalente à une représentation sous forme arborescente :



La prise de notes associée à chaque nœud d'une étape constitue la deuxième forme d'aide ; elle permet de limiter le problème de "saturation intellectuelle". Grâce à cette prise de notes l'utilisateur documente ses différentes recherches afin de ne pas oublier le pourquoi de ses différentes démarches.

VI Conclusion

L'utilisation des concepts liés à l'hypertexte nous a permis de réaliser un dialogue mieux adapté à la recherche progressive et interactive d'images. Le dialogue est représenté sous la forme d'une structure hypertexte que l'utilisateur construit dynamiquement au gré de ses différentes recherches.

Ce nouveau dialogue a été réalisé sur la version de RIVAGE qui a été décrite dans le chapitre précédent. La modification des structures des objets existants a été possible grâce aux qualités d'évolutivité et de contrôle des mises à jour qu'offre le langage Smalltalk 80.

La structure hypertexte proposée constitue une première approche, d'autres formes de parcours peuvent y être ajoutées comme par exemple la recherche d'images par l'image, cette

forme de parcours nécessite une meilleure intégration de l'image dans l'environnement hypertexte.

Conclusion

Cette dernière partie a montré comment le processus EXPRIM pouvait être assimilé à un processus d'apprentissage. L'étude de quelques techniques d'apprentissage symbolique nous a permis d'extraire les qualités que nous souhaitons donner à cet apprentissage : représentation d'un concept dans une taxonomie, et utilisation d'une logique majoritaire pour mettre en évidence les caractéristiques des exemples.

L'originalité de la méthode proposée se situe dans l'adaptation de méthodes de bouclage de pertinence avec des techniques de représentation de concept dans des taxonomies. Dans notre méthode le thesaurus joue le rôle d'une taxonomie dans laquelle le système tente de représenter le besoin de l'utilisateur. Les méthodes de bouclage de pertinence interviennent dans l'analyse des choix de la phase "Après-Visualisation" où un poids de pertinence est calculé pour chaque terme présent dans les descriptions. La propagation des poids dans le thesaurus permet au système d'avoir une représentation dans le thesaurus du concept à apprendre. Cette propagation utilise une certaine logique majoritaire qui permet de dire qu'un concept est appris lorsqu'une partie de ces représentations est pertinente dans les choix de l'utilisateur. La représentation dans le thesaurus du besoin de l'utilisateur permet la reformulation par le système d'une nouvelle demande. Cette nouvelle demande est montrée à l'utilisateur qui peut alors la modifier pour la rendre plus proche de ces besoins. Cette méthode permet de représenter et de décrire le processus EXPRIM dans sa totalité en mettant en évidence les différentes étapes existantes. L'ensemble de la méthode a été décrite dans le deuxième chapitre, cette description a été effectuée à l'aide d'un formalisme permettant d'expliquer les algorithmes de la méthode indépendamment de toute forme de réalisation.

Le troisième chapitre a été consacré à la description du prototype RIVAGE qui a permis de valider la méthode proposée pour la réalisation du processus EXPRIM. La forme du dialogue ainsi que les structures des objets manipulés ont été décrites. Cette expérimentation nous a permis de mettre au point les parties dites "heuristiques" comme la déformation de la demande, la propagation des poids... L'évaluation, menée par une stagiaire de l'IUT carrière de l'information, a montré que RIVAGE était capable de répondre à une variété de besoin d'une manière similaire à celle d'un documentaliste connaissant la base d'images, le système et la formulation booléenne.

L'adaptation d'une structure hypertexte pour la représentation du dialogue, décrite dans le

chapitre 4, a montré que RIVAGE est ouvert à toutes formes d'évolutions. Ainsi les perspectives émises à court terme pourront être intégrées dans le système assez aisément. Quant aux autres perspectives une étude plus approfondie est nécessaire, mais leurs expérimentations dans le prototype RIVAGE ne poseront aucun problème.

CONCLUSION

Cette thèse a été consacrée à l'étude de la problématique de la recherche d'images et à la proposition d'un système interactif et progressif de recherche d'images, RIVAGE, utilisant le processus EXPRIM et réalisé sous la forme d'un apprentissage à partir d'exemples.

La problématique de la recherche d'informations a été étudiée à travers une présentation de la recherche d'informations en général, et à travers une présentation de l'image et de ses spécificités. Cette étude a mis en évidence les points suivants :

au sujet des documents :

- la réalisation d'un tel système doit passer par une phase d'indexation manuelle des images,
- une analyse de l'image est nécessaire, cette analyse doit être menée en suivant une méthode qui doit permettre de ne rien oublier et d'obtenir une certaine homogénéité dans les descriptions,
- l'utilisation d'un thesaurus lors de la phase d'indexation apporte une cohérence sémantique entre les différentes descriptions, en réduisant le nombre de cas de synonymie et de polysémie, sources de bruit et de silence dans les recherches,

au sujet de la recherche :

- un système de recherche d'images ne peut pas comme certains systèmes de recherche textuelle être entièrement automatique,
- la place du dialogue dans un processus de recherche d'images est très importante, celui-ci doit être très interactif, cohérent et progressif,
- l'arrivée des supports optiques, et surtout celle du vidéodisque, a permis aux systèmes de recherche d'images d'offrir un dialogue plus interactif car l'image peut être vue en résultat d'une recherche, et donc être utilisée dans le dialogue avec l'utilisateur en lui permettant d'effectuer des tris visuels,
- l'utilisation d'une connaissance sémantique de type thesaurus peut nous permettre de proposer un processus de recherche cohérent,
- l'utilisation du thesaurus conjointement à une formulation booléenne offre un pouvoir d'expression important mais qui n'est pas facilement contrôlable par un utilisateur non initié,
- l'apprentissage a un rôle important à jouer, que ce soit à l'échelle de la vie du système ou à l'échelle d'une session d'interrogation. Cette thèse illustre fortement ce dernier point.

Notre contribution à la réalisation du projet EXPRIM a débuté par une participation à un projet ESPRIT qui s'est concrétisée par la définition et la réalisation d'un système "intelligent"

de recherche d'images basé sur les idées du projet EXPRIM et utilisant l'Imageur Documentaire de la SEP.

Notre approche décisive concernant la réalisation du processus EXPRIM s'est concrétisée avec le système RIVAGE. Dans ce système le processus EXPRIM est assimilé à un processus d'apprentissage ayant les caractéristiques suivantes :

- représentation du concept à apprendre (le besoin de l'utilisateur) par un marquage dans une taxonomie (le thesaurus)
- utilisation d'une logique majoritaire pour mettre en évidence les caractéristiques des exemples.

L'originalité de la méthode proposée se situe dans l'adaptation de méthodes de bouclage de pertinence avec des techniques de représentation de concept dans des taxonomies. Dans cette méthode, le thesaurus joue le rôle d'une taxonomie dans laquelle le système tente de représenter le besoin de l'utilisateur en attachant des **poids d'expressivité** aux termes du thesaurus. L'ensemble des poids d'expressivité forme un **niveau d'expression** de la demande que le système peut parcourir pour montrer à l'utilisateur une expression de ses besoins. Cette méthode permet de représenter et de décrire le processus EXPRIM dans sa totalité et, d'une façon générale les processus de recherche d'images.

Le prototype RIVAGE a permis de valider la méthode proposée. Il a été réalisé d'une manière incrémentale grâce au langage Smalltalk 80 qui représente un outil efficace et performant de prototypage d'applications interactives. Avec cet outil, la structure des objets complexes de l'application a pu être décrite et mise à l'épreuve aisément. Cette expérimentation nous a permis de mettre au point les parties dites "heuristiques" de la méthode (la déformation de la demande, la propagation des poids...)

Une évaluation des performances a montré que RIVAGE, grâce à son processus de recherche interactif et progressif, était capable de répondre à une variété de besoins d'une manière similaire à celle d'un documentaliste connaissant la base d'images, le système et la formulation booléenne, le processus de recherche étant le plus souvent amorcé avec, au départ, une demande assez pauvre. L'utilisation prochaine de RIVAGE dans le milieu industriel, permettra de mesurer sa réelle capacité à satisfaire des besoins concrets d'images.

De nombreuses perspectives ont été envisagées, celles définies à court terme pourront être implantées dans RIVAGE aisément car la réalisation d'une des perspectives à moyen terme qui consistait à proposer une structure hypertexte pour la représentation du dialogue, a montré que

RIVAGE était ouvert à toutes formes d'évolutions.

Les autres perspectives envisagées sont les suivantes :

- la réalisation d'outils d'aide à l'indexation et d'aide à la construction de thesaurus ; cette évolution peut faire baisser le coût de mise en œuvre d'une application de recherche d'images qui constitue un réel blocage à la commercialisation sur le marché économique d'un système de recherche d'images tel que le système RIVAGE,
- la généralisation du processus d'apprentissage mis en œuvre dans RIVAGE, afin de définir un modèle de recherche progressive de documents fondé sur une connaissance sémantique ; cette perspective est avant tout liée à la recherche sur les modèles utilisés dans les systèmes de recherche d'informations. Cette perspective permettrait d'utiliser notre méthode pour des recherches d'informations autres que la recherche d'image.
- l'intégration de capacités d'explicabilité dans le processus de recherche ; cette perspective a pour but d'améliorer le dialogue homme-machine lors du déroulement du processus de recherche,
- la réalisation d'un apprentissage à long terme ; cette perspective est sûrement celle qui prolonge le mieux notre travail car elle devrait nous permettre d'adapter d'autres formes d'apprentissage afin de donner au système la capacité d'évoluer au fur et à mesure des utilisations.

BIBLIOGRAPHIE

- [And 88] Andraws A.
Système d'aide à la construction de thesaurus (SACT).
Rapport de DEA, CRIN, Nancy, Septembre 1988.
- [Ant 88] Antoniadis G., Lulich Badin G., Polity Y., Raoult J.
A french words recognition for Information retrieval system.
Proceedings ACM IR 88, pp 67-84.
- [Ato 84] Atouri M.
Utilisation du système Expert pour la déformation de requêtes dans le projet EXPRIM.
Rapport de DEA, CRIN, Nancy 1984.
- [Bar 83] Baryla C.
Un vidéodisque interactif à la Bibliothèque Sainte-Geneviève.
Bulletin des Bibliothèques de France, Paris, t 28, n° 4, 1983, pp 373-382.
- [Bar 88] Barthes C., Glize P.
Planning in an expert system for Automated Information retrieval.
Proceedings ACM IR 88, Grenoble pp 535-550.
- [Bas 86] Bassano J.C.
DIALECT : Un système expert pour la recherche documentaire.
Sixièmes journées internationales SYSTEMES EXPERT, Avignon 86.
- [Bas 87] Bassano J.C.
Systèmes experts et systèmes documentaires intelligents. Etat de l'art et prospective.
Proceeding of Avignon 87. Les systèmes experts et leur applications, pp 491-510.
- [Bla 88] Black W.J.
Les systèmes intelligents basés sur la connaissance.
Editions Masson 1988.

- [Ble 81] Blery G.
La mémoire photographique.
Bulletin interphototèque n° 41, "Analyse de l'image fixe", Décembre 81.
- [Bus 45] Bush V.
As we may think.
Atlantic Monthly, July 1945, pp 101-108.
- [Cau 89] Cauvet C., Rolland C.
O* : Un modèle pour la conception de bases de données orientées objets.
Actes du congrès INFORSID, Nancy, 1989, pp 265-284.
- [Cha 77] Chaumier J.
Le traitement linguistique de l'information documentaire : l'analyse documentaire.
Entreprise Moderne d'édition, 1977.
- [Chi 86] Chiramella Y., Defude B. Bruandet M.F. Kerkouber D.
IOTA : A full text information retrieval system.
Proceeding of ACM 1986 : Information Retrieval, PISE, pp 207-213.
- [Chi 87] Chiramella Y., Defude B.
IOTA : Un prototype de système expert en recherche d'information.
Proceeding of Avignon 87. Les systèmes experts et leur applications, pp 511-526.
- [Cle 88] Clemencin G.
Querying the French Yellow Pages : Natural language Acces to the directory.
Proceeding of RIAO 88, Boston, pp 286-312.
- [Clos 64] Cros R.C., Gardin J.C., Levy F.
L'automatisation des recherches documentaires.
Gauthier-Villars. PARIS GU, 1964.

- [Clu 88] Cluzeau-Ciry M.
Typologie des utilisateurs et des utilisations d'une banque d'images.
Le documentaliste, vol 25, n°3 Mai-Juin 1988, pp115-120.
- [Con 87] Conklin J.
Hypertext : An introduction and Survey.
Computer, september 1987, pp 17-41.
- [Coc 88] Cochet Y., Paget G.,
Formal Epigenis of Neural Networks.
IEEE ICNN 88, San Diego, Mars 1988.
- [Cou 86] Coulon D, Kayser D.
Informatique et langage naturel : présentation générale des méthodes d'interprétation du texte écrits.
Techniques et Sciences Informatiques, Vol 5, n°2, mars-avril 1986, pp 103-108.
- [Cre 85] Créhange M., David J.M., Foucaut O., Heully B.
Le point sur EXPRIM (Système Expert pour la recherche d'IMages)
Rapport interne CRIN, Nancy n° 85-R-019 1985.
- [Cre 86] Créhange M., David J.M., Foucaut O., Halin G., Thiéry O.
Les structures de données dans le projet EXPRIM.
Proceedings of INFORSID, Fontevraud, Mai 1986.
- [Cre 88a] Créhange M.
Bases d'images et Intelligence Artificielle
Dans "Images et Vidéodisque" la Documentation Française et la DBMIST, Paris 88.
- [Cre 88b] Créhange M.
Par et pour la recherche d'Images : EXPRIM
Le médiéviste et l'ordinateur, n°19, Printemps 1988.
- [Cre 89a] Créhange M., Halin G.
Image progressive retrieval from a Videodisk : A Machine Learning Problem.
Conférence on Applications of Artificial Intelligence, Orlando, USA, Mars 89.

- [Cre 89b] Créhange M., Halin G.
Machine Learning Technics fo Progressive Retrieval in an Image Database.
Congrès à BTW89 (Datenbanksysteme in Büro, Technik und Wissenschaft) Zurich, Mars 1989.
- [Cre 89c] Créhange M., Foucaut O., Halin G., Mouaddib N., Foucaut J.F.
Semantics of user Interface for Image Retrieval : possibility theory and learning techniques.
A paraître dans Information Processing and Management.
- [Cro 64] Cros R.C., Gardin J.C., Levy F.
L'automatisation des recherches documentaires, un modèle général : le SYNTHOL.
Gauthier-Villard, Paris, 1975.
- [Cro 86] Croft W.B.
User-specified domain knowledge for document retrieval.
Proceeding of ACM on Development in IR. Pise 86 pp 201-206.
- [Dav 89a] David A., Thiéry O., Créhange M.
Intelligent Image Base Computer Aided Education (IICAE).
Conference on Applications of Artificial Intelligence, Orlando, USA, Mars 89.
- [Dav 89b] David A., Thiéry O., Créhange M.
Intelligent Hypermedia for Education.
ICCAL (International Conference on Computer-Assisted Learning), Dallas, USA, Mai 1989, paru chez Springer-Verlag, edition H. Mauer.
- [Deb 88] Debili F., Radassa P., Fluhr C.
About reformulation in Full Text IRS.
Proceeding of RIAO 88, Boston, pp 343-357
- [Del 82] Delobel C. Adiba M.
Bases de données et systèmes relationnels.
Dunod, Paris 1982.

- [Del 86] Delcroix J.C.
Semantic relationships and general framework for expert informations retrieval systems.
Rapport de la tâche D2/I du projet Esprit n° 901, juillet 86.
- [Del 88] Delcroix J.C., Ailloud B., Bigeard C.
Economical feasibility of expert systems for image bases.
Rapport de la tâche K1 du projet Esprit n° 901, janvier 1988.
- [Den 79] Denis M.
Les images mentales.
PUF 79.
- [Dil 80] Dillon M., Desper J.
Automatic relevance feedback in boolean retrieval systems
Journal of Documentation 1980, 36, pp 197-208.
- [Dre 84] Dreyfus, Hubert L.
Intelligence artificielle, mythes et limites.
Flammarion 1984.
- [Dur 79] Durand J.C., Moitrieux G.
Lire les images.
CRDP Nancy 79.
- [Dur 87] Dureau J.M.
Un vidéodisque d'images régionales en Rhône-Alpes : un travail collectif des bibliothèques, musées et archives.
Bulletin d'information de l'Association des Bibliothécaires Français, n°135, 1987, pp 25-32.
- [Elh 87] Elhana A.
Coopération entre SBC et systèmes de données. Application à l'implantation d'heuristiques d'EXPRIM en OPS5+.
Rapport de DEA, CRIN Nancy, septembre 1987.

- [Far 87] Farrenny H., Ghallab M.
Eléments d'intelligence artificielle.
Editions Hermes, 1987.
- [Fuh 86] Fuhr N.
Two Models of Retrieval with Probabilistic indexing.
Proceeding of ACM Information Retrieval, Pise 1986 pp 249-257.
- [Gar 84] Garnier François
Thesaurus iconographique, système descriptif des représentations.
Le leopard d'or 1984.
- [Gau 81] Gautier G.
Vingt leçons sur l'image.
Interphotothèque. Décembre 81.
- [God 88] Godot N., Guimier-Sorbets A.M., Mordini P.
Interrogation simultanée de banques de données et d'images, les possibilités offertes par SIGMINI.
Proceeding of IMACOM' 88, Besançon.
- [Gol 83] Goldberg A., Robson D.,
Smalltalk 80 ; the language and its implementation.
Addison-Wesley reading, Massachussets, 1983.
- [Goo 88] Goodman D.
HyperCard - HyperTalk, guide d'utilisation et de programmation.
Edition Dunod, 1988.
- [Hal 86] Halin G. & al.
Specification of the functions of the dialogue and piloting module.
Rapport CRIN n°87-R-115. Rapport Task T2/J du contrat Esprit N°901.
- [Hal 88a] Halin G., Mouaddib N., Foucaut O., Créhange M.
Semantics of user Interface for Image Retrieval : possibility theory and learning techniques applied on two prototypes.
Proceedings of Conference RIOA 88, Boston (USA), Mars 1988.

- [Hal 88b] Halin G., Leonard D.
Spécifications externes et internes du système EXPRIM.
Rapport interne CRIN, 1988, n° 88-R-075.
- [Hal 89a] Halin G., Créhange M.,
Image Progressive Retrieval from a Videodisc.
Proceedings of Opticalinfo '89, Amsterdam, Avril 1989.
- [Hal 89b] Halin G., Hamon C.
An hypermedia structure for an interactive and progressive image retrieval system.
Proposé au 13ième "International Online Information Meeting", Londres, décembre 1989.
- [Ham 89] Hamon C.
Réalisation d'une structure hypertexte sur le système RIVAGE basée sur le processus EXPRIM.
Rapport de DEA, CRIN Nancy, juin 1989.
- [Hat 88] Hatala A.
Edition électronique et Culture.
Proceeding of IMACOM' 88, octobre 1988, Besançon.
- [Hud 83] Hudrisier H.
L'iconothèque - Documentation audio-visuelle et banques d'images.
Thèse d'Etat, la documentation Française, Paris 1982.
- [Hud 84] Hudrisier H.
L'imageur documentaire SEP - SYGMA : un outil complémentaire nécessaire au vidéodisque d'archives.
BUROTICA 84. Edition ? Paris 1984.
- [Hud 85] Hudrisier H.
L'imageur documentaire.
Le documentaliste, vol 22,n°4-5. Juillet - Octobre 1985, pp 155-160.

- [Hul 87] Hull R., King R.
Semantic Database Modelling : Survey, Applications and research Issues.
ACM Computing Survey, 19(3) : pp 202-260.
- [Hum 89] Humbert S.
L'expérimentation du système RIVAGE.
Rapport interne CRIN, Nancy, 1989.
- [Ima 88] **Image et Vidéodisque.**
DBMIST, La documentation française, Paris 88.
- [Jac 88] Jacob P.S., Rau L.F.
Natural Language Techniques for Intelligent Information Retrieval.
Proceeding of the 11th international in information retrieval, Grenoble 88, pp 85-95.
- [Jul 88] Julien C.
Bases d'informations généralisées : contribution à l'étude des mécanismes de consultation d'objets multimedia.
Thèse de l'Université Paul Sabatier, Toulouse, Octobre 1988.
- [Kor 87] Korchia R.
Les archives photographiques.
Extrait n°148 "Monuments Historiques", Editions de la Caisse des Monuments Historiques et des Sites, 1987.
- [Kod 85] Kodratoff Y.
Une théorie et une méthodologie de l'apprentissage symbolique.
Proceeding Cognitiva 85 Paris, Juillet 85.
- [Kod 86] Kodratoff Y.
Leçons d'apprentissage symbolique.
Cepadues edition 1986.

- [Lar 77] Larson J., Michalski R.S.
Inductive inference of VL decision rules.
Proceeding of the Workshop on Pattern Directed Inference Systems, SIGART Newsletter 63, June 1977, pp 38-44.
- [Lâa 88] Lâasri H., Maître B., Haton J.P.
Organisation, coopération et exploitation des connaissances dans les architectures de blackboard : ATOME.
Proceeding of Avignon 88, les systèmes experts et leur applications, pp 371-390.
- [Léo 86] Léonard D.
Contribution à la spécification d'une interrogation assistée dans un système iconographique utilisant un thesaurus.
Rapport de DEA, CRIN, Nancy, septembre 1986.
- [Léo 89a] Léonard D, Chrisment C., Zurfluh G., Halin G.
Bases de données multimedia et langages orientés objets.
Support de cours de congrès INFORSID, Nancy, Mai 1989.
- [Léo 89b] Léonard D
Conception et représentation des objets à comportement complexe.
Actes du congrès INFORSID, Nancy 89, pp 285-298.
- [Los 86] Losse R., Bookstein A., Yu C.
Probabilistic Models for Document Retrieval : A comparaison of performance on Experimental & Synthetic databases.
Proceeding of ACM Information Retrieval, Pisa 1986 pp 258-264.
- [Maz 85] Mazerolles D.
Réalisations des interfaces du projet EXPRIM.
Rapport DESS, CRIN, Nancy, Septembre 1985.
- [Mic 83a] Michalski R.S., Carbonell J.G., Mitchell T.M.
Machine Learning, an Artificial Intelligence Approach.
Tioga Publishing Company, Palo Alto 1983.

- [Mic 83b] Michalski R.S., Carbonell J.G., Mitchell T.M.
An overview of Machine Learning.
In Machine Learning, an Artificial Intelligence Approach, 1983 pp 3-23.
- [Mic 83c] Michalski R.S., Diettrich T.G.
An comparative review of selected methods for learning from examples.
In Machine Learning, an Artificial Intelligence Approach, 1983 pp41-81.
- [Min 77] Minsky M.
The society theory of thinking.
Proceedings of IJCAI, Cambridge, August 1977.
- [Mis 75] **Mistral.**
Manuel de presentation, CII, 1975.
- [Mit 83] Mitchell T.M., Utgoff P.E., Banerji R.
Learning by experimentation, acquiring and refining problem-solving heuristic.
In Machine Learning, an Artificial Intelligence Approach, 1983 pp 163-190.
- [Mor 88] Mordini M., Jarmouni idrissi M., Guimier-Sorbets A.M.
Self-structured data banks semantic integrity and query assistance interface.
Proceeding of RIAO 88, Boston, pp 439-449.
- [Mou 88] Mouaddib N.
Gestion des informations nuancées dans les systèmes d'informations. Application à une base de connaissance en sciences naturelles.
Congrès BD3, Bénodet, 1988.
- [Mou 89a] Mouaddib N.
Gestion des informations nuancées : une proposition de modèle et de méthode pour l'identification nuancée d'un phénomène.
Thèse de l'Université de Nancy I, Mai 1989.

- [Mou 89b] Mouaddib N., Halin G.
Deux modèles sémantiques pour la mise en correspondance entre requêtes et documents.
Proceeding of INFORSID, Nancy, Mai 1989, pp 451-476.
- [Odd 77] Oddy Robert N.
Information retrieval from man machine dialogue.
Journal of Documentation. Mars 1977 vol 33 n° p1-14.
- [Opt 87] **OPTICA 87**
Proceedings of the international meeting for Optical Publishing & Storage
Amsterdam April 1987.
- [Opt 89] **OPTICA 89**
Proceedings of the international meeting for Optical Publishing & Storage
Amsterdam April 1989.
- [Pel 88] Pelletier F.
Les mémoires optiques en 1988.
Dans la revue "Mémoires optiques & systèmes" n° 68 Septembre Octobre 1988.
- [Pie 87] Pierrel J.M.
Dialogue oral Homme-Machine.
Editions Hermès, Paris, 1987.
- [Pla 88] Plande J.J.
Lecture optique par laser et interactivité.
In Proceedings of IMACOM 88. Salon international sur la vidéo interactive et ses applications. Besançon.
- [Pol 88] Pollitt A.S.
Men USE for Medicine : End User Browsing and Searching of MEDLINE via the MeSH Thesaurus.
Proceeding of RIAO 88 pp 547-573.

- [Por 76] Porcher L.
Introduction à une sémiotique des images.
Editions Credif Didier, Paris, 1976.
- [Qui 84] Quinqueton J., Sallantin J.
Généralisation par points de vue et apprentissage de concept
Rapport de recherche INRIA, Janvier 1984.
- [Qui 87] Quinqueton J., Boucheron S., Sallantin J, Soldano H.
CALM: Contester pour apprendre en logique Modale
Intellectica Vol 1, n°2-3 1987 pp 144-177.
- [Rau 88] Rau Lisa F.
Conceptual Information Extraction and Retrieval From Natural Language input.
Proceeding of RIAO 88 pp 424-437.
- [Reb 88] Rebuffet M., Halin G., Créhange M.
Technical feasibility report.
Rapport de la tâche K2 du projet ESPRIT n° 901, janvier 1988.
- [Roc 65] Rocchio J.J. Jr.
Relevance feedback in information retrieval.
Chap 14, in the Smart System Experiments in Automatic Document Processing.
Prentice-Hall, Englewood Cliffs, New Jersey, 1971.
(reprinted from Scientific Report ISR-9, Computation Laboratory, Haward University, August 1965)
- [Rol 87] Rolland C., Foucaut O., Benci G.
Conception des systèmes d'informations : la méthode REMORA.
Editions Eyrolles, 1987.
- [Roo 86] Roosels G., David A.
Enseignement par l'image assistée par ordinateur.
Rapport de la tâche I3 du projet Esprit n°901, Septembre 1986.

- [Rua 85] Ruaud Y.
Système expert pour la partie avant visualisation du projet EXPRIM.
Rapport de DEA, CRIN, Nancy 1986.
- [Sal 83a] Salton G., Fox E.A., Wu H.
Extended Boolean Information Retrieval.
Communication of th ACM 26:11, Novembre 1983, pp 1022-1036.
- [Sal 83b] Salton G., Mc Gill M.J.
Introduction to Modern Information Retrieval.
Mc Graw-Hill, International company New-York 1983.
- [Sal 83c] Salton G., Fox E.A., Buckley C., Voorhees E.
Boolean Query Formulation with relevance feedback.
Tech Rep TR 83-539, Department of computer Science, Cornell University, January 1983.
- [Sal 84] Salton G., Voorhees E.
A comparison of two methods for boolean query relevancy feedback.
Information Processing & Mangement Vol 20, n°5/6 pp637-651, 1984.
- [Sal 85] Salton G.
A note on Information Retrieval Models & Theories.
Proceeding of RIAO, Grenoble 1985, pp 1-27.
- [Sal 86] Salton G.
Recent Trends in Automatic Information Retrieval.
Proceedings of ACM Information Retrieval, Pisa 1986, pp 1-10.
- [Sin 87] Sinno S., Degez-Vataire D.
Les Logiciels documentaires de pilotage de vidéodisques.
BPI, DBMIST, La documentation française, Paris, 1987.
- [Smi 77] Smith J.M., Smith C.P.
Database Abstraction. Aggregation & Generalization.
ACM Transaction Database Systems, june 1977.

- [Smi 80] Smith L.C.
Artificial intelligence application in information systems.
In annual review of Information Science of Technology, 15, Knowledge Industry
Publication, Inc, White Plains, New York 1980, pp 67-105.
- [Tab 76] Tabani V.
A Fuzzy Model of Document Retrieval Systems.
Information Processing and Management, 12:3, 1976, pp 177-188.
- [Thi 89] Thiéry O., David A., Créhange M.
**Images and Artificial Intelligence in Computer Aided Education : Our
prototype BIRDS.**
International Conference on Technology and Education, Orlando, USA, Mars 89.
- [Vic 87] Victory Alissa, Hellen Brooks, Bruce Robinson, Victory Brian.
A reference and referral system using expert systems techniques.
Journal of documentation, vol 43, n° 1, Mars 1987, pp 1-23.
- [v Kev 87] van Kevel S.J.
Laservision ROM, Systems for Digital Data & Video Storage.
In Proceedings of OPTICA 87, Amsterdam april 1987.
- [v Rij 77] van Rijsbergen
A theoretical Basis for the use of co-occurrence in retrieval.
Journal of Documentation. Vol 33, 1977 pp106-119
- [v Rij 79] van Rijsbergen C.J.
Information Retrieval,
Second Edition, Butterworths, London 1979.
- [v Sly 79] van Slype G.
Conception et gestion des systèmes documentaires.
Les éditions d'organisations 79.
- [v Sly 86] van Slype G., Maniez J.
Les langages documentaires.
Les éditions d'organisations 1986.

- [Wal 79] Walter W.G. & Kraft D.h.
A mathematical Model of a weighted Boolean Retrieval System.
Information Processing and Management 12:3 vol 15 1979 PP 239-245.
- [Won 86a] Wong SKM. & Ziarko W.
A machine learning approach to information Retrieval.
ACM Pix 86 pp 228-233.
- [Won 86b] Wong SKM. & Ziarko W.
**An adaptive decision support system based on the probabilistic
approximate classification.**
Proceeding of Avignon 87. Les systèmes experts 87, pp 713-726.
- [Zah 65] Zadeh L. A.
Fuzzy sets
Information and control, vol 8, 1965, pp 338-353.

LES ANNEXES

Annexe 1 : Bordereau de saisie et notice d'indexation de l'application pilote.

Annexe 2 : Les principales fonctions de l'Imageur Documentaire de la SEP.

Annexe 3 : Le détail des recherches ayant servi à l'évaluation de RIVAGE.

Annexe 4 : Rapport d'activité du projet ESPRIT (Volume séparé).

Annexe 1 : Bordereau de saisie et notice d'indexation de l'application pilote.

Direction
du Patrimoine

Archives Photographiques

29/12/1986

BORDEREAU DE SAISIE

COTE

Zone obligatoire
Valeur unique
Champ inversé

Transcrire le numéro d'inventaire ou la cote, lorsque celle-ci est différente du numéro d'inventaire (fonds Gilletta, de Brébisson, Nadar ...)

TITRE ET OU LEGENDE

Zone facultative (mais recommandée)

Le titre est celui donné par le photographe : quand il existe, le transcrire entre guillemets et le faire suivre éventuellement de la légende. La légende est une notation du contenu : elle peut être transcrite à partir de la légende figurant sur la plaque, de la description inscrite sur l'inventaire ou elle peut être créée par le documentaliste : dans ce cas, on évitera qu'elle soit redondante avec les autres champs, dans la mesure du possible. On pourra y rendre compte de notions difficiles à traduire en descripteurs (position d'un objet, relation entre personnes, description d'une action ...)

AUTEUR DE LA PHOTOGRAPHIE

Zone facultative
Valeurs multiples
Champ inversé

Nom suivi du prénom entre parenthèse de l'auteur de la prise de vue. Lorsque la mention d'auteur ne peut être attribuée de façon certaine à l'un ou l'autre photographe d'un même atelier ou studio, le nom du photographe est affixé du terme ATELIER. Exemples : BRAUN (ATELIER) / GILLETTA (ATELIER) / NADAR (ATELIER) / NEURDEIN (ATELIER) / SEEBERGER (ATELIER)

Si le prénom d'un photographe n'est pas connu, seul le nom sera mentionné : exemple : DOMINI. Si seule l'initiale du prénom est connue, elle sera indiquée : exemple : MIEUSEMENT (E.)

Quand l'auteur a un pseudonyme, indiquer le pseudonyme suivi du nom véritable et du prénom : exemple : NADAR (FELIX TOURNACHON PSEUD. FELIX)

Pour l'orthographe des noms, se reporter à la liste des photographes des fonds des Archives Photographiques.

S'il y a plusieurs auteurs, le séparateur utilisé est le point-virgule.

OBJET DE LA PHOTOGRAPHIE, DESCRIPTEURS PRINCIPAUX

Zone obligatoire
Valeurs multiples
Champ inversé

Représentation du sujet principal de la photographie, par, en moyenne, 4 descripteurs figurant dans le thésaurus. Séparateur utilisé : le point-virgule.

.../...

OBJET DE LA PHOTOGRAPHIE, DESCRIPTEURS SECONDAIRES

Zone facultative
Valeurs multiples
Champ inversé

Représentation des autres sujets par 0 à 10 descripteurs figurant dans le thésaurus : séparateur : le point-virgule.

OBJET DE LA PHOTOGRAPHIE, MORPHOLOGIE

Zone facultative (mais recommandée)
Valeurs multiples
Champ inversé

Descripteurs figurant dans la liste "analyse morphologique de l'image" (micro-thésaurus : caractéristique générales de l'image) ; séparateur : le point-virgule.

OBJET DE LA PHOTOGRAPHIE, CONNOTATION

Zone facultative
Valeurs multiples
Champ inversé

Descripteurs figurant dans le micro-thésaurus "connotation". Pour l'indexation des premiers bordereaux, on considérera ce micro-thésaurus comme une liste ouverte.

Le séparateur est le point-virgule

PRECISION SUR L'OBJET DE LA PHOTOGRAPHIE

Zone facultative
Valeurs multiples
Champ inversé

Informations présentées sous la forme de noms propres ou de descripteurs qui entreront dans les listes annexes II (listes ouvertes) du thésaurus.

les noms des édifices, des parcs ... Lorsque l'objet de la photographie sera une oeuvre d'art conservée dans un musée, un autre édifice ou un site on ne fera pas mention du nom du musée, de l'édifice ou du site dans cette zone sauf une salle ou une autre partie de cet édifice ou une partie du site était représentée.

les sujets géographiques (géographique physique : noms de fleuves, de montagnes... et les noms de lieux lorsqu'ils ne se confondent pas avec le lieu de prise de vues)

(par exemple une maquette de Lille, une allégorie...),

les noms de personnes physiques (pour les pseudonymes, observer les mêmes règles que celles indiquées pour l'auteur de la photographie) et morales, les noms des troupes de théâtre,

le titre des pièces de théâtre,

.../...

- Les thèmes iconographiques représentés : la Nativité ...
 - Le titre donné par l'auteur de l'oeuvre photographiée (exemple : "le radeau de la Méduse) ou par l'usage. Cette zone est à utiliser quand la photographie représente une peinture, une sculpture, un manuscrit, la création d'un artisan d'art ou d'un couturier.
 - Les noms de manifestations - (expositions, congrès, fêtes...).
- Le séparateur est le point-virgule.

COMMENTAIRE

Zone facultative

On pourra apporter des informations complémentaires sur le contenu de la photographie : utilisation actuelle d'un édifice, état d'un objet, détail d'une enseigne ...

AUTEUR DE L'OBJET DE LA PHOTOGRAPHIE

Zone facultative
Valeurs multiples
Champ inversé

Indiquer selon le cas, le nom du peintre, du sculpteur, de l'architecte, de l'artisan d'art, du couturier, de l'inventeur ... A distinguer de l'auteur de la photographie, c'est à dire du photographe.

Pour l'orthographe des noms de peintres, de sculpteurs, d'architectes ..., il est conseillé de se reporter aux listes des bases de données du Ministère ou aux ouvrages de référence (THIEME-BECKER, BENEZIT...)

Nom suivi du prénom entre parenthèse ; séparateur ; le point-virgule

DATATION DE L'OBJET DE LA PHOTOGRAPHIE

Zone facultative
Champ inversé

Date de réalisation de l'oeuvre photographiée ; formats possibles (XXXX = millésime en quatre chiffres) :

XXXX

avant XXXX
après XXXX
entre XXXX et XXXX
(non connu)
XXXX ?
vers XXXX

Remarque : pour désigner un siècle : entre XX00 et XX99 ; exemple : 19^{ème} siècle entre 1800 et 1899. Pour désigner une fourchette de siècles : entre XX00 et XX99, exemple 16^{ème} et 17^{ème} siècle : entre 1500 et 1699. Les dates avant J.-C. seront précédées de -

.../...

LIEU DE PRISE DE VUE

Zone facultative
Champ inversé

Dans le cas d'un pays étranger, pays, éventuellement nom de la ville, ou nom de la région si on ne connaît pas la ville.
Pour la France, numéro du code postal : exemple 75004 pour le quatrième arrondissement de Paris.

PRECISION SUR LE LIEU DE PRISE DE VUE

Zone facultative
Champ inversé

Mettre dans l'ordre les groupes de données ci-dessous :

-lieu-dit
-site ou édifice (exemple : Jardin des Tuileries
-adresse : rue, numéro dans la rue
-appellation historique : ancienne dénomination du lieu où le sujet est photographié : ancien nom d'une rue débaptisée

Séparateur : le point-virgule

DENOMINATION DE L'EDIFICE DE CONSERVATION DE L'OBJET DE LA PHOTOGRAPHIE

Zone facultative

Pour l'intitulé d'un musée, on se reportera au Répertoire des Musées de France de G. Barnaud, ou à l'"International Directory of Art" (pour les musées étrangers)

Zone à utiliser quand le lieu de conservation est distinct du lieu de prise de vue. Cependant, quand il n'y a pas de certitude sur le lieu de prise de vue d'un objet d'art (photo prise à l'occasion d'une restauration ...), il est préférable de remplir cette zone et de ne pas utiliser la zone "lieu de prise de vue"

LIEU DE CONSERVATION DE L'OBJET DE LA PHOTOGRAPHIE

Zone facultative

Même format que pour la zone lieu de prise de vue.

Zone à utiliser lorsque le lieu de conservation est distinct du lieu de prise de vue

PRECISION SUR LE LIEU DE CONSERVATION DE L'OBJET DE LA PHOTOGRAPHIE

Zone facultative

Zone à utiliser lorsque le lieu de conservation est distinct du lieu de prise de vue. Indiquer éventuellement le département du musée où l'objet est conservé...

.../...

DATATION DE LA PRISE DE VUE

Zone facultative
Champ inversé

Format :

Si le jour est connu : XXXX,MM, JJ

Si seul le mois est connu : XXXX, MM

Dans ces deux cas le point d'interrogation peut être utilisé

Si le jour ou le mois ne sont pas connus, mêmes formats que pour la datation de l'objet de la photographie.

APPARTENANCE DU CLICHE A UN ENSEMBLE

Zone facultative

Indiquer si le cliché a été pris lors d'une mission ou lors d'un reportage. Préciser éventuellement le nom de la mission (exemple : mission héliographique). Indiquer éventuellement le nom du commanditaire

PROPRIETE DU CLICHE

Zone obligatoire

Celle-ci peut être publique (Etat, établissement public, collectivité territoriale) ou privée (personne morale, personne physique)

Format pour les clichés des Archives photo :
ETAT ; CULTURE ; ARCHIVES PHOTOGRAPHIQUES

Format pour les clichés de la Caisse :
ETAT ; CULTURE ; CAISSE NATIONALE DES MONUMENTS HISTORIQUES ET DES SITES

DATE D'ENTREE DANS LES COLLECTIONS

Zone facultative
Champ inversé

Même format que pour l'objet de la photographie

MODE D'ENTREE DANS LES COLLECTIONS

Zone facultative

Le mode d'acquisition peut être : un achat, un achat en vente publique, un achat par préemption, un achat par préemption en vente publique, un achat sous réserve d'usufruit, une commande, une dation, un dépôt, une donation, une donation sous réserve d'usufruit, un don, un legs, un legs sous réserve d'usufruit, une mission exécutée par les photographes du service... Dans le cas d'un achat, d'une donation ou d'un legs sous réserve d'usufruit, on précisera, dans la zone "date d'entrée dans les collections", la date d'acquisition et la date d'entrée matérielle.

DIFFUSION

Zone obligatoire

Organisme auquel le public doit s'adresser pour obtenir un tirage :
CNMHS, 4, RUE DE TURENNE 75004

MENTION OBLIGATOIRE

Zone obligatoire

Mention que l'utilisateur s'engage à faire figurer sur toute représentation ou reproduction du document photographique faite par lui :
Selon les cas indiquer :

c ARCH. PHOT. PARIS / SPADEM
ATGET / cARCH. PHOT. PARIS SPADEM
CHOUMOFF / cARCH. PHOT. PARIS SPADEM
DOUCET / cARCH.PHOT. PARIS SPADEM
GILLETTA / cARCH.PHOT. PARIS SPADEM
LEFEVRE-PONTALIS / cARCH. PHOT. PARIS SPADEM
MARTIN-SABON / cARCH. PHOT. PARIS SPADEM
MOREAU / cARCH.PHOT. PARIS SPADEM
MOREAU-NELATON / cARCH.PHOT. PARIS SPADEM
NADAR / cARCH. PHOT. PARIS SPADEM
FELIX NADAR / cARCH. PHOT. PARIS SPADEM
PAUL NADAR / cARCH. PHOT. PARIS SPADEM
SEEGERGER / cARCH. PHOT. PARIS SPADEM

...

LIEUX DE CONSERVATION DES EPREUVES ANCIENNES

Zone facultative

Selon les cas, indiquer: Bibliothèque nationale, Bibliothèque historique de la ville de Paris, Bibliothèque du patrimoine, Musée Carnavalet, Museum of modern art (New-York) ...

BIBLIOGRAPHIE

Zone facultative

Références d'ouvrages, de catalogues, de colloques, d'articles se rapportant à la photographie : exemple : SZARKOWSKI (John) MORRIS HAMBURG (Maria).- The work of Atget .- London : Gordon Fraser, 1983.- 4 vol.

HISTORIQUE DU CLICHE

Zone facultative

Préciser si le cliché appartient à un fonds (lorsque le nom du fonds ne se confond pas avec celui du photographe), indiquer le nom du client, du dernier propriétaire... Préciser si le cliché appartenant aux collections des archives photographiques a fait l'objet d'un dépôt dans une autre institution (calotypes au musée d'Orsay)

NUMERO ATTRIBUE PAR LE PHOTOGRAPHE

Zone facultative
Champ inversé

Exemple : numéro ATGET : 262

NUMERO VIDEOSISQUE

Zone obligatoire
Champ inversé

Exemple : 334

Direction
du Patrimoine

Archives Photographiques

29/12/86

INDEXATION DES 1000 PHOTOGRAPHIES
VIDEODISQUE ATGET

"ANALYSE DES IMAGES OU DELIRES ET CAUCHEMARS DES ICONOGRAPHES" (1)

La "révolution iconique" (2), l'apparition du vidéodisque, la naissance des banques d'images ont suscité un intérêt et un regard neufs sur la documentation audiovisuelle. Plusieurs travaux sur le traitement documentaire de l'image ont fait l'objet d'une publication récente (3).

Les pratiques des institutions sont fort diverses : les bibliothèques, les services d'archives, les musées, les agences de presse n'ont pas la même approche de l'image. Choisira-t-on d'analyser de façon approfondie et détaillée le document iconographique à l'aide d'un vocabulaire iconographique et d'une syntaxe ou l'indexera-t-on par quelques mots-clés extraits de langages documentaires déjà utilisés pour le traitement d'une documentation textuelle ? Effectuera-t-on un dépouillement image par image ou cataloguera-t-on des lots plus ou moins importants d'images ? L'utilisation de technologies permettant la visualisation interactive des images modifiera-t-elle les règles de l'analyse

L'analyse d'une centaine de photographies effectuée en 1986 par le Service des Archives Photographiques en collaboration avec le Bureau Marcel Van Dijk, l'Université de Nancy et l'Université de Dijon a permis de dégager un certain nombre de règles que l'équipe des documentalistes tentera d'appliquer à l'indexation des 1000 images du vidéodisque ATGET dans le cadre du projet ESPRIT (4).

(1) Titre d'un article de P. Richard paru dans ethnologie française XIII, 3, 1983

(2) L'expression est de Serge Moscovici

(3) Citons entre autres :

-Analyse de l'image fixe : réflexions et guide bibliographique ; - Paris : La Documentation française, 1981 . - 167p. - (Interphotothèque 41)

-Le Traitement documentaire de l'image fixe . - Paris : Centre Pompidou, Bibliothèque Publique d'Information, 1985 . - (Dossier technique ; 3)

-HUDRISIER (Henri) . - L'Iconothèque : documentation audiovisuelle et banque d'images . - Paris : La Documentation française, 1982

-GARNIER (François) . - Thésaurus iconographique : système descriptif de représentations . - Paris : Le Léopard d'or, 1984 . - 240p.

(4) Les indications qui suivent portent sur trois zones bordereau : Objet de la photographie. Descripteurs principaux. Objet de la photographie. Descripteurs secondaires. Objet de la photographie. Morphologie de l'imagé. Elles complètent les prescriptions qui figurent déjà dans le mode d'emploi du bordereau. Il est possible que ces deux

LES ARCHIVES PHOTOGRAPHIQUES

Rappelons que les fonds conservés au Service des Archives Photographiques sont fortement spécialisés dans l'architecture et les beaux-arts, mais que d'autres genres de clichés sont présents : paysage, portrait, spectacles, expositions, événements divers ... Ces collections sont utilisées par des "spécialistes" (inspecteurs et architectes des monuments historiques, chercheurs, professeurs, étudiants), par le secteur de l'édition, par des collectivités diverses qui sont amenées à monter des expositions et par une petite minorité d'amateurs de la photographie.

Le système documentaire des Archives Photographiques doit être en mesure de répondre aux besoins de toutes les catégories d'utilisateurs et aux recherches les plus diverses.

Le contenu de l'image doit faire l'objet de toutes les approches possibles : technique, historique, sociologique ... On trouvera donc dans le thésaurus des Archives Photographiques des termes appartenant à des vocabulaires spécialisés (de l'architecture, des beaux-arts ...) aussi bien que des termes se rapportant à l'univers de la vie quotidienne.

LES OUTILS

On aura recours aux :

-documents permettant d'identifier la photographie : légende(s), catalogues, ouvrages, guides, dictionnaires, correspondance ...

-documents techniques : par exemple les "vocabulaires" de l'Inventaire sont d'une aide précieuse pour l'identification des formes architecturales et des objets.

-notes du thésaurus iconographique de F. Garnier pour ce qui concerne l'iconographie religieuse (p. 122 et suivantes)

ANALYSE DE L'OBJET PHOTOGRAPHIE

Identification et sélection
des éléments présents sur la photographie

On peut indiquer, parmi les critères qui permettent d'effectuer la sélection des "items" à retenir :

- La lisibilité de l'élément
- La présence d'un descripteur ou de son générique dans le thésaurus correspondant à l'élément identifié
- Le nombre de descripteurs autorisés pour l'indexation d'un document : environ une quinzaine en tout pour les zones DE1 et DE2.

Certains éléments secondaires pourront être retenus dans l'analyse : dans une photographie d'architecture, on notera la présence de personnages ; dans une scène de rue, on pourra indiquer l'existence d'affiches... Cela ne doit pas empêcher qu'une sélection s'opère et que des éléments très secondaires, peu lisibles ne soient pas "traduits" en descripteurs sauf s'il s'agit de représentations très rares ou jugées importantes sur le plan historique ou dans l'oeuvre d'un photographe. L'identification des différents éléments d'une image ne doit pas, bien entendu, faire perdre de vue la signification principale de la photographie

Le choix des descripteurs

De façon générale, c'est le terme le plus spécifique qui doit être choisi comme descripteur :

exemples :

-pour une façade de boutique, on utilisera le descripteur DEVANTURE et non le descripteur BOUTIQUE qui est son générique.
-on cherchera à préciser le type d'un chapiteau : CHAPITEAU CORINTHIEN

Dans certains cas précis, cependant, on pourra avoir recours à un descripteur et à son générique :

exemple :

-des roues de voitures à bras apparaissent en évidence sur la photographie ; d'autres pièces de ces véhicules sont présentes : on utilisera ROUE et ACCESSOIRE DE VEHICULE (qui est son générique)

Lorsque l'on ne trouve pas, dans le thésaurus, de descripteur correspondant à l'item retenu, on sélectionnera un descripteur qui sera son générique ou un terme proche. Si l'absence d'un descripteur pose problème, ce terme sera noté comme candidat descripteur et pourra être intégré au vocabulaire. Au cours de cette phase test, le vocabulaire pourra faire l'objet de modifications

Certains descripteurs ont des significations différentes selon qu'ils sont au singulier ou au pluriel :

exemples :

CISEAU(X), BUREAU(X), LUNETTE(S) ...

L'article permet de distinguer entre deux descripteurs dans certains cas :

exemples :

L'EGLISE (allégorie) et EGLISE (édifice)

LE CONSULAT (époque) et CONSULAT (batiment)

De façon générale, avant d'utiliser un descripteur, on vérifiera son sens (par son environnement sémantique et éventuellement par sa note d'application) (1)

L'image ne se réduit pas à un assemblage d'éléments : dans la mesure où le thésaurus le permet, on utilisera, outre les descripteurs se rapportant aux différents éléments lisibles sur l'image, un (des) descripteur(s) recouvrant une notion plus globale ; ce descripteur devra bien entendu, être distinct du ou des génériques des descripteurs déjà utilisés et éventuellement des termes qui lui sont associés :

exemples :

-pour la calèche du Tzar en voyage à Paris, on utilisera RELATION INTERNATIONALE

-pour les poseurs de rail, on pourra utiliser TRAVAUX PUBLICS

-MODE pourra être employé pour des vêtements ou des accessoires d'habillement très lisibles et intéressants.

On notera, la plupart du temps en descripteur secondaire, l'ensemble auquel appartient l'objet représenté :

exemples :

-le terme CHATEAU pour le salon de musique de Fontainebleau

-TRESOR RELIGIEUX pour la nef de Sainte Ursule

(1) Il est possible d'utiliser un descripteur même si l'on a des doutes sur la représentation, à condition de faire mention

-VILLAGE pour la photographie de Domini représentant des maisons au bord d'une rivière

Si, sur la photographie d'un objet mobilier, l'environnement (salle, partie de bâtiment ...) n'est pas visible sur la photographie, il ne faudra pas indiquer cet environnement par un (des) descripteur(s).

Mémento

On caractérisera, lorsque cela s'imposera, le genre de d'image (thésaurus AP, p. 142) (1)

Quelques définitions :

-paysage : représentation de la nature où les figures d'hommes ou d'animaux et les constructions ne sont que des accessoires. (Petit Robert)

-paysage urbain : vue générale d'une ville ou d'une partie de ville.

-petit métier : descripteur qu'on pourra utiliser en complément de termes tels que : ramoneur, vitrier, remouleur, rempailleur, marchand de jounaux, marchand de quatre saisons, marchand d'oublies, marchand de marrons, fleuriste, marchand ambulancier, marchand de souvenirs, marchand forain, porteur, d'eau ...

-portrait à la manière de : portrait imitant une peinture, un dessin, une caricature ... On indiquera, dans la zone commentaire, l'auteur de l'oeuvre qui a servi de modèle

-portrait charge : portrait caricatural dans lequel l'artiste a outré certains traits du personnage ou l'a gratifié d'attributs qui enrichissent l'image de significations, souvent satiriques, qui concernent sa nature, ses qualités, ses défauts, ses fonctions ou ses activités (thésaurus iconographique Garnier)

-photographie événementielle : il peut s'agir d'un événement historique, d'un fait d'actualité d'époque, d'une catastrophe naturelle (inondation ...)

-scène de rue : scène d'extérieur où les êtres humains jouent un certain rôle (soit par leur nombre, soit par l'action qu'il effectuent, soit par leur présence au premier plan ...)

Deux de ces descripteurs peuvent être utilisés pour un même document, par exemple SCENE DE RUE et PETIT METIER

On fera mention de la saison, dans la mesure du possible.

On fera mention de l'époque et du style dans la zone descripteurs secondaires (thésaurus AP, pp.135-138)

(1) La pagination du thésaurus AP est celle de l'édition par champs sémantiques, version juillet 1986

-la photographie d'architecture

on n'utilisera pas les descripteurs AVANT , PENDANT, APRES , CLASSEMENT, INSCRIPTION sauf si cette information est mentionnée dans l'inventaire.

On pourra noter l'état de l'édifice (thésaurus AP, p.140-141 + RUINE DEMOLITION, REMPLOI)

On indiquera le matériau (thésaurus AP p. 146 .(1))

La photographie d'oeuvre d'art

Pourront figurer dans l'indexation l'édifice, partie d'édifice, le meuble, l'objet support du décor (micro-thésauri architecture, ameublement, listes mobilier, mobilier religieux ...), la catégorie technique ou la technique (micro-thésauri art graphique, sculpture, et liste technique), le matériau (liste matériau) et l'iconographie : pour l'iconographie religieuse, on aura recours à la classe SUJET BIBLIQUE du thésaurus iconographique Garnier. On notera également l'état de l'objet. (thésaurus p.140-141)

Le portrait

On utilisera les descripteurs HOMME, FEMME, ENFANT. On mentionnera éventuellement la profession. La position sera indiquée dans la zone descripteurs secondaires (thésaurus AP p.139)

LA MORPHOLOGIE

Les descripteurs VUE D'EXTERIEUR et VUE D'INTERIEUR ne seront pas utilisés pour les paysages, les scènes de rue, les scènes d'intérieur la photographie en atelier, et, dans certains cas, pour le portrait. Dans les autres cas, il faudra les mentionner.

En ce qui concerne les définitions des plans, on se reportera au Larousse photographie, pp.115-116

On ne mentionnera pas les plans pour les photographies d'objets isolés.

Pour les photographies en couleur, on utilisera le descripteur "photographie couleur".

(1) Le plus souvent dans la zone Objet de la photographie. Descripteurs secondaires

**Annexe 2 : Les principales fonctions de l'Imageur
Documentaire de la SEP.**



REF DTI/DTE/43002/86

ID	TR	31210	01
----	----	-------	----

EDITION 0

DATE 20/10/86

IMAGEUR**DOCUMENTAIRE**

MANUEL D'UTILISATION
- IMAGEUR DOCUMENTAIRE -

ORIGINE : S.E.P.SIGNATURE **COMPOSE [F 6]**

Cet ordre visualise dans la ZONE désignée la ou les premières images de la liste dont on doit fournir le NOM.

Les images de cette liste ne sont pas protégées ; on va pouvoir composer la liste, en y insérant, en y déplaçant ou en y retranchant des images par COPIE, DEPLACE ou DETRUIRE.

La couleur de la zone en ETAL est : gris

> COMPOSE zone liste (? Nom de la liste)

Causes d'erreur : . l'espace visuel est totalement occupé,
 . la touche frappée ne désigne pas une ZONE,
 . la touche désigne la CORBEILLE,
 . la touche désigne une ZONE déjà occupée,
 . le NOM donné ne correspond pas à une liste existante,
 . le NOM donné désigne une liste déjà visualisée.

SOURCE [F 9]

Cet ordre visualise dans la ZONE désignée la ou les premières images de la liste dont on doit fournir le nom.

La structure de cette liste est protégée contre toute modification;
 Les ordres DEPLACE ou DETRUIRE y sont interdits.

On peut seulement dupliquer ces images par l'ordre COPIE.
 Cette protection se traduit par un fond d'image bleu.

>SOURCE zone LISTE (? NOM DE LA LISTE)

Causes d'erreur : . l'espace visuel est totalement occupé,
 . la touche frappée ne désigne pas une ZONE
 . la touche désigne la CORBEILLE
 . la touche désigne une ZONE déjà occupée
 . le NOM donné ne correspond pas à une liste existante,
 . le NOM donné désigne une liste déjà visualisée.

SEGMENT [F 7]

Cet ordre génère une liste composée d'images consécutives du vidéo-disque, et la visualise dans la ZONE désignée.

Cette nouvelle liste d'images n'a pas de NOM ; elle est visualisée, sans protection, dans la ZONE désignée,

>SEGMENT zone (?) image-début (?) image-fin (?) pas (?)

Image-début et image-fin sont les numéros des images extrêmes du SEGMENT, pas est l'intervalle entre les images.

Cause d'erreur : . l'espace visuel est totalement occupé.
 . la touche frappée ne désigne pas une ZONE,
 . la touche désigne la CORBEILLE,
 . la touche désigne une ZONE déjà occupée,
 . les bornes ou le pas sont extérieurs au domaine 1 à 54000.

ZOOM [DEL]

Cet ordre visualise une image de la MOSAIQUE sur l'écran VIDEO.

Il s'applique aussi à la CORBEILLE ; dans ce cas, c'est la dernière image placée dans la CORBEILLE qui est visualisée.

On peut ZOOMer successivement plusieurs images en les désignant tour à tour par la sélection sur le 'touch-pad' sans refrapper l'ordre ZOOM.

Pour supprimer l'image zoomée de la case VIDEO taper DZoom. (INSERT)

Cet ordre est inopérant si on désigne la case VIDEO.

>ZOOM case (?)

Cause d'erreur : .la touche frappée ne désigne ni une case de la MOSAIQUE,
 ni la CORBEILLE.
 .la case désignée est vide.

DEZOOM [INSERT]

Cet ordre supprime l'image ZOOMée de la case VIDEO.

Il restaure l'image de la case VIDEO si la case VIDEO était occupée.

Cet ordre est inopérant s'il n'y avait pas d'image ZOOMée.

>DEZOOM

? AVANCE [--->]

Cet ordre permet de se déplacer en avant de la liste dans la ZONE désignée, et du nombre d'images demandées (1, 4, 16, 64, jusqu'au bout ou par page initialement définie).

Pour AVANCER à nouveau sur la même ZONE, il suffit de frapper un nouveau pas.

Pour AVANCER sur une nouvelle ZONE, il suffit de désigner une autre ZONE, puis un pas.

>AVANCE zone(?) pas(?)

Causes d'erreur : . aucune ZONE n'a été désignée
 . la touche désigne la CORBEILLE
 . la touche désigne une ZONE vide,
 . aucune ZONE n'a encore été désignée,
 . la LISTE est positionnée à sa fin.

? RECULE [<----]

Cet ordre permet de se déplacer en arrière de la liste dans la ZONE désignée, et du nombre d'images demandées (1, 4, 16, 64, jusqu'au bout ou par page ...).

Pour RECULER à nouveau sur la même ZONE, il suffit de frapper un nouveau pas.

Pour RECULER sur une nouvelle ZONE, il suffit de désigner une autre ZONE, puis un pas.

>RECOULE zone(?) pas(?)

Causes d'erreur : . la touche frappée ne désigne pas une ZONE,
 . la touche désigne la CORBEILLE
 . la touche désigne une ZONE vide,
 . aucune ZONE n'a encore été désignée,
 . la LISTE est positionnée à son début

PAS1 (1) 'touch-pad'

Déplacement image par image

PAS4 (4) 'touch-pad'

Déplacement quatre images par quatre images

PAS16 (16) 'touch-pad'

Déplacement seize images par seize images

PAS64 (64) 'touch-pad'

Déplacement soixante-quatre images par soixante-quatre images

PAGE (PAGE) 'touch-pad'

Déplacement page par page

La page correspond à la taille de la ZONE désignée :

- 1 image pour une ZONE en PILE,
- 4, 8, 12 ou 16 images par les ZONES en ETAL.

INFINI [BUTEE] 'Touch-pad'

Positionnement en extrémité de LISTE :

- en début de LISTE pour RECOULE,
- en fin de LISTE pour AVANCE.

COPIE [IMAGE]

Cette ordre insère une copie de l'image désignée dans la case de destination,

> COPIE image(?) destination()

Causes d'erreur : la touche frappée ne désigne pas une image,
 la touche frappée désigne une CASE vide,
 l'image désignée est protégée,
 la destination est une ZONE en SOURCE,
 la destination désigne la VIDEO alors qu'une image est zoomée.
 la destination désigne la CORBEILLE.

DEPLACE IMAGE [F1 ou SHIFT F1]

Cette ordre déplace l'image désignée dans la case de destination.

> DEPLACE image(?) destination(?)

Causes d'erreur : la touche frappée ne désigne pas une image,
la touche frappée désigne une CASE vide,
l'image désignée appartient à une ZONE en SOURCE,
l'image désignée est protégée,
la destination est une ZONE en SOURCE,
la destination désigne la VIDEO et un ZOOM est en cours,
la destination désigne la CORBEILLE.

ECHANGE IMAGE [SHIFT F3]

Cet ordre permet d'échanger deux images appartenant à des ZONES en COMPOSITION.

En particulier, lorsqu'appliquée à deux images d'une même ZONE en COMPOSITION, cet ordre permet de réordonner une LISTE.

> ECHANGE image1(?) image2(?)

Causes d'erreur : la touche frappée ne désigne pas une image,
la touche frappée désigne une CASE vide,
l'image désignée appartient à une ZONE en SOURCE,
l'image désignée est protégée,
la destination désigne la VIDEO et un ZOOM est en cours.
la CASE destination est au delà de la fin de LISTE,
la destination désigne la CORBEILLE.

ECHANGE ZONE [CTRL F3]

Cet ordre permet d'échanger entre elles deux ZONES d'affichage ainsi que les LISTES qu'elles contiennent. Cet ordre diffère d'un ordre DEPLACE LISTE qui constituerait à échanger le contenu de deux LISTES mais pas leurs attributs (nqm, mode etc.)

> ECHANGE ZONE zone1 ? zone2 ?

Causes d'erreur : la touche frappée ne désigne pas une image,
la zone désigne la VIDEO et un ZOOM est en cours,
la zone désigne la CORBEILLE.

DETRUIT [SHIFT F5]

Cet ordre supprime l'image désignée. Cette image prend place dans la CORBEILLE, par dessus les autres.

> DETRUIT image (?)

Causes d'erreur : la touche frappée ne désigne pas une image,
la touche frappée désigne une CASE vide,
la touche frappée désigne la CORBEILLE,
l'image désignée appartient à une zone en source,
ou en source,
l'image désignée est protégée.

COPIE LISTE [CTRL F2]

Cete ordre insère une copie de l'ensemble des images de la LISTE désignée devant la case de destination.

> COPIE LISTE source destination ?

Causes d'erreur : la touche frappée ne désigne pas une image,
la touche frappée désigne une CASE vide,
l'image désignée est protégée,
la destination est une ZONE en SOURCE,
la destination désigne la VIDEO et un ZOOM est en cours,
la CASE destination est au delà de la fin de LISTE,
la destination désigne la CORBEILLE,
la source et la destination désignent la même LISTE.

MASQUE [F 4]

Cet ordre protège une image de toutes actions sauf celle de COPIE
Il permet de sélectionner également une image avant un envoi vers le logiciel d'exploitation de la banque de données textuelles.

> MASQUE case (?)

DEMASQUE [F 5]

Cet ordre annule l'ordre précédent.

> DEMASQUE case (?)

DEPLACE [CTRL F1]

Cet ordre déplace une liste visualisée d'une zone pré-définie vers une autre zone (PILE ou ETAL)

> DEPLACE zone (?) devant case (?)

TEXTE [F 2]

Cet ordre permet de visualiser sur l'écran vidéo le document Micro-Questel associé à l'image choisie.

> DOCUMENT case (?)

LEGENDE [F 3]

Cette fonction permet de visualiser la légende (2 lignes de texte) en bas de l'écran vidéo.

> LEGENDE case (?)

SEPARÉ [F 6]

Cette fonction permet de séparer à l'intérieur d'une zone des images sélectionnées (masquées) avec des cours sélectionnés.

REPERTOIRE [F 8]

Cette fonction visualise sur l'écran vidéo les listes sauvegardées antérieurement.

Annexe 3 : Le détail des recherches ayant servi à l'évaluation
de RIVAGE.

Plan des requêtes

1. Plan très large d'une vue du ciel et d'un fleuve avec des monuments et le ciel brumeux.
2. Images sur lesquelles on voit des animaux dans un parc ou dans un lieu avec de la végétation.
3. Images d'une grande rue avec beaucoup de monde avec un temps peu clair.
4. Tapisseries en couleur représentant des êtres humains.
5. Images de ruelles sombres et tristes.
6. Façades de maisons en plan rapproché avec des personnes devant.
7. Edifice religieux en extérieur dans un milieu rural.
8. Images de restaurants, auberges, cafés ... en ville.

1. Plan très large d'une vue du ciel et d'un fleuve avec des monuments et le ciel brumeux.

17 images dans la base sont pertinentes (73, 651, 527, 24, 25, 153, 155, 601, 159, 228, 229, 38, 373, 181, 630, 631, 533)

Formulation booléenne :

Etape 1:

[1] contenu = (FLEUVE_D OU RIVIERE)

ET

[2] contenu = (MONUMENT OU ARC DE TRIOMPHE OU OBELISQUE OU CENOTAPHE OU TROPHEE OU STELE OU PYRAMIDE OU EDIFICE OU CHATEAU OU PALAIS OU MANOIR OU HOTEL PARTICULIER OU BAPTISTERE OU COUVENT OU ABBAYE OU PAGODE OU TEMPLE OU CHAPELLE ABASIALE OU CHAPELLE CASTRALE OU CHAPELLE RAYONNANTE OU CHAPELLE SYNODALE OU CATHEDRALE OU BASILIQUE OU AQUEDUC OU PONT_D OU VIDAUC)

ET

[3] contenu = (BROUILLARD OU NUAGE)

OU

morphologie = (SOMBRE OU GRISAILLE)

[4] morphologie = (VUE PERSPECTIVE OU PLONGEE OU PLAN D'ENSEMBLE)

-> 2 images pertinentes(228, 651)

Etape 2 :

L'élément 4 est enlevé -> même résultat.

Etape 3 :

L'élément 3 est enlevé -> 19 images :

- 11 pertinentes (159, 228, 38, 229, 73, 651, 527, 533, 630, 24, 155),

- 8 mauvaises (227, 186, 165, 464, 565, 250, 507, 95).

Etape 4 :

L'élément 3 et 4 sont enlevés -> 25 images :

Les 19 images de l'étape précédente plus :

- 3 images pertinentes (25, 601, 631),
- 3 mauvaises (589, 220, 246).

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = FLEUVE, MONUMENT.

-> 36 images :

- 16 images choisies ou pertinentes (651, 527, 533, 24, 25, 153, 155, 601, 159, 228, 229, 38, 373, 630, 631, 181)
- 18 images rejetées (322, 323, 8, 529, 471, 223, 227, 490, 236, 301, 365, 366, 309, 310, 311, 250, 312, 565)
- 2 images neutres (464, 586)

Etape 1 :

Le système propose :

Contenu de la photographie = PONT_D

- > 21 nouvelles images en plus des images choisies dans l'étape précédente :
- 14 images rejetées (589, 12, 525, 336, 343, 220, 95, 481, 98, 165, 307, 246, 507, 63),
- 7 images neutres (138, 16, 23, 154, 283, 351, 494).

Etape 2 :

Le système propose :

Contenu de la photographie = FLEUVE

- > 9 nouvelles images en plus des images dans les étapes précédentes :
- 1 images choisies (73)
- 6 images rejetées (512, 82, 587, 272, 102, 40)
- 2 images neutres (315, 319).

RECHERCHE PRECISE :

Etape 0:

Contenu de la photographie = FLEUVE, MONUMENT

-> 3 images :

- 3 choisies (24, 25, 155).

Etape 1 :

Le système propose :

Contenu de la photographie = FLEUVE, PONT_D

-> 12 nouvelles images :

- 8 choisies (228, 229, 38, 533, 630, 631, 601, 159),
- 2 rejetées (227, 250),
- 2 neutres (586, 484).

Etape 2 :

Le système propose :

Contenu de la photographie = QUAI, ARCHE

-> 2 nouvelles images :

- 1 choisies (527),
- 1 neutres (283).

Etape 3 :

Le système propose :

Contenu de la photographie = FLEUVE

-> 17 nouvelles images :

- 4 choisies (651, 153, 181, 373),
 - 13 rejetées (322, 323, 8, 236, 490, 301, 365, 366, 309, 310, 311, 312, 565).
-

2. Images sur lesquelles on voit des animaux dans un parc ou dans un lieu avec de la végétation.

5 images sont pertinentes (210, 271, 146, 34, 240)

Formulation booléenne :

Etape 1 :

[1] contenu = (COMBAT D'ANIMAUX OU ANIMAL_D OU COLOMBE OU ESPECE ANIMAL_D OU CROCODILE OU REPTILE_D OU ECHASSIER_D OU COQ OU GALLINACE_D OU POULE OU OISEAU_D OU CYGNE OU AIGLE OU RAPACE_D OU BOVIN_D OU CHIEN OU BICHE OU CERF OU CERVIDE_D OU ANE OU CHEVAL OU LION OU MOUTON_D OU ELEPHANT PACHYDERME_D OU SINGE)

ET

[2] contenu = (PARC OU PARC NATUREL OU ALLEE OU CHEMIN OU JARDIN PUBLIC OU PARC OU ARBRE_D OU CONIFERE OU PALMIER OU HERBE OU VEGETATION_D OU CAMPAGNE_D OU ARBUSTE_D)

-> 17 images :

- 3 pertinentes (34, 271, 240)

Etape 2 :

L'élément 2 est enlevé -> 76 images :

- 5 choisies (210, 271, 146, 34, 240)

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = ANIMAL, VEGETATION

-> 272 images que l'utilisateur refuse de visualiser, le système propose alors une liste de termes à l'intérieur de laquelle l'utilisateur choisit les plus pertinents :

- ANIMAL_D, ESPECE ANIMAL, ARBRE, VEGETATION_D, ELEMENT VEGETAL.

-> 238 images, l'utilisateur refuse de les visualiser.

-> 16 images :

- 3 choisies (34, 271, 240),

- 11 rejetées (576, 609, 579, 387, 70, 550, 459, 147, 187, 636, 157),

- 2 neutres (489, 281).

Etape 1 :

Le système propose beaucoup de termes :

Contenu de la photographie = ELEVAGE, CLOCHER, CHAUSSEE,

L'utilisateur garde les termes suivants : PROFESSION LIEE AUX ANIMAUX, CANIN, OVIN, PROFESSION DE L'AGRICULTURE, ROCHER, CHEMIN, ELEVAGE.

-> 20 nouvelles images :

- 1 choisie (210)

- 20 rejetées (126, 367, 88, 440, 22, 405, 277, 20, 370, 16, 400, 47, 622, 72, 167, 39, 5, 163, 35)

Etape 2 :

Le système propose :

Contenu de la photographie = ARBRE_D

L'utilisateur ajoute ANIMAL.

-> 143 images non visualisées, l'utilisateur choisit dans une liste de termes : ANIMAL_D

-> 5 nouvelles images rejetées (103, 474, 205, 206, 127).

Etape 3 :

Le système propose :

Contenu de la photographie = CANIN, HOMME, ARBRE_D

-> trop d'images, l'utilisateur choisit CHIEN dans la liste de termes proposée.

-> 37 images :

- 1 choisies (146).

RECHERCHE PRECISE :

Etape 0:

Contenu de la photographie = ANIMAL, VEGETATION

-> 272 images non visualisées, l'utilisateur choisit dans une liste : ANIMAL_D, ESPECE ANIMALE, ARBRE, VEGETATION_D, ELEMENT VEGETAL.

-> 16 images :

- 3 choisies (34, 271, 240),

- 11 rejetées (576, 609, 579, 387, 70, 550, 459, 147, 187, 636, 157)

- 2 neutres (489, 281).

Etape 1 :

Le système propose :

Contenu de la photographie = VOIE, ESPECE ANIMAL

-> 9 nouvelles images rejetées (392, 649, 363, 622, 14, 405, 469, 375, 26)

Etape 2 :

Le système propose :

Contenu de la photographie = ARBRE_D

L'utilisateur ajoute ANIMAL,

-> 93 images non visualisées, le système propose une liste de termes, l'utilisateur décide d'arrêter sa recherche.

3. Images d'une grande rue avec beaucoup de monde avec un temps peu clair.

17 images sont pertinentes dans la base (1, 3, 32, 35, 295, 359, 425, 233, 4, 34, 222, 223, 219, 234, 14, 189, 299)

Formulation booléenne :

Etape 1:

[1] contenu = (RUE_D OU BOULEVARD OU AVENUE OU VILLE_D OU QUARTIER OU VILLAGE OU CARREFOUR OU PLACE_D OU ESPLANADE_D)

ET

[2] contenu = (ENFANT OU PERSONNE AGEE OU FOULE OU FEMME OU HOMME OU FAMILLE_D OU FILLE)

ET

[3] contenu = (PLUIE OU NUAGE OU BROUILLARD)

OU

morphologie = (SOMBRE)

-> 0 IMAGES

Etape 2 :

L'élément 3 est enlevé -> 2 images :

- 1 pertinente (234),

- 1 mauvaise (412).

Etape 3 :

Les éléments 2 et 3 sont enlevés -> 17 images :

- 5 pertinentes (219, 234, 14, 189, 235),

- 12 mauvaises (419, 422, 263, 565, 405, 472, 441, 218, 27, 412, 125, 29)

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = RUE, FOULE, ADULTE

-> 242 images non visualisées,

-> 7 images :

- 2 choisies (219, 189),

- 5 rejetées (419, 422, 218, 27, 125).

Etape 1 :

Le système propose beaucoup de termes l'utilisateur sélectionne le terme FOULE,

Contenu de la photographie = FOULE,

-> 11 nouvelles images :

- 3 choisies (234, 222, 223),

- 4 rejetées (43, 123, 180, 36),

- 4 neutres (33, 104, 190, 235).

Etape 2 :

Le système propose Contenu de la photographie = RUE_D, ARC DE TRIOMPHE, FACADE, OBSEQUES,

l'utilisateur garde seulement le terme RUE_D,

Contenu de la photographie = RUE_D,

-> 81 nouvelles images :

- 8 choisies (1, 14, 32, 295, 425, 219, 299, 359),

- 1 neutre (412)

- 68 rejetées (393, 394, 395, 524, 398, 15, 400, 656, 657, 277, 405, 154, 411, 28, 416, 421, 424, 427, 428, 303, 304, 431, 432, 435, 436, 437, 565, 441, 572, 318, 447, 450, 452, 453, 333, 461, 462, 463, 81, 465, 340, 341, 342, 468, 216, 469, 470, 600, 603, 604, 607, 608, 353, 354, 355, 356, 357, 358, 360, 361, 362, 484, 487, 234, 372, 629, 503, 505, 378).

Etape 3 :

Le système propose :

Contenu de la photographie = FOULE,

l'utilisateur décide de s'arrêter.

RECHERCHE PRECISE :**Etape 0:**

Contenu de la photographie = FOULE, RUE, ADULTE,

-> 7 images :

- 2 choisies (219, 189),

- 5 rejetées (419, 422, 218, 27, 125).

Etape 1 :

Le système propose beaucoup de termes l'utilisateur sélectionne le terme FOULE,

Contenu de la photographie = FOULE,

-> 11 nouvelles images :

- 3 choisies (234, 222, 223),

- 4 rejetées (43, 123, 180, 36),

- 4 neutres (33, 104, 190, 235).

Etape 2 :

Le système propose Contenu de la photographie = RUE_D, ARC DE TRIOMPHE, FACADE, OBSEQUES,

l'utilisateur garde seulement le terme RUE_D,

Contenu de la photographie = RUE_D,

-> 81 nouvelles images :

- 8 choisies (1, 14, 32, 295, 425, 219, 299, 359),

- 1 neutre (412)

- 68 rejetées (393, 394, 395, 524, 398, 15, 400, 656, 657, 277, 405, 154, 411, 28, 416, 421, 424, 427, 428, 303, 304, 431, 432, 435, 436, 437, 565, 441, 572, 318, 447, 450, 452, 453, 333, 461, 462, 463, 81, 465, 340, 341, 342, 468, 216, 469, 470, 600, 603, 604, 607, 608, 353, 354, 355, 356, 357, 358, 360, 361, 362, 484, 487, 234, 372, 629, 503, 505, 378).

Etape 3 :

Le système propose :

Contenu de la photographie = FOULE,

l'utilisateur décide de s'arrêter.

4. Tapisseries en couleur représentant des êtres humains.

11 images sont pertinentes dans la base (70, 128, 129, 538, 101, 634, 103, 127, 205, 206, 126)

Formulation booléenne :

Etape 1 :

[1] contenu = (TAPISserie_D OU TABLEAU_D OU PEINTURE_D OU PEINTURE MURALE_D OU FRESQUE OU PEINTURE A L'HUILE OU CADRE OU PEINTURE A L'EAU OU ESTAMPE OU BRODERIE)

ET

[2] contenu = (PERSONNAGE OU HOMME OU FEMME OU ENFANT)

ET

[3] morphologie = (PHOTOGRAPHIE COULEUR)

-> 3 images pertinentes (127, 205, 206).

Etape 2 :

L'élément 3 est enlevé -> 16 images :

- 7 pertinentes (128, 129, 101, 70, 205, 206, 127),

- 9 mauvaises (160, 255, 584, 170, 119, 124, 125, 258, 463).

Etape 3 :

Les éléments 2 et 3 sont enlevés -> 29 images :

- 11 pertinentes (70, 128, 129, 538, 101, 634, 103, 127, 205, 206, 126),

- 18 mauvaises.

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = TAPISserie, PEINTURE, PERSONNAGE

-> 108 images non visualisées, l'utilisateur choisit dans une liste : TABLEAU,

PEINTURE_D, TAPISserie_D

-> 14 images :

- 6 choisies (128, 129, 101, 70, 538, 634),

- 7 rejetées (160, 302, 562, 632, 57, 443, 639)

- 1 neutre (258).

Etape 1 :

Le système propose :

Contenu de la photographie = SAINT, FEMME,

l'utilisateur ajoute TABLEAU,

Contenu de la photographie = SAINT, FEMME, TABLEAU,

-> 70 images non visualisées, l'utilisateur choisit dans une liste : SAINT_D, TABLEAU_D,

-> 6 nouvelles images rejetées (546, 261, 102, 270, 496, 473)

Etape 2 :

Le système propose :

Contenu de la photographie = TABLEAU

-> 2 nouvelles images rejetées (436, 539).

Etape 3 :

Le système propose :

Contenu de la photographie = TABLEAU,

l'utilisateur ajoute PERSONNAGE,

Contenu de la photographie = TABLEAU, PERSONNAGE,

l'utilisateur choisit dans une liste : TABLEAU_D

-> 3 nouvelles images rejetées (125, 573, 584)

Etape 4 :

Le système propose :

Contenu de la photographie = TABLEAU,

-> 13 nouvelles images rejetées,

l'utilisateur décide de s'arrêter.

RECHERCHE PRECISE :

Etape 0:

Contenu de la photographie = TAPISserie, PERSONNAGE

-> 1 image choisie (70)

Etape 1 :

Le système propose beaucoup de termes, l'utilisateur choisit TAPISSERIE et PERSONNAGE,

Contenu de la photographie = TAPISSERIE, PERSONNAGE,

-> nouvelle image rejetée (61).

Etape 3 :

Le système propose beaucoup de termes, l'utilisateur choisit TAPISSERIE, SAINT, ANGE, CADAVRE,

Contenu de la photographie = TAPISSERIE, SAINT, ANGE, CADAVRE,

-> 2 nouvelles images rejetées (624, 539).

Etape 4 :

Le système propose beaucoup de termes, l'utilisateur choisit SAINT, PERSONNAGE et il ajoute TABLEAU

Contenu de la photographie = SAINT, PERSONNAGE, TABLEAU,

-> 1 nouvelle image rejetée (546),

l'utilisateur décide de s'arrêter.

5. Images de ruelles sombres et tristes.

25 images sont pertinentes dans la base (526, 275, 555, 303, 603, 604, 352, 354, 357, 622, 368, 375, 415, 418, 427, 433, 465, 487, 503, 450, 453, 485, 421, 431)

Formulation booléenne :

Etape 1:

[1] contenu = (RUE_D OU RUELLE OU IMPASSE)

ET

[2] connotation = (DELABRE OU ANCIEN)

OU

contenu = (DETERIORATION OU DEMOLITION)

OU

morphologie = (SOMBRE)

-> 12 images :

- 10 pertinentes (352, 419, 453, 487, 427, 555, 465, 603, 604, 415)

- 2 mauvaises.

Etape 2 :

L'élément 2 est enlevé -> 97 images :

- 25 pertinentes

- 72 mauvaises.

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = RUE,

Connotation = TRISTE,

-> 17 images :

- 12 choisies (450, 418, 419, 357, 453, 485, 555, 526, 368, 465, 375, 503),

- 3 rejetées (428, 656, 378),

- 2 neutres (425, 405).

Etape 1 :

Le système propose :

Contenu de la photographie = RUELLE, IMMEUBLE,

Connotation = DELABRE, VETUSTE,

-> 2 nouvelles images :

- 1 choisie (415),

- 1 neutre (553).

Etape 2 :

Le système propose :

Contenu de la photographie = PAVE,

Connotation = TRISTE,

-> 4 nouvelles images rejetées (653, 332, 610, 513).

Etape 3 :

Le système propose :

Contenu de la photographie = RUELLE, RUE_D, PERSONNAGE

Connotation = DELABRE,

-> 1 nouvelle image neutre (549)

Etape 4 :

Le système propose :

Contenu de la photographie = PAVE,

Connotation = TRISTE,

-> 22 nouvelles images :

- 5 choisies (487, 427, 622, 431, 433),

- 9 rejetées (316, 321, 452, 414, 608, 34, 618, 633, 441),

- 8 neutres (391, 467, 470, 600, 424, 361, 430, 372).

RECHERCHE PRECISE :

Etape 0:

Contenu de la photographie = RUE,

Connotation = TRISTE,

-> 17 images :

- 12 choisies (450, 418, 419, 357, 453, 485, 555, 526, 368, 465, 375, 503),

- 3 rejetées (428, 656, 378),

- 2 neutres (425, 405).

Etape 1 :

Le système propose :

Contenu de la photographie = RUELLE, IMMEUBLE,

Connotation = DELABRE, VETUSTE,

-> 2 nouvelles images choisies (415, 352).

Etape 2 :

Le système propose :

Contenu de la photographie = PAVE,

Connotation = TRISTE,

-> 4 nouvelles images rejetées (653, 332, 610, 513).

Etape 3 :

Le système propose :

Contenu de la photographie = PERSONNAGE

Connotation = DELABRE,

-> 1 nouvelle image neutre (549)

Etape 4 :

Le système propose :

Contenu de la photographie = PAVE,

Connotation = TRISTE,

-> 22 nouvelles images :

- 5 choisies (487, 427, 622, 431, 433),

- 9 rejetées (316, 321, 452, 414, 608, 34, 618, 633, 441),

- 8 neutres (391, 467, 470, 600, 424, 361, 430, 372).

6. Façades de maisons en plan rapproché avec des personnes devant.

29 images sont pertinentes dans la base (393, 15, 400, 405, 410, 412, 435, 299, 304, 437, 441, 453, 7, 440, 484, 358, 361, 485, 571, 117, 655, 270, 326, 349, 356, 419, 425, 218, 219)

Formulation booléenne :

Etape 1:

- [1] contenu = (FACADE_D OU MAISON_D OU VILLE OU IMMEUBLE)
 ET
 [2] contenu = (PERSONNAGE OU HOMME OU FEMME OU ENFANT)
 ET
 [3] morphologie = (PLAN RAPPROCHE OU PLAN MOYEN)

-> 30 images :

- 13 pertinentes (453, 326, 270, 400, 405, 349, 419, 356, 484, 361, 299, 304, 440),
- 17 mauvaises (577, 130, 4, 14, 655, 661, 415, 33, 486, 360, 621, 565, 630, 373, 502, 635, 637)

Etape 2 :

l'élément 2 est enlevé -> 57 images dont 16 pertinentes (655, 400, 270, 405, 425, 299, 304, 440, 453, 218, 219, 349, 356, 484, 358, 361).

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = FACADE, PERSONNAGE,
 Morphologie de la photographie = PLAN RAPPROCHE, PLAN MOYEN,
 -> 97 images non visualisées,
 -> 15 images :
 - 10 choisies (419, 484, 356, 361, 299, 270, 400, 304, 655, 349),

- 4 rejetées (621, 360, 661, 415),
- 1 neutre (14).

Etape 1 :

le système propose :

Contenu de la photographie = HOTEL PARTICULIER, PORTAIL_D,
 l'utilisateur ajoute PERSONNAGE,

Contenu de la photographie = HOTEL PARTICULIER, PORTAIL_D, PERSONNAGE,

-> 25 nouvelles images :

- 4 choisies (326, 393, 258, 571),
- 18 rejetées (577, 456, 457, 460, 399, 404, 280, 408, 409, 545, 610, 490, 619, 428, 559, 438, 316, 574)
- 3 neutres (590, 32, 556).

Etape 2 :

le système propose :

Contenu de la photographie = PERSONNAGE,

l'utilisateur ajoute MAISON,

Contenu de la photographie = MAISON, PERSONNAGE,

-> 31 nouvelles images :

- 2 choisies (412, 425),
- 24 rejetées (640, 321, 514, 515, 448, 325, 450, 327, 451, 650, 332, 656, 340, 277, 342, 472, 544, 352, 418, 431, 629, 439, 570, 510)
- 5 neutres (453, 469, 353, 430, 561).

Etape 3 :

le système propose :

Contenu de la photographie = PERSONNAGE,

l'utilisateur ajoute MUR,

Contenu de la photographie = MUR, PERSONNAGE,

-> 4 nouvelles images rejetées (489, 367, 497, 306).

RECHERCHE PRECISE :

Etape 0:

Contenu de la photographie = FACADE, PERSONNAGE,

Morphologie de la photographie = PLAN RAPPROCHE, PLAN MOYEN,

-> 15 images :

- 10 choisies (419, 484, 356, 361, 299, 270, 400, 304, 655, 349),
- 4 rejetées (621, 360, 661, 415),
- 1 neutre (14).

Etape 1 :

le système propose :

Contenu de la photographie = RUE_D,

l'utilisateur ajoute PERSONNAGE,

Contenu de la photographie = RUE_D, PERSONNAGE,

-> 5 nouvelles images :

- 1 choisies (412),
- 3 rejetées (431, 450, 656),
- 1 neutre (469).

Etape 2 :

le système propose :

Contenu de la photographie = FACADE,

l'utilisateur ajoute HOMME,

Contenu de la photographie = FACADE, HOMME,

-> 1 nouvelle image rejetées (577).

Etape 3 :

le système propose :

Contenu de la photographie = PERSONNAGE,

-> 47 nouvelles images :

- 9 choisies (393, 410, 571, 453, 326, 219, 351, 485, 356),
- 30 rejetées (3, 5, 8, 650, 269, 399, 18, 404, 277, 408, 409, 160, 549, 438, 186, 570, 188, 192, 327, 472, 600, 352, 354, 482, 228, 229, 230, 490, 373, 123)
- 8 neutres (1, 289, 297, 556, 430, 561, 353, 233).

7. Edifice religieux en extérieur dans un milieu rural.

17 images sont pertinentes dans la base (385, 514, 651, 647, 138, 527, 271, 147, 46, 186, 188, 195, 452, 75, 340, 341, 359)

Formulation booléenne :

Etape 1:

[1] contenu = (CHAPELLE_D OU EGLISE_D OU BASILIQUE OU CATHEDRALE OU CLOCHER_D OU BAPTISTERE OU ABBAYE OU COUVENT OU CLOITRE OU PAGODE OU TEMPLE)

ET

[2] contenu = (CAMPAGNE_D OU VILLAGE)

ET

[3] morphologie = (VUE D'EXTERIEUR)

-> 1 image pertinente (340)

Etape 2 :

l'élément 2 est enlevé -> 39 images dont 15 pertinentes (385, 452, 514, 647, 651, 75, 271, 527, 147, 340, 341, 359, 46, 186, 188).

Etape 3 :

les éléments 2 et 3 sont enlevés -> 57 images dont 17 les pertinentes.

Avec RIVAGE :

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = EGLISE, CAMPAGNE,

Morphologie de la photographie = VUE D'EXTERIEUR,

-> 28 images :

- 10 choisies (359, 46, 385, 514, 452, 651, 527, 271, 340, 341),
- 11 rejetées (642, 644, 71, 648, 274, 541, 542, 553, 60),

- 7 neutres (386, 79, 470, 217, 543, 500, 629).

Etape 1 :

le système propose :

Contenu de la photographie = ARBRE_D, CLOCHER,

-> 21 nouvelles images :

- 1 choisie (195),

- 19 rejetées (320, 321, 448, 515, 516, 327, 650, 656, 657, 338, 471, 665, 425, 369, 565, 314, 570, 509, 510),

- 1 neutre (451).

Etape 2 :

le système propose :

Contenu de la photographie = EGLISE_D,

-> 25 nouvelles images rejetées (1, 3, 5, 6, 520, 9, 525, 82, 84, 535, 536, 87, 26, 83, 544, 416, 98, 161, 165, 166, 295, 297, 43, 307, 183).

Etape 3 :

le système propose :

Contenu de la photographie = EGLISE_D,

-> 33 nouvelles images rejetées (381, 382, 383, 264, 266, 334, 529, 19, 473, 346, 538, 539, 540, 29, 290, 546, 547, 623, 624, 496, 562, 498, 499, 626, 627, 56, 568, 57, 59, 569, 443, 380, 575).

RECHERCHE PRECISE :

Etape 0:

Contenu de la photographie = EGLISE, CAMPAGNE,

Morphologie de la photographie = VUE D'EXTERIEUR,

-> 6 images rejetées (19, 37, 38, 264, 425, 29).

Etape 1 :

le système propose beaucoup de termes, l'utilisateur choisit :

Contenu de la photographie = CATHEDRALE, BASILIQUE, COUVERT VEGETAL,

-> 2 nouvelles images rejetées (87, 89).

Etape 2 :

le système propose beaucoup de termes, l'utilisateur choisit :

Contenu de la photographie = BASILIQUE,

-> 2 nouvelles images :

- 1 choisie (452),

- 1 rejetée (470).

Etape 3 :

le système propose beaucoup de termes, l'utilisateur choisit :

Contenu de la photographie = CLOCHER, EGLISE_D,

-> 9 nouvelles images :

- 2 choisies (385, 514),

- 7 rejetées (386, 644, 626, 629, 648, 541, 543).

Etape 4 :

le système propose :

Contenu de la photographie = CLOCHER,

-> 6 nouvelles images :

- 1 choisies (527),

- 7 rejetées (47, 73, 496, 81, 464).

8. Images de restaurants, auberges, cafés ... en ville.

25 images sont pertinentes dans la base (395, 398, 397, 401, 15, 148, 400, 399, 404, 408, 409, 405, 406, 407, 410, 414, 415, 291, 422, 424, 429, 299, 442, 440, 484).

Formulation booléenne :

Etape 1 :

[1] contenu = (DEBIT DE BOISSON OU RESTAURANT OU HOTEL)

ET

[2] contenu = (VILLAGE OU VILLE_D OU QUARTIER)

ET

[3] morphologie = (VUE D'EXTERIEUR)

-> 0 images.

Etape 2 :

l'élément 2 est enlevé -> 30 images :

- 21 pertinentes (397, 398, 399, 400, 401, 148, 404, 405, 406, 407, 408, 409, 410, 414, 415, 484, 422, 424, 429, 440, 442),

- 9 mauvaises.

Etape 3 :

les éléments 2 et 3 sont enlevés -> 35 images :

- 23 pertinentes (397, 398, 399, 400, 401, 148, 404, 405, 406, 407, 408, 409, 410, 414, 415, 484, 422, 424, 429, 440, 442, 15, 291),

- 12 mauvaises.

RECHERCHE LARGE :

Etape 0 :

Contenu de la photographie = RESTAURANT,

Morphologie de la photographie = VUE D'EXTERIEUR,

-> 13 images :

- 10 choisies (484, 424, 398, 401, 148, 408, 409, 442, 414, 415),

- 2 rejetées (614, 426),

- 1 neutres (17).

Etape 1 :

le système propose :

Contenu de la photographie = GRILLE,

-> 8 nouvelles images :

- 2 choisies (291, 399),

- 6 rejetées (449, 619, 556, 655, 306, 316),

Etape 2 :

le système propose :

Contenu de la photographie = RESTAURANT,

l'utilisateur ajoute VILLE,

Contenu de la photographie = RESTAURANT, VILLE,

-> 2 nouvelles images rejetées (39, 652).

Etape 3 :

le système propose :

Contenu de la photographie = DEVANTURE,

l'utilisateur enlève DEVANTURE et demande d'"autres termes", le système propose :

Contenu de la photographie = GRILLE

-> 38 nouvelles images :

- 6 choisies (395, 400, 407, 397, 405, 410),

- 32 rejetées (329, 373, 628, 485, 483, 480, 216, 467, 339, 591, 461, 77, 588, 578, 441, 308, 303, 557, 425, 548, 420, 412, 278, 532, 403, 402, 649, 646, 387, 514, 513).

RECHERCHE PRECISE :

Etape 0 :

Contenu de la photographie = RESTAURANT,

Morphologie de la photographie = VUE D'EXTERIEUR,

-> 13 images :

- 10 choisies (484, 424, 398, 401, 148, 408, 409, 442, 414, 415),

- 2 rejetées (614, 426),

- 1 neutres (17).



Etape 1 :

le système propose :

Contenu de la photographie = GRILLE,

-> 8 nouvelles images :

- 2 choisies (291, 399),

- 6 rejetées (449, 619, 556, 655, 306, 316),

Etape 2 :

le système propose :

Contenu de la photographie = RESTAURANT,

-> 50 nouvelles images :

- 6 choisies (395, 15, 400, 405, 422, 299)

- 44 rejetées (1, 391, 393, 394, 14, 275, 411, 412, 421, 485, 363, 362, 361, 399, 360, 358, 357, 461, 295, 425, 427, 428, 431, 304, 432, 433, 435, 356, 355, 354, 353, 352, 463, 462, 436, 437, 438, 441, 66, 450, 451, 452, 453, 503).

Etape 3 :

le système propose :

Contenu de la photographie = RESTAURANT,

-> 27 nouvelles images :

- 7 choisies (397, 404, 406, 407, 410, 429, 440),

- 20 rejetées (448, 388, 389, 390, 7, 392, 10, 396, 13, 16, 402, 403, 285, 286, 287, 420, 423, 434, 439, 377).

NOM DE L'ETUDIANT : HALIN Gilles

NATURE DE LA THESE : DOCTORAT DE L'UNIVERSITE DE NANCY I en INFORMATIQUE

VU, APPROUVE ET PERMIS D'IMPRIMER

NANCY, le 05 OCT. 1989 n°1833

LE PRESIDENT DE L'UNIVERSITE DE NANCY I



RESUME

La recherche d'images est une recherche d'informations à part entière, mais les modèles utilisés dans les systèmes de recherche d'informations (SRI) existants ont été créés pour rechercher de l'information textuelle. Ces modèles doivent être adaptés pour prendre en compte ce nouveau type d'information qu'est l'image. L'arrivée des supports optiques a permis à l'ordinateur de manipuler l'image en tant que document au même titre qu'un document textuel. Un système de recherche d'images doit prendre en compte les caractéristiques de l'image et utiliser des modèles de recherche d'informations adaptés.

Cette thèse propose un processus interactif et progressif de recherche d'images basé sur le processus EXPRIM et inspiré des techniques d'apprentissage symbolique développées en IA. Le processus EXPRIM itère sur trois phases non réellement distinctes:

- la phase Avant-Visualisation : l'utilisateur formule une demande que le système interprète pour sélectionner un ensemble d'images à partir de leur description se trouvant dans une base descriptive.

- la phase Visualisation : l'utilisateur visualise les images sélectionnées par le système et effectue un choix en composant deux ensembles d'images : les images choisies et les images rejetées.

- la phase Après-Visualisation : le système analyse les choix de l'utilisateur pour essayer de comprendre ses besoins afin de les formuler sous la forme d'une nouvelle demande.

La méthode proposée assimile le processus EXPRIM à un processus d'apprentissage où le concept à apprendre est le besoin de l'utilisateur et où les exemples et contre-exemples du concept sont respectivement les images choisies et les images rejetées. Le concept est représenté dans le thesaurus en attachant des poids aux termes mesurant la capacité des termes à exprimer les besoins de l'utilisateur ; l'ensemble de ces poids forme ce qu'on appelle le **niveau d'expression** de la demande.

Cette méthode contient : un modèle de représentation du concept à l'intérieur du thesaurus, un mécanisme d'évolution de ce concept en fonction des choix de l'utilisateur, une méthode de construction du concept à partir de sa représentation dans le thesaurus, des mécanismes d'interprétation et de déformation du concept appris pour la sélection d'images dans la base.

Le prototype RIVAGE met en œuvre la méthode avec comme champ d'expérimentation une base d'images du Ministère de la Culture. Ce prototype a été réalisé dans un environnement objet et il a été évalué à l'aide de techniques classiques d'évaluation des systèmes de recherche d'informations. Cette réalisation constitue la deuxième expérience de mise œuvre d'un tel système car au cours du projet ESPRIT, une première approche a été proposée et une amorce de système a été réalisée.

Cette thèse se termine en présentant une évolution du système RIVAGE, qui est en cours d'expérimentation. Cette évolution consiste à implanter une structure Hypertexte adaptée à la recherche progressive et interactive d'images, offrant à l'utilisateur tout le confort nécessaire à une telle recherche.