



INSTITUT UNIVERSITÉ PARIS VI
CNRS
UNIVERSITÉ PARIS VII
BLAISE PASCAL

**METIS : un système et une Méthode d'Explication
de Taxinomies destinés à l'Identification de
Structures conceptuelles**

Thèse de doctorat de l'Université Paris VI

E. AÏMEUR

LAFORIA TH94/09

21 Juin 1994

**LABORATOIRE FORMES
ET INTELLIGENCE
ARTIFICIELLE**

unité associée 1095

LAFORIA

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS 6

Spécialité: Informatique

**Présentée pour obtenir le titre de
DOCTEUR DE L'UNIVERSITÉ PARIS 6**

par

Esma AÏMEUR

**METIS : un système et une Méthode d'Explicitation de
Taxinomies destinés à l'Identification de Structures
conceptuelles**

soutenue le : 21 juin 1994

devant le jury composé de :

MM.	François Rechenmann	<i>Président</i>
	Jean-Gabriel Ganascia	<i>Directeur</i>
	Guy William Mineau	<i>Rapporteur</i>
	Amedeo Napoli	<i>Rapporteur</i>
	Jacques Lebbe	<i>Examineur</i>
	Jean-Pierre Treuil	<i>Examineur</i>

A mon père et à ma mère,

REMERCIEMENTS

Je tiens en tout premier lieu à remercier vivement mon Directeur de thèse Jean-Gabriel Ganascia, Professeur à l'université Paris 6. Son encadrement permanent, la qualité de ses exigences scientifiques et la confiance qu'il m'a témoignée, m'ont permis de mener à terme ce travail. Qu'il veuille bien trouver ici, l'expression de ma profonde gratitude.

Je remercie Guy William Mineau, Professeur à l'université Laval de Québec pour avoir accepté d'être rapporteur de cette thèse malgré son éloignement et ses multiples responsabilités. Je le remercie également pour les conseils judicieux dont il m'a fait part.

Je remercie Amedeo Napoli, Chargé de Recherches au C.N.R.S. pour l'intérêt qu'il a porté à mon travail, pour toute l'attention qu'il m'a accordée et enfin pour avoir gentiment accepté d'être rapporteur de ma thèse.

Je remercie François Rechenmann, Directeur de Recherches à l'I.N.R.I.A. de m'avoir fait l'honneur de présider le jury et aussi de s'être intéressé à ce travail et de le juger.

Je remercie très chaleureusement Jacques Lebbe, Maître de Conférences à l'université Paris 6, de faire partie de mon jury ainsi que pour les discussions enrichissantes et parfois vives que nous avons eues. La précision de ses commentaires et ses nombreuses suggestions m'ont maintes fois permis d'approfondir mes réflexions.

Je suis très reconnaissante à Jean-Pierre Treuil, Ingénieur de Recherches à l'ORSTOM et Directeur du Laboratoire d'Informatique Appliquée de l'ORSTOM pour sa disponibilité, et ses judicieux conseils. Je le remercie également de faire partie du jury.

Je remercie Christian Mullon, Directeur de Recherches à l'ORSTOM et ancien Directeur du Laboratoire d'Informatique Appliquée de l'ORSTOM, de m'avoir accueillie dans son laboratoire et pour les précieux conseils qu'il m'a prodigués.

Je remercie Didier Paugy Responsable de l'Antenne ORSTOM auprès du Muséum National d'Histoire Naturelle de Paris (Laboratoire d'Ichtyologie), qui m'a permis de comprendre les problèmes liés à la systématique.

Je ne saurais trop remercier Olivier Gey, Philippe Laublet et Saliha Khouas pour leur aide précieuse, leurs multiples lectures de la thèse et leurs qualités humaines tout au long de ces années de recherche. Je leur témoigne ici, mon amitié et ma reconnaissance.

Je remercie tous les membres du LIA de l'ORSTOM Bondy et de l'équipe ACASA du LAFORIA pour l'ambiance chaleureuse qui régnait lors de nos réunions hebdomadaires.

Mes discussions avec R. Vignes, M. Guerdoud et D. Goujet à propos de la systématique m'ont beaucoup apporté, qu'ils soient chaleureusement remerciés.

J'exprime ma profonde gratitude à tous mes relecteurs, que ce soit pour des articles scientifiques ou pour une première version du manuscrit, et plus précisément à M. Baron, C. Ben-Yelles, M-C. d'Erceville, C. Faron, O. Gey, H. Giroire, Q. Kieu, B. Leroux, C. Mering, I. Moulinier, F. Ouzzani, M. Sebag, A. Shambrook, P. Shambrook, J. Thomas et C. Vicat.

je remercie aussi le Laboratoire d'Informatique Appliquée de l'ORSTOM Bondy et la Mission Technique Informatique pour leur soutien financier ainsi que S. Vallat et A. Aing pour leur soutien amical.

Je tiens à remercier H. Mouri , A. Blot, V. Simmoneau et Y. Chatelin pour leur aimable collaboration.

Je n'oublie pas de remercier très chaleureusement ma famille et mes amis. Grâce à leur soutien permanent et à leur sollicitude, j'ai mieux supporté la nostalgie de mon pays.

Le mot de la fin est pour une personne qui bien que n'étant pas du domaine, a su me faire garder confiance en moi. Ses encouragements et critiques constructives durant les tournants difficiles de ma thèse, m'ont été salutaires. Je lui exprime ici toute ma reconnaissance.

Résumé

Dans les sciences d'observation telles que l'archéologie, la minéralogie ou la zoologie, de grandes quantités d'objets sont répertoriées sous forme de taxinomies et exprimées dans des catalogues au moyen d'un langage de description varié et nuancé. La lourde gestion de ces catalogues et les problèmes liés à la description et à la caractérisation des classes d'objets, ont fait naître le besoin de construire d'une part une base de connaissances permettant d'intégrer des informations cohérentes, homogènes et non redondantes, d'autre part une structure de discrimination fiable, permettant de caractériser les classes de manière pertinente. Le principe d'explicitation des connaissances retenu dans le cadre de notre travail repose sur une approche incrémentale. Notre système METIS explicite progressivement les connaissances de l'expert et raffine le langage de description (appelé modèle structurel) à partir d'une suite de descriptions de classes qui lui sont présentées. Chaque nouvelle classe fournie au système, constitue un enrichissement et une nouvelle vision du modèle structurel, soit par spécialisation d'une description déjà acquise, soit par l'ajout d'une nouvelle description. METIS modifie ainsi dynamiquement le modèle structurel des descriptions de classes explicitées. La taxinomie du domaine sert de fil conducteur pour guider le processus d'explicitation des connaissances. Elle est progressivement enrichie, permettant ainsi au système de présenter les traits discriminants et d'évaluer leur pertinence par apprentissage, ou avec le concours de l'expert.

Mots clés : explicitation de connaissances, discrimination, classification, identification, apprentissage inductif, taxinomie, graphes conceptuels, modèle structurel.

Abstract

Observation sciences such as archeology, mineralogy or zoology, describe a wide variety of objects with taxinomies, using a textual description language for the definition of catalogs. The complexity of catalogs management and problems related to the description and characterization of objects classes led to the need of two structures : a knowledge base allowing the integration of coherent, homogeneous and non redundant information, and a discriminating structure allowing a pertinent characterization of classes. The principle of knowledge elicitation adopted in the present work is based on an incremental approach. Our system, METIS, elicits progressively the expert knowledge and refines the description language (called structural model) from a serie of class descriptions presented to the expert. Each new class presented to the system is a new view of the expert model, either by specialization of a previously acquired description or by adding a new description. METIS dynamically modifies the structural model of elicited classes descriptions. The taxinomy of the domain is used to guide the knowledge elicitation process. It is progressively completed, allowing the presentation of discriminating features and the evaluation of their pertinence by machine learning or by interacting with the expert.

Keywords : knowledge elicitation, discrimination, classification, identification, inductive learning, taxinomy, conceptual graphs, structural model.

SOMMAIRE

1.INTRODUCTION	1
1.1.Présentation de la problématique.....	1
1.2.Objectifs.....	3
1.3.Présentation de l'approche.....	4
1.4.Présentation du mémoire.....	6
2.CLASSIFICATION.....	9
2.1.Définition de la classification.....	9
2.2.Classification en biologie.....	11
2.2.1.Problèmes de discrimination et d'identification.....	12
2.2.2.Méthodes informatiques appliquées en systématique	16
2.3.La classification vue comme un processus de construction de classes.....	19
2.3.1.Apprentissage symbolique automatique (classification conceptuelle).....	19
2.3.2.Analyse des données.....	21
2.3.2.1.Méthodes non hiérarchiques (partitionnement).....	21
2.3.2.2.Méthodes hiérarchiques (descendantes et ascendantes).....	22
2.4.La classification vue comme un processus de raisonnement sur les classes.....	22
2.4.1.Apprentissage de concepts	23
2.4.1.1.Arbres de décision (ID3).....	23
2.4.1.2.Construction de systèmes de règles (CHARADE).....	25
2.4.2.Le raisonnement par classification dans les systèmes de représentation à base d'objets	27
2.4.3.Analyse des données.....	27

2.4.3.1. <i>Analyse discriminante</i>	28
2.5. Conclusion	28
3. EXPLICITATION DES CONNAISSANCES	31
3.1. Introduction	31
3.2. Méthodes non contraintes	34
3.2.1. Technique d'interview	34
3.2.2. Technique d'analyse de protocoles.....	34
3.2.3. Observation directe.....	35
3.3. Méthodes contraintes	36
3.3.1. Grilles-répertoires.....	36
3.3.2. Tri de cartes.....	40
3.3.3. Technique de "laddering"	43
3.4. Conclusion	46
4. REPRÉSENTATION DES CONNAISSANCES	49
4.1. Représentations logiques	49
4.2. Réseaux sémantiques	50
4.3. Les logiques terminologiques (KL-ONE)	53
4.4. Les graphes conceptuels	56
4.4.1. Éléments fondamentaux	57
4.4.2. Opérations fondamentales.....	62
4.4.3. Jointure maximale.....	64
4.4.4. Projection	65
4.4.5. Généralisation et spécialisation	67
4.4.6. Hiérarchie de généralisation.....	68
4.4.7. Abstraction et définition	69
4.4.8. Contraction.....	71
4.4.9. Expansion.....	71
4.4.10. Hiérarchie Aristotélicienne.....	72

4.5.Conclusion	72
5.L'EXPLICITATION DES CONNAISSANCES DANS METIS.....	75
5.1.Base de la taxinomie du domaine	77
5.2.Représentation des connaissances structurées dans METIS.....	78
5.2.1.Formalisme de représentation des connaissances de METIS	79
5.2.2.Structuration de la base de connaissances.....	80
5.2.2.1.Graphe primitif.....	82
5.2.2.2.Base des graphes.....	83
5.2.2.3.Modèle structurel.....	85
5.2.2.4.Taxinomie des classes.....	85
5.3.Acquisition des connaissances structurées de METIS....	87
5.4.Phase d'initialisation	88
5.4.1.Niveaux du domaine.....	88
5.4.2.Explicitation du noyau G0	89
5.4.3.Informations sur le noyau.....	90
5.4.4.Cohérence.....	91
5.5.Phase de classification.....	91
5.5.1.Parcours de la taxinomie	92
5.5.2.Cohérence : validation des traits hérités	93
5.6.Phase de discrimination	94
5.6.1.Traits discriminants.....	95
5.6.1.1.Traits discriminants suffisants.....	97
5.6.1.2.Traits discriminants insuffisants.....	97
5.6.2.Recherche de traits discriminants par METIS.....	98
5.6.2.1.Procédure d'apprentissage	105
5.6.2.2.Évaluation de la pertinence des traits discriminants	119
5.6.3.Traits donnés par l'expert	120
5.6.4.Contrôle de cohérence et mise à jour des données	123

5.7.Phase de saisie de données intrinsèques.....	124
5.7.1.Acquisition de la description.....	124
5.7.2.Raffinement du modèle structurel	125
5.7.3.Cohérences des données	126
5.7.4.Traitement des négations	129
5.7.4.1.Négation de relations descriptives.....	129
5.7.4.2.Négation de relations structurelles	130
5.7.5.Traitement des exceptions.....	130
5.7.5.1.Exception sur les relations descriptives.....	130
5.7.5.2.Exception sur les relations structurelles	131
5.8.Architecture du système.....	131
5.8.1.Schéma	131
5.8.2.Fonctionnalités	132
5.8.2.1.Initialisation.....	133
5.8.2.2.Classification	133
5.8.2.3.Discrimination.....	134
5.8.2.4.Saisie	135
5.8.2.5.Interaction avec l'expert	135
5.9.Conclusion	136
6.METIS : MISE EN OEUVRE DE LA MÉTHODE.....	139
6.1.Organisation des connaissances.....	139
6.1.1.Hiérarchie de généralisation.....	139
6.1.2.Taxinomie des classes	143
6.1.3.Concepts	143
6.1.4.Treillis des types de concepts.....	144
6.1.5.Treillis de composition	145
6.1.6.Relations	146
6.1.7.Référents.....	147
6.1.8.C-modèle.....	148
6.1.9.Base canonique.....	151

6.1.10.Canon	152
6.1.11.Opérations utilisées.....	152
6.2.Implantation et Expérimentation.....	156
6.2.1.Expérimentation en minéralogie	156
6.2.2.Expérimentation en pédologie	165
6.3.Conclusion	170
7.CONCLUSION GÉNÉRALE	173
RÉFÉRENCES.....	179
ANNEXE 1	193
ANNEXE 2	199
ANNEXE 3	207
ANNEXE 4	219

1. INTRODUCTION

Que font les zoologistes, les botanistes, les médecins, les microbiologistes, les pathologistes, lorsqu'ils s'intéressent à l'anatomie, à la morphologie, à la physiologie, à la génétique, à la biologie moléculaire, ou à l'écologie, de l'homme, des animaux et des plantes ? Que font-ils pour regrouper et recouper leurs résultats respectifs en un ensemble cohérent et homogène ? Ils créent des taxinomies.

Une taxinomie est une hiérarchie dont les classes de même niveau et descendant d'un même ancêtre s'excluent mutuellement. Les niveaux d'une taxinomie correspondent aux *rangs* taxinomiques (par exemple, dans la classification de Charles Linné le règne, l'embranchement, la classe, l'ordre, la famille, le genre, l'espèce, la variété et la race) et les éléments correspondent aux *taxons*. À un rang on associe plusieurs taxons, mais à un taxon on ne fait correspondre qu'un rang. Ainsi par exemple, au taxon "Citharinidae" correspond le rang "Famille" et au rang "Famille" correspondent plusieurs taxons tels que : "Citharinidae", "Cyprinidae" et "Characidae".

1.1. PRÉSENTATION DE LA PROBLÉMATIQUE

Dans les sciences d'observation telles que l'archéologie, la minéralogie, la zoologie, ou la botanique, de grandes quantités d'objets sont répertoriées sous forme de taxinomies et exprimées dans des catalogues au moyen de langages de description variés et nuancés.

Plus précisément, la tâche des taxinomistes consiste à observer, à nommer, à décrire et à organiser les objets sous forme d'une taxinomie. La description des objets, leur caractérisation et leur classification posent un certain nombre de **problèmes** liés soit à la construction soit à l'exploitation des classes. Avant d'énoncer ces problèmes, précisons qu'ils ont fait l'objet de plusieurs travaux en classification dans le domaine de l'intelligence artificielle.

Le *premier problème* concerne la construction de la **taxinomie**. C'est une tâche très complexe, car même si des méthodes plus ou moins générales existent, elles comportent encore beaucoup de subjectivité. Ainsi, les résultats obtenus à partir d'un même ensemble d'observations diffèrent la plupart du temps d'un taxinomiste à l'autre, en raison de l'interprétation particulière qu'ils accordent à ces données. Ces différences d'interprétation sont en réalité le fait d'expériences personnelles. De plus, la taxinomie n'est pas indépendante de l'usage qui en est fait ultérieurement. Ainsi, en médecine, la classification des maladies peut être entreprise dans l'optique de recenser tous les cas de maladies, en espérant soigner les malades, ou encore pour observer les manifestations morbides chez les malades. A chacun de ces buts correspond un choix d'organisation particulier. En ce qui concerne la biologie par exemple, le but visé est celui de l'identification des spécimens et de l'organisation des connaissances.

Le problème de la construction de la taxinomie se pose en intelligence artificielle, en terme de construction de classes (*conceptual clustering ou concept formation*) [MICHALSKI & STEPP, 83], [STEPP & MICHALSKI, 86] [LEBOWITZ, 87] [FISHER, 87], [GENNARI & al, 89], [MINEAU, 90], [DECAESTECKER, 91], [GEY, 91], [ANQUETIL & VAUCHER, 93], [BISSON, 93]. Il se pose également en terme de catégorisation¹ dans le domaine de la représentation par objets [NAPOLI, 92a] [NAPOLI, 94].

Le *deuxième problème* concerne la **caractérisation**. Caractériser un taxon, par exemple une espèce, revient à rechercher par comparaison avec les autres taxons ce qui le discrimine, c'est à dire les propriétés qui, si elles sont présentes chez un individu, attestent que celui-ci appartient au taxon. La caractérisation manuelle des taxons est une tâche difficile. De manière générale, les taxinomistes essaient d'avoir les caractérisations les plus courtes possibles.

Les travaux de classification en intelligence artificielle se sont intéressés à la caractérisation en termes d'extraction de caractéristiques discriminantes sous forme d'arbres de décision ou de systèmes de règles [QUINLAN, 86] [GANASCIA, 87]. En effet, à partir de description d'exemples appartenant à des classes connues et dont le langage de description est figé à l'avance, il s'agit de construire des fonctions de reconnaissance capables de caractériser les différentes classes.

¹ Le lecteur intéressé par les travaux en catégorisation (surtout en psychologie cognitive) pourra consulter [ROSCH & LLYOD, 78] et [VAN MECHELEN & al, 93].

Le *troisième problème* concerne la **description**, car il existe souvent une incertitude sur la pertinence des descripteurs choisis. Par ailleurs, le langage de description peut ne pas être homogène et présenter des ambiguïtés, des redondances, voire des incohérences. La description des taxons peut être abondante ou au contraire insuffisante, dépendant du niveau de granularité de la description donnée par différents experts.

1.2. OBJECTIFS

Ce travail a pour objectif de faciliter la tâche des taxinomistes qui créent manuellement des descriptions et des caractérisations de classes. Ainsi, nous proposons un outil qui construit une base de connaissances en manipulant des taxinomies dans des domaines tels que l'ichtyologie², la pédologie ou la minéralogie.

Notre approche fait appel à des techniques d'*explicitation*³ des connaissances [BOOSE, 84] [BOOSE, 89] [MAJOR & REICHGELT, 90] et d'apprentissage⁴ de caractéristiques discriminantes [QUINLAN, 83], [QUINLAN, 86], [GANASCIA, 87].

Ces techniques sont issues du domaine de l'acquisition des connaissances qui consiste à extraire les connaissances et à les organiser au moyen de techniques manuelles, semi-automatiques (interactives) ou automatique (apprentissage).

La plupart des travaux réalisés dans le domaine de la classification (du point de vue exploitation des classes), ont peu abordé l'amélioration du langage de description d'une part et l'amélioration de la caractérisation des classes d'autre part. Pourtant, ces problèmes se posent de façon cruciale en acquisition des connaissances. Notre objectif dans ce travail est justement de fournir un outil informatique qui puisse reconstruire les taxinomies existantes dans les catalogues en **améliorant la description et la caractérisation des classes** à l'aide d'une **technique d'apprentissage symbolique**.

2 Plus précisément la systématique des poissons d'eaux douces et d'eaux saumâtres de l'Afrique occidentale.

3 Une définition précise de l'explicitation des connaissances est donnée dans le chapitre 3.

4 Il s'agit d'apprentissage par induction.

Ayant une vision constructiviste de l'acquisition des connaissances [KRIVINE & DAVID, 91], [LINSTER, 92], nous pensons que l'explicitation des connaissances n'est pas simplement un transfert des connaissances de l'expert vers la machine, mais aussi une activité d'interaction avec l'expert qui permet de créer de nouvelles connaissances.

Ainsi, notre approche intègre une technique d'induction dans un outil interactif d'explicitation des connaissances, ce qui constitue en fait, un apport aux méthodes traditionnelles d'acquisition des connaissances.

1.3. PRÉSENTATION DE L'APPROCHE

La présente thèse propose un système d'explicitation de taxinomies destiné à l'identification de structures conceptuelles appelé METIS⁵. Le principe d'explicitation des connaissances retenu dans le cadre de notre travail repose sur une approche incrémentale. METIS va progressivement expliciter les connaissances de l'expert et raffiner un modèle structurel minimal du domaine à partir d'une suite de descriptions de classes qui vont lui être présentées. Chaque nouvelle description fournie au système constitue un enrichissement et une nouvelle vision du modèle de l'expertise soit par spécialisation d'une description déjà acquise soit par l'ajout d'une nouvelle description. METIS modifie ainsi dynamiquement le modèle structurel des objets explicités. La taxinomie du domaine va servir de fil conducteur pour guider le processus d'explicitation des connaissances. En effet, elle va être progressivement enrichie (en caractérisant les classes une à une et en les discriminant de leur classes sœurs), permettant ainsi au système, de présenter les traits discriminants [AIMEUR & GANASCIA 93a] et d'évaluer leur pertinence, soit automatiquement, soit avec le concours de l'expert.

⁵ Un Système et une Méthode d'Explicitation de Taxinomies destinés à l'Identification de Structures conceptuelles.

METIS comporte plusieurs caractéristiques qui visent à solutionner les deux derniers types de problèmes (caractérisation et description) :

En ce qui concerne la caractérisation :

- Une procédure d'apprentissage appelée IDICE utilise un algorithme de type ID3 afin d'apprendre de nouveaux traits discriminants. Contrairement à ID3, IDICE prend en compte des descriptions incomplètes et permet à l'expert, d'imposer des contraintes quant au choix et à l'utilisation de certains traits. Par ailleurs, une étape de complétion d'exemples favorise la recherche de traits potentiellement discriminants et homogénéise dans la mesure du possible la description des taxons. L'expert est seul juge quant à la pertinence des traits : METIS suggère mais ne décide pas à la place de l'expert.

En ce qui concerne la description :

- Le langage de description n'est pas figé et évolue au fur et à mesure que l'on considère un nouveau taxon ou que l'on spécialise la description de ceux qui existent déjà. Il est suffisamment concis pour ne contenir que les informations pertinentes et suffisamment riche pour décrire tous les taxons considérés.
- La représentation des connaissances à l'aide du formalisme des graphes conceptuels permet une meilleure lisibilité que celle du formalisme logique, une bonne visualisation graphique de la description des objets, une richesse d'expression sémantique une représentation claire de la structuration des connaissances et enfin de nombreuses facilités de manipulation dues à la souplesse des opérations définies sur les graphes conceptuels.
- Enfin la non-redondance de la base et la cohérence statique des informations entrées dans le système sont assurées par la prise en compte d'un modèle de cohérence (C-Modèle) construit lui aussi de manière incrémentale. METIS s'y réfère en permanence pour valider interactivement ses connaissances. La validation se fait grâce à un dialogue avec l'expert qui se trouve ainsi confronté à ses propres connaissances.

Une amélioration conjointe des deux problèmes se fait dans METIS de la manière suivante :

- L'interaction avec l'expert permet l'élaboration de la base de connaissances. Sa structuration fait appel aux notions de graphe primitif, de base de graphes, de modèle structurel, et de taxinomie de classes. Le processus d'explicitation est interactif et itératif, il s'arrête quand l'expert a entré la description de la dernière classe.

- La description et la caractérisation des classes utilisent la taxinomie du domaine comme fil conducteur. Partant d'un noyau de description générique choisi avec l'expert, et constituant le point de départ du processus d'explicitation incrémental, METIS modélise petit à petit le domaine d'application. L'obtention du modèle descriptif est une étape fondamentale dans le processus cognitif qui permet de passer de l'observation à la description.
- L'explicitation des connaissances se fait par étapes selon l'organisation hiérarchique des connaissances. Le mécanisme d'héritage permet une économie en place mémoire, une restriction de l'espace de recherche et donc une diminution du nombre de questions à poser à l'expert.

METIS a été expérimenté en ichtyologie et dans deux autres domaines qui ne relèvent pas des sciences naturelles mais où des taxinomies préalablement connues⁶ existent : la minéralogie et la pédologie. Ceci montre que la méthode METIS est facilement généralisable.

L'acronyme de notre système possède un sens bien précis dans la mythologie grecque, comme le confirme la phrase empruntée à [DETIENNE & VERNANT, 74] dans *Les ruses de l'intelligence : la mêtis des grecs*.

"La mêtis est bien une forme d'intelligence et de pensée, un mode du connaître ; elle implique un ensemble complexe, mais très cohérent, d'attitudes mentales, de comportements intellectuels qui combinent le flair, la sagacité, la prévision, la souplesse d'esprit, la feinte, la débrouillardise, l'attention vigilante, le sens de l'opportunité, des habiletés diverses, une expérience longuement acquise"

Tout comme la mêtis des Grecs, nous avons souhaité que notre système puisse expliciter le savoir-faire des taxinomistes en utilisant une méthode intelligente, souple et efficace.

1.4. PRÉSENTATION DU MÉMOIRE

Le chapitre 2 a pour but de situer notre approche de la classification. Il aborde la classification en biologie, puis illustre les différentes acceptions de la notion de

⁶ Il n'existe pas toujours un consensus international sur les taxinomies élaborées : c'est le cas de la pédologie.

classification en donnant des exemples en apprentissage symbolique automatique, en analyse des données et en représentation à base d'objets. Ainsi, nous nous intéressons à la classification aussi bien du point de vue du processus de construction de classes que du point de vue du raisonnement sur ces classes.

Compte tenu des problèmes rencontrés en classification, le chapitre 3 vise à détailler les techniques d'*explicitation des connaissances*⁷, en montrant en particulier que l'acquisition des connaissances intègre deux aspects importants et inextricablement liés : la *modélisation*⁸ et l'explicitation des connaissances. Nous examinons dans ce chapitre, d'une part les techniques directes qui comprennent l'interview, l'analyse de protocoles, l'observation directe ..., et d'autre part les techniques indirectes comme les grilles répertoires, le tri de cartes et la technique de "Laddering". Nous présentons en même temps les systèmes KRITON [DIEDERICH & al, 88], AQUINAS [BOOSE, 89], CATO [MAJOR, 91] et ALTO [MAJOR & REICHGELT, 90] qui sont des réalisations informatiques des techniques d'explicitation des connaissances précitées. Nous indiquons brièvement la technique d'explicitation de connaissances employée dans METIS et la comparons à celles utilisées dans les systèmes ETS [BOOSE, 84], AQUINAS et ALTO.

Comme il est nécessaire de bien représenter les connaissances explicitées afin de pouvoir les exploiter et les manipuler, le chapitre 4 décrit quelques formalismes de représentation des connaissances tels que la logique des propositions, la logique des prédicats, les réseaux sémantiques et les graphes conceptuels. Nous développons dans ce chapitre une réflexion autour des réseaux sémantiques et plus précisément autour des graphes conceptuels en montrant l'avantage de ces représentations par rapport au formalisme logique du point de vue de la lisibilité (visualisation graphique de la description des objets), de la richesse d'expression sémantique, de la clarté dans la structuration des connaissances et enfin de la facilité de manipulation.

L'analyse précédente des méthodes d'explicitation des connaissances et des mécanismes de représentation de ces connaissances nous permet alors de présenter en détail, au chapitre 5 la méthode d'explicitation des connaissances de METIS. Nous montrons comment se réalise la construction interactive de la base de connaissances structurées en taxinomie. METIS explicite progressivement les connaissances de l'expert

⁷ Plus précisément celles qui s'intéressent aux classifications.

⁸La modélisation comporte deux aspects imbriqués : une modélisation du domaine et une modélisation du processus de résolution de problèmes [WIELINGA & BREUKER, 86], [STEELS, 90].

et raffine un modèle structurel minimal du domaine à partir d'une suite de descriptions de classes qui lui sont présentées. La taxinomie du domaine va servir de fil conducteur pour guider le processus d'explicitation des connaissances. En effet, elle va être progressivement reconstruite (en caractérisant les classes une à une et en les discriminant de leur classes sœurs), permettant ainsi au système, de présenter les traits discriminants et d'évaluer leur pertinence, soit automatiquement (par apprentissage), soit avec le concours de l'expert. Enfin, une architecture générale du système est donnée ainsi que les différentes fonctionnalités (initialisation, classification, discrimination, saisie et interaction avec l'expert).

Le chapitre 6 décrit la mise en œuvre de METIS (implémentation en termes de graphes conceptuels). Nous présentons tout d'abord, les types de connaissances pris en compte par METIS et leur organisation. Nous indiquons ensuite comment sont implémentées les opérations qui manipulent ces connaissances. Ainsi, nous étendons le formalisme de SOWA par de nouvelles règles de formation canonique. Le dernier aspect concerne l'expérimentation effectuée dans les domaines de l'ichtyologie, de la pédologie et de la minéralogie.

Enfin le chapitre 7 clôt ce travail en présentant un bilan de la contribution que nous avons apportée aux techniques d'explicitation de connaissances (sous forme de taxinomie) et en indiquant les perspectives possibles de notre recherche.

2. CLASSIFICATION

L'usage de la classification est fondamental pour la science. Ainsi, le scientifique qui n'établit pas une classification dans le domaine qu'il étudie, utilise toutefois une classification élaborée par d'autres. Ce chapitre aborde dans un premier temps, la classification en biologie puis, illustre dans un second temps les différentes acceptions de la notion de classification en donnant des exemples en apprentissage symbolique automatique en analyse des données ou encore dans les systèmes de représentation à base d'objets. Ce chapitre a pour but de nous permettre de situer le cadre dans lequel s'inscrit le mot "classification" dans notre approche.

2.1. DÉFINITION DE LA CLASSIFICATION

Il n'est pas possible de définir de manière unique ce qu'est la classification, car selon les personnes et selon les communautés le terme classification reçoit des acceptions différentes. Pour certains, il est utilisé au sens de *classement* (positionnement d'une classe dans une hiérarchie existante), pour d'autres il est utilisé au sens de construction de classes (création de classes et éventuellement leur regroupement en unités hiérarchisées). Le terme classification est aussi utilisé pour définir le résultat du processus de regroupement [MARCUS, 93] (exemple : la classification des plantes) .

Remarque :

Le classement consiste à attribuer une classe à un objet non identifié, une fois la classification construite. Si l'identification est parfois appelée classification, cela provient de la traduction malheureuse du mot anglais "classification" comme le soulignent [SNEATH & SOKAL, 73].

J. Lebbe (communication personnelle) constate que le terme classification est employé selon quatre sens différents : trois processus et un résultat. Les trois processus

sont : la construction de classes, la discrimination et l'identification. Quant au résultat c'est la structure classificatoire qui découle du processus de construction des classes.

Quant à Napoli et Rechenmann [NAPOLI & RECHENMANN, 93] ils parlent de deux acceptions différentes de la classification : la première est liée à l'organisation hiérarchique des connaissances: construction de catégories et organisation de l'ensemble de ces catégories en hiérarchies. La seconde est liée à l'exploitation de ces représentations hiérarchiques.

En ce qui nous concerne, nous retenons le point de vue suivant, à savoir que la classification peut être vue soit comme un processus de *construction de classes*, soit comme un processus de *raisonnement sur ces classes* (discrimination et identification). Dans notre travail, ce sera plutôt le deuxième sens qui sera pris en considération [AIMEUR & GANASCIA 93c].

Remarque :

La discrimination permet, étant donné un ensemble, de distinguer un objet de tous les autres. L'identification quant à elle permet de trouver la classe d'appartenance d'un individu donné.

On peut maintenant se poser la question de savoir quand faut-il classer et pourquoi ?

Dès que l'on sélectionne [LAKOFF, 87] quelque chose par rapport au reste, dès que l'on a un *tropisme* positif vis à vis d'un objet par rapport au reste, on *classe* (D. Goujet : communication personnelle). Au demeurant tous les êtres vivants classent (exemple : la paramécie sélectionne son hôte, donc elle classe).

On s'aperçoit ici que le biologiste D. Goujet emploie le terme classification dans le sens d'identification et que le raisonnement de la nature vivante est par essence classificatoire. R. Pankhurst [PANKHURST, 91] défend le point de vue selon lequel les classifications seraient fondamentales, non seulement pour la biologie, mais pour tout le langage humain. L'utilisation concrète des noms dans le langage dépend d'une classification et d'une dénomination préalable de ces objets. L'utilisation de la notion de "table" dans un discours dépend de l'existence d'un consensus entre le locuteur et les auditeurs sur le concept de "table". La meilleure illustration en est le fameux passage de "*Through the Looking Glass*" de Lewis Carroll dans lequel Alice se retrouve dans une

forêt où les choses n'ont pas de noms et se rend compte qu'il lui devient impossible d'engager une conversation.

Pour Ermine [ERMINE, 93], la classification sert à planifier un savoir ; un groupe d'experts peut établir une classification pour mettre en commun leurs connaissances. C'est alors l'idée d'une norme qui prédomine, la recherche d'un noyau stable de connaissances. La classification constitue les fondations sur lesquelles tout un chacun peut s'appuyer. D'autre part, on peut l'envisager comme un outil pratique pour un utilisateur plus ou moins familiarisé avec un domaine de savoir. Elle est alors une source documentaire précieuse, permettant de multiples recherches d'informations. Elle peut également servir à l'enseignement, car du fait de sa forme arborescente, une classification permet de replacer une entité dans un contexte plus large, ou bien de pouvoir la détailler en des termes plus précis.

En conclusion, nous sommes tout à fait convaincus que la classification est utile et importante aussi bien pour les personnes qui construisent les classes que pour celles qui les manipulent.

Après avoir tenté de fournir quelques éléments de réponse à la question que nous nous sommes posée, regardons maintenant l'impact de la classification sur les sciences biologiques.

2.2. CLASSIFICATION EN BIOLOGIE

La classification est fondamentale pour la biologie et ne cesse de ce fait de prêter à controverse. Le débat a été particulièrement animé sur le rôle que doit jouer l'histoire par rapport à la structure. La question est de savoir si les classifications d'ordre biologique sont purement structurales, ou, si d'une certaine manière, elles se conforment à l'histoire. Un deuxième conflit a porté sur la place et l'importance relative de la théorie implicite associée à l'observation. Les classifications biologiques peuvent-elles rester neutres ou libres de toute théorie, tributaires seulement de l'observation directe ou, au contraire, sont-elles intimement liées à une ou plusieurs théories scientifiques ? Or, ces deux problèmes sont inextricablement liés. Par ailleurs, si les classifications structurales ne sont fondées que sur ce qui est observable directement, les classifications historiques se heurtent à des difficultés puisque le passé n'est pas observable. Toutes ces questions ont

conduit à plusieurs courants ou pensées¹ sur la façon de considérer ces classifications. Ainsi on distingue la *systématique évolutionniste* (phylogénétique), la *phénétique numérique* et enfin l'*analyse cladistique*.

2.2.1. Problèmes de discrimination et d'identification

On fait appel à la discrimination (définie dans § 2.1.) lorsque l'on veut construire une *diagnose* ou une *clé d'identification*.

La diagnose est le terme employé par les biologistes pour désigner le résultat d'une caractérisation [AIMEUR, 93a], [AIMEUR, 93b]. Caractériser un concept, c'est rechercher par comparaison avec les autres concepts ce qui le discrimine, c'est-à-dire les propriétés qui, si elles sont présentes chez un individu, attestent que celui-ci appartient au concept [VIGNES, 91]. Selon D. Goujet (communication personnelle) la diagnose contient les caractères distinctifs d'un taxon uniquement au niveau où on le considère (par exemple les caractères qui distinguent une espèce de toutes les autres espèces du même genre, et les caractères appartenant au genre ne sont pas considérés) de plus, elle doit être *pure*, c'est à dire qu'elle ne doit pas contenir de caractères dont les valeurs sont des disjonctions (exemples : taille-fleur : grande ou petite, Espèce : herbacée ou ligneuse).

La réalisation manuelle de ces diagnoses représente une part importante de la *systématique*², comme le soulignent [MATILE & al 87] en donnant cette définition de la systématique :

"L'étude et la description de la diversité des êtres vivants, la recherche de la nature et des causes de leurs différences et de leurs ressemblances, la mise en évidence des relations de parenté existant entre eux et l'élaboration d'une classification traduisant ces relations de parenté".

Les systématiciens ont créé les clés d'identification de façon empirique. Le principal inconvénient de ces clés tient au fait qu'elles sont déterministes, c'est-à-dire que l'on doit répondre à toutes les questions imposées par la clé d'identification sous peine de blocage

1 Nous ne décrivons pas ces pensées classificatoires, car cela dépasse le cadre de ce travail. Le lecteur intéressé pourra consulter les ouvrages suivants : [HENNING, 49], [SOKAL & SNEATH, 63], [SNEATH & SOKAL, 73] [MAYR, 86], [HULL, 86], [TASSY, 86], [MATILE & al, 87] [PAJAUD, 90a], [PAJAUD, 90b].

2 "Cette science est à la fois un préalable, car toute la biologie se fonde sur elle, et un aboutissement car elle représente une synthèse des connaissances acquises pour les êtres vivants" [GUERDOUD, LEBBE, 94].

du processus. De plus, le risque de faire des erreurs en répondant aux questions est grand.

Les méthodes traditionnelles d'identification procèdent soit par élimination (l'élimination se fait par dichotomie) soit par appariement³ [ABBOTT & al, 85]. Elles sont nombreuses et comprennent : les clés-diagnostiques (appelées aussi clés d'identification) les clés à base de cartes perforées ("Punched Card Key"), les méthodes d'appariement (basées sur les coefficients de similarités ou sur le théorème de Bayes) et enfin les méthodes interactives ("On-line") [PANKHURST, 91]. Nous allons examiner deux méthodes⁴ de construction de clés-diagnostiques : les *clés dichotomiques* et les *clés tabulaires*.

Clé dichotomique :

C'est la méthode traditionnellement la plus utilisée. On obtient une hiérarchie dite linéenne (du nom de son concepteur : Linné). Chaque séparation d'un niveau en sous-niveaux de la hiérarchie s'effectue par la mise en évidence de deux sous-classes⁵ ; l'une regroupe les espèces possédant un caractère donné et l'autre rassemble le reste des espèces pour lesquelles ce caractère est, soit faux, soit inapplicable.

Le succès de cette méthode auprès des systématiciens tient au fait que les clés sont le moyen le plus rapide d'identifier une espèce parce qu'elles éliminent à chaque ligne de larges fractions du groupe. A condition de ne pas faire d'erreurs, une clé bien faite apporte la réponse souhaitée en utilisant le nombre minimum de caractères.

Dans l'approche dichotomique [FORTUNER, 93] l'inconvénient de la clé tient au fait qu'elle est déterministe. Pour une clé le monde est soit noir soit blanc ; ainsi si le caractère C présente l'état E1, allez à la ligne 4, s'il présente l'état E2, allez à la ligne 27... En réalité l'échantillon à identifier comprend souvent des spécimens⁶ ayant l'état E1 et d'autres, l'état E2. La question est de savoir si cela est causé par un mélange d'espèces ou bien si la variabilité intra-spécifique est plus grande que ne le suppose l'auteur de la clé. On parle de variabilité intra-spécifique quand les caractères varient pour des individus

3 En réalité il ne s'agit pas d'appariement au sens habituel car dans ce cas précis, cela revient à faire du raisonnement à partir de cas.

4 Ce sont des méthodes manuelles.

5 En intelligence artificielle, cela reviendrait à raisonner en logique booléenne.

6 Ceci est un cas particulier, car dans le cas général on n'identifie qu'un caractère à la fois.

appartenant à une même espèce. D'autre part l'arrangement de la clé, avec des renvois à des lignes qui peuvent être éloignées l'une de l'autre, fait que toute erreur peut causer des conséquences graves.

Exemple : clé des genres de la famille *Clupeïda* [LEVEQUE & al, 91]

- 1- Présence de deux supramaxillaires (antérieur et postérieur) à la mâchoire supérieure.....**Sardinella**
 - Un seul supramaxillaire (postérieur) à la mâchoire supérieure**2**
- 2- Une échancrure médiane nette à la mâchoire supérieure.....**Ethalmalosa**
 - Pas d'échancrure médiane ; écailles à bord distal non laciné.....**3**
- 3- Ecussons prépelviens et postpelviens carénés**4**
 - Ecussons prépelviens non carénés.....**5**
- 4- Dents prémaxillaires et dentaires ne comprenant pas de canines plus grandes**Pellonula**
 - Dents prémaxillaires et dentaires comprenant des canines plus grandes**Odaxothrissa**
- 5- Dorsale inséréé au milieu du corps et au-dessus des pelviennes**Laeviscutella**
 - Dorsale insérée en arrière des pelviennes**Sierrathrissa**

Clé tabulaire :

Une clé tabulaire est un tableau dont les lignes correspondent aux espèces et dont les colonnes correspondent aux caractères, parfois les valeurs associées aux caractères sont binarisées et l'on obtient une table de vérité. Les risques présentés par la clé dichotomique sont réduits dans la clé tabulaire car tous les caractères y sont présentés simultanément pour chaque espèce (la clé est **polytomique**)⁷. Une erreur faite sur un caractère peut être corrigée grâce aux autres caractères. De plus, l'utilisateur a une certaine liberté dans le choix des caractères qu'il va utiliser. Alors que dans la clé dichotomique, le caractère employé à chaque ligne est choisi une fois pour toutes par l'auteur de la clé, dans une clé tabulaire l'utilisateur est libre d'écarter certains des caractères, soit parce

⁷ En intelligence artificielle, cela reviendrait à raisonner en termes d'attributs valeurs.

dans une clé tabulaire l'utilisateur est libre d'écarter certains des caractères, soit parce qu'il ne leur fait pas confiance, soit parce que ces caractères n'étaient pas observables dans le spécimen à identifier. De plus, l'arrangement en tableau permet à l'auteur de la clé de noter une certaine variabilité spécifique. La clé tabulaire est facile à mettre à jour lorsque de nouvelles espèces sont décrites dans le genre. Il suffit d'ajouter les lignes au tableau. Le gros inconvénient de ce type de clé est qu'elles cessent d'être d'un emploi pratique quand elles comportent plus d'une ou deux douzaines d'espèces. En effet, on aurait un nombre de lignes trop important dans le tableau et cela aurait pour conséquence de ralentir considérablement le processus d'identification.

Exemple : Clé tabulaire de quelques espèces de la famille *Citharinidae* [LEVEQUE & al, 91]

	Présence- écailles	Couleur- Corps	Présence- barbillons	Nombre- dorsales	Struct-Nag- antérieure
chelaethiops	oui	argenté	non	1	?
eburneensis	oui	argenté	non	2	rayonnée
ansorgii	oui	jaunâtre	non	2	rayonnée
distichodoïdes	oui	jaunâtre	non	2	rayonnée

Conclusion :

Pour pallier aux inconvénients des clés dichotomiques et tabulaires, il s'avère nécessaire d'avoir recours aux méthodes informatisées⁸. Il s'agit en fait, d'une *identification assistée par ordinateur* [LEBBE & al, 89].

⁸ On obtient ainsi des méthodes automatiques basées sur l'utilisation de clés dichotomiques et tabulaires.

2.2.2. Méthodes informatiques appliquées en systématique

De nombreux systèmes informatiques ont été construits dans un but de discrimination et d'identification. Les programmes destinés à la *construction de clés d'identification*⁹ sont des programmes de discrimination, ceux qui *utilisent les clés d'identification*¹⁰ sont des programmes d'identification. Ainsi nous pouvons citer pour les programmes de construction de clés d'identification [PANKHURST, 70], [DALLWITZ, 74], [MORSE, 74], et [PANKHURST, 88]. Quant aux programmes d'identification interactive nous pouvons citer [PANKHURST & AITCHISON, 75] [LEBBE, 84] et [DALLWITZ, 89]. Il existe, cependant, des systèmes intégrés dans le cadre des bases de données PANDORA [PANKHURST, 93] qui font aussi bien de la discrimination que de l'identification. Nous allons à titre d'exemple, décrire le système MAKEY [VIGNES, 91] pour la partie discrimination et les systèmes NEMAID [FORTUNER, 86] et INTKEY [DALLWITZ, 89] pour la partie identification.

Le système MAKEY

MAKEY [VIGNES 91], [LEBBE & VIGNES 91] est un programme de construction de clés d'identification. En effet il a pour but de créer des graphes d'identification de spécimens à partir de descriptions d'espèces déjà acquises. Il utilise comme entrée, des bases de connaissances au format XPER [LEBBE, 84]. Ces bases sont constituées d'un ensemble de descriptions de concepts et de connaissances supplémentaires. Ces descriptions acceptent des disjonctions de valeurs (prise en compte de la variabilité intra-classe), des relations entre descripteurs et des valeurs de typicité d'une valeur dans un taxon.

L'algorithme utilisé est descendant, et procède pas à pas sans retour arrière. La méthode de création du graphe est une heuristique consiste à sélectionner pas à pas le "meilleur" descripteur pour former un nœud du graphe selon les critères à optimiser, ensuite à créer les branches partant de ce nœud et à continuer le processus dans chaque branche jusqu'à ce qu'un critère d'arrêt soit atteint.

⁹ En apprentissage, les programmes de construction de clé correspondent aux systèmes d'apprentissage supervisés tels que ID3 [QUINLAN, 83] ou CHARADE [GANASCIA, 87]. Étant donné des descriptions d'exemples (dont les classes sont connues) il s'agit d'apprendre des caractéristiques discriminantes.

¹⁰ Les programmes d'identification correspondent en intelligence artificielle, aux systèmes experts destinés au diagnostic [SHORTLIFFE, 76] ou encore aux systèmes de raisonnement à partir de cas [KOLODNER & al, 85].

Selon les buts que souhaite atteindre l'utilisateur, MAKEY peut créer différents graphes d'identification (par exemple : identifier uniquement les genres à partir de la description des espèces) à partir desquels il peut élaborer des diagnostics.

Remarque :

Dans [NIQUIL, 93] on utilise MAKEY pour guider l'acquisition des connaissances, grâce à une technique de discrimination à partir de descriptions de concepts.

Le système NEMAID

NEMAID [FORTUNER, 86] est un programme d'identification. Il utilise un coefficient de similarité entre les spécimens à identifier et toutes les espèces du genre auquel ils appartiennent. D'une façon générale, un coefficient de similarité permet pour chaque caractère, la comparaison de sa valeur chez le spécimen à identifier et chez une espèce du même genre et leur attribue une note de 0 (totalement différents) à 1 (parfaitement semblables). L'idée fondamentale de NEMAID est qu'il faut tenir compte de la variabilité intra-spécifique dans le calcul de la similarité. Celle-ci est prise en compte différemment selon que les caractères soient quantitatifs ou qualitatifs.

Les avantages de NEMAID sont surtout ce que l'on appelle en anglais "*the graceful degradation*"¹¹ : une erreur n'entraîne pas la dégradation brutale des performances du système. Si par exemple, l'utilisateur fait une erreur sur dix caractères utilisés, l'espèce correcte sera encore donnée comme étant 90% semblable au spécimen à identifier. Un autre avantage est la grande liberté donnée à l'utilisateur dans les choix de caractères qu'il veut utiliser, et dans la définition des limites de la variabilité.

Les inconvénients sont, d'abord la nécessité d'entrer un nombre relativement élevé de caractères pour bénéficier de la "graceful degradation". D'autre part, la réponse correcte n'est pas nécessairement en tête de liste. Une fois la liste obtenue, il est indispensable que l'utilisateur aille étudier dans la littérature la description complète des espèces les plus semblables afin de décider qu'elle est la bonne. Ceci implique qu'il ait l'expertise nécessaire pour faire ce choix en toute connaissance de cause.

¹¹ Également appelée "traitement du bruit"

Afin de pallier aux inconvénients de NEMAID, la station de travail experte NEMISYS a été créée avec l'aide de deux informaticiens Jim Diederich et Jack Milton (de l'université de Californie à Davis) [FORTUNER, 93a].

Le système INTKEY

INTKEY est un système d'identification interactive¹². Il peut être utilisé de deux façons, soit pour l'identification, soit pour la recherche d'informations. La différence n'est pas importante et réside dans le traitement des valeurs manquantes et des valeurs qui se chevauchent. Ainsi, quand on utilise, lors de l'identification, un caractère inconnu pour un taxon¹³ donné, ce dernier ne doit pas être éliminé. En revanche, durant la recherche d'information, si le caractère considéré n'est pas effectivement enregistré le taxon doit être éliminé. Il existe des procédures qui créent des disjonctions ou des conjonctions de caractères lors de la recherche d'informations. Par exemple, il est tout à fait possible de rechercher des taxons qui ne possèdent pas telle valeur pour un caractère donné.

Dans INTKEY, on associe à certains caractères des poids qui sont utilisés dans le calcul de similarité et on peut aussi bien lister les similarités que les différences entre taxons. On peut aussi trouver les descriptions caractéristiques des taxons si on le désire.

Conclusion

Après avoir fait un tour d'horizon sur la classification en biologie, examinons les différentes acceptions du terme classification en apprentissage symbolique automatique, d'une part, et en analyse des données, d'autre part.

¹² INTKEY est issu de la version 3 du système ONLINE de Richard Pankhurst [PANKHURST & AITCHISON, 75] à laquelle ont été ajoutées des caractéristiques supplémentaires.

¹³ Rappelons qu'un taxon est une unité taxinomique reconnue par les codes internationaux, par exemple famille, genre ou espèce.

2.3. LA CLASSIFICATION VUE COMME UN PROCESSUS DE CONSTRUCTION DE CLASSES

Le nombre et la diversité des objets d'un domaine donné nécessitent de procéder à des regroupements, c'est à dire *définir des classes* d'objets, leur *attribuer un nom* et *produire une description caractérisant la classe*.

2.3.1. Apprentissage symbolique automatique (classification conceptuelle)

En *apprentissage non supervisé*, le système dispose d'un ensemble d'observations (des exemples non classifiés) qu'il doit structurer hiérarchiquement en fonction de leur ressemblance. Il construit ainsi, un ensemble de classes dont il donne une définition intensionnelle. Cette tâche est appelée *Classification conceptuelle*.

La classification conceptuelle [MICHALSKI & STEPP, 83] non incrémentale peut se définir de la façon suivante.

Étant donné un ensemble d'observations, on veut construire :

- Un ensemble de classes (concepts) regroupant ces observations et qui maximisent les similitudes intra-classes et les différences inter-classes,
- Une définition symbolique (intensionnelle) de chacune des classes,
- Une organisation de ces classes sous une forme hiérarchique.

On peut citer de manière non exhaustive les systèmes suivants : CLUSTER/2 [MICHALSKI & STEPP, 83], WITT [HANSON & BAUER, 89], COCLUSH [GEY, 91], etc.

Par ailleurs, après les travaux de Lebowitz [LEBOWITZ, 87] et de Fisher [FISHER, 87] sur les systèmes UNIMEM et COBWEB, on s'est intéressé à l'incrémentalité, ce qui a conduit à l'émergence de nouveaux systèmes traitant de la *classification conceptuelle incrémentale*¹⁴. Nous citons de manière non exhaustive, les

¹⁴ La classification conceptuelle incrémentale est aussi appelée formation de concepts.

systèmes CLASSIT [GENNARI & al, 89]¹⁵ et ADECLU [DECAESTECKER, 91] qui sont des descendants de UNIMEM . Dans la méthode de classification conceptuelle, l'ensemble des objets à traiter est immédiatement disponible, alors qu'en formation de concepts les objets sont au contraire traités les uns après les autres et au fur et à mesure de leur apparition. La formation de concepts s'appuie alors sur un processus de classification suivi de la restructuration de la hiérarchie des concepts en cours de construction.

Outre le fait qu'elles soient capables de traiter les exemples un à un, ces méthodes déterminent elles même le nombre de classes "optimal" à un niveau donné de la hiérarchie.

Nous choisissons maintenant de décrire le système COBWEB, un des premiers représentants de la famille des systèmes de formation de concepts et, de ce fait, l'archétype maintes fois cité.

Le système COBWEB

Le système COBWEB construit des classifications, dont les classes sont à un niveau donné deux à deux exclusives. La prédiction et la mise en place des attributs (donc la construction des classes) s'appuient sur des critères probabilistes. En fait, la description de chaque classe consiste à mémoriser, pour chacun des attributs, les fréquences d'apparition de ces différentes valeurs et ce, pour l'ensemble des exemples couverts par cette classe.

Lorsqu'un nouvel exemple est présenté au système, on essaie tout d'abord de le classer et si on n'y parvient pas, on envisage alors de modifier la classification, soit en *créant* une nouvelle classe, soit en *fusionnant* des classes ou encore, en *divisant* des classes. Le choix entre ces diverses possibilités est effectué à l'aide d'une heuristique (la *category utility*) qui permet d'évaluer la qualité des différentes solutions.

L'exploration de l'espace de recherche repose sur une méthode de gradient ("hill-climbing") et la présence d'opérateurs (Fusion/Division) permet de "simuler" un retour arrière et donc de diminuer la dépendance de la classification obtenue vis à vis de l'ordre de présentation des exemples.

¹⁵ Cet article contient une étude générale puis particulière de ces systèmes (y compris Unimem, Cobweb et Classit)

2.3.2. Analyse des données

L'analyse des données est une discipline dont le but est de mettre en relief les structures pertinentes de grands ensembles de données [CRAUSER & al, 89]. L'analyse conduira souvent à situer des objets les uns par rapport aux autres et à mesurer l'influence d'un certain nombre de paramètres sur cette répartition.

Parmi les méthodes d'analyse des données en classification numérique, nous distinguons les méthodes non hiérarchiques qui produisent directement une partition en un nombre fixé de classes et les méthodes hiérarchiques qui produisent des partitions en classes de plus en plus spécialisées.

2.3.2.1. Méthodes non hiérarchiques (partitionnement)

La classification non hiérarchique consiste à regrouper n individus en k classes, le nombre de classes étant généralement connu à l'avance [BOUROCHE & SAPORTA, 87], de sorte que les individus d'une même classe se ressemblent le plus possible et que les classes soient séparées au mieux. Il est cependant nécessaire de disposer d'une distance entre les individus afin de les partitionner en classes, ainsi que d'un critère global mesurant la qualité de la partition.

Le même algorithme est réutilisé, successivement et plusieurs fois, afin de tendre vers une solution pertinente. Les différentes méthodes non-hiérarchiques comprennent les étapes suivantes [CRAUSER & al, 89] :

- fixer le nombre de classes k souhaité,
- choisir une configuration initiale en k classes. Cette configuration est souvent construite à partir d'un choix aléatoire de k individus (appelés noyaux). Les autres individus sont regroupés en fonction du noyau le plus proche,
- on calcule ensuite les distances entre un individu et les centres de gravité des groupes de la configuration précédente. Chaque individu est ensuite réaffecté à la classe du centre de gravité le plus proche (d'autres critères peuvent être utilisés). Par la méthode des centres de gravité (centroïdes) fixes, les centres de gravité des classes sont calculés une fois que tous les individus ont été réaffectés à un groupe. Par la méthode des centres de gravité mobiles, ils sont recalculés après chaque allocation d'un individu à un groupe,
- on réitère la même opération jusqu'à obtenir un certain optimum ou une configuration stable.

La méthode des nuées dynamiques [DIDAY & al, 80] généralise ces méthodes en permettant aux noyaux d'être formés de plusieurs individus. Les noyaux sont ensuite recalculés en minimisant une fonction d'inertie du groupe (par exemple la somme des variances). Cette méthode permet une reconnaissance efficace de la structure réelle des classes.

2.3.2.2. Méthodes hiérarchiques (*descendantes et ascendantes*)

Les analyses hiérarchiques cherchent à représenter les individus par un ensemble de parties hiérarchiquement emboîtées [CRAUSER & al, 89]. Il est possible d'obtenir une hiérarchie de partitions par deux types de méthodes (la deuxième étant plus utilisée que la première) :

- soit par une méthode *descendante* qui procède par *divisions* successives allant du groupe aux individus. Au départ, tous les individus forment un groupe. Par la suite, on forme différents groupes en scindant l'ensemble des individus dans le but d'obtenir des groupes les plus hétérogènes possible entre eux. A chaque étape, on reproduit de façon similaire une division des groupes obtenus à l'étape précédente.

- soit par une méthode *ascendante* qui procède par *regroupements* successifs allant des individus vers le groupe. On obtient les premiers groupes en réunissant les individus les plus proches, puis on fusionne les groupes les plus semblables etc.

Conclusion

Nous constatons que selon les domaines considérés, la classification vue comme un processus de construction de classes fait appel à de nombreux systèmes qui produisent des taxinomies plus ou moins "bonnes" . Il nous faut maintenant, mettre en évidence les mécanismes d'exploitation de ces structures : c'est l'objet de la section suivante.

2.4. LA CLASSIFICATION VUE COMME UN PROCESSUS DE RAISONNEMENT SUR LES CLASSES

Après avoir passé en revue certaines méthodes de construction de classes, nous allons nous intéresser au processus d'exploitation et de raisonnement sur ces classes et plus précisément au problème de discrimination de classes déjà construites. La discrimination revient à trouver un jeu de traits discriminants qui permet de distinguer au

mieux une classe par rapport aux autres. Pour cela, on construira, comme exemples, des arbres de décision ou des systèmes de règles .

2.4.1. Apprentissage de concepts

En *apprentissage supervisé*, le système dispose d'un ensemble d'exemples (et éventuellement de contre-exemples), dans lequel chaque exemple correspond à la description d'une situation du domaine étudié et à l'expertise qui lui est associée (appelée classe ou concept). Le but de l'apprentissage est de construire, à l'aide d'algorithmes de généralisation, des fonctions de reconnaissances capables de *caractériser* les différents concepts, c'est *l'apprentissage de concepts*. Citons comme exemples ID3 [QUINLAN, 83], AQ [MICHALSKI & STEPP, 83], CHARADE [GANASCIA, 87] qui utilisent une représentation en logique des propositions et ALLY [NICOLAS, 87], AGAPE [GANASCIA, 87], FOIL [QUINLAN, 90], KBG [BISSON, 93] qui utilisent une représentation en logique des prédicats.

2.4.1.1. Arbres de décision (ID3)

A partir de la donnée d'une collection d'exemples, ID3 apprend des procédures de classification exprimées sous la forme d'*arbres de décision* [QUINLAN, 83].

Les exemples sont décrits par un ensemble de couples (<attribut> <valeur>). Chaque attribut possède un nombre fini de valeurs mutuellement exclusives. Supposons que l'on ait une collection d'individus décrits par trois attributs ("taille", "cheveux", "yeux"). Les ensembles de valeurs possibles des attributs "taille", "cheveux", "yeux" étant respectivement {petit, grand}, {blonds, bruns, roux}, {bleus, marrons}.

Chaque exemple est de plus associé à un concept particulier encore appelé "classe" de cet exemple. Les individus peuvent appartenir à la classe nordique notée "N" ou à la classe méditerranéenne notée "M". Notons que les classes sont données au départ.

Il s'agit donc d'élaborer à partir de ces huit exemples une procédure de classification exprimée sous forme d'arbre de décision qui permette de déterminer la classe "nordique" ou "méditerranéenne" d'un individu pour lequel la taille ainsi que la couleur des yeux et celle des cheveux sont connues.

Exemples :

I1 = petit, blonds, bleus, N

- I2 = petit, bruns, bleus, M
- I3 = grand, bruns, marrons, M
- I4 = grand, blonds, marrons, M
- I5 = grand, bruns, bleus, M
- I6 = petit, blonds, marrons, M
- I7 = grand, roux, bleus, N
- I8 = grand, blonds, bleus, N

A chaque étape de l'algorithme, on choisit un attribut présent dans les descriptions d'exemples auquel on associe un nœud de l'arbre de décision, puis on crée autant de fils à ce nœud qu'il y a de valeurs associées à cet attribut. Dans ID3 l'attribut qui partitionne l'ensemble des exemples est celui qui optimise *localement* le gain d'information (celui qui discrimine le plus). Le processus est réitéré sur les nœuds ainsi créés, jusqu'à ce que les feuilles de l'arbre obtenu correspondent toutes à des ensembles d'exemples appartenant à la même classe (voir figure ci-dessous).

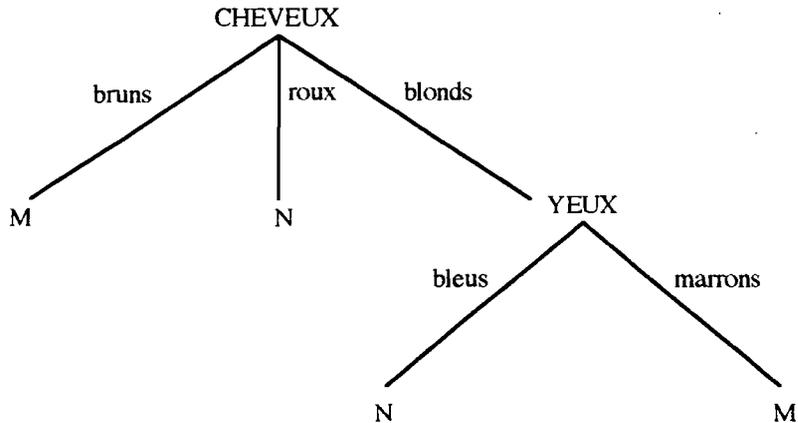


Figure 2.1. Arbre de décision

Si les arbres de décision sont faciles à construire avec cette méthode, ils sont malheureusement difficiles à lire dès que leur taille est importante.

L'algorithme ID3 reste limité à des problèmes de classification très simples à cause de la difficulté qu'il y a à introduire des attributs numériques, et à cause de l'aspect déterministe des procédures apprises [GANASCIA, 87]. Les arbres de décision obtenus

avec ID3 ne peuvent être enchaînés les uns aux autres, il s'agit d'un ensemble de règles isolées et non d'un système de règles.

Contrairement au système ID3, le système CHARADE engendre des systèmes de règles, il peut néanmoins être vu comme la généralisation dans un espace multidimensionnel de la méthode de QUINLAN. Le prochain paragraphe aura pour objet la description de ce système.

2.4.1.2. Construction de systèmes de règles (CHARADE)

Un système de règles est constitué d'un ensemble de règles qui s'enchaînent. Il possède des propriétés globales telles que la complétude, l'absence de redondances et de circularités, etc. Le système d'apprentissage CHARADE génère automatiquement un système de règles à partir d'un ensemble d'exemples, ces règles vont servir à caractériser les classes auxquelles appartiennent les exemples. CHARADE est basé sur l'utilisation de deux treillis distributifs: l'un pour la base d'apprentissage (i.e., l'ensemble des exemples) et l'autre pour l'espace de description.

Outre l'économie de représentation qu'elles procurent, les propriétés de ces treillis sont mises à profit lors de la phase d'apprentissage, en vue de faciliter la détection de régularités. En effet, une régularité correspond à une "corrélation empirique" (corrélation entre descripteurs) observée sur l'ensemble d'apprentissage. Si tous les exemples qui possèdent un descripteur d1 dans leur description possèdent aussi le descripteur d2 dans leur description, on peut déduire que d1 implique d2 sur l'ensemble d'apprentissage. Le principe d'induction utilisé consiste en une généralisation de cette relation à l'espace de description entier.

Même si on sait détecter et qualifier les régularités, il n'est pas question de détecter toutes les régularités possibles car, d'une part, la procédure d'exploration serait exponentielle (2^n conjonctions de descripteurs, n représentant le nombre de descripteurs), d'autre part le nombre de règles engendrées serait important et certaines de ces règles seraient redondantes ce qui conduirait à la fois à des duplications inutiles et à des erreurs.

L'algorithme explore l'espace de recherche du général au spécifique, on commence par les descriptions conjonctives les plus simples en allant vers les descriptions conjonctives les plus complexes. Ainsi, l'espace de description formé par les trois descripteurs d1, d2 et d3, sera exploré dans l'ordre suivant $\{\emptyset, d1, d2, d3, d1 \text{ et } d2, d1 \text{ et } d3, d2 \text{ et } d3, d1 \text{ et } d2 \text{ et } d3\}$.

Contraintes sur l'exploration

- Utilisation des propriétés de **règles** :

Dans la mesure où les propriétés des règles traduisent l'équivalence entre règles, leur mise en oeuvre permet d'éliminer les règles inutiles. C'est donc là ce qui permet d'éviter les redondances tout en limitant l'exploration de l'espace de description.

- Utilisation des propriétés du **système de règles** que l'on désire construire:

De même que l'on introduit les propriétés des règles pour éviter l'exploration des descriptions qui sont a priori infécondes et pour restreindre la création de règles à celles qui sont vraiment utiles, on utilise les propriétés de la base de connaissances que l'on souhaite engendrer. C'est là l'aspect le plus original du système CHARADE car, grâce à ces propriétés, il donne une sémantique claire aux heuristiques de généralisation habituellement utilisées tout en introduisant de nouvelles heuristiques. De fait, ces propriétés dérivent des caractéristiques opérationnelles de la base de connaissances que l'on veut obtenir. Nous pouvons quelques propriétés :

le **bruit** : nombre d'exemples minimum que doit couvrir un descripteur ou une conjonction de descripteurs.

le **nombre de couches** : ou limitation de la profondeur de recherche ; cette propriété fixe le nombre maximal de descripteurs que peut comporter une conjonction de descripteurs en position de prémisse de règles.

le **fractionnement des classes** : cette propriété introduit un coefficient de fractionnement de classes qui indique la proportion minimale d'exemples de chaque classe C couverts par la prémisse d'une règle émise.

la **couverture des exemples** : cette propriété limite le jeu de règles acquises à celles qui permettent de classer les exemples.

Sans ces propriétés, on obtiendrait une collection de règles dont beaucoup seraient inutiles, certaines correspondant à des coïncidences accidentelles, d'autres n'offrant aucun intérêt pratique. Ces caractéristiques sont parfois utilisées pour construire des outils d'aide à l'acquisition des connaissances, vérifiant la pertinence des règles fournies par l'utilisateur. CHARADE traduit ces caractéristiques sous forme de contraintes de la procédure d'exploration de l'espace de description.

2.4.2. Le raisonnement par classification dans les systèmes de représentation à base d'objets

Dans une représentation à base d'objets, le raisonnement par classification [RECHENMANN, 88], [AGUIRRE, 89], [MARINO & *al*, 90], [CAPPONI, 93] opère sur la hiérarchie des objets et cherche à mettre en évidence les dépendances implicites qui existent entre les objets [HATON & *al*, 91]. Dans les nombreux langages de représentation qui permettent d'organiser des connaissances en hiérarchie d'héritage, comme KL-ONE [BRACHMAN & SCHMOLZE, 85] et ses descendants ou encore SHIRKA [RECHENMANN, 88], le raisonnement par classification permet la recherche, la gestion et la maintenance des informations ou encore l'acquisition dynamique de nouvelles connaissances.

La mise en œuvre du raisonnement par classification repose sur un cycle comprenant trois étapes. Dans ce qui suit, un objet est générique ou spécifique, et il appartient à une hiérarchie H qui repose sur un ordre partiel noté \leq et appelé *subsumption* [NAPOLI, 94] :

- *Instanciation* : création d'un nouvel objet, noté X , en s'appuyant sur un nouveau résultat ou sur des informations externes ; X peut représenter une requête à satisfaire ou encore un objet à reconnaître,
- *Classification* : parcours de la hiérarchie H et mise en place de X dans H , en accord avec l'ordre partiel \leq ; X a pour ascendants directs ses subsumants les plus spécifiques (SPS), et pour descendants immédiats ses subsumés les plus généraux (SPG),
- *Propagations* : la mise en place de X provoque des modifications dans H , notamment au niveau des liens qui existent entre les objets. Ces modifications déclenchent des opérations de mise à jour d'objets interdépendants et/ou la production de nouveaux objets, ce qui ramène le cycle à sa première étape. Le cycle continue tant que les informations disponibles permettent de produire de nouveaux objets à classifier et que le but associé au problème à résoudre n'est pas atteint.

2.4.3. Analyse des données

La section suivante abordera une méthode d'analyse des données particulière, l'analyse discriminante.

2.4.3.1. Analyse discriminante

Quand on cherche à discriminer des classes (groupes d'individus) déjà données et caractérisées par un grand nombre de variables, on utilise un ensemble de techniques appelé *analyse discriminante*. On se trouve face à un double problème : d'une part, définir les variables qui permettent le mieux de distinguer les différences entre groupes d'individus, et d'autre part, connaissant les caractéristiques d'un individu non classé, trouver son groupe d'appartenance. La forme du problème posé et les méthodes qui en résultent dépendent du but de la discrimination [CELEUX & al, 89].

L'analyse discriminante se situe au carrefour des méthodes descriptives et des méthodes explicatives [CRAUSER & al, 89]. Elle a un but *descriptif* lorsqu'il s'agit de déterminer si les groupes diffèrent sur l'ensemble des variables quantitatives et si le nombre de ces dernières peut être réduit. Elle a un but *explicatif (décisionnel)* lorsqu'il s'agit d'affecter un individu supplémentaire à un des groupes sur la base de ces variables quantitatives.

Il existe plusieurs méthodes d'analyse discriminante. Nous citerons par exemple, les méthodes métriques qui comprennent l'analyse discriminante linéaire et la discrimination quadratique.

Cas de l'analyse discriminante linéaire :

Supposons que les exemples soient décrits au moyen de p attributs numériques (i.e. que les exemples sont des points de \mathbb{R}^p). Si l'on suppose que les classes sont linéairement séparables, l'analyse discriminante linéaire va chercher des hyperplans permettant de séparer les différentes classes.

2.5. CONCLUSION

Nous avons abordé dans ce chapitre la classification en biologie en examinant les aspects discrimination et identification et en abordant successivement les méthodes traditionnelles et les méthodes informatisées.

Nous nous sommes intéressés ensuite à la classification en tant que processus de construction de classes et en tant que processus de raisonnement sur ces classes, aussi

bien dans le domaine de l'apprentissage symbolique que dans celui de l'analyse des données.

Pour terminer ce bref tour d'horizon sur les différentes acceptions données au terme de "classification", nous dirons que dans le cadre de notre système METIS la classification est vue comme un processus de raisonnement sur les classes. C'est à dire que l'on part d'une classification déjà existante et que l'on va expliciter incrémentalement une base de connaissances en construisant interactivement des caractérisations de classes, la taxinomie du domaine servant de fil conducteur pour le processus d'explicitation des connaissances et d'élément de dialogue avec l'expert. Le graphe de discrimination construit par METIS va permettre l'identification des individus appartenant aux classes.

Les différentes techniques d'explicitation des connaissances et plus précisément celles qui s'intéressent aux classifications sont présentées au chapitre suivant.

3. EXPLICITATION DES CONNAISSANCES

Ce chapitre a pour objectif d'abord de situer de manière claire le statut de l'explicitation des connaissances dans le cadre de l'acquisition des connaissances, ensuite de présenter les différentes techniques d'explicitation des connaissances. Nous examinerons donc, de manière non exhaustive¹, les méthodes contraintes et non contraintes d'explicitation des connaissances ; nous nous intéresserons aussi aux outils qui découlent de l'implémentation de ces méthodes, en mettant l'accent sur leurs caractéristiques et limites. Enfin dans le but de nous positionner par rapport à ces outils nous les comparerons avec le système METIS (voir chapitre 5).

3.1. INTRODUCTION

L'acquisition des connaissances consiste à extraire les connaissances et à les organiser au moyen de techniques manuelles, semi-automatiques (interactives) ou automatiques (apprentissage). L'acquisition des connaissances fait appel aux notions d'explicitation² des connaissances et de modélisation³ (voir figure 3.1.).

Pour nous, l'explicitation des connaissances consistera à acquérir les connaissances de manière semi automatique, c'est-à-dire à construire conjointement le langage de description⁴ (extraction des concepts et relations du domaine) et le modèle des

¹ Nous ne nous intéresserons pas aux techniques de test de mémorisation d'échelonnage multidimensionnel, le brainwriting, de regroupements hiérarchiques de Johnson etc. Le lecteur trouvera une description de toutes ces techniques notamment dans [AUSSENAC, 89] [BOOSE, 89], [BOY, 91] et [DIENG, 92].

² C'est ainsi que nous traduisons le mot anglais "elicitation".

³ La modélisation comporte deux aspects imbriqués : une modélisation du domaine et une modélisation du processus de résolution, du raisonnement [WIELINGA & BREUKER, 86], [STEELS, 90].

⁴ "Parmi les tâches qui préludent à la construction d'une base de connaissances, il en est une de première importance, c'est la définition du langage de description : tous les concepts opératoires du micro-domaine d'expertise doivent être isolés puis nommés" [GANASCIA, 90].

connaissances du domaine (organisation des concepts et relations du domaine au sein d'un graphe) en interagissant avec l'expert.

Nous avons délibérément choisi de traduire le terme anglais *Knowledge elicitation*⁵ par l'expression *explicitation des connaissances* car à notre avis, c'est celle qui correspond le mieux, même si certains auteurs parlent de recueil des connaissances ou d'extraction des connaissances [DIENG, 92].

Par ailleurs, l'activité de modélisation consiste à construire un modèle d'expertise comprenant un modèle de la méthode de résolution de problèmes et un modèle des connaissances du domaine.

Le modèle de résolution de problèmes appelé "modèle conceptuel" est destiné à modéliser les inférences, les tâches et les stratégies propres au domaine d'application indépendamment de toute implémentation. Nous citons, pour mémoire, et sans rentrer dans les détails, les travaux de Clancey sur la *classification heuristique* [Clancey, 85], ceux de Chandrasekaran sur les *tâches génériques* [CHANDRASEKARAN, 87], ceux de Breuker et Wielinga sur les *modèles d'interprétation* de la méthodologie KADS [WIELINGA & BREUKER, 86] et enfin ceux de McDermott sur les "Roles limiting Methods"⁶ [McDermott, 88].

Le modèle de l'expertise s'attache à mettre en évidence les différents types de connaissances et les relations les liant, en construisant un modèle non exécutable de l'application [HAMROUN & CARRE, 94]. Les travaux sur les connaissances du domaine et leur réutilisabilité sont plus récents. Nous distinguons de manière non exhaustive, les travaux de [WEILINGA & al, 92] sur CommonKads et les travaux de [GRUBER, 93].

⁵ En effet selon le dictionnaire (HARRAPS) le verbe *to elicit* possède plusieurs sens qui sont :

- obtenir une réponse,
- provoquer une réponse,
- tirer les faits au clair,
- découvrir la vérité.

Dans le cadre de cette étude nous ne retiendrons que les trois premiers sens. Quant au verbe *explicititer*, il a pour sens : énoncer formellement (voir petit Robert). Par ailleurs si on examine l'étymologie latine du mot *explicititer* (ex : hors de, *plicare* : plier), on s'aperçoit que *explicititer* veut dire *déplier*. Ainsi, *explicititer* des connaissances revient à les déplier, à les rendre claires, formelles, précises, non ambiguës...

⁶ Appelée méthode à limitation de rôles. Elle consiste à donner à chaque entité du domaine un rôle bien précis.

Au demeurant, les notions d'explicitation des connaissances et de modélisation sont inextricablement liées, car l'explicitation des connaissances fait appel à la modélisation et vice-versa.

Ainsi, par exemple, l'outil dit d'explicitation des connaissances ALTO [MAJOR & REICHGELT, 90] basé sur la technique de "laddering"⁷ (voir § 3.3.3.) modélise le domaine sous forme de hiérarchie de concepts, alors que l'outil d'acquisition de connaissances SALT [MARCUS & al, 85] [MARCUS & McDERMOTT, 93] modélise des problèmes de configuration et d'ordonnement basés sur la méthode "propose and revise" en utilisant des techniques d'explicitation des connaissances.

La distinction entre l'explicitation des connaissances et la modélisation joue sur l'importance accordée à l'explicitation des connaissances lors de l'organisation (représentation et structuration) des connaissances. On parlera alors d'outils d'explicitation de connaissances si la modélisation de la méthode de résolution de problèmes n'est pas l'aspect le plus important⁸.

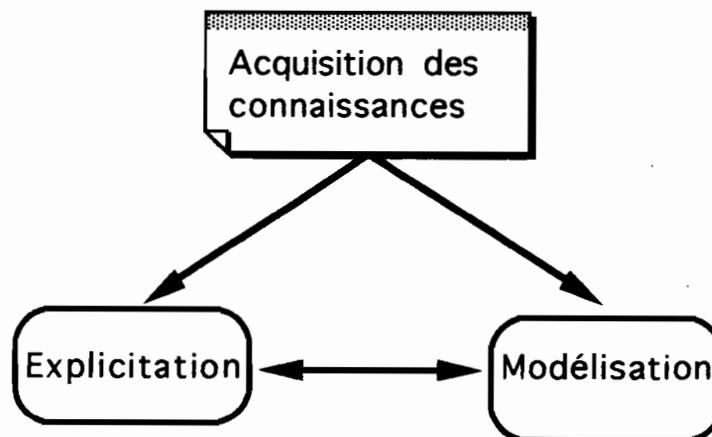


Figure 3.1. Schéma d'acquisition des connaissances

⁷ Il n'existe pas de traduction française dans la littérature. Pour notre part, si nous devons lui donner une traduction, ce serait "technique de l'échelle".

⁸ D'autres mettent plus l'accent sur la modélisation du processus de résolution de problèmes. De sorte que lorsque l'on parle d'acquisition guidée par les modèles [LEROUX & LAUBLET, 93b], on s'appuie en fait, sur la modélisation préalable de la méthode de résolution de problèmes

3.2. MÉTHODES NON CONTRAINTES

Les méthodes non contraintes sont appelées méthodes directes [AUSSENAC, 89] [FIRLEJ & HELLENS, 91] méthodes standards [MAJOR & REICHGELT, 90] ou encore "non contrived methods" [SHADBOLT & BURTON, 90]. Elles permettent à l'expert d'énoncer son savoir de manière "naturelle", c'est-à-dire sans passer par la contrainte d'une structure préalable définie par le cogniticien. On parlera alors de transfert direct de la connaissance.

De manière non exhaustive, nous retrouvons dans ces méthodes : les techniques d'interview, les techniques d'analyse de protocole, l'observation directe, l'utilisation de questionnaires, les verbalisations rétrospectives, l'introspection, le "brainwriting" etc.

3.2.1. Technique d'interview

C'est sans doute, la technique la plus communément utilisée pour expliciter la connaissance de l'expert ; elle permet, par interaction directe, d'obtenir rapidement les concepts généraux du domaine. L'interview peut suivre une ligne directrice fixée à l'avance. Il s'agit alors, d'une interview structurée. De nombreux conseils sont donnés dans [GALLOUIN, 88] [HARMOND & KING, 87] sur la façon de conduire un entretien.

Limites de la technique d'interview

Les limites de la technique d'interview résident dans le choix de l'expert et dans le manque d'objectivité de l'information fournie par l'expert. En effet, l'expert décrit ce qu'il pense être son raisonnement et les objets ou les concepts qu'il manipule avec un décalage dans le temps et l'action, il peut donc avoir lui-même une vision déformée. De plus, le cogniticien doit se familiariser avec le domaine, car sans cela, la conduite de l'interview devient très difficile.

3.2.2. Technique d'analyse de protocoles

Ici, le cogniticien cherche à extraire, puis à étudier tout ce que pense l'expert en situation de travail. Pour cela, l'expert est enregistré pendant qu'il simule la résolution d'un problème réel en s'efforçant, ainsi qu'on le lui a demandé, de "penser tout haut" [NEWELL & SIMON, 72]. Le cogniticien analyse ensuite ses notes et les informations recueillies sur bande magnétique.

La technique d'analyse de protocole passe par quatre étapes : la collecte du protocole, sa transcription, son codage et enfin, son interprétation.

Limites de la technique d'analyse de protocole

L'intérêt majeur de cette technique est la mise en valeur des hésitations de l'expert et de ses réponses spontanées. Ceci dit, le processus mis en œuvre est très long surtout lors du codage et donc, très coûteux en durée et en charges de travail. Par ailleurs il y a peu de spécialistes susceptibles de faire des analyses rigoureuses, car la formation aux techniques d'analyse de protocole est très longue [ZACKLAD, 93].

Au demeurant, même si selon [ERICSSON & SIMON, 80], le fait de penser tout haut ne perturbe guère le processus de raisonnement, cette technique introduit quand même du bruit (risques d'omissions, d'incomplétude, ...) lors du codage.

3.2.3. Observation directe

Cette méthode repose sur un enregistrement vidéo de l'expert en situation de travail réel. L'observation peut porter sur les tâches effectuées, sur les données et documents utilisés et sur les actions produites [BOY & al, 88]. Cette technique est inspirée de l'analyse du travail en psychologie du travail. Elle permet de recueillir une grande quantité d'informations et met en évidence un modèle de comportement humain [BOY & RAPPAPORT, 87].

Limites de l'observation directe

La méthode présente plusieurs avantages : il y a une bonne appréhension de l'expertise, un reflet de l'activité réelle et enfin un accès à des connaissances non verbalisables. Néanmoins, la méthode nous semble très coûteuse et relève plus du travail d'un ergonome ou d'un psychologue que d'un cogniticien. De plus, elle ne convient pas à toutes les expertises. Ainsi, si l'on observe un analyste financier en train de traiter un dossier en vue d'accorder un crédit, on n'apprend rien. En effet, le regarder lire des chiffres et donner des notes ne nous renseigne en rien sur sa façon de penser.

Exemple : l'outil KRITON

KRITON [DIEDERICH & al, 88] est un outil destiné aux tâches d'analyse et permettant de créer des bases de connaissances pour l'environnement BABYLON [PRIMIO & BREWKA, 85]. Il permet l'explicitation des connaissances grâce à trois

techniques : l'analyse de protocoles pour acquérir la connaissance *procédurale* et *associative*, l'interview et l'analyse de textes pour acquérir la connaissance *statique*.

- *Analyse de textes* : les textes sont lus à partir d'un fichier, les noms y sont mis en valeur. Le cognicien peut inclure des noms dans la hiérarchie décrivant la façon dont le texte présente l'organisation des concepts du domaine.
- *Interview* : elle permet de compléter l'information recueillie à partir de l'analyse de textes ; elle est basée sur la technique des grilles-répertoires. Elle permet d'expliciter les attributs décrivant les concepts du domaine.
- *Analyse de protocoles* : quand l'expert résout un problème, il décrit ses actions, conclusions ou observations.

Le résultat de cette explicitation est une représentation intermédiaire sous forme de réseaux, où les nœuds correspondent aux concepts et les liens aux relations entre eux. On distingue des relations structurelles (connaissance statique explicitée à partir des interviews et de l'analyse de texte) et des relations associatives (explicitées grâce à l'analyse de protocole). Le tout est ensuite traduit dans le formalisme de BABYLON (schémas et règles de production).

3.3. MÉTHODES CONTRAINTES

Les méthodes contraintes correspondent aux méthodes indirectes [AUSSENAC, 89] [FIRLEJ & HELLENS, 91] [SHADBOLT & BURTON, 90] [MAJOR & REICHGELT 90]. Elles contraignent l'expert à énoncer son savoir à l'aide de structures établies au préalable telles que des tableaux, des grilles d'évaluation, des arbres, des réseaux, etc. Elles visent ainsi, à faire émerger des connaissances difficilement verbalisables.

Les méthodes contraintes offrent des solutions en complétant les méthodes non contraintes. on distingue les grilles-répertoires, le tri de cartes, l'échelonnage multidimensionnel, le regroupement hiérarchique etc.

3.3.1. Grilles-répertoires

Kelly [KELLY, 55] a conçu une technique d'entretien de psychothérapie clinique pour aider les patients à prendre conscience de la façon dont ils structuraient leur univers, c'est la technique des *grilles-répertoires*. Selon KELLY, dès qu'un patient perçoit

l'organisation des catégories sous lesquelles il classe ses expériences et son environnement, ce patient maîtrise mieux ces catégories, ce qui lui permet d'agir plus efficacement et, éventuellement de modifier l'organisation de ces catégories.

Appliquée à l'explicitation des connaissances, cette technique permet à un expert de prendre conscience des concepts selon lesquels il structure sa connaissance.

Principe :

L'expert commence par fournir un ensemble de 10 à 20 objets ou éléments à classer ou à structurer et ces objets doivent être bien définis. Pour chaque triplet possible des objets de la liste, l'expert cite une caractéristique qui distingue deux des objets du troisième, ainsi que la propriété opposée : prenons l'exemple de la température ; on obtient ainsi la paire <chaud, froid>. Le cognicien note cette *dimension* (couple formé par la caractéristique et son opposée) et l'évaluation de tous les objets sur une échelle de valeur de un à cinq dont les extrêmes correspondent à la caractéristique (5) et à son opposé (1).

Lorsque l'ensemble des caractéristiques applicables a été acquis, les résultats sont rassemblés dans une grille dont les lignes sont les dimensions et les colonnes, les objets. L'expert continue, alors, à évaluer chaque caractéristique pour chacun des objets jusqu'à ce que la grille soit complète.

Vient ensuite, une étape de formalisation, dans le sens où une matrice des "distances" entre objets est construite en sommant pour chaque couple d'objets les écarts des évaluations de toutes les caractéristiques. Cette matrice permet la création d'un arbre des "ressemblances" dont les nœuds représentent les distances calculées entre objets ou sous-ensembles d'objets proches et dont les feuilles sont les objets eux-mêmes.

La dernière étape est celle de l'analyse et du raffinement des connaissances : quand l'expert commente les résultats de la matrice et en cas de désaccord sur les liens entre objets ou caractéristiques, il peut décider de modifier les coefficients affectés à chaque objet. Il peut aussi, ajouter une nouvelle caractéristique pour différencier deux objets associés ; il peut aussi ajouter un objet comme contre-exemple, pour montrer que deux caractéristiques ne sont pas liées. Il est à noter que toute modification implique une réévaluation de l'ensemble des objets pour une caractéristique, ou de l'ensemble des caractéristiques pour un objet.

Exemple :

Dans le cas de l'explicitation d'une expertise de vendeur d'automobiles, on obtient la grille suivante :

	V1	V2	V3	V4	V5	V6	V7	V8
Maniable/Peu maniable	-		+		+			
Sportive/Non sportive					+	-	+	
Sûre/Légère		+		-		+		
Habitable/Peu habitable	+				-		+	
Confortable/Inconfortable			+			+		-
Silencieuse/Bruyante		+	+	-				
Nerveuse/Lourde					+	-	+	
Esthétique/Inesthétique	+						+	
Peu coûteuse/Chère	-							+

Remarque :

Le "+" sur la ligne i à la colonne j indique que le véhicule Vi possède la qualité Qj, un signe "-" indiquant le contraire.

Une fois la première grille établie, on demande à l'expert de noter sur une échelle de 1 à 5 tous les objets, pour toutes les qualités. On obtient la grille suivante :

	V1	V2	V3	V4	V5	V6	V7	V8
Maniable/Peu maniable	2	3	5	2	5	3	3	2
Sportive/Non sportive	2	2	2	1	5	2	3	1
Sûre/Légère	4	5	3	2	4	5	4	1
Habitable/Peu habitable	5	4	3	2	1	5	5	2
Confortable/Inconfortable	5	5	4	1	3	5	4	1
Silencieuse/Bruyante	4	4	5	2	2	4	4	1
Nerveuse/Lourde	2	2	3	1	5	2	4	1
Esthétique/Inesthétique	4	2	3	3	4	2	4	2
Peu coûteuse/Chère	1	1	4	4	2	1	1	5

Les étapes de formalisation et d'analyse vont opérer sur la dernière grille

Limites des grilles-répertoires

Bien que la technique des grilles-répertoires soit facilement automatisable, et que l'on obtienne rapidement des prototypes (maquettes) on observe les limites suivantes :

- la contrainte d'utiliser une cotation numérique,
- les attributs doivent être applicables à tous les éléments,
- il n'y a pas de vérification de la complétude,
- la méthodologie s'applique particulièrement bien aux problèmes de diagnostic d'interprétation et de classification qui utilisent une démarche analytique, mais elle n'est pas adaptée aux problèmes de synthèse tels que la planification, la configuration ou le contrôle de processus.

Les outils ETS et AQUINAS

La technique des grilles-répertoires a donné naissance à tout un panel d'outils d'explicitation des connaissances (ETS, AQUINAS) très efficaces mais restreints à la classification heuristique [CLANCEY, 85], [CLANCEY, 89a], [CLANCEY, 89b].

ETS [BOOSE, 84], [BOOSE, 89], [BOOSE & al, 89] est un outil d'explicitation des connaissances basé sur la technique des grilles-répertoires. Il permet de faire une interview automatique d'expert en vue de la construction automatique de règles de production. Tous les concepts doivent être de même niveau de granularité et les caractéristiques ne peuvent être multivaluées, en d'autres termes, le langage est décrit par une seule grille-répertoire proche du langage propositionnel.

AQUINAS [BOOSE, 89], [BOOSE & BRADSHAW, 93] est une généralisation et une amélioration d'ETS. Une fois qu'une grille est construite interactivement, AQUINAS aide l'expert à raffiner et à étendre la base de connaissances en appelant différents modules. Par exemple, les similitudes entre solutions et caractéristiques sont analysées pour aider l'expert à raffiner des distinctions utiles et à éliminer les redondances ou inconséquences. AQUINAS dispose aussi de grilles-répertoires hiérarchiques permettant de décomposer des problèmes complexes.

AQUINAS offre diverses fonctionnalités en plus de son ancêtre ETS : il intègre, en réalité, plusieurs sous systèmes destinés à différentes tâches et combinant plusieurs modes de résolution de problèmes. Parmi ces tâches, on distingue la décomposition de problèmes, le traitement de données incertaines, l'utilisation de multiples sources d'expertise, l'intégration de différents types de données etc.

Entre autres applications, AQUINAS a été utilisé pour classer les différents outils d'acquisition des connaissances. On peut donc le voir comme un outil d'acquisition dans lequel a été implantée une application particulière permettant de guider le cognitif dans le choix entre plusieurs méthodes ou outils d'acquisition. Il sert, ainsi, d'atelier pour la conception de systèmes à base de connaissances [BOOSE & al, 89].

3.3.2. Tri de cartes

Dans la technique du tri de cartes (card sort) on dispose d'un ensemble de cartes qui représentent les objets particuliers du domaine d'expertise. Ces éléments doivent se situer à un niveau d'abstraction proche, pour que des comparaisons soient possibles et pour que les cartes puissent être mises en paquets. On demande à l'expert de choisir un sous-

ensemble de cartes qu'il souhaite comparer, l'expert pouvant ajouter ou supprimer des cartes tant que le processus de tri n'a pas été entamé. A partir d'un critère que l'expert se donne, on lui demande de faire autant de partitions du sous-ensemble de cartes sélectionnées qu'il le juge nécessaire. On aura par exemple, les groupes "maigre", "moyen" "gros" et "inapplicable" si le tri est effectué par rapport au "poids".

On demande à l'expert de rattacher à chaque groupe de carte sélectionné, une étiquette et tous les éléments appartenant à un groupe sont reliés à ce groupe par un lien *est-un*. [LAUBLET & al, 91]. Une fois que la partition est réalisée, elle est enregistrée et le processus est itéré jusqu'à ce qu'il n'y ait plus de nouveaux critères pour l'expert. Ces critères portent entre-autres sur des relations de parenté, de dépendances, ou encore de similarité [FIRLEJ & HELLENS, 91].

La méthode du tri de cartes, permet d'utiliser des valeurs discrètes au lieu d'échelles numériques binaires et contrairement aux grilles-répertoires, les critères utilisés ne sont pas forcément applicables à tous les éléments.

La classification est formalisée en une forêt d'arbres. Chaque arbre a pour racine le nom d'un groupe. Un nœud de l'arbre a pour fils le nom des sous-groupes ou des éléments dont il est constitué. Les éléments constituent les feuilles de l'arbre (voir figure 3.2.).

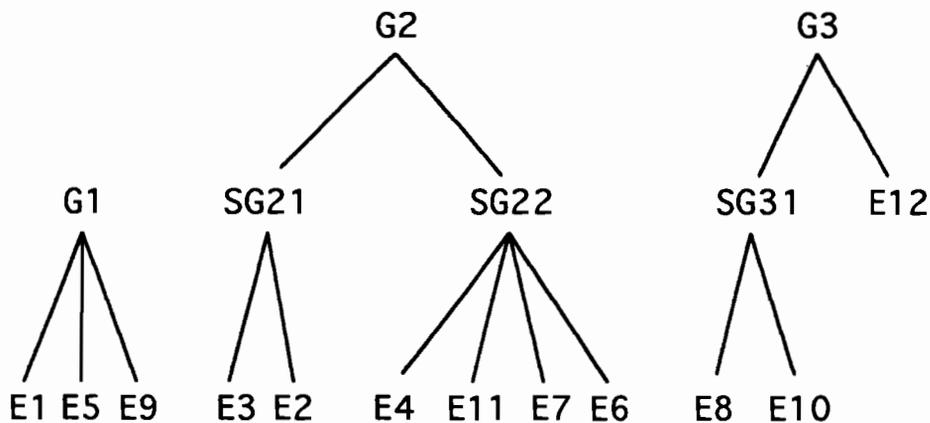


Figure 3.2. un exemple de tri de cartes

Limites du tri de cartes

La méthode du tri de cartes est dédiée à des problèmes d'analyse et permet une classification rapide. En revanche, elle n'est utilisable que dans des domaines où il y a un nombre fini de solutions.

L'outil CATO

CATO [MAJOR, 91] a été conçu afin d'être intégré tout comme ALTO⁹, à un atelier d'ingénierie de la connaissance (KEW), dans le cadre du projet ACKnowledge [MARTY & al, 89]. Le but de KEW est d'intégrer dans un même environnement plusieurs outils d'explicitation des connaissances et d'apprentissage symbolique. On espère ainsi, que le cognicien utilisera les différents outils afin de produire une base de connaissances commune.

De la même façon, on souhaite que KEW soit en mesure de guider l'utilisateur dans le choix d'un outil afin d'accroître l'efficacité du processus d'acquisition des connaissances.

Avant d'entamer le processus de tri, CATO pose quelques questions initiales à l'expert sur le domaine d'application ainsi que sur les concepts qu'il souhaite trier. A partir de cette étape d'initialisation, le contrôle devient effectif de la part de l'utilisateur. En nommant les concepts, les cartes correspondantes s'affichent à l'écran et l'utilisateur a tout le loisir d'ajouter des cartes ou d'en supprimer tant que l'opération de tri n'est pas commencée.

Par ailleurs, l'utilisateur possède un large champ de manœuvre concernant la manière dont le tri doit être effectué. Il peut, par exemple, trier les cartes sans mentionner le critère sur lequel il a opéré le tri. Ceci s'avère particulièrement utile quand l'expert éprouve une difficulté à penser à un tri et veut donc simplement s'exercer un moment afin de se fixer les idées. Une fois les cartes triées, les groupes sont formés et le critère de tri est nommé. Si l'expert est content du résultat, le tri est explicitement accepté et CATO génère des règles à partir des informations fournies puis il les stocke. Il est à noter que le tri peut être reproduit à n'importe quelle étape jusqu'à son acceptation finale.

Si l'expert a déjà un quelconque tri en tête, les groupes et le critère de tri sont alors fournis avant que l'opération de tri ne commence, ceci facilitant largement l'attribution d'une carte à un groupe.

Pendant le déroulement du tri, CATO peut montrer à l'expert les critères de tri déjà pris en compte afin de lui rappeler quels sont les tris déjà utilisés.

Quand l'expert choisit l'option "compte rendu" (*Report option*), CATO affiche aussitôt la matrice des tris et des concepts avec le groupe auquel appartient chaque

⁹ Cet outil sera décrit dans la prochaine section.

concept dans chaque tri. L'expert a ainsi, accès aux résultats sous différentes formes et ceci peut se comparer à la grande flexibilité des langages de requêtes (query language) sans avoir à apprendre une syntaxe particulière.

Le but de CATO est de pouvoir discriminer de manière unique tous les concepts considérés : il arrive cependant qu'à la fin d'un tri, la matrice montre que certains concepts ne sont pas différenciés de manière unique ou qu'il ne le sont pas du tout et il faut alors penser à une autre technique d'explicitation des connaissances, à moins d'améliorer la description des concepts qui n'était, peut être pas, assez riche.

3.3.3. Technique de "laddering"

Dans cette méthode d'explicitation des connaissances, le cogniticien "pousse" l'expert à gravir les barreaux de "laddering" afin d'explicitier des sens plus profonds entre les éléments et leurs caractéristiques. Cette technique exploratoire vise à *abstraire* une connaissance pertinente du domaine ; la méthode de laddering peut être utilisée d'une manière informelle au cours d'un interview ou alors dans une situation plus formelle comme nous allons le voir.

Quand la technique est utilisée de manière plus formelle, les questions sont posées indépendamment d'un contexte donné, par exemple "Pouvez-vous penser à tous les objets que possède x ?". Le risque de ce genre de question est de conduire à un foisonnement d'informations non pertinentes et donc peu exploitables. La stratégie d'exploration du cogniticien doit donc être subtile et il ne doit pas hésiter à faire répéter à l'expert sa réponse ou à lui demander de la reformuler. Ceci afin d'explicitier de nouvelles abstractions et mettre en évidence leurs distinctions propres.

Dans la technique de laddering, l'expert et le cogniticien construisent une représentation graphique du domaine en termes de relations entre éléments [RUGG & al, 90]. Le résultat est souvent fourni sous forme de réseau sémantique. Compte tenu de la nature de la connaissance à explicitier on distingue deux sortes de techniques de laddering: la technique de laddering *orientée tâche* et la technique de laddering *orientée domaine*. Dans la première, on essaie d'atteindre la connaissance que possède l'expert sur la tâche considérée ; ainsi, les arcs du réseau sémantique comprennent des sous-tâches et des super-tâches. Dans la technique d'échelle orientée domaine, l'accent est mis sur "la connaissance structurée", c'est-à-dire la connaissance concernant les types d'entités utilisées dans le domaine.

De manière générale, dans la technique de laddering, on cherche à faire expliciter à l'expert ses connaissances sous forme de hiérarchies d'abstraction. Cette technique implique les étapes suivantes :

1ère étape

Demander à l'expert de choisir le point de départ (l'entité noyau)

2ième étape

Parcourir la connaissance du domaine de l'expert à l'aide des questions suivantes :

1- Se déplacer en "descendant" dans la hiérarchie des connaissances de l'expert:

- pouvez-vous me donner des exemples de cette entité ?

2- Se déplacer à travers la hiérarchie des connaissances de l'expert :

- dans quelles classes possibles peut-on ranger cette entité ?

3- Se déplacer en "remontant" dans la hiérarchie des connaissances de l'expert :

- quelles sont les entités de même niveau que l'on peut regrouper ensemble ?
- quelles sont les entités de même niveau dont elle est exemple ?
- quelle est la différence entre entité1 et entité2 ?

3ième étape

Expliciter les attributs des différentes entités.

Limites de la technique de laddering

L'intérêt majeur de cette technique réside dans la représentation hiérarchique ou graphique des connaissances et dans l'effort d'abstraire les concepts du domaine. Ceci dit, nous pensons que le contexte doit être pris en compte car sans cela, il est difficile de repérer les questions pertinentes.

L'outil ALTO

ALTO est un outil automatisé basé sur la technique de laddering et plus précisément sur la technique de la grille en échelle *grid-laddering* [MAJOR & REICHGELT, 90]. Il est restreint à la technique de laddering orientée domaine (§ 3.3.3).

ALTO construit un arbre de généralisation où les liens entre concepts sont de type *est-un* (is-a tree). La stratégie employée consiste à demander soit un exemple de concept pour générer un concept plus spécifique, soit un ensemble de concepts de même niveau qui seraient des exemples du concept considéré afin de générer un concept plus général (abstrait). De même pour expliciter les caractéristiques d'un attribut, le cognicien peut demander dans quelle mesure des concepts de même niveau diffèrent et quels sont leurs points communs.

L'outil ALTO intègre deux modes d'interaction : l'un pour les débutants, l'autre pour les utilisateurs plus expérimentés (de la technique). L'utilisateur doit déterminer quelle connaissance ajouter et où l'ajouter dans l'arbre. En outre, l'analyse repère les nœuds de l'arbre non discriminés et demande à l'utilisateur s'il veut les différencier.

Il est à noter que durant l'analyse de l'arbre, un contrôle de redondance est effectué. Ainsi, si à un certain nœud parent de l'arbre tous les nœuds fils partagent une caractéristique commune, celle-ci est aussitôt associée au nœud père.

L'un des points forts de ALTO est sa capacité d'analyser par différents contrôles la connaissance une fois explicitée. En revanche, ses limites proviennent de celles de la technique de laddering orientée domaine elle-même, où il n'y a pas de prise en compte de la tâche et où seules les connaissances statiques sont prises en compte.

Comparaison des méthodes contraintes et non contraintes

La différence entre les méthodes non contraintes et les méthodes contraintes [FIRLEJ & HELLENS, 91] tient au fait que dans les premières, l'expert est la seule source d'information pour le cognicien qui dépend totalement de ce qu'il lui affirme i.e. ("dites moi ce que vous savez"), alors que dans les secondes le cognicien interagit avec l'expert i.e. ("laissez moi tester ce que je crois que vous savez").

Nous sommes convaincus que l'utilisation des méthodes contraintes constitue une aide importante pour l'expert quand il veut fournir une certaine forme de connaissances qu'il n'est pas capable de formuler dans les méthodes non contraintes surtout dans le cas d'expert ayant du mal à transmettre son expertise à autrui, ce qui est un cas très courant.

Ces méthodes sont plus fiables que les méthodes non contraintes, car l'expert devient plus précis et réduit ainsi le facteur temps qui est crucial pour l'ingénieur de la connaissance : elles permettent de tester rapidement la connaissance déjà explicitée.

Ces méthodes sont bien "rentabilisées" dans des domaines où l'information est en majorité numérique ou alors très structurée. Enfin, dernier aspect qui mérite toute notre attention, elles sont facilement automatisables.

3.4. CONCLUSION

Nous arrivons au terme de notre étude qui consistait, rappelons le, à passer en revue quelques méthodes manuelles ou semi-automatiques d'explicitation des connaissances afin de situer la méthode d'explicitation des connaissances utilisée dans METIS.

Les méthodes non contraintes définies dans la section 3.2, sont plus souples et plus naturelles pour l'expert et sont, donc, facilement utilisées au prix, cependant, de risques d'omission très importants. Les méthodes contraintes, quant à elles, portent sur les formes sous-jacentes de la représentation des objets et relations chez l'expert ; pour cela, elles présupposent une structure de représentation des connaissances définie au préalable. Les méthodes informatisées constituent des versions automatisées des méthodes, qu'elles soient contraintes ou non ; bien souvent elles intègrent les deux ; on dispose alors d'outils dits hybrides.

Si nous devons situer la méthode utilisée dans METIS (voir chapitre 5) par rapport aux méthodes non contraintes, nous dirions qu'elle s'appuie sur une technique *généralisée* des grilles-répertoires (voir § 3.3.1.), dans le sens où l'on n'est pas contraint par une cotation numérique des attributs, et que les attributs eux mêmes ne sont pas forcément applicables à tous les objets.

La tentation est forte de rapprocher la méthode utilisée dans METIS avec la technique de laddering orientée domaine, implémentée dans ALTO (§ 3.3.3) : ceci se justifie par le fait que les éléments manipulés peuvent être aussi bien des objets que des classes d'objets [MAJOR & REICHGELT 90].

De plus dans METIS, tout comme dans ALTO, nous raisonnons à différents niveaux d'abstraction du fait que l'on se base sur une taxinomie réelle du domaine d'application. Ceci dit, la différence entre les deux méthodes tient au fait que l'analyse

des connaissances dans ALTO se fait a posteriori pour engendrer une nouvelle phase d'explicitation des connaissances, alors que dans METIS la généralisation des traits discriminants (contrôle de redondance dans ALTO) se fait en même temps que l'explicitation des connaissances car la méthode est incrémentale.

Il nous faut souligner le fait qu'ETS et AQUINAS intègrent tous deux un outil de "laddering". Ceci va nous amener à faire aussi la comparaison de notre méthode avec ETS et AQUINAS.

METIS, comme AQUINAS est un système d'explicitation des connaissances qui construit interactivement et "petit à petit" un langage de description et qui est bien adapté à la résolution de problèmes de classification [AIMEUR & GANASCIA, 93c]. Dans les deux systèmes on ne fait pas appel au cognicien : en effet, on a un nombre d'objets prédéfini à caractériser et il y a mise en évidence des relations entre les objets et leurs caractéristiques.

L'une des différences entre le système METIS et le système AQUINAS tient au fait que les objets à caractériser dans METIS ne sont pas des objets individuels (instances) mais des classes d'objets organisées selon une taxinomie.

Dans AQUINAS on prend en compte les hiérarchies. Ainsi, lors de la construction de la grille d'évaluation, quatre sortes de hiérarchies sont utilisées [BOOSE, 89] : la hiérarchie de traits, la hiérarchie de solutions, la hiérarchie d'experts et enfin la hiérarchie de cas. Dans METIS, les traits et les solutions ne sont pas séparés mais sont organisés dans une même hiérarchie : la hiérarchie de généralisation qui est calquée sur la taxinomie des classes et dans laquelle chaque nœud est lui même décrit de manière hiérarchique sous forme de *graphes conceptuels* (voir chapitre 4). Ainsi les connaissances sont directement structurées et sont donc facilement exploitables.

Pour illustrer la prise en compte de hiérarchies, prenons l'exemple d'une taxinomie des fruits du point de vue vitaminique. Dans ETS, on essaiera de discriminer l'orange, la pomme, l'abricot, les pruneaux..., sans tenir compte de la hiérarchie qui existe, alors que dans METIS et dans AQUINAS on discrimina d'abord les fruits frais des fruits séchés et des fruits secs, puis dans les descendants des fruits frais on discrimina les agrumes des fruits à noyau et des fruits à pépins et enfin, dans les descendants des agrumes on discrimina l'orange du citron, de la mandarine etc.

Dans ETS et AQUINAS l'expert n'est pas aidé dans la recherche de traits discriminants. Dès que le système détecte que deux objets ont des "profils" similaires l'expert doit obligatoirement fournir au système une caractéristique permettant de les

différencier. Ce n'est pas le cas pour le système METIS où on fait référence aux propriétés structurelles du langage de description afin de déduire un jeu de traits discriminants. On le suggère alors à l'expert qui est seul juge quant à la pertinence des traits. Lorsque METIS n'arrive pas à découvrir par lui-même un jeu de traits potentiellement discriminants, il pose clairement la question à l'expert.

En définitive, l'explicitation des connaissances nécessite de disposer d'un formalisme de représentation qui soit clair, suffisamment riche et pertinent pour pouvoir exploiter et manipuler les connaissances aisément. Par conséquent, le chapitre suivant a pour objet l'étude de quelques formalismes de représentation des connaissances afin de montrer les avantages et les limites de chacun d'eux et de présenter le formalisme de représentation des connaissances que nous avons choisi.

4. REPRÉSENTATION DES CONNAISSANCES

La représentation des connaissances fait appel à divers formalismes pour organiser les connaissances, les stocker et offrir des opérations formelles qui facilitent leur manipulation. Nous allons examiner dans ce chapitre quelques formalismes de représentation des connaissances en indiquant leurs avantages et inconvénients et essayer de dégager, par une analyse, les formes de représentation qui sont les plus pertinentes pour notre problématique.

4.1. REPRÉSENTATIONS LOGIQUES

La logique classique joue un rôle majeur en représentation des connaissances. Son formalisme est dépourvu d'ambiguïté et permet de décrire des situations du monde réel par des formules concises¹.

Le calcul propositionnel ou *logique des propositions* a pour objet l'étude des formes de raisonnement dont la validité est indépendante de la structure des propositions composantes et résulte uniquement de leurs propriétés d'être vraies ou fausses [DELAHAYE, 86]. La logique des propositions est la composante la plus simple de la logique classique.

Exemple :

Chacune des phrases suivantes est une proposition :

"Zola est un auteur"

"Germinal est le titre d'un livre"

"Zola a écrit Germinal"

¹ Nous n'allons pas décrire le formalisme logique. Le lecteur intéressé peut se référer à [CORI & LASCAR, 93].

Remarque :

on n'utilise en logique des propositions ni les *quantificateurs* ni les *variables* ainsi on ne peut représenter les phrases suivantes de façon précise : "Tous les philosophes sont sages", "Il existe des gens qui supportent les revers de fortune avec philosophie". En effet on devra traiter chaque phrase dans son entier et on ne pourra donc pas traiter les divers composants des phrases.

La *logique des prédicats* (ou calcul des prédicats) appelée aussi logique du premier ordre est un langage formel au moyen duquel des expressions très diverses peuvent être formulées. Elle constitue une extension de la logique des propositions, car elle prend en compte une certaine variabilité et une certaine quantification.

Exemple :

Pour représenter la phrase "l'auteur Zola a écrit *Germinal*", on pourrait utiliser la formule suivante :

Auteur (Zola) \wedge Livre (Germinal) \wedge Ecrire (Zola, Germinal)

4.2. RÉSEAUX SÉMANTIQUES

Les réseaux sémantiques [DOMINE, 88] décrivent la réalité sous forme d'un graphe (réseau) composé de *nœuds* reliés par des *arcs*.

Structurant hiérarchiquement la connaissance, les réseaux sémantiques font appel à trois catégories d'objets :

- des entités : ce sont des mots qui représentent des choses. Exemple : dans la figure 3.1. *OISEAU* et *SAUMON* sont des entités.
- des propriétés : ce sont des mots qui qualifient les entités. Exemple : dans la figure 3.1. *comestible* est une propriété.
- des liens : ce sont des pointeurs étiquetés qui permettent soit de relier des entités entre elles, soit de qualifier des entités à l'aide de propriétés. Exemple : dans la figure 3.1. le lien *est-un* .

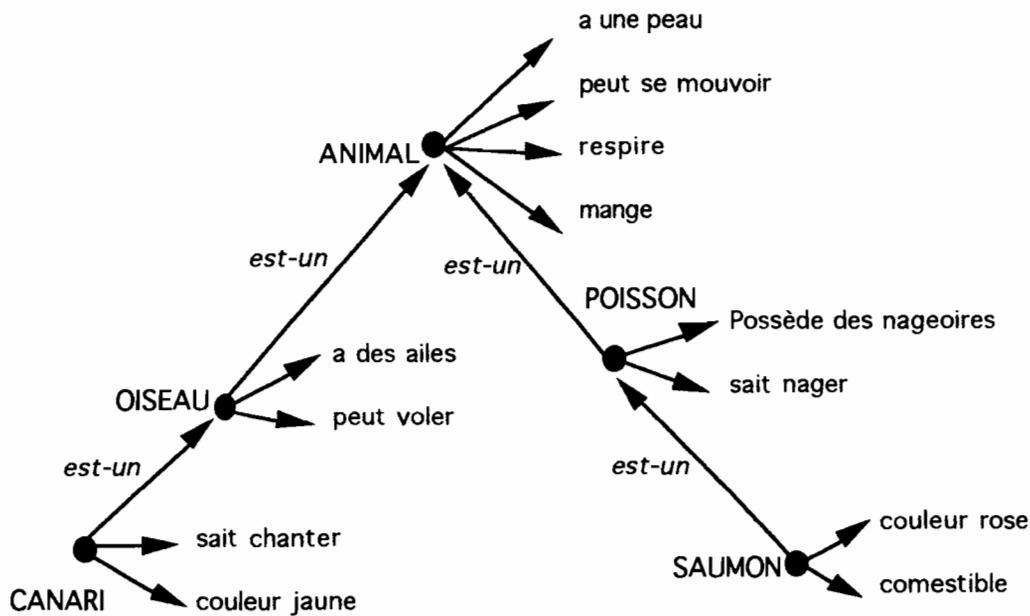


Figure 4.1. Un exemple de réseau sémantique tiré² de Collins et Quillian (1969)

Remarque :

Le lien "est-un" entre classes d'entités traduit des relations d'héritage entre propriétés. Par exemple, la classe SAUMON hérite des propriétés de la classe POISSON. Ainsi on sait que le saumon est un poisson de couleur rose, comestible, qui sait nager et qui possède des nageoires.

Le premier informaticien à avoir introduit la notion de réseaux sémantiques est Quillian [QUILLIAN, 68]. Ses réseaux sont constitués d'une taxinomie, dont les nœuds sont des concepts génériques ou individuels, auxquels sont attachés des propriétés qui définissent le concept. Quillian a émis l'hypothèse selon laquelle la structuration de la mémoire humaine pourrait s'apparenter à ces réseaux sémantiques. Un peu plus tard, Simmons [SIMMONS & BRUCE, 71] développait un réseau dans lequel le choix des arcs était fondé sur la notion de cas sémantique. Lié à un ATN (Réseau de Transition Augmenté) son système pouvait analyser une phrase, traduire son sens en une structure du réseau et engendrer des réponses à des questions simples portant sur le contenu de ces réseaux.

Parmi les travaux de la même période, nous citerons le réseau structurel actif de Norman et Rumelhart [RUMELHART & NORMAN, 73] et le programme HAM (Human

² Nous n'avons pas présenté l'exemple intégralement .

Associative Memory) de Anderson et Bower [ANDERSON & BOWER, 73] qui cherchaient tous les deux à modéliser la mémoire humaine.

L'absence de méthodologie dans les principes généraux des réseaux sémantiques et les confusions relatives aux définitions des nœuds et des liens, ont fait naître des réflexions d'ordre épistémologique. Les plus connues sont celles de Woods ("*what's in a link*") [WOODS, 75] et de Brachman ("*what's in a concept*") [BRACHMAN, 77], ("*On the Epistemological Status of Semantic Networks*") [BRACHMAN, 79].

Ainsi, comment décrire le *contenu* d'un concept ou plus précisément sa définition ? Quelle serait la *structure* de ce type de description ?

Si un concept est plus spécifique qu'un autre, alors comment cette spécialisation est-elle formulée en termes de description de concepts ?

Brachman a apporté des réponses à ces questions et a proposé un formalisme de représentation des connaissances KL-ONE [BRACHMAN, 79] [BRACHMAN & SCHMOLZE, 85] que nous décrivons dans le paragraphe suivant.

Intérêt des réseaux sémantiques :

L'intérêt majeur que représentent les réseaux sémantiques, c'est l'économie de la représentation qu'ils sous-tendent [GANASCIA, 90] et une plus grande facilité d'explicitation grâce à la représentation graphique. Les réseaux sémantiques constituent un moyen de structuration des bases de connaissances ainsi, grâce au mécanisme d'héritage de propriétés, un nœud d'un réseau sémantique peut hériter des propriétés des nœuds qui le subsument [HATON & al, 91].

Les réseaux sémantiques constituent une version graphique du calcul des prédicats. Contrairement à la représentation logique, ils fournissent des informations groupées autour des nœuds. Ainsi la vision de l'information est plus globale. Ils sont utilisés actuellement dans les systèmes de compréhension des langues naturelles (représentation de lexiques sémantiques) ou de vision par ordinateur (description de scènes). On les retrouve, aussi, imbriqués dans des formalismes plus complexes où ils servent de support à la description et à la classification des entités.

Par ailleurs, les réseaux sémantiques peuvent aussi servir de représentation intermédiaire ; ainsi la représentation graphique est couramment utilisée comme représentation intermédiaire entre la langue naturelle et un langage de programmation.

Remarque :

Une comparaison entre les différents formalismes que nous présentons sera faite dans la dernière section de ce chapitre.

4.3. LES LOGIQUES TERMINOLOGIQUES (KL-ONE)

Les logiques terminologiques sont une autre forme de représentation de connaissances structurées.

Un courant de recherche sur la représentation des connaissances s'est créé à partir des idées qui sont à la base de KL-ONE [BRACHMAN & SCHMOLZE, 85]. Il a donné naissance à une famille de systèmes de représentation appelés logiques terminologiques^{3,4}. BACK [NEBEL, 90], CLASSIT [BRACHMAN & *al*, 91] et LOOM [MacGREGOR, 91] en sont des représentants typiques.

Le système KL-ONE

KL-ONE se situe à la frontière entre les réseaux sémantiques et les langages de *frames*. Il permet de représenter les concepts d'un domaine d'application sous forme d'unités d'information appelées *concepts*, qui sont munis d'attributs ou de *rôles* exprimant les relations que le concept entretient avec les autres concepts. Des contraintes sur le type et la cardinalité de sa valeur sont associées au rôle, à l'instar des facettes d'un langage de *frames*. L'ensemble des concepts est organisé en un graphe d'héritage multiple de racine THING. Les rôles d'un concept sont alors hérités par ses spécialisations [HATON & *al*, 91].

³Les logiques terminologiques reposent sur les principes suivants [MacGREGOR, 91] :

- Les entités d'un domaine d'application sont représentées par des objets généralement appelés concepts. L'organisation des concepts s'appuie sur la relation de subsomption : un concept *C* subsume un concept *D* si tout individu décrit par *D* est nécessairement décrit par *C*. Les concepts sont partiellement ordonnés par une hiérarchie semblable à une hiérarchie d'héritage.
- Un *classifieur* gère la hiérarchie des concepts et trouve la place des nouveaux concepts. L'insertion de nouveaux concepts dans la hiérarchie peut provoquer un ensemble d'inférences.
- Il existe un niveau de langage pour décrire les concepts (*TBox*), il s'agit de la *connaissance terminologique* et un niveau de langage pour déclarer les faits dans lesquels interviennent ces concepts (*ABox*), il s'agit en fait de la *connaissance assertionnelle*.

⁴Un point de vue différent sur la subsomption est abordés dans [NAPOLI, 92b].

Il existe dans KL-ONE deux types de concepts. Les concepts *primitifs* dénotent de grandes catégories naturelles comme les personnes, les animaux, les plantes, etc. et servent à construire les concepts *définis*. Par exemple, *Personne* est un concept primitif, spécialisation du concept primitif *Mammifère* (voir figure ci-dessous). Il a un rôle *date-de-naissance* de cardinalité $(1, 1)$, sa valeur ayant pour type le concept primitif *Date*. Une cardinalité égale à $(1, 1)$ exprime que le rôle a exactement une valeur atomique et une seule. *Parent* est un concept défini, spécialisation de *Personne*. Il possède un rôle *enfants* de cardinalité $(1, \infty)$, dont la valeur est une liste de représentants de *Personne* ou individus de type *Personne*.

Un concept primitif n'est pas défini, au sens où l'ensemble de ses propriétés exprime des conditions nécessaires mais non suffisantes pour qu'un individu appartienne à la classe. Ainsi, si un individu est de type *Personne*, il possède le rôle *date-de-naissance*, mais l'inverse n'est pas vrai : un individu possédant le rôle *date-de-naissance* n'est pas forcément de type *Personne*, les animaux pouvant aussi avoir un tel rôle. Un concept défini sera lui complètement décrit par ses relations avec les autres concepts, ses propriétés constituant un ensemble nécessaire et suffisant pour déterminer l'appartenance d'un individu à cette classe. Si un individu est de type *Parent*, il possède le rôle *enfants* et réciproquement, posséder le rôle *enfant* suffit pour déterminer qu'un individu est de type *Parent*.

Exemple :

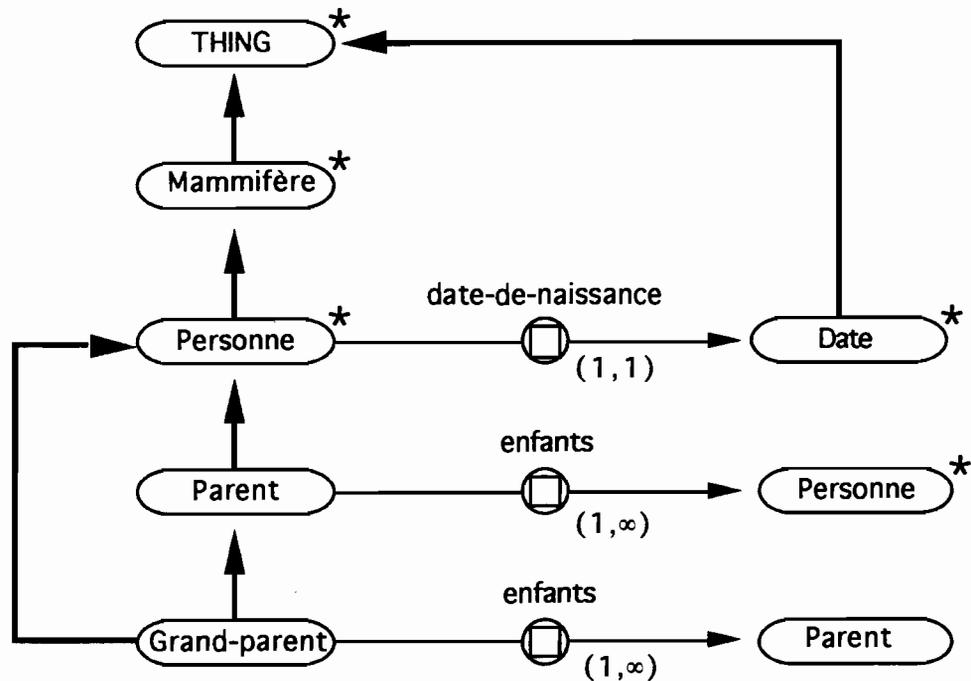


Figure 4.2. Exemple d'une hiérarchie de concepts et leurs rôles en KL-ONE, tiré de [HATON & al, 91]

Remarque :

les astérisques marquent les concepts primitifs.

La classification

La notion de concept défini, c'est à dire défini par un ensemble de propriétés nécessaires et suffisantes, est à la base du mécanisme fondamental de KL-ONE, le classifieur. Ce dernier détermine les relations hiérarchiques optimales entre les concepts à partir de leurs propriétés. Il permet donc d'insérer un concept à la place adéquate dans la taxinomie [BIEBOW, 92].

Exemple :

Un parent est une personne qui a des enfants qui sont aussi des personnes; un grand parent est une personne qui a des enfants qui sont des parents. Le classifieur classe lui-même grand-parent comme spécialisation de parent.

Les logiques terminologiques n'ont pas le pouvoir expressif de la logique du premier ordre, cependant, la recherche dans ce domaine est très active. Elle reflète une tendance prometteuse dans les études actuelles sur la représentation des connaissances.

4.4. LES GRAPHES CONCEPTUELS

Sowa a voulu définir un formalisme général de représentation en réseaux pour la représentation des connaissances. Il développa ainsi les *graphes conceptuels* [SOWA, 84] qui ne sont a priori conçus pour aucune application particulière.

Sowa émet l'hypothèse que les graphes conceptuels sont isomorphes aux structures mentales sous-jacentes à la pensée humaine et propose que sa théorie soit considérée comme un langage universel de représentation des connaissances [SOWA, 92].

Les graphes conceptuels, comme le système KL-ONE, appartiennent à la lignée des réseaux sémantiques à héritage. Leur notoriété respective sont à peu de choses près, comparables; leurs applications portent sur les mêmes domaines (langues et bases de connaissances) ; de nombreux systèmes les utilisent, avec des variantes multiples [BIEBOW, 92]. Le lecteur intéressé par une comparaison entre les deux systèmes peut se référer à [BIEBOW & CHATY, 93].

Depuis 1984, plusieurs groupes de recherche ont utilisé la théorie des graphes conceptuels pour développer différentes applications en traitement du langage naturel, raisonnement, en ingénierie des connaissances, etc. [SOWA, 92]. D'autres groupes de recherches (notamment le groupe de recherche français sur les graphes conceptuels) tentent d'améliorer la théorie de Sowa en proposant des extensions du formalisme. On trouvera dans [CHEIN, 92], un noyau pour les graphes conceptuels simples destiné à servir de canevas lors de l'introduction d'une nouvelle notion augmentant les capacités expressives ou opératoires du langage.

Plusieurs systèmes d'acquisition des connaissances utilisent aujourd'hui le formalisme des graphes conceptuels⁵. On distingue notamment les travaux en extraction des connaissances de [ANTONACCI & al, 88] au centre scientifique d'IBM de Rome et ceux de [BERARD-DUGOURD & al, 88] et [HERNERT, 93] en acquisition de connaissances à partir de textes au centre scientifique d'IBM de Paris. Par ailleurs, les

⁵ On trouve dans [SOWA, 92] un inventaire des applications basées sur le formalisme des graphes conceptuels.

travaux de [KABBAJ & FRASSON, 93a] [KABBAJ & FRASSON, 93b] à l'université de Montréal s'intéressent à l'acquisition des connaissances dans le cadre des systèmes tutoriels intelligents. Enfin, nous pouvons citer les travaux de l'équipe de Brian Garner [GARNER & TSUI, 89] [GARNER & al, 89] à l'université de Deakin (Australie) qui s'intéressent à l'acquisition des connaissances et au raisonnement à partir de cas pour développer des systèmes experts.

Nous allons maintenant détailler certains aspects de la théorie de Sowa en nous inspirant de [SOWA, 84].

4.4.1. Éléments fondamentaux

Un graphe conceptuel [NOGIER, 91] est défini comme un graphe biparti ayant deux types de nœuds :

- les *concepts*, représentés par des boîtes ou placés entre crochets [...] correspondent à des contenus de la pensée et obéissent à la notation suivante : [<type> : <réfèrent>]. Le réfèrent du concept précise son sens, il donne le degré de quantification du concept, que celui-ci soit quantifié explicitement par un nombre ou bien de manière qualitative, par exemple : [Personne : 'Max']. Dans le cas d'un concept générique, par exemple : [Personne : *] ou [Personne], le réfèrent est non instancié et peut prendre toutes les valeurs possibles en accord avec le type du concept;
- les *relations*, représentées par des ovales avec une flèche entrante et une autre sortante ou entre parenthèses : -> (...) ->, symbolisent les liens qui existent entre les concepts du graphe.

La définition suivante issue de [SOWA, 84] établit plus précisément la syntaxe des graphes conceptuels :

Définition

Un graphe conceptuel est un graphe biparti, connexe et fini.

- Les deux types de nœuds d'un graphe conceptuel sont les concepts et les relations conceptuelles,
- De toute relation conceptuelle part ou arrive au moins un arc, chacun de ces arcs devant être lié à un concept,
- Si une relation a n arcs, ses arcs reçoivent les étiquettes $1, 2, \dots, n$,
- Un concept peut former à lui seul un graphe conceptuel, mais toute relation conceptuelle doit être liée à au moins un concept.

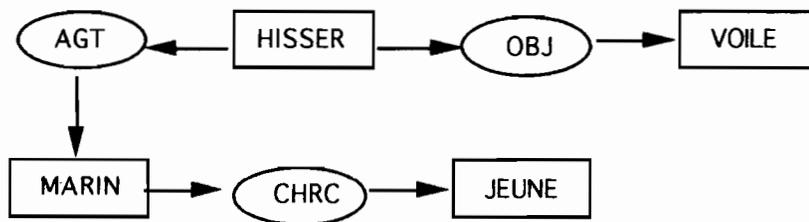
Remarque :

Un graphe conceptuel représente une seule formule logique [SOWA, 84], alors qu'un réseau sémantique représente une collection (ou produit) de formules et décrit leurs connections mutuelles en les immergeant dans leur contexte : le domaine du discours.

Exemple :

Soit la phrase "Un jeune marin hisse la voile" :

Le graphe conceptuel correspondant à cette phrase est :



On peut employer la représentation condensée suivante, dite représentation linéaire, un peu moins lisible que la précédente, mais nettement plus pratique :

[HISSER]-

(AGT) -> [MARIN] -> (CHRC) -> [JEUNE]

(OBJ) -> [VOILE]

Remarque :

CHRC, AGT et OBJ veulent respectivement dire caractéristique, agent et objet.

Le référent du concept

Les concepts peuvent avoir différents types de référents :

- référence générique :

[Personne : *] "une personne"

- référence générique nommée :

[Personne :*x] "une personne x"

- référence individuelle :

[Personne : #36] "cette personne"

- référence nommée :

[Personne : Sarah] "Sarah"

- référence numérique :

[Nombre : 15] "15"

- référence ensembliste :

[Personne : {Jean, Paul, Marc}] " Jean, Paul, et Marc"

- référence ensembliste de cardinalité 3 :

[Personne : {Jean, Paul, Marc} @3] " Jean, Paul, et Marc"

- référence ensembliste générique :

[Personne : {*}] "plusieurs personnes"

- référence ensembliste générique de cardinalité 3 :

[Personne : {*}@3] "trois personnes"

- référence ensembliste partielle :

[Personne : {Jean, Paul, *}] " Jean, Paul, et d'autres"

- référence ensembliste partielle de cardinalité 4 :

[Personne : {Jean, Paul, *}@4] " Jean, Paul, et 2 autres"

- référence associant une mesure et une unité de mesure :

[Mesure : @75cl] "75 cl"

- référence sous forme de graphe conceptuel :

[Conclusion : [Proposition : [bradycardie]-> (et) -> [infection].].]

Lorsque le référent est un graphe conceptuel, on utilise alors le type label particulier "proposition" avec pour référent un graphe conceptuel.

La hiérarchie des concepts

La hiérarchie des concepts permet au modèle sémantique de généraliser et de spécialiser des concepts. C'est-à-dire que l'on peut représenter les concepts à différents niveaux d'abstraction. La hiérarchie des concepts est composée de diverses familles de concepts du même type, c'est-à-dire dépendant du même hyperonyme. Par exemple,

imaginons qu'il existe un type Meuble désignant l'ensemble des concepts relatifs au mobilier . Meuble est un hyperonyme de chaise, bureau, armoire, etc. Ces différents concepts sont de type Meuble.

Comme tous les concepts ne sont pas comparables, c'est une relation d'ordre partielle : "<" qui va ordonner les concepts selon leur degré de généralité. La hiérarchie des concepts admet un héritage de propriétés du *super-type* au *sous-type*. Les termes de super-type et sous-type sont employés pour désigner la position respective de deux concepts dans la hiérarchie.

Exemple :

Yacht < Navire signifie qu'un yacht est un navire ou plutôt une sorte de navire (navire de plaisance).

Soit $x < y$ signifie x est du type de y .

x est un sous-type de y .

y est un super-type de x .

La hiérarchie des concepts décrite par Sowa comporte deux autres types de concepts qui, bien que n'ayant pas vraiment de signification, assurent la complétude du système. C'est ce qui confère une structure de treillis à la hiérarchie des types, nous distinguons donc :

\perp l'élément absurde, quelque soit le type $t > \perp$

T l'élément universel, quelque soit le type $t < T$

Exemple :

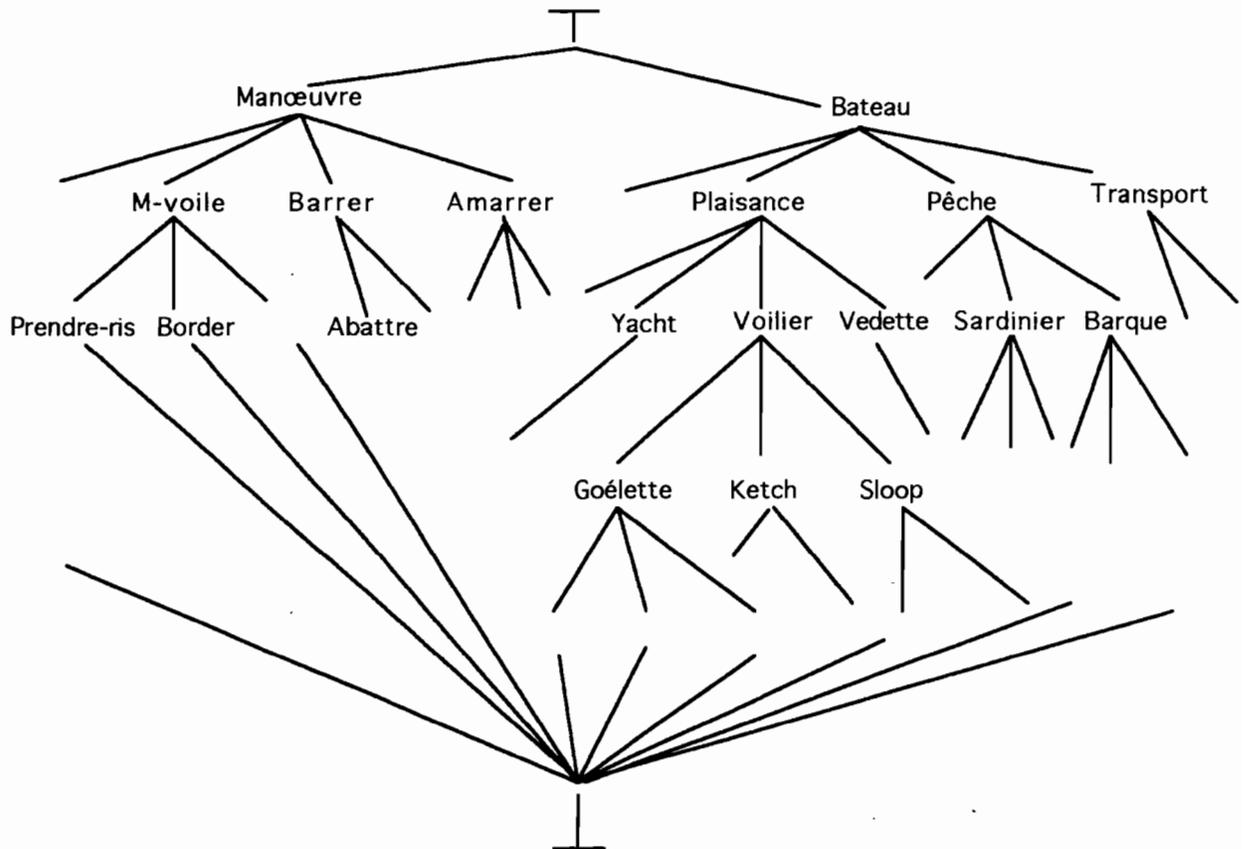


Figure 4.3. Un exemple de hiérarchie de concepts

La hiérarchie des relations

L'ensemble des relations conceptuelles [NOGIER, 91] se scinde en plusieurs groupes de relations de même type : localisation spatiale, localisation temporelle, valeur, etc. Il existe deux notations différentes pour les relations :

- lorsque la relation désigne une famille de relations, elle est dite à 1 niveau, et se note (REL), par exemple (FAIRE-POINT).
- lorsque la relation est spécifique à un certain type, elle est appelée relation à 2 niveaux et se note (REL.X). Le deuxième argument sert à préciser le sens donné par le père de la famille (REL), par exemple : (FAIRE-POINT.LONG) (faire le point en longitude pour un bateau).

De la même manière que le référent des concepts génériques est non instancié, une relation à un niveau est en réalité une relation à deux niveaux où le deuxième niveau est non instancié. La variable muette (*) joue le rôle de deuxième niveau : (REL) est équivalent à (REL.*)

Remarque :

John Sowa propose des relations conceptuelles à n niveaux [SOWA, 84]. Cependant, la mise en œuvre informatique de telles relations accroît les risques d'explosion combinatoire lorsque deux relations s'unifient. Elles sont donc peu utilisées.

Le Canon

Le **canon** contient tous les éléments nécessaires pour dériver l'ensemble de *graphes canoniques*. Il se compose d'une *hiérarchie* de concepts et de relations, de l'ensemble des référents, d'un *prédicat de conformité* établissant des contraintes d'association entre types de concepts et référents et enfin d'un ensemble fini de graphes conceptuels appelés *base canonique*.

Graphes canoniques

Les graphes canoniques (ou graphes sémantiquement corrects représentant une combinaison permise de mots) sont des graphes conceptuels particuliers au sens où ils ont pour fonction d'éviter le traitement des phrases non valides du point de vue sémantique. Ils garantissent que les graphes conceptuels manipulés, correspondent à des situations non absurdes dans l'univers du discours.

Par exemple, on peut définir ainsi le graphe canonique de l'une des acceptions du verbe **naviguer** :

[Marin] <- (Agt) <- [Naviguer] -> (loc) -> [Mer]

Ce graphe canonique permet d'éviter les phrases absurdes, telles que : *un Camembert naviguait sur un lac*.

4.4.2. Opérations fondamentales

Différentes opérations sont réalisables sur les graphes canoniques pour en construire de nouveaux. Elles sont décrites par les règles de formation canonique que nous allons présenter. Soient u et v deux graphes canoniques quelconques :

1. la **copie** : une copie exacte de u est un graphe canonique.
2. la **restriction** : pour tout concept c de u , $type(c)$ peut être remplacé par un sous-type de $type(c)$; si c est un concept générique, son référent peut être remplacé par un

marqueur individuel. Ces changements sont permis seulement si le référent de c est conforme au type de c , avant et après restriction.

Exemple :

[Bateau : *] peut être restreint à [Bateau : 'Espérance'] car on a individualisé le concept générique Bateau en spécifiant son référent.

[Bateau] peut être restreint à [Catamaran] car on a remplacé le type du concept Bateau par le sous-type Catamaran

3. la **jointure** : si un concept c de u est identique à un concept d de v , alors le graphe w obtenu par effacement de d et par rattachement à c de tous les arcs des relations conceptuelles qui étaient liées à d , est appelé jointure des graphes u et v sur le concept c .

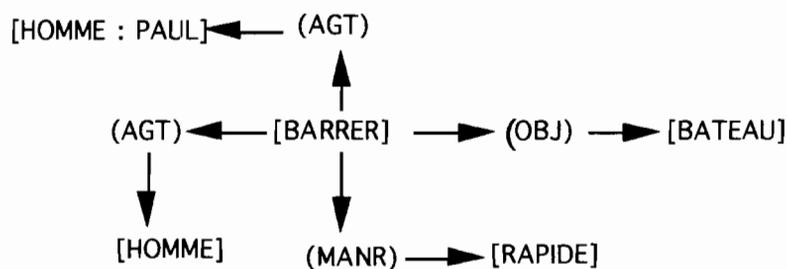
Exemple :

Soient les graphes G_u et G_v suivant:

G_u : [HOMME] <- (AGT) <- [BARRER] -> (OBJ) -> [BATEAU]

G_v : [HOMME : PAUL] <- (AGT) <- [BARRER] -> (MANR) -> [RAPIDE]

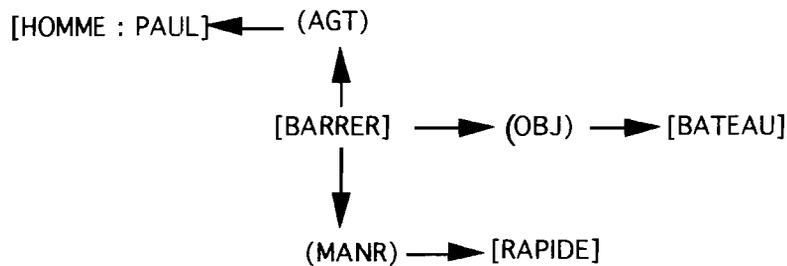
La jointure de G_u et G_v sur le concept [BARRER] nous donne le graphe G_w suivant :



4. la **simplification** : si les relations conceptuelles r et s du graphe u sont dupliquées, alors l'une d'elles peut être effacée dans u avec tous les arcs.

Exemple :

Si on avait restreint dans le graphe G_w le concept [Homme] à [Homme : Paul] il y aurait eu une redondance et la règle de simplification nous aurait donné le graphe suivant :



Les règles de formation canonique [THAYSE & al, 90] constituent une grammaire générative pour les graphes conceptuels de la même manière que les règles de réécriture de Chomsky constituent une grammaire générative pour les structures syntaxiques des langues naturelles.

Remarque :

les règles de formation canonique ne sont pas des règles d'inférence, elles garantissent simplement la correction des graphes générés. Elles peuvent être considérées comme des règles de réfutation (ainsi, un moyen de remettre en cause un graphe est de montrer qu'il est canoniquement dérivable d'un graphe faux).

La définition suivante précise la notion de graphe canoniquement dérivable.

Définition :

Soit A un ensemble de graphes canoniques. Un graphe w est dit canoniquement dérivable de A , si et seulement si :

- w est un élément de A , ou bien,
- w peut être dérivé par application des règles de formation canonique, à partir de deux graphes u et v qui sont eux-mêmes canoniquement dérivables de A .

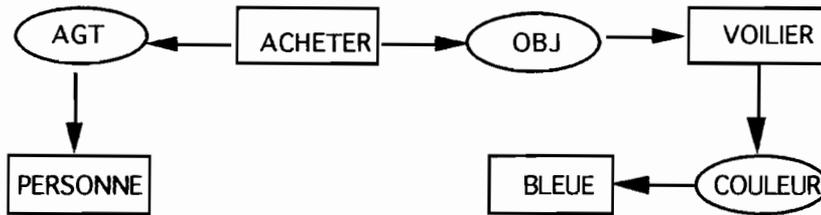
4.4.3. Jointure maximale

La jointure maximale combine les quatre règles canoniques de formation. A partir de deux graphes, elle effectue des restrictions (règle de restriction) en cherchant pour chaque couple de concepts le plus grand spécialisé commun. Puis elle les joint sur un concept commun aux deux graphes (règle de jointure). Et finalement, elle ôte les relations dupliquées qui apparaissent (règle de simplification).

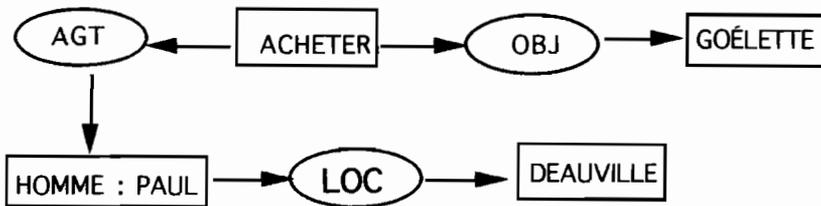
Exemple :

Partant de deux graphes G_U et G_V :

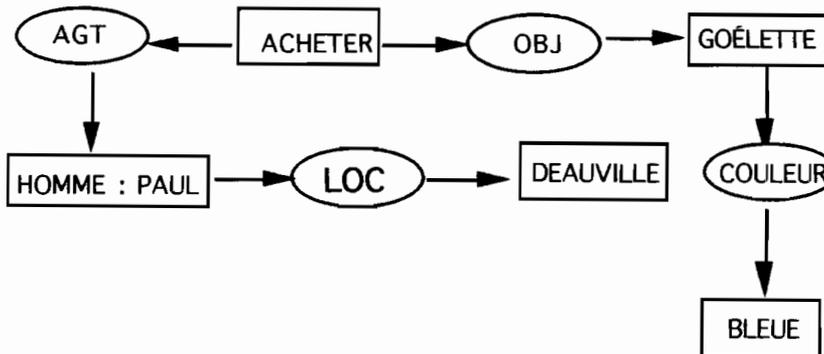
Une personne a acheté un voilier bleu



Paul a acheté une goélette à Deauville



On construit le graphe G_W suivant :



Le concept [PERSONNE] est restreint à [HOMME : PAUL], le concept [VOILIER] est restreint à [GOÉLETTE], puis les deux graphes sont joints sur les concepts [ACHERETER] et [HOMME : PAUL]. Les relations (AGENT) et (OBJET) qui apparaissent deux fois ne sont pas dupliquées.

4.4.4. Projection

L'opération de projection [NOGIER, 91] d'un graphe G_V dans le graphe G_U consiste à exhiber un sous-graphe G_W de G_U tel que :

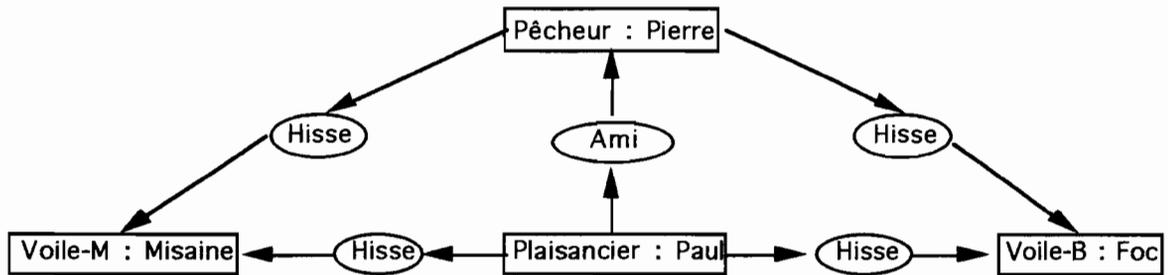
- les relations conceptuelles de G_w sont des instances de celles de G_v , c'est à dire qu'elles sont soit identiques soit plus spécifiques,
- si une relation r_{ij} de G_v relie les concepts v_i et v_j alors elle (ou l'une de ses spécialisations) reliera les concepts correspondants w_i et w_j de G_w , et ces concepts obéissent à la dernière règle,
- tous les concepts w_i de G_w sont des restrictions des concepts v_j de G_v , c'est à dire que quel que soit i , où i détermine la position du concept dans le graphe $w_i < v_j$;
- le graphe G_w est appelé noyau de la projection de G_v dans G_u ,
- le graphe obtenu en retirant de G_u les arêtes de G_w est nommé partie pendante de G_u dans la projection de G_v dans G_u .

La projection permet d'extraire d'un graphe donné (G_u) un sous-graphe (G_w) qui respecte un certain modèle (G_v). Cette opération autorise donc le filtrage (pattern matching) sur les graphes conceptuels.

La projection peut être considérée sous un autre angle. Plutôt que de chercher à extraire d'un ensemble d'informations une information précise, elle sert à vérifier que cet ensemble correspond à un schéma donné. L'apport essentiel de cette opération au modèle des graphes conceptuels est la possibilité de reconnaître un motif au sein d'un graphe.

Exemple :

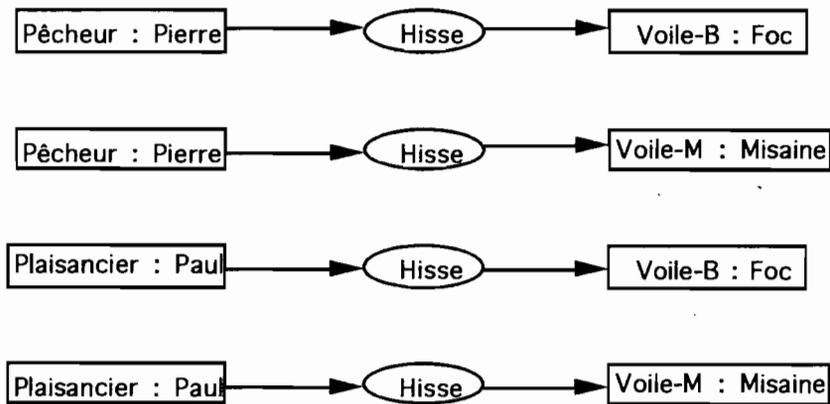
Graphe GU :



Graphe GV :



Les noyaux de projection de GV dans GU :



Partie pendante de GU :



Figure 4.4. Projection

4.4.5. Généralisation et spécialisation

Définition [SOWA, 84]

Soient u et v deux graphes conceptuels. Si w est canoniquement dérivable de u et v , w est appelé *spécialisation* de u et de v , et u et v sont appelés *généralisation* de w . On écrit : $w \leq u$ et $w \leq v$.

Généralisation

A l'inverse de la jointure maximale qui fusionne deux graphes tout en les spécialisant, la généralisation nous permet d'obtenir un résultat inverse, c'est à dire un graphe plus général.

Cette opération permet d'obtenir le plus petit graphe général commun à deux graphes. Pour chaque couple de concepts, elle exhibe leur plus petit généralisé commun. Puis elle élimine les arcs pour lesquels il est impossible de généraliser. Le graphe obtenu est donc plus général que les graphes de départ.

Exemple : soient les graphes suivants :

(G1 [HOMME : JIM] <- (AGT) <- [BARRER] -> (OBJ) -> [SLOOP] ->
(CHRC) -> [GRAND]

(G2) [BARRER]-
(AGT) -> [FEMME : PATRICIA]
(OBJ) -> [GOÉLETTE]
(MANR) -> [PRUDENCE]

La généralisation de G1 et G2 nous donne :

G3 [BARRER]-
(AGT) -> [PERSONNE]
(OBJ) -> [VOILIER]

Ici [PERSONNE] est généralisé à partir de [HOMME : JIM] et [FEMME: PATRICIA] et [VOILIER] à partir de [GOÉLETTE] et [SLOOP].

4.4.6. Hiérarchie de généralisation

La généralisation [SOWA, 84] définit une relation d'ordre partiel sur les graphes conceptuels, appelés *hiérarchie de généralisation*. Pour tous graphes conceptuels u , v , w on a :

- réflexivité : $u \leq u$,

- transitivité : si $u \leq v$ et $v \leq w$ alors $u \leq w$,
- antisymétrie : si $u \leq v$ et $v \leq u$ alors $u = v$,
- si v est un sous-graphe de u , alors $u \leq v$,
- si u est identique à v , à l'exception d'une ou plusieurs étiquettes de type de v qui sont restreintes à des concepts individuels de même type, alors $u \leq v$,
- le graphe [T] est une généralisation de tous les graphes conceptuels.

Remarque :

la relation \leq est une relation de "sous-graphe".

4.4.7. Abstraction et définition

Ce que nous avons vu jusqu'ici était statique mais, le formalisme permet également la définition de nouveaux types ainsi que la déduction de nouveaux graphes.

Définition 1: une abstraction n -adique $\lambda a_1 a_2 \dots a_n \mu$ consiste en un graphe canonique μ , appelé corps, associé à une liste de concepts génériques $a_1 a_2 \dots a_n$ dans μ , appelés paramètres formels. La liste des paramètres $a_1 a_2 \dots a_n$ formels suivant λ distingue les paramètres formels des autres concepts présents dans μ .

Ainsi l'abstraction est l'équivalent d'une procédure dans un langage de programmation. La lettre λ introduit les paramètres formels et dans le corps μ , chaque concept utilisé comme paramètre formel possède l'une des variables comme référent.

Exemple :

$\lambda x,y$ [PÊCHER]-

(AGT) -> [PÊCHEUR : *x]

(OBJ) -> [POISSON : *y] -> (COULEUR) -> [BLEUE]

Dans cet exemple $\lambda x,y$ identifie [PÊCHEUR : *x], [POISSON : *y] comme paramètres formels. Les concepts [PÊCHER] et [BLEUE] sont comme des variables locales dans une procédure. Le corps μ est constitué de tout le graphe.

Définition2 : Une *définition de type* déclare qu'une étiquette de type t est définie par une *abstraction monadique* (une seule variable) $\lambda a \mu$.

On l'écrit : **type** t (a) **is** μ . Le corps μ est appelé *différence* de t , et *type* (a) est appelé *gène* de t . L'abstraction $\lambda a \mu$ peut être introduite dans le champ *type* de n'importe quel concept où l'étiquette t peut être introduite.

Exemple :

"Un poisson est un animal qui vit dans la mer" se traduit par :

type POISSON (x) **is**

[ANIMAL : $*x$] <- (AGT) <- [VIVRE] -> (LOC) -> [MER]

type (a) est : ANIMAL

Le corps μ est [ANIMAL : $*x$] <- (AGT) <- [VIVRE] -> (LOC) -> [MER]

La définition des types simplifie considérablement l'écriture des graphes.

Les deux éléments du modèle, les concepts et les relations, peuvent être définis par des graphes conceptuels [SOWA, 84]. Le processus de définition de relations conceptuelles par des graphes est similaire aux définitions de type de concepts.

Exemple :

[PERSONNE : SARAH] -> (PROPRIÉTAIRE) -> [FRÉGATE : ESPÉRANCE]

On peut définir la relation Propriétaire de la manière suivante :

Relation PROPRIÉTAIRE (x, y) **is**

[APPARTENIR] -

-> (SOURCE) -> [PERSONNE: $*X$]

-> (OBJ) -> [BATEAU : $*y$]

Remarque :

Dans cet exemple l'abstraction n'est pas monadique.

Une fois un type de concept défini, il est utilisable par deux nouvelles opérations : la contraction de type et l'expansion de type.

4.4.8. Contraction

La contraction de type consiste à remplacer un sous-graphe correspondant à un type de concepts par le concept de ce type.

Exemple :

" le bateau : 'Yasmine', qui est amarré dans une marina est vert" dont le graphe u est :

$$u : [\text{VERTE}] \leftarrow (\text{COUL}) \leftarrow [\text{BATEAU : YASMINE}] \leftarrow (\text{AGT}) \leftarrow [\text{AMARRER}] \rightarrow (\text{LOC}) \rightarrow [\text{MARINA}]$$

Soit la définition de type suivante :

type BATEAU-DE-PLAISANCE (x) **is**

$$[\text{BATEAU : } *x] \leftarrow (\text{AGT}) \leftarrow [\text{AMARRER}] \rightarrow (\text{LOC}) \rightarrow [\text{MARINA}]$$

Cette définition de type est de la forme **type** $t(a)$ **is** μ

Le graphe v obtenu par contraction de type est le suivant :

$$v : [\text{BATEAU-DE-PLAISANCE : YASMINE}] \rightarrow (\text{COUL}) \rightarrow [\text{VERTE}]$$

4.4.9. Expansion

Dans un graphe conceptuel, si un concept correspond à un type défini, alors l'opération qui consiste à substituer à ce concept le graphe conceptuel qui le définissait est appelée *expansion de type*. L'expansion permet ainsi de déployer un graphe conceptuel en représentant les connaissances implicitement contenues dans ces concepts.

Remarque :

L'expansion de type n'est pas la réciproque de l'opération de contraction de type [SOWA, 84], c'est à dire que le résultat d'une contraction de type suivie d'une expansion de type n'est pas identique au graphe initial, puisque l'étiquette de type du concept contracté n'est pas restaurée. C'est pourquoi Sowa définit une autre expansion de type qu'il nomme *expansion de type maximale* et qui serait la réciproque de l'opération de contraction. Finalement la première expansion est une *expansion de type minimale*.

Dans notre cas, seule l'expansion de type minimale sera considérée.

4.4.10. Hiérarchie Aristotélicienne

Définition

Une hiérarchie de types T est dite *Aristotélicienne* si chaque étiquette de type t , qui est lui même un sous-type d'un autre type, est définie par une abstraction $t = \lambda a \mu$

En d'autres termes, une *hiérarchie Aristotélicienne* est une hiérarchie de types dans laquelle tous les types de concepts autres que les types primitifs peuvent être introduits par des définitions substantielles spécifiant un genre commun (voir définition précédente) et une ou plusieurs différences.

Exemple :

les types (Espèce, Genre et Famille) qui ne sont pas des types primitifs dans le système METIS (voir chapitre 5) définissent une hiérarchie Aristotélicienne [AïMEUR, 93c]. Ainsi par exemple, une espèce donnée sera définie à partir du genre auquel elle appartient. C'est-à-dire que si le genre est caractérisé entre autres par le fait d'avoir des écailles, l'espèce elle, sera caractérisée non seulement par le fait qu'elle ait des écailles, mais en plus que celles-ci ont une forme cycloïde.

4.5. CONCLUSION

Nous nous sommes intéressés dans ce chapitre à quelques formalismes de représentation des connaissances en vue de déterminer celui qui est le plus adéquat pour contribuer à notre problématique. Ainsi, nous avons présenté dans un premier temps le formalisme logique incluant la logique des propositions et la logique des prédicats, puis examiné dans un second temps, quelques formalismes de représentation sous forme de réseaux, comme les réseaux sémantiques (avec une de leurs variantes, le système KL-ONE), et les graphes conceptuels de Sowa. Même si nous ne les avons pas abordés notons que d'autres formalismes puissants de représentation des connaissances existent tels que la représentation par objets qui permet de mettre en œuvre plusieurs modes de raisonnement, notamment le raisonnement par classification (voir chapitre 2).

Nous allons maintenant analyser la pertinence et les limites des formalismes de représentation abordés vis à vis de nos objectifs en dégageant ainsi le choix que nous avons adopté.

Les langages logiques sont définis uniquement de manière syntaxique, et leur sémantique résulte du processus d'interprétation des éléments du langage selon un domaine étudié. De plus si la logique mathématique est particulièrement bien adaptée à la représentation de concepts mathématiques [HATON & al, 91], elle ne constitue toutefois pas un formalisme universel. En effet il n'est pas possible d'utiliser des connaissances de nature incertaine, contextuelles ou sujettes à révision sans quitter le cadre de la logique classique.

L'inconvénient du formalisme logique est qu'il n'est pas structuré : il faut par exemple, parcourir l'ensemble des formules logiques d'une banque de données pour rassembler toutes les informations relatives à un objet (une instance). Ce qui n'est pas le cas des représentations graphiques qui structurent l'information.

Les réseaux sémantiques ont pour principal inconvénient de ne permettre de représenter que des prédicats binaires. En effet, les prédicats d'arité 3 ou plus, sont convertis en forme binaire [RICH, 87]. Un réseau sémantique sera donc plus complexe pour exprimer suffisamment d'informations. De plus les réseaux sémantiques sont trop rigides, car l'accès aux informations n'est pas assez souple.

En ce qui concerne les graphes conceptuels, les points faibles selon SABAH [SABAH, 88], portent plutôt sur le manque de détails de certains opérateurs (en particulier la correspondance entre graphes conceptuels et formules logiques) et sur les possibilités d'applications. Les points forts de la théorie des graphes conceptuels sont nombreux : ils constituent un modèle de représentation général très clair, ils permettent d'intégrer une culture générale très vaste et apportent enfin une rigueur logique de la représentation ainsi que des aspects sémantiques riches. De plus, des mécanismes lourds, difficiles à expliquer et à mettre en œuvre, s'expriment très facilement dans ce type de langage grâce aux opérations définies par Sowa (telles que la projection, la jointure etc.). Notons par ailleurs, que dans le formalisme des graphes conceptuels, il y a peu de primitives prédéfinies et il y a séparation des liens de premier type (relations) des liens de deuxième type ("is a" et "a kind of"), de plus les inférences effectuées sont contrôlées en permanence par le canon.

Nos domaines d'application étant la biologie, la pédologie et la minéralogie (la description des objets est textuelle, variée et cependant assez structurée), et notre

problématique étant la mise en œuvre d'une méthode d'explicitation de ces connaissances structurées, nous avons besoin d'un formalisme de représentation des connaissances qui permette une meilleure lisibilité⁶ que celle du formalisme logique, une bonne visualisation graphique de la description des objets, une richesse d'expression sémantique une représentation claire de la structuration des connaissances et enfin de nombreuses facilités de manipulation de ces structures de connaissances.

Nous sommes convaincus que le formalisme des graphes conceptuels répond à nos exigences et s'avère de ce fait, être le plus pertinent pour notre problématique. Nous l'avons choisi et mis en œuvre dans METIS, ce système est décrit au chapitre 5.

⁶ Ce que nous voulons c'est qu'une structure soit éclatée en n formules.

5. L'EXPLICITATION DES CONNAISSANCES DANS METIS

Ce chapitre portera sur le principe général d'explicitation des connaissances du système METIS. Nous aborderons dans un premier temps l'organisation des connaissances dans les sciences de l'observation en général et dans METIS en particulier. La structuration de la base de connaissances fait appel aux notions de graphe primitif, de base de graphes, de modèle structurel, et de taxinomie de classes. Nous examinerons dans un second temps, les phases d'explicitation des connaissances en les illustrant à l'aide d'exemples. Ainsi, nous verrons successivement les phases d'initialisation, de classification de discrimination et de saisie de données. Nous présenterons enfin, l'architecture générale de METIS ainsi que les fonctionnalités qui interviennent dans cette architecture.

Preliminaires

Dans les sciences d'observation telles que l'archéologie, la minéralogie, la zoologie, ou la botanique, de grandes quantités d'objets sont répertoriées sous forme de taxinomies et exprimées dans des catalogues au moyen d'un langage de description textuel. Ces catalogues sont établis de longue date et possèdent une diffusion internationale (avec les validations qui en découlent). La lourde gestion (accès aux informations, mise à jour ...) de ces catalogues a fait naître le besoin de construire des bases de connaissances permettant d'intégrer toutes sortes d'informations existant sous forme de classifications livresques.

Face à la masse de données accumulée dans les catalogues, la question est de savoir quel est le meilleur moyen d'extraire, de caractériser, de conserver et d'exploiter automatiquement une connaissance intéressante. De plus, le temps consacré à la

description des taxons¹ d'une part, et à l'accès aux informations d'autre part, ainsi que la révision des familles et des genres rend nécessaire le besoin d'outils d'aide à l'acquisition des connaissances.

Notre travail consiste à représenter et exploiter ces hiérarchies documentaires. Pour ce faire, nous nous proposons de faire appel à une technique d'explicitation des connaissances pour caractériser et organiser ces informations et au formalisme des graphes conceptuels pour construire des *descriptions structurées*.

Le principe d'explicitation des connaissances retenu dans le cadre de notre travail repose sur une approche incrémentale. METIS va progressivement aider à expliciter les connaissances de l'expert et raffiner un modèle structurel minimal du domaine à partir d'une suite de descriptions de classes qui vont lui être présentées. Chaque nouvelle description fournie au système constitue un enrichissement et une nouvelle vision du modèle de l'expertise soit par spécialisation d'une description déjà acquise, soit par l'ajout d'une nouvelle description. METIS modifie ainsi dynamiquement le modèle structurel des objets explicités. La taxinomie du domaine va servir de fil conducteur pour guider le processus d'explicitation des connaissances. En effet, elle va être progressivement enrichie (en caractérisant les classes une à une et en les discriminant de leur classes sœurs), permettant ainsi au système, de présenter les traits discriminants [AIMEUR & GANASCIA 93a] et d'évaluer leur pertinence, soit automatiquement, soit avec le concours de l'expert.

En définitive, on inverse le mécanisme utilisé pour la construction des systèmes experts destinés à résoudre des tâches de classification. Dans notre cas, partant d'un ensemble de classes connues, il s'agit de reconstruire la base de connaissances qui est implicitement utilisée par l'expert.

Plus précisément, le principe d'explicitation des connaissances est basé sur l'utilisation successive de quatre phases (voir § 5.3.) : l'initialisation, la classification, la discrimination et enfin la saisie de données.

La phase *d'initialisation* consiste à initialiser le noyau du domaine d'application ainsi que les informations concernant le noyau et la taxinomie considérée (nom et nombre de niveaux²). Le noyau du domaine est en fait, le langage de description initial du

¹ Dans la suite du chapitre nous emploierons indifféremment le terme taxinomie du domaine ou taxinomie des classes et le terme taxon ou classe. Un taxon dans METIS, c'est une famille, un genre ou une espèce. le terme taxinomie du domaine ou taxinomie des classes

domaine que METIS va raffiner progressivement. Le noyau est représenté par un graphe appelé "graphe primitif".

La phase de *classification* positionne une nouvelle classe (fournie par l'expert) dans la taxinomie en validant et en propageant la description qui a permis de discriminer ses classes ascendantes de leurs classes sœurs. La phase de *discrimination* intervient lorsqu'il faut distinguer une classe donnée de ses classes sœurs. La phase de *saisie de données* acquiert auprès de l'expert des descriptions supplémentaires², cette saisie se fait librement, contrairement à la saisie de traits ou de valeurs discriminantes où la saisie est guidée.

5.1. BASE DE LA TAXINOMIE DU DOMAINE

Nous allons montrer à l'aide d'un bref historique comment l'être humain a depuis longtemps pris l'habitude d'organiser ses connaissances sous forme de hiérarchies de classes où des classes plus générales, dominent des classes plus spécifiques.

Un peu d'Histoire

On ne saurait connaître précisément le moment où l'homme a commencé à systématiser ses connaissances de la nature mais cela s'est probablement réalisé très tôt. Selon [AUROI, 92] ce sont les Grecs qui, les premiers ont commencé à répertorier de manière ordonnée et systématique les plantes qu'ils trouvaient. Les expéditions d'Alexandre le Grand en direction des Indes (VI^e siècle avant J.-C.) avaient permis de découvrir de nombreuses plantes jusqu'alors inconnues en Europe, comme le bananier. Si Aristote a été à la base d'un effort de systématisation³ de ces connaissances nouvelles, c'est son disciple Théophraste (370-287 avant J.-C.), appelé le *père de la botanique* qui écrivit le premier traité sur les plantes (*Recherches sur les plantes*). Il créa aussi le premier jardin botanique avec 450 espèces, et établit un classement des plantes en trois catégories: arbres, arbustes et herbes. Cette classification est restée en vigueur pendant près de ... dix huit siècles.

² Selon l'expert, ces descriptions risquent de jouer plus tard, un rôle dans la discrimination.

³ Il a créé le premier système de classification. Ce dernier permettait de distinguer les vertébrés des non-vertébrés.

Lorsque les grandes découvertes commencèrent, le système de Théophraste ne suffit plus. Plusieurs systèmes de classification virent le jour. Entre 1623 et 1750, pas moins de 25 nouveaux systèmes de classification furent inventés, différents selon les pays (Angleterre, France, Espagne etc.).

Il appartient au naturaliste suédois Charles Linné (1707-1778), de présenter en 1735, dans *Systema naturae* un système de classification qui s'imposera universellement. Il propose un regroupement hiérarchique des espèces selon la classification suivante: Règne (animal ou végétal), Phylum ou embranchement (vertébrés ou invertébrés), Classe, Ordre, Famille, Genre, Espèce, Variété-Race.

Le système linéen avait l'avantage d'être très pratique, aussi fut-il adopté rapidement par les Anglais et à la fin du XVIII^e siècle par la France. Le système linéen a été revu au XIX^e siècle, en ce qui concerne les classes et ordres, par divers naturalistes comme A.L. de Jussieu.

Dans METIS et en accord avec Didier Paugy, l'un des spécialistes, de la Faune des poissons d'eaux douces et saumâtres de l'Afrique de l'Ouest [LEVEQUE & al, 91], nous nous sommes limités dans notre application ichthyologique aux trois niveaux suivants : Famille, genre et espèce, bien que nous aurions pu considérer un nombre quelconque de niveaux⁴.

5.2. REPRÉSENTATION DES CONNAISSANCES STRUCTURÉES DANS METIS

METIS explicite des taxinomies de classes et de ce fait, les connaissances qu'il manipule sont structurées⁵. Par ailleurs, la représentation de connaissances structurées nécessite l'utilisation d'un formalisme de représentation approprié au stockage et à la manipulation de ces informations.

⁴ Nous verrons dans le chapitre 6 que ce nombre de niveaux change dans les autres applications (minéralogie et pédologie).

⁵ Nous entendons par connaissances structurées aussi bien la structure interne des classes (les attributs) que la structure externe (hiérarchie qui organise l'ensemble des classes)

5.2.1. Formalisme de représentation des connaissances de METIS

Afin de représenter les connaissances structurées, nous avons utilisé le formalisme des graphes conceptuels tel que présenté et justifié au chapitre 4.

L'exemple ci-après montre comment représenter, sous forme d'un graphe conceptuel, des descriptions d'espèces de poissons.

Exemple :

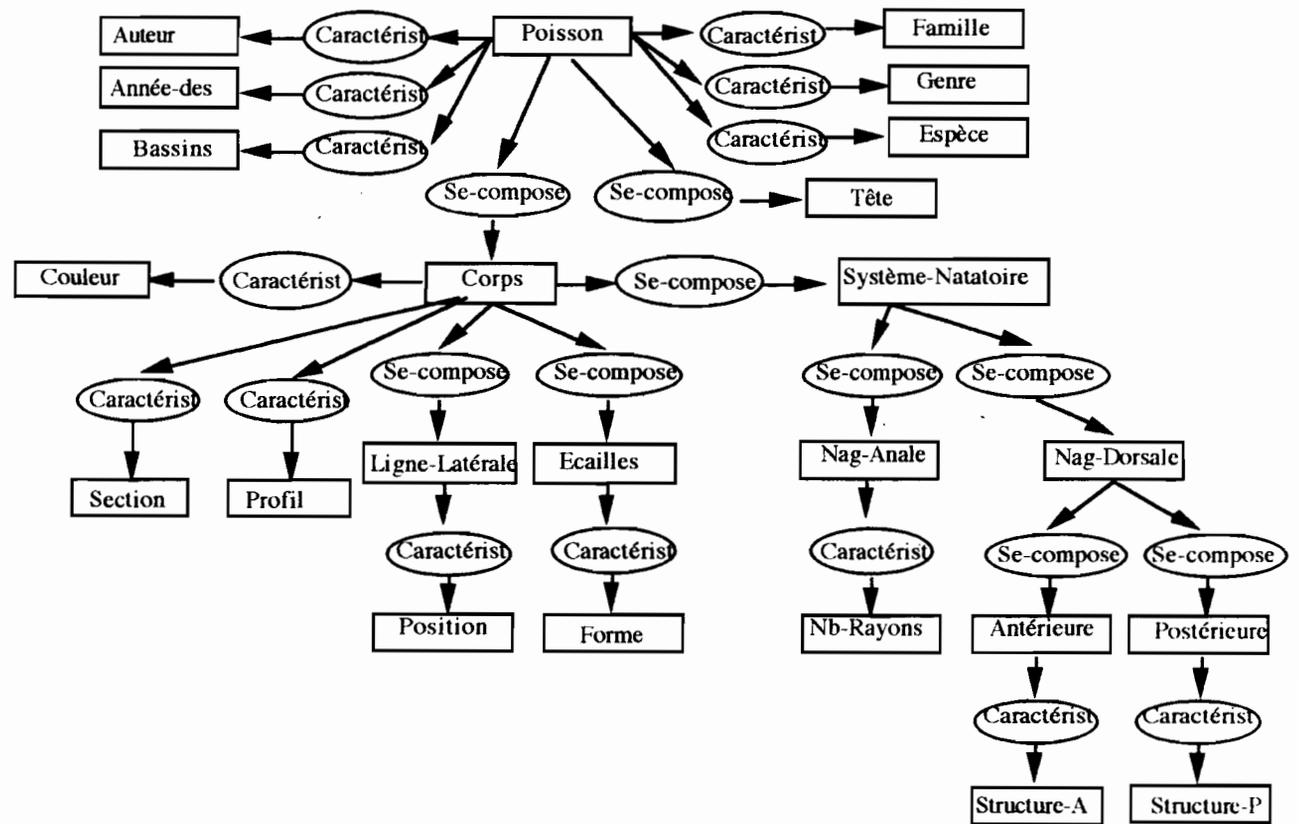


Figure 5.1. Exemple d'un graphe conceptuel en Ichtyologie

Dans ce graphe "Auteur" désigne la personne qui a décrit l'espèce, "Année-des" désigne l'année de description et "Bassins" désigne les noms de bassins dans lesquelles on trouve l'espèce. "Famille", "Genre", "Espèce" sont des noms de classe.

Limitations sur le type de graphes conceptuels utilisés

La recherche de similarités entre objets structurés est directement reliée à l'application d'une technique d'appariement de graphes (plus grand sous-graphe commun). Or, nos exemples (descriptions de poissons) sont totalement appariables et il n'y a qu'un seul appariement possible, la structure générale de base d'un poisson étant toujours la même. Nous nous limitons aux graphes conceptuels existentiels où chaque concept n'apparaît qu'une fois et où les relations sont binaires [MINEAU & al, 90] [GEY, 91]. Cette restriction correspond exactement au cas d'objets appariables⁶ sans ambiguïté et de façon unique. De plus, ce sont des formes simples qui facilitent le traitement et qui permettent de formaliser clairement les restrictions initiales. Dans ce cadre, généraliser deux graphes revient tout simplement à trouver leurs arcs conceptuels communs (triplets de la forme : concept, rel, concept). En restreignant le type de graphes utilisables, on facilite le processus de généralisation. Car pour généraliser deux graphes quelconques, il faut trouver l'ensemble de leurs sous-graphes communs, ce qui est en toute généralité un problème NP-complet⁷.

5.2.2. Structuration de la base de connaissances

Notre méthode d'explicitation des connaissances s'appuie sur deux structures hiérarchiques: d'une part la **base des graphes** et d'autre part la **taxinomie des classes** associées aux objets.

Définition 1

On définit sur chacune de ces deux structures, une relation d'ordre partiel notée "*sous-graphe-de*" [LEVINSON, 85] qui permet d'ordonner les graphes. Rappelons qu'un graphe est un couple (S, A) où S est un ensemble de sommets et A un ensemble d'arêtes. $G_1 = (S_1, A_1)$ est un sous-graphe de $G_2 (S_2, A_2)$ si et seulement si $S_1 \subseteq S_2$, $A_1 \subseteq A_2$ et si pour tout $a \in A_1$, les extrémités de a sont dans A_1 .

Un graphe G_1 est un *prédécesseur* du graphe G_2 si et seulement si G_1 est un *sous-graphe-de* G_2 , c'est à dire si l'ensemble des sommets (respectivement des liens qui

⁶ L'appariement n'est pas multiple.

⁷ La complexité dans le domaine d'apprentissage à partir d'objets structurés (sous forme de graphes et d'arbres) a été traitée dans [LIQUIERE, 90].

les reliant) dans G_1 est un sous-ensemble de l'ensemble des sommets (respectivement des liens qui les relient) dans G_2 ; ce dernier est alors un *successeur* de G_1 .

Avant de définir le graphe primitif, introduisons les notions de **concepts structurels** et de **concepts descriptifs**. En biologie et plus précisément en ichtyologie, la description des objets, bien qu'elle soit exprimée en langue naturelle, est cependant assez structurée. Nous distinguons d'une part, les *concepts et relations structurels* qui définissent les composants de l'objet et les liens qui les relient et, d'autre part, les *concepts et relations descriptifs* qui décrivent les attributs relatifs à ces composants.

Exemple :

[corps] -> (se-compose) -> [écailles] (connaissance structurelle)

corps, écailles : concepts structurels.

se-compose : relation structurelle.

[écailles] -> (caractéristique) -> [couleur] (connaissance descriptive)

couleur : concept descriptif.

caractéristique : relation descriptive.

Remarque :

Les concepts structurels sont composés d'autres concepts ils sont dits *concepts composites*, les autres concepts sont dits *non composites*. La relation qui les relie est une relation de *composition*.

Il existe différents types de composition [WINSTON & al, 87] :

Composant-composite : c'est la forme de composition la plus courante, elle décrit le rapport d'un ensemble de parties à un tout. Cette relation est fonctionnelle au sens où chaque partie a un rôle à jouer dans la composition et où les parties peuvent être différentes. Exemple : l'anse d'une tasse.

Membre-collection : il s'agit de l'appartenance d'un élément à une collection d'objets ou d'individus. Exemple : l'arbre fait partie de la forêt.

Portion-Masse : ce type de composition fait référence à la relation portion-masse ou sous-unité-unité. Contrairement à la relation Composant-composite, les parties sont toutes identiques. Exemple : les grains de sel, les minutes d'une heure.

Matière-Objet : ce type de composition fait référence à la relation matière-objet et modélise le rapport qui existe entre un objet et ce dont il est fait, constituants et ingrédients. Si un des constituants est supprimé l'objet perd son identité. Exemple : H₂O sans hydrogène (H) n'est plus de l'eau.

Phase-processus : il s'agit dans ce type de relation de la composition temporelle. Exemple : le matin, midi, l'après-midi et la soirée font partie de la journée.

Lieu-espace : il s'agit dans ce type de relation de la composition spatiale, la situation géographique d'un endroit. Exemple : l'oasis est dans le désert.

En ce qui nous concerne, nous ne retenons que la relation de composition (Composant-composite), car la description d'un organisme vivant nécessite la description des parties qui le constituent et qui sont différentes. De plus, si une partie manque l'organisme ne perd pas son identité. Par ailleurs, nous avons délibérément choisi de formuler autrement la relation Lieu-espace. Par exemple, "le poisson se trouve dans le bassin du Tchad" est exprimé par : [Poisson] -> (Caractérist) -> [Bassin : Tchad].

Définition2

Un graphe dans METIS, représente une description de taxon, c'est-à-dire, une description de famille, de genre ou encore d'espèce (figure 5.3.). Ce graphe est formé aussi bien de concepts et de relations descriptifs que de concepts et de relations structurels.

5.2.2.1. Graphe primitif

Le **graphe primitif** G_0 sert de cadre descriptif des connaissances de l'expert. La figure 5.2 illustre le graphe primitif G_0 de notre application. Il est composé essentiellement de structures constantes (présentes chez tous les poissons), des concepts classes ([Famille], [Genre], [Espèce]) et de quelques concepts descriptifs servant à désigner l'espèce. Ces derniers sont discriminants pour les espèces mais ne peuvent servir à l'identification des spécimens⁸, (exemples : [Auteur] et [Année-de-description]). Par exemple, on serait incapable d'identifier par le graphe de discrimination un spécimen

⁸ Le graphe de discrimination est destiné à l'identification de spécimens.

en connaissant simplement le nom de la personne qui a décrit l'espèce à laquelle il appartient.

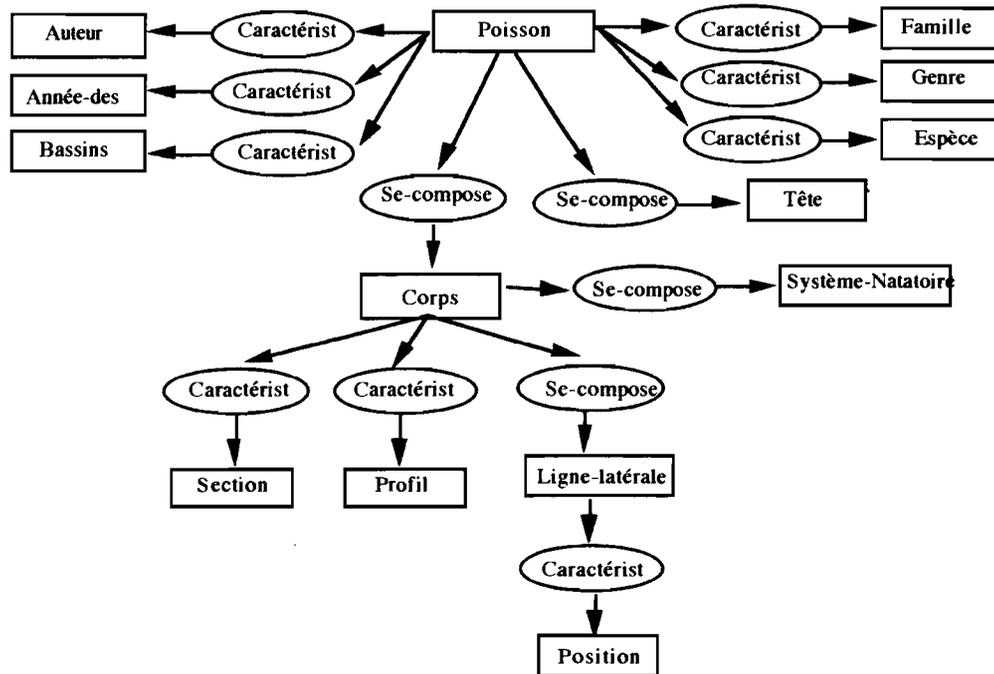


Figure 5.2. Graphe primitif G_0

5.2.2.2. Base des graphes

La **base des graphes** BG (figure 5.3) regroupe l'ensemble des graphes associés à tous les objets stockés dans METIS. Celle-ci est représentée par un sup-demi-treillis défini sur l'ordre partiel "sous-graphe-de" dont la racine (élément maximal) correspond au graphe primitif G_0 et dont les autres niveaux correspondent aux classes de la taxinomie considérée. La description d'un nœud donné de la base des graphes se compose d'une description héritée du nœud père, ainsi que d'une description propre au nœud considéré.

La description associée à chaque nœud comporte une partie *discriminante* (description qui le différencie de ses nœuds frères) et une partie *intrinsèque* (qui n'a pas encore joué de rôle dans la discrimination⁹) (voir exemple décrit en § 5.6.1). La

⁹ Elle peut devenir discriminante par la suite pour distinguer l'espèce courante de nouvelles espèces que l'expert décrira dans le futur. En fait, la description discriminante est essentielle, c'est-à-dire que lorsque la procédure d'apprentissage de METIS ne trouve de description discriminante, le système la demande d'autorité à l'expert, alors que la description intrinsèque est accessoire et est fournie spontanément au système par l'expert. En effet, l'expert peut penser que la description intrinsèque est potentiellement discriminante pour la suite.

construction de la base des graphes étant guidée par la taxinomie du domaine (figure 5.4.), le nombre de niveaux est le même dans les deux structures. Dans notre exemple, il y a trois niveaux aussi bien dans la base des graphes que dans la taxinomie du domaine.

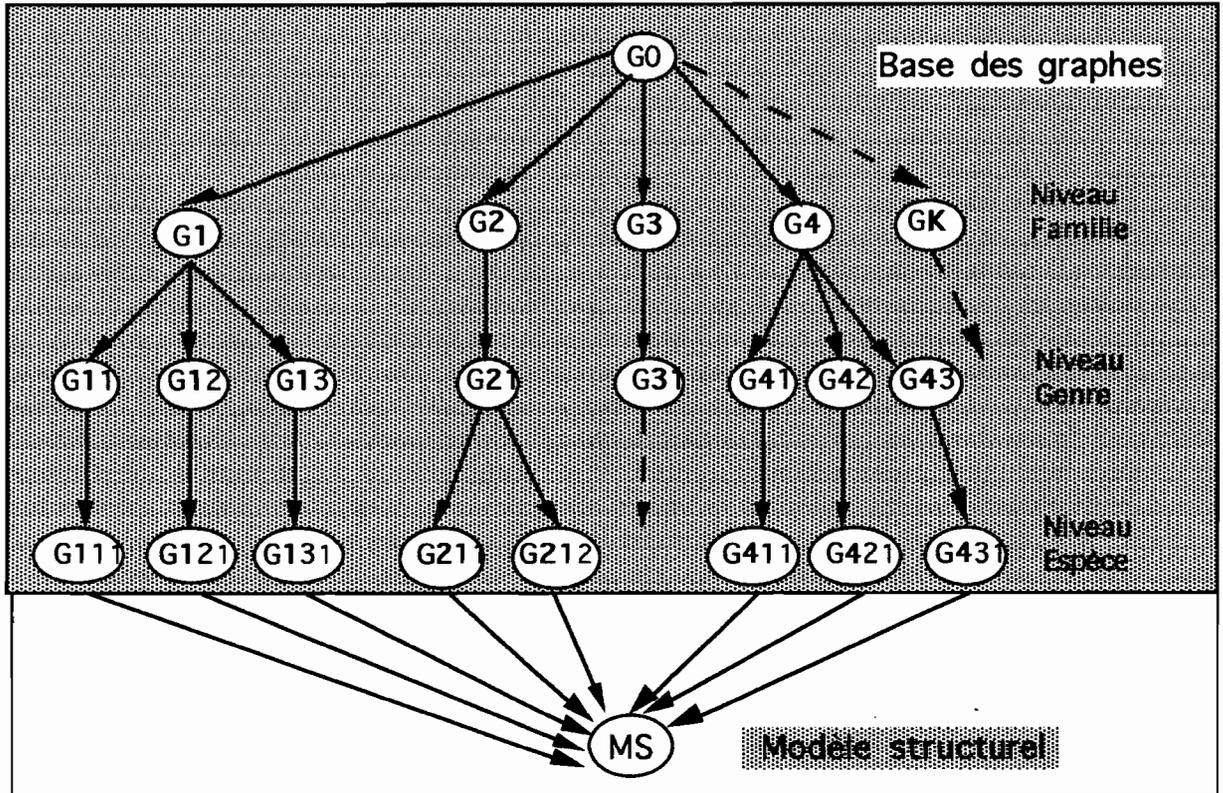


Figure 5.3. Base des graphes et Modèle structurel

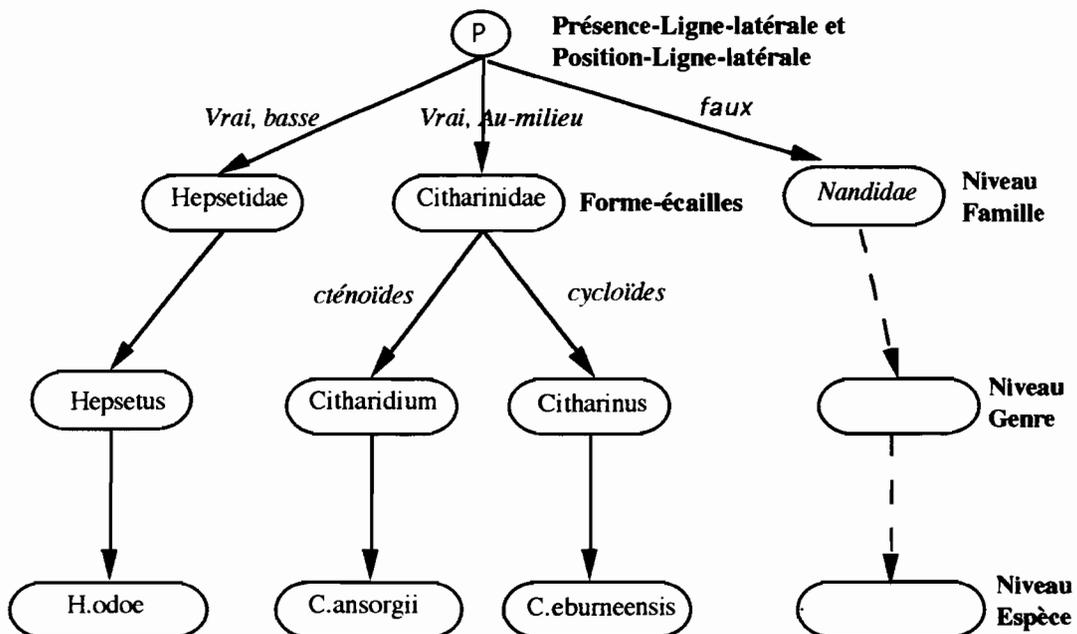


Figure 5.4. Graphe de discrimination simplifié

Remarque1 :

les flèches dans les figures 5.3. et 5.4. sont des liens de spécialisation.

Dans le cadre de notre application le graphe primitif G_0 constitue la description commune à tous les poissons, les nœuds feuilles contiennent des descriptions d'espèces, et les nœuds intermédiaires contiennent des descriptions de familles et de genres.

Remarque2 :

Dans la base des graphes de la figure 5.3., l'héritage entre deux classes est *simple*¹⁰ car une classe ne peut avoir qu'un seul ascendant direct et les classes sont exclusives.

5.2.2.3. *Modèle structurel*

Le modèle structurel est un graphe qui représente le langage de description du domaine¹¹. Il dérive de la base de graphes : c'est "la plus grande spécialisation" de tous les graphes (descriptions de classes) contenus dans la base (figure 5.3). Il contient tous les concepts, toutes les relations et toutes les valeurs qui apparaissent dans la description des graphes acquis. En d'autres termes, pour chaque graphe de la base des graphes¹², il existe une projection du modèle structurel.

5.2.2.4. *Taxinomie des classes*

Le but de la discrimination est de pouvoir discriminer une classe par rapport à d'autres grâce à une conjonction de traits appelés *traits discriminants*.

De manière générale, un trait est un arc conceptuel de la forme [Concept1 : Ref1] -> (Relation) -> [Concept2 : Ref2] où Ref1 et Ref2 peuvent être génériques. Le trait est valué quand au moins un des référents n'est pas générique. Plus précisément un trait est un élément de description qui peut prendre le statut de trait discriminant quand il permet de distinguer un taxon d'un autre.

¹⁰ C'est le cas de la méthode de classification interactive et incrémentale de [FININ, 86].

¹¹ Dans [CONRUYT & al, 93], on parle de modèle descriptif.

¹² Nous aurions pu considérer la base des graphes comme étant un treillis avec pour racine inférieure le modèle structurel. Nous avons quand même préféré les séparer car le traitement du modèle structurel est trop spécifique notamment pour le contrôle de cohérence (voir § 5.7.3.).

Un trait discriminant est formé de l'arc conceptuel suivant :

[concept-*origine*] -> (relation) -> [concept-*destination*], où le référent associé au concept-origine est générique¹³ et le référent associé au concept-destination est soit générique (cas d'un trait structurel) soit restrictif (cas d'un trait descriptif) (voir chapitre 4). Un trait discriminant est soit structurel soit descriptif. Dans le premier cas (trait structurel), il suffit qu'il soit présent chez un taxon et absent chez tous les autres pour que la discrimination soit faite. Dans le second cas (trait descriptif), il faut que les référents que peut prendre le concept destination chez les différents taxons soient exclusifs.

Exemple :

le trait [Écailles] -> (Caractérist) -> [Forme] est discriminant car il permet de distinguer entre le genre *Citharinus* et le genre *Citharidium* (Cf. figure 5.11).

Le **graphe de discrimination** GD_i (figure 5.4) reflète la taxinomie des classes du domaine selon différents niveaux (familles, genres, espèces) et se présente sous la même forme que la base des graphes. On retrouve à chaque nœud du graphe de discrimination des traits discriminants hérités du nœud père et des traits discriminants (éventuellement des conjonctions de traits) permettant la distinction entre les nœuds fils. On peut voir le graphe de discrimination comme étant la base des graphes "épurée", c'est à dire ne contenant que les noms de classes et les traits discriminants. En fait, même si la base de graphes permet aussi de discriminer les taxons, le graphe de discrimination lui, permet de le faire plus rapidement et à moindre coût puisqu'il tend à être minimal (voir § 5.6.2.1.). Les traits discriminants sont mis en relief et ne sont plus noyés dans la description des taxons.

Remarque 1 :

La notation adoptée en biologie pour désigner les espèces est la notation binômiale. C'est à dire que l'on répète le nom du genre au niveau de l'espèce. Dans la figure 5.4., pour une commodité d'écriture nous avons simplement représenté la première lettre du nom de genre.

¹³ Rappelons que quand le référent est générique on peut ne pas le mettre.

Remarque 2 :

La classe P dans la figure 5.4. correspond au graphe primitif (G_0 dans la figure 5.3.) et les traits discriminants qui lui sont associés lui permettent de distinguer les familles.

5.3. ACQUISITION DES CONNAISSANCES STRUCTURÉES DE METIS

De manière générale, le principe d'explicitation des connaissances de METIS, repose sur l'utilisation tour à tour de la base des graphes, et du graphe de discrimination. Il comporte quatre phases (Figure 5.5) que nous allons détailler successivement.

La phase d'**initialisation** consiste à initialiser le noyau (graphe primitif) du domaine d'application ainsi que les informations concernant le noyau et la taxinomie considérée (nom et nombre de niveaux). La phase de **classification** positionne une nouvelle classe (fournie par l'expert) dans la taxinomie en validant¹⁴ et en propageant la description discriminante de son ascendance. En fait, la description discriminante de l'ascendance d'une classe change de statut au niveau de la classe considérée car elle devient description intrinsèque héritée, à l'opposé de la description intrinsèque qui lui est propre. La phase de **discrimination** intervient lorsqu'il faut distinguer une classe donnée de ses classes sœurs. La phase de **saisie de données intrinsèques** acquiert auprès de l'expert des descriptions intrinsèques supplémentaires (qui peuvent jouer un rôle par la suite dans la discrimination), cette saisie se fait librement, contrairement à la saisie de traits ou de valeurs discriminantes où la saisie est guidée. Une fois la phase d'initialisation achevée, METIS passera successivement par les phases de classification, de discrimination, et enfin de saisie de données intrinsèques (voir figure 5.5.). Un contrôle de cohérence est assuré dans chaque phase. Le processus s'arrête lorsqu'il n'y a plus de description de classes à expliciter.

¹⁴ METIS demande à l'expert de confirmer que les traits discriminants ainsi que les valeurs associées à ces traits sont toujours valables.

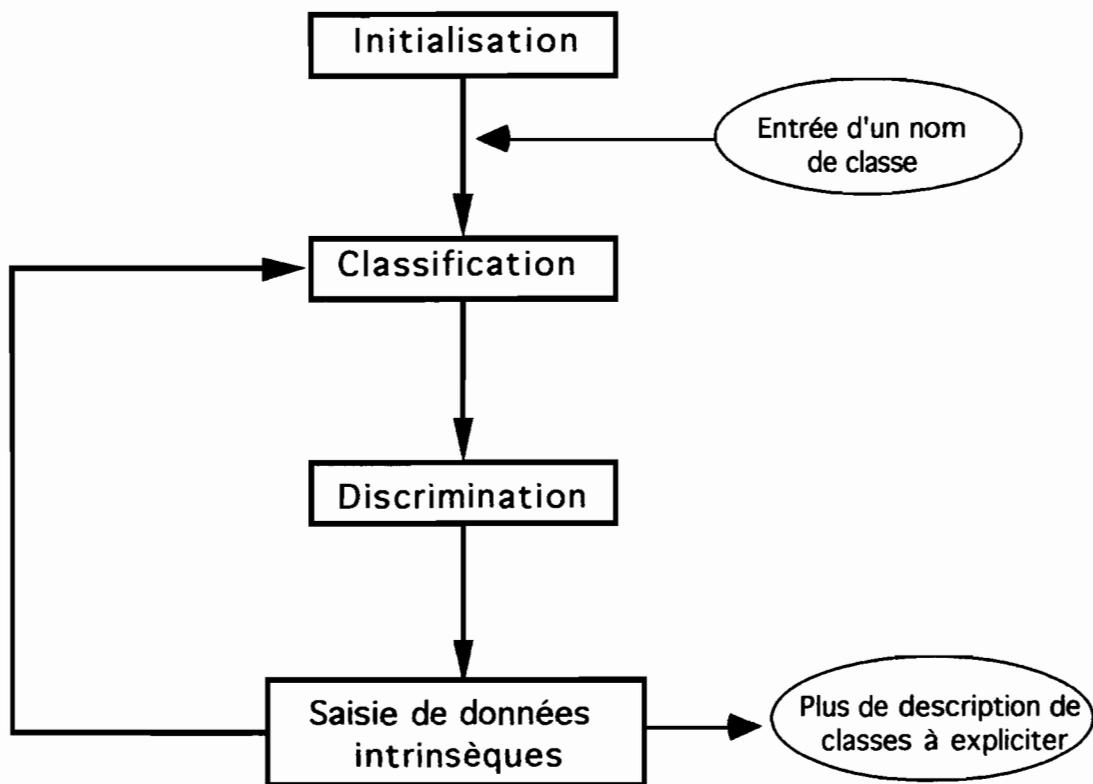


Figure 5.5. Phases d'explicitation des connaissances

Dans cette section, nous allons simuler le principe d'explicitation des connaissances sur trois exemples (description d'espèces) que nous présentons au fur et à mesure dans les différentes phases. L'explicitation d'une description d'espèce revient à expliciter des descriptions partielles de la famille, du genre, puis de l'espèce. On procède ainsi, taxon par taxon.

5.4. PHASE D'INITIALISATION

Dans cette phase, METIS demande à l'expert d'introduire le nombre et les noms des niveaux de la taxinomie que l'on va considérer ainsi que le noyau minimal du domaine d'application.

5.4.1. Niveaux du domaine

Dans notre application, le nombre de niveau est de 3 et les noms de niveaux, du plus général au plus spécifique, sont : Famille, Genre, Espèce.

5.4.2. Explicitation du noyau G_0

METIS demande à l'expert d'introduire un premier concept C_i ainsi que les couples (relations, concepts non composites) et les couples (relations, concepts composites) qui lui sont associés, en prenant bien soin d'indiquer à chaque fois si le concept est composite ou non. Ensuite, il demande de manière récursive pour chaque concept composite lié à C_i les couples (relations, concepts non composites) et les couples (relations, concepts composites) qui lui sont associés. L'explicitation du noyau se termine quand l'expert a introduit le dernier concept composite.

METIS initialise la base des graphes avec G_0 et la taxinomie des classes avec une classe racine (appelée P dans notre cas car on traite des Poissons).

Exemple :

Dans notre application, le noyau (figure 5.2.) est explicité de la manière suivante : l'expert commence par introduire le concept [Poisson], puis les six couples (relations, concepts non-composites) et les deux couples (relations, concepts composites) suivants :

[Poisson] -> (Caractérist) -> [Famille]

[Poisson] -> (Caractérist) -> [Genre]

[Poisson] -> (Caractérist) -> [Espèce]

[Poisson] -> (Caractérist) -> [Auteur]

[Poisson] -> (Caractérist) -> [Année-des]

[Poisson] -> (Caractérist) -> [Bassins]

[Poisson] -> (Se-compose) -> [Corps]

[Poisson] -> (Se-compose) -> [Tête]

Ensuite, METIS demande à l'expert de faire de même pour le premier concept composite [Corps], ce qui va amener l'expert à introduire les couples suivants :

[Corps] -> (Caractérist) -> [Section]

[Corps] -> (Caractérist) -> [Profil]

[Corps] -> (Se-compose) -> [Ligne-Latérale]

[Corps] -> (Se-compose) -> [Système-Natatoire]

Le concept [Ligne-Latérale] étant encore un concept composite, l'expert introduit le couple suivant :

[Ligne-Latérale] -> (Caractérist) -> [Position]

METIS a fini d'expliciter le noyau car le concept [Position] n'est pas composite et l'expert a décidé de ne plus spécialiser les concepts [Système-Natatoire] et [Tête].

5.4.3. Informations sur le noyau

L'expert peut fournir à METIS des informations concernant le noyau. Il indique, par exemple, que les concepts [Auteur] (nom de la personne qui a décrit l'espèce), [Année-description] et [Bassins] (noms des bassins où l'on peut trouver l'espèce) ne sont demandables qu'au dernier niveau (niveau espèce), car ils servent à décrire la fiche signalétique de l'espèce. C'est-à-dire qu'il ne faut pas questionner l'expert aux deux premiers niveaux¹⁵ sur des informations qui n'ont de sens qu'au dernier niveau. De plus, les arcs qui les comprennent ne sont pas généralisables ; par exemple, si toutes les espèces d'un même genre ont été décrites la même année, on ne peut généraliser cette information au niveau du genre. De même, l'expert peut interdire l'utilisation de certains arcs conceptuels pour la discrimination (voir § 5.6.) tels que : ([Poisson] -> (Caractérist) -> [Famille], [Poisson] -> (Caractérist) -> [Genre], [Poisson] -> (Caractérist) -> [Espèce] car ce sont des noms de taxons) et ([Poisson] -> (Caractérist) -> [Auteur], [Poisson] -> (Caractérist) -> [Année-des], [Poisson] -> (Caractérist) -> [Bassins]...) car ils n'apportent aucune information discriminante pour identifier les spécimens d'une espèce.

¹⁵ Rappelons qu'un niveau est un rang taxinomique (voir chapitre 1)

5.4.4. Cohérence

Lors de la saisie du noyau, METIS contrôle les erreurs de syntaxe (exemple : [Poisson] -> [Caractérist] -> [Famille] n'est pas admis car la relation est toujours entre parenthèses) et vérifie qu'un nom de concept n'apparaît qu'une fois (compte tenu de la restriction imposée). Ceci assure l'absence de circuits.

Quand l'expert déclare que l'initialisation est terminée, METIS est en mesure d'expliciter la première description, c'est-à-dire le premier exemple. L'explicitation d'une description se fait par explicitation de ses descriptions partielles (METIS demande à l'expert de décrire la famille, puis le genre et enfin l'espèce).

5.5. PHASE DE CLASSIFICATION

Comme nous l'avons dit au chapitre 2, le terme classification est utilisé dans le sens du processus de raisonnement sur les classes. Plus précisément ces dernières étant données par l'expert, il s'agit de construire un arbre de discrimination, permettant de distinguer les individus appartenant aux classes, de façon incrémentale.

METIS commence par questionner l'expert afin de *prendre connaissance* des classes de l'objet dont la description est à expliciter (nomenclature du taxon [famille], [genre], [espèce]), puis essaie de *positionner* la ou les nouvelle(s) classe(s) à l'aide du graphe de discrimination en validant les traits discriminants qui ont permis de caractériser l'ascendance de la ou des classe(s) considérée(s). Après avoir créé une classe, METIS lui associe une étiquette qui désigne son numéro d'ordre dans la base des graphes. Cette étiquette sert lors du calcul de *couverture de classes* (voir § 5.7.3) pour le contrôle de cohérence, et elle est mise à jour dès que la classe a au moins deux classes filles (ce sera justifié au § 5.7.3.). L'algorithme de numérotation et de mise à jour des étiquettes de classes est donné dans (§ 5.7.3) et en Annexe 1.

Remarque :

Dans la phase de classification, le graphe de discrimination est enrichi uniquement par les classes (voir figure 5.6.). Les traits discriminant les nouvelles classes n'interviennent que dans la phase de discrimination¹⁶.

5.5.1. Parcours de la taxinomie

Dans le but de positionner la nouvelle classe, la taxinomie est parcourue de manière descendante en accord avec l'ordre partiel existant, c'est à dire que l'on se déplace des classes les plus générales vers les plus spécifiques .

Exemple :

La première description que l'expert désire fournir à METIS est celle de l'espèce *Citharinus eburneensis* appartenant au genre *Citharinus* de la famille *Citharinidae*. METIS ayant pris connaissance de l'espèce dont la description est à expliciter et de sa position dans la taxinomie¹⁷, construit le graphe de discrimination et la base de graphes initiale constituée de G_0 , G_1 , G_{11} et G_{111} (figure 5.6).

Remarque :

A ce stade les graphes G_1 , G_{11} et G_{111} ne diffèrent de G_0 que par la spécialisation du concept classe ([Famille], [Genre] ou [Espèce]). C'est-à-dire que nous aurons le concept [Famille : Citharinidae] dans G_1 et les concepts [Famille : Citharinidae] et [Genre : Citharinus] dans G_{11} et les concepts [Famille : Citharinidae] [Genre : Citharinus] et [Espèce : eburneensis] dans G_{111} .

¹⁶ Dans la phase de classification, METIS vérifie que les traits discriminants déjà présents dans l'ascendance de la classe considérée sont toujours valides.

¹⁷ La prise en compte de la description de la première classe est un peu particulière, car on suppose que METIS dispose au préalable d'un graphe de discrimination formé que d'un seul nœud : la racine. En effet le lecteur pourrait se demander comment on peut classer s'il n'y a pas de graphe de discrimination. Étant donné que METIS ne remet pas en cause les positions des classes dans la taxinomie, l'algorithme d'explicitation commence par positionner la classe avant de la discriminer de ses classes sœurs si nécessaire.

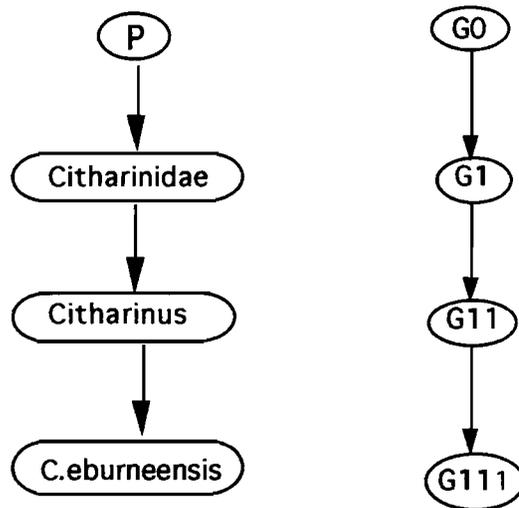


Figure 5.6. Graphe de discrimination 1 et base de graphes initiale

5.5.2. Cohérence : validation des traits hérités

Nous allons décrire le principe de contrôle de cohérence dans la phase de classification sans toutefois l'illustrer. Car ce serait plus parlant quand il y a la phase de classification suivie de la phase de discrimination (cas d'au moins deux taxons). Nous y reviendrons donc dans la phase de discrimination lors de l'introduction du second exemple que l'on classera avant de le discriminer du premier (voir § 5.6.3.1.).

A chaque passage avec succès dans le graphe de discrimination (on passe par exemple du niveau famille au niveau genre en *validant le trait discriminant* quand il existe), le niveau correspondant dans la base des graphes est aussi validé (METIS valide les traits intrinsèques du taxon considéré). Puis le processus revient sur le graphe de discrimination et ainsi de suite jusqu'à arriver au niveau où l'on a positionné la classe.

Quand un trait discriminant hérité n'est pas valide, il se produit une incohérence. C'est à dire que la valeur qui est associée au concept destination (voir § 5.2.2.4.) n'est pas vérifiée, ou tout simplement que le trait est *inapplicable* (on ne peut décrire la *couleur* de la nageoire adipeuse si cette dernière est absente). Il faut alors, soit étendre le domaine de valeurs du concept avec la nouvelle valeur, en admettant que l'expert a omis de la donner, soit faire appel à l'expert pour remettre en question le trait et le "redescendre", aux niveaux inférieurs. Si le trait intrinsèque hérité est structurel, il faut redescendre aussi toute la sous-structure dépendant de la présence de ce trait.

METIS effectue ensuite les mises à jour nécessaires dans le graphe de discrimination et dans la base des graphes.

5.6. PHASE DE DISCRIMINATION

La phase de discrimination est illustrée par la figure 5.7. Durant cette phase, deux cas se présentent : soit le trait (ou la conjonction de traits) discriminant est *suffisant* pour discriminer la nouvelle classe de ses classes sœurs, soit, il ne l'est pas. Un trait est dit suffisant quand la valeur du trait discriminant disponible pour la nouvelle classe à discriminer est différente de toutes celles que prend le trait pour les classes sœurs.

Dans le cas où le trait est insuffisant, METIS doit d'abord apprendre un nouveau trait et ce n'est qu'en cas d'échec qu'il doit s'adresser à l'expert pour obtenir un nouveau trait. L'échec se produit quand la procédure d'apprentissage de METIS ne fournit aucun trait (voir § 5.6.2.1.) ou quand l'expert rejette tous les traits proposés par METIS.

Remarque :

METIS ne traite que les classes monothétiques, c'est-à-dire que l'attribut discriminant est possédé par tous les individus appartenant à la classe considérée.

Le problème de recherche de traits discriminants d'un ensemble de classes sœurs peut être présenté de la manière suivante :

- Étant donné une classe c et ses sous-classes $\{c_1, c_2, \dots, c_n\}$, telle que chacune possède une description (obtenue à partir des nœuds de la base des graphes).
- Trouver un ensemble minimal de traits $\{t_1, t_2, \dots, t_d\}$ permettant de discriminer les classes $\{c_1, c_2, \dots, c_n\}$, les unes par rapport aux autres. Ces traits sont appelés traits discriminants.

Les traits discriminants sont à rechercher parmi les traits appartenant à la description intrinsèque des classes.

Nous allons simuler maintenant la phase de discrimination sur trois exemples (Figure 5.8 à 5.15).

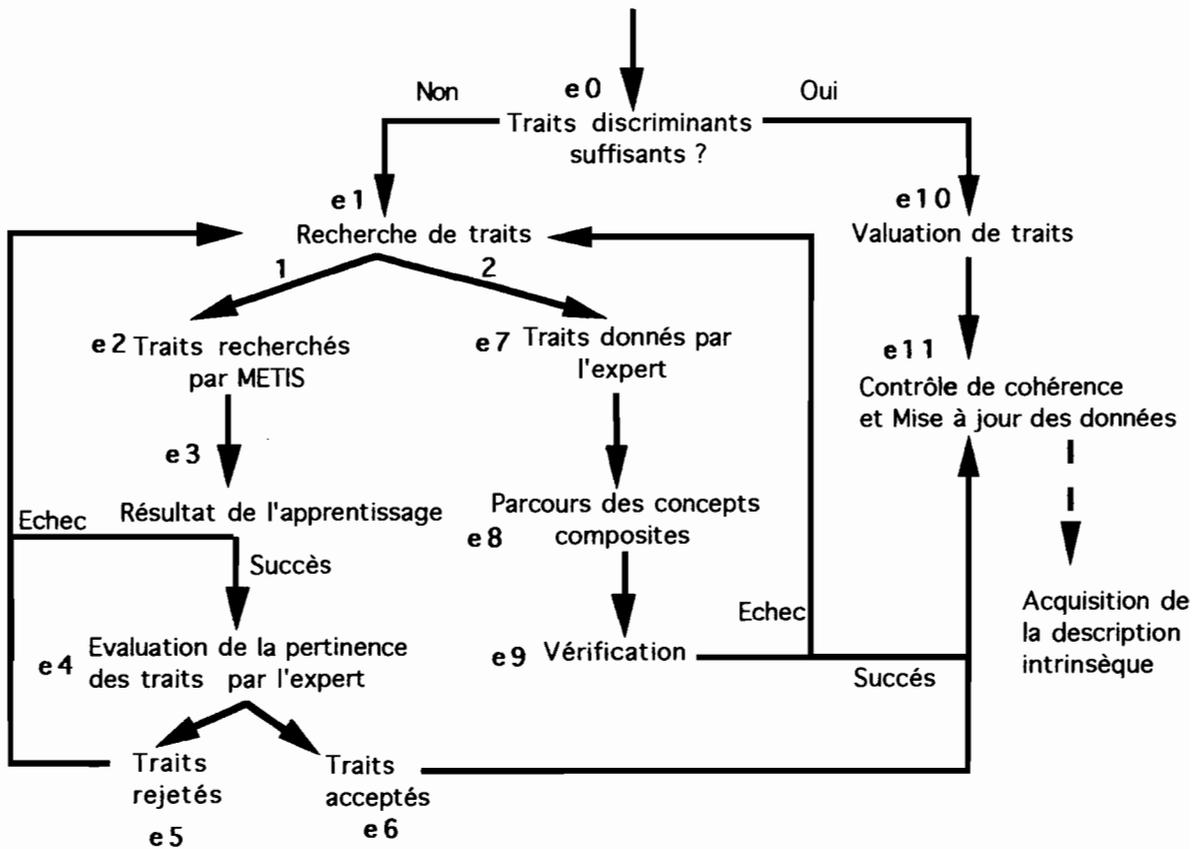


Figure 5.7. Phase de discrimination

Remarque :

L'étape *valuation* du trait (e10) signifie l'ajout d'une nouvelle valeur au trait.

Dans l'étape (e1) la recherche de traits se fait d'abord par METIS et ce n'est qu'en cas d'échec que l'on fait appel à l'expert.

La flèche en pointillés signifie que l'étape acquisition de la description intrinsèque n'est pas obligatoire et qu'elle appartient à la dernière phase (**saisie de données intrinsèques**).

5.6.1. Traits discriminants

Il existe deux types de traits discriminants : les traits discriminants descriptifs et les traits discriminants structurels (voir § 5.2.2.4.).

Un trait discriminant est un arc conceptuel "non valué" c'est à dire que le concept *destination* tout comme le concept *origine* sont génériques (exemples : [Corps] -> (Se-

Compose) -> [Écailles] ou [Écailles] -> (Caractérist) -> [Forme]). Quand un trait discriminant est associé à un niveau donné, (par exemple niveau famille) il devient description discriminante au niveau immédiatement inférieur (niveau genre).

Si on associe à un nœud N1 du graphe de discrimination un trait discriminant descriptif lui permettant de différencier ses nœuds fils, ces derniers auront alors pour description discriminante dans la base de graphes, le trait valué¹⁸.

Exemple :

Nous allons montrer quelle est la description intrinsèque et la description discriminante des nœuds (Citharinidae, Citharinus, Citharidium) de la base des graphes et quels sont les traits discriminants associés aux mêmes nœuds dans le graphe de discrimination. Cette illustration correspond à l'étape où il y a eu description du genre Citharidium (voir figure 5.13).

Pour ne pas répéter la description intrinsèque de Citharinidae, nommons la "desc".

Citharinidae

Description intrinsèque : desc

Description discriminante : { \emptyset } (car cette classe ne possède pas de classes sœurs)

Traits discriminants : {[Écailles] -> (Caractérist) -> [Forme]}

Citharinus

Description intrinsèque : desc (héritage de desc mais pas de traits intrinsèques propres)

Description discriminante : {[Écailles] -> (Caractérist) -> [Forme : Cycloïde]}

Traits discriminants : { \emptyset } (car pas de fils à discriminer)

¹⁸ Si on associe dans le graphe de discrimination un trait discriminant structurel, à un taxon famille par exemple, il ne sera pas valué quand il appartiendra à la description discriminante des taxons fils dans la base des graphes. En effet un trait structurel n'est jamais valué qu'il soit discriminant ou pas, en fait, il est soit présent soit absent.

Citharidium

Description intrinsèque : desc (héritage de desc mais pas de traits intrinsèques propres)

Description discriminante : {[Écailles] -> (Caractérist) -> [Forme : Cténoïde]}

Traits discriminants : { \emptyset } (car pas de fils à discriminer)

5.6.1.1. Traits discriminants suffisants

Quand le trait discriminant présent est suffisant pour discriminer la nouvelle classe par rapport à ses classes soeurs (Figure 5.7, étape e0), il n'y a pas de recherche de nouveaux traits mais simplement une adjonction de la valeur du trait qui a permis de différencier la nouvelle classe. Ceci se fait d'abord au niveau du graphe de discrimination puis au niveau de la base des graphes (Figure 5.7, étape e10). Dans l'étape e11 METIS effectue un contrôle de cohérence (voir § 5.3.4.3.) et les mises à jour qui s'imposent au niveau de la base des graphes et au niveau du graphe de discrimination.

A un niveau donné (niveau genre, par exemple), il y a échec dans la phase de discrimination (Figure 5.7, étape e1), quand le trait discriminant s'avère insuffisant.

On parle de trait discriminant insuffisant quand le trait suffisait à discriminer un ensemble de classes E1 à une étape donnée et qu'il devient insuffisant pour discriminer une nouvelle classe par rapport à ses classes sœurs en l'occurrence E1.

Dans le cas d'un échec, il faut pouvoir apprendre au moins un trait permettant de le distinguer (figure 5.7, étape e2), soit à défaut, le demander de manière explicite à l'expert (figure 5.7, étape e7).

5.6.1.2. Traits discriminants insuffisants

Quand les traits discriminants sont insuffisants, METIS recherche un trait ou une conjonction de traits discriminants à l'aide d'une procédure d'apprentissage (étape e2). Il demande ensuite à l'expert d'évaluer la pertinence des traits trouvés (voir § 5.6.2.1.). Si l'expert les rejette, METIS recherche de nouveaux traits. En cas d'échec, il fait appel à l'expert afin qu'il lui fournisse de nouveaux traits.

5.6.2. Recherche de traits discriminants par METIS

Pour apprendre de nouveaux traits METIS fait appel à un algorithme d'apprentissage inspiré d'ID3 (voir chapitre 2) appelé IDICE et que nous présenterons après les exemples, ci-dessous.

Pour illustrer le cas de traits discriminants insuffisants nous allons présenter deux descriptions d'espèces. Ceci revient à expliciter les taxons suivants : la famille Citharinidae les genres Citharinus et Citharidium et enfin les espèces Citharinus-eburneensis et Citharidium-ansorgii.

Exemples :

Première description

L'expert désirant commencer par décrire l'espèce Citharinus-eburneensis (voir phase de classification), METIS va expliciter la description de l'espèce en explicitant auparavant celle de la famille Citharinidae puis du genre Citharinus auxquels elle appartient. C'est à dire que l'expert ne peut donner les informations spontanément, il faut qu'il réponde aux questions de METIS.

Remarque :

METIS se sert du graphe primitif (modèle structurel de base) pour entamer le dialogue. C'est-à-dire qu'il prend un à un les concepts présents dans le graphe primitif et demande la valeur qui est associée à chacun d'eux.

Explicitation de la famille Citharinidae

METIS pose des questions¹⁹ à l'expert par l'intermédiaire du *graphe primitif* G_0 , afin d'expliciter les informations concernant la famille (exemples : Comment est la section du corps, comment est le profil du corps ?, quelle est la position de la ligne latérale ?). Si d'après l'expert, le graphe primitif est insuffisant pour caractériser la famille, METIS lui demande²⁰ de fournir un trait ou une structure conceptuelle intrinsèque susceptible de la

¹⁹ METIS pose des questions à propos des concepts non composites avant de s'intéresser aux concepts composites

²⁰ Comme lors de l'explicitation du noyau, METIS demande à l'expert de lui introduire en priorité les concepts caractéristiques avant les concepts composites. Dans le cas de plusieurs concepts composites il lui demande d'introduire en priorité ceux qui ont le moins de composants.

discriminer par rapport à toutes les autres familles. En effet dans ce cas, METIS ne peut plus guider l'acquisition étant donné que le graphe primitif auquel il se réfère est insuffisant.

Par ailleurs, sachant que le nom du descripteur, l'année de description ainsi que les noms de bassins sont des informations relatives à la fiche signalétique de l'espèce (données sous forme d'informations sur le noyau lors de la phase d'initialisation, voir § 5.4.), METIS attend d'arriver au stade de la description de l'espèce pour poser les questions concernant les valeurs associées à ces informations (Figure 5.10).

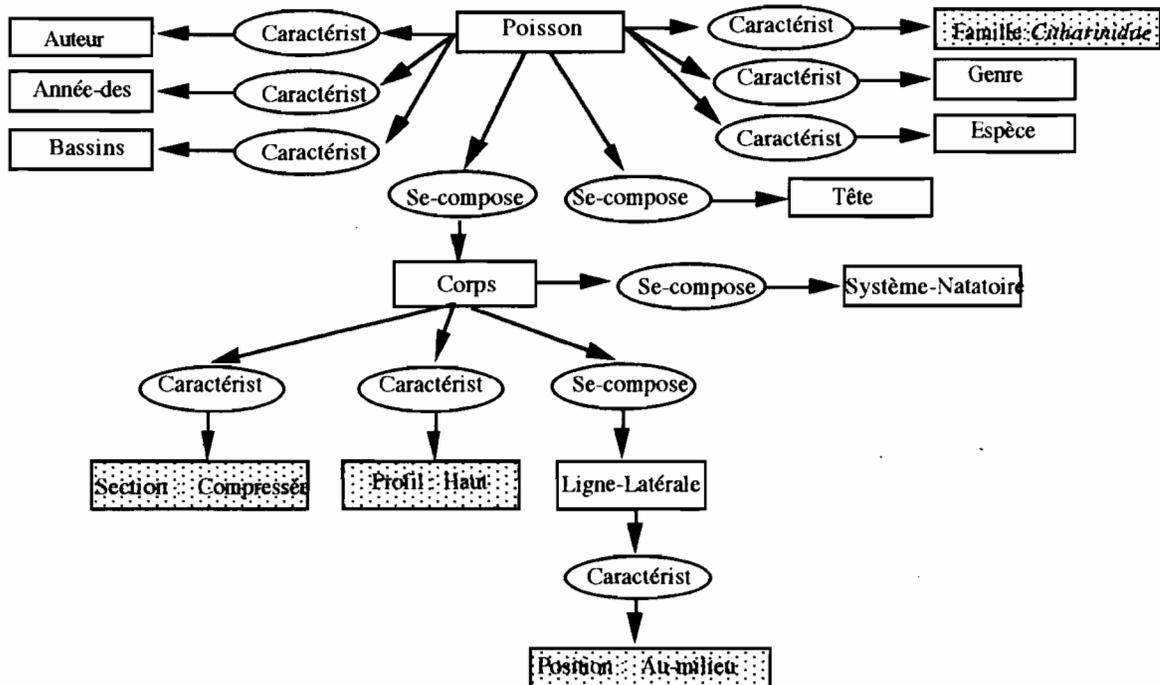


Figure 5.8. Graphe de la famille Citharinidae (G₁)

Explicitation du genre Citharinus

A cette étape, METIS demande à l'expert un trait décrivant le genre *Citharinus*, car la description de famille héritée de l'étape précédente ne suffit pas d'après l'expert. Il faut donc raffiner le graphe de la famille ce qui amène l'expert à introduire les traits suivants : [Corps] -> (Se-compose) -> [Écailles] et [Écailles] -> (Caractérist) -> [Forme : Cycloïde]. Le nouveau graphe est illustré par la figure 5.9.

Remarque

Dans les graphes ci-dessous, les structures en gras sont propres au graphe considéré, alors que le reste du graphe est hérité.

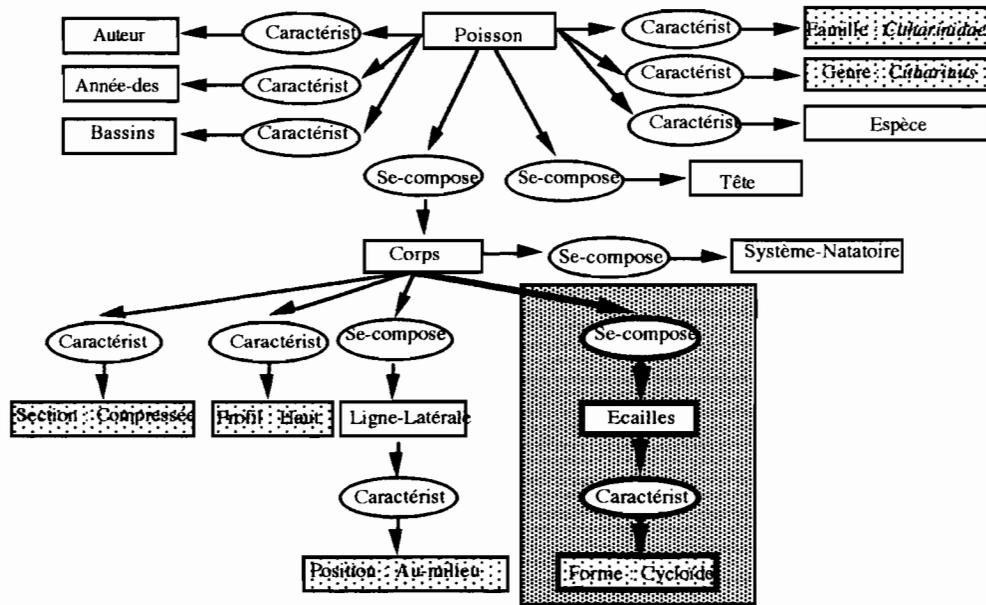


Figure 5.9. Graphe du genre *Citharinus* (G₁₁)

Explicitation de l'espèce *Citharinus eburneensis*

Dans cette dernière étape, METIS commence par compléter le graphe primitif en demandant à l'expert les valeurs associées aux concepts utiles à l'identification : le nom de l'auteur, l'année de description et les noms de bassins où on trouve cette espèce, puis il demande à l'expert ce qui, selon lui, décrit cette espèce et ce qui est susceptible de la distinguer de toutes les autres. L'expert introduit alors toute la description relative au système natatoire.

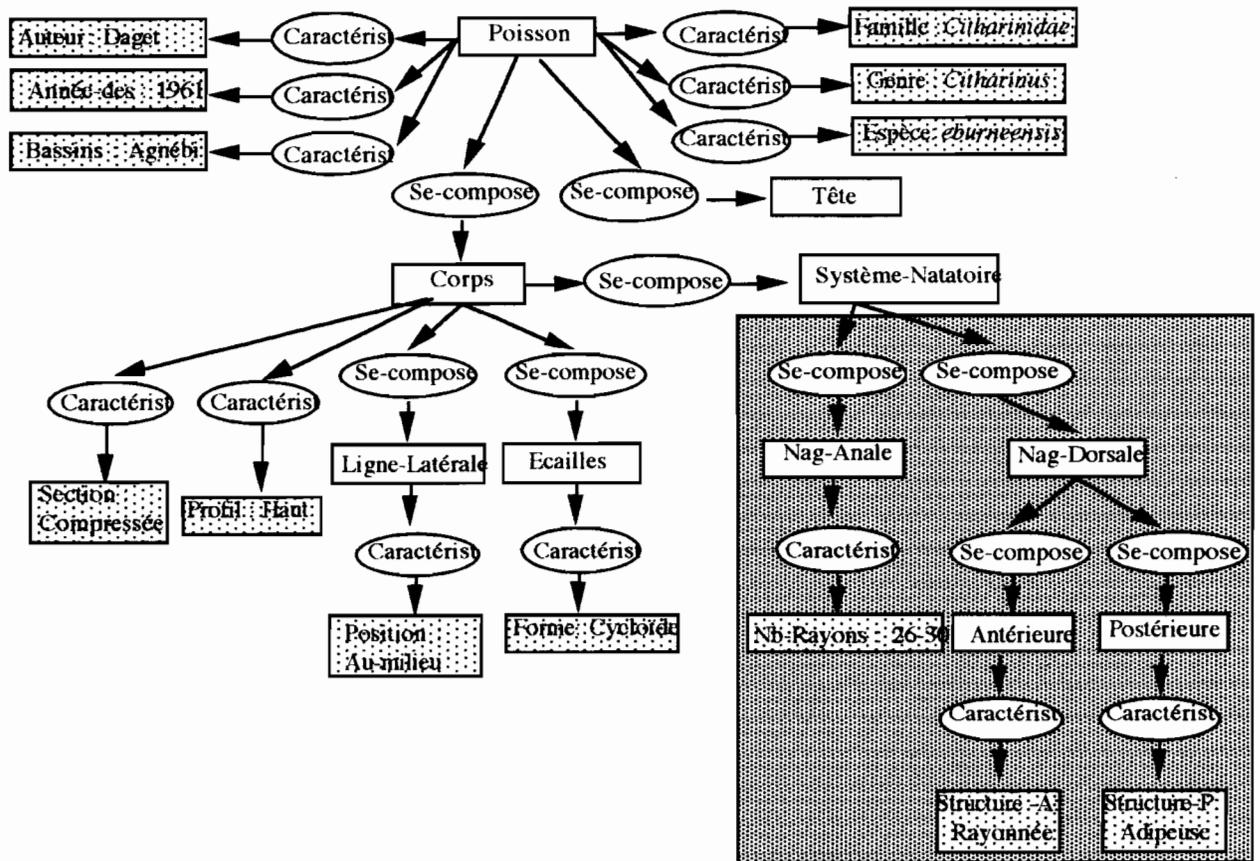


Figure 5.10. Graphe de l'espèce *Citharinus eburneensis* (G111)

Dans cette première description, il n'y a pas de discrimination à faire car il n'y a qu'une seule espèce, et donc absence de traits discriminants.

Deuxième description

La deuxième description à expliciter est celle de l'espèce *Citharidium ansorgii* appartenant au genre *Citharidium* de la famille *Citharinidae*. En effet, l'expert souhaite raffiner la taxinomie en décrivant les classes *Citharidium* et *Citharidium ansorgii*.

Classification

Avant de rattacher la classe *Citharidium* à la classe *Citharinidae*, METIS doit vérifier auprès de l'expert que *Citharidium* est une spécialisation de *citharinidae*, c'est-à-dire que le profil du corps est bien haut, que la section est compressée et que la position de la ligne latérale est au milieu (voir figure 5.8.). Après cela, il rattachera la classe *Citharidium ansorgii* à celle de *Citharidium*.

Discrimination

La description de la famille devant être commune avec celle de l'espèce *Citharinus eburneensis*, METIS va examiner les traits décrivant le genre *Citharidium* afin de trouver un trait potentiellement discriminant. METIS ayant le choix entre le trait [Corps] -> (Se-compose) -> [Écailles] et le trait [Écailles] -> (Caractérist) -> [Forme] (voir figure 5.9), commence par demander à l'expert si le genre *Citharidium* possède des écailles. L'expert répond par l'affirmative, et METIS ne peut donc pas prendre ce trait pour discriminer les genres.

METIS va ensuite s'enquérir de la forme des écailles²¹ de *Citharidium ansorgii*. Aussitôt, l'expert indique qu'elle n'est plus cycloïde, mais cténoïde. METIS "décide alors" que la forme des écailles est un trait discriminant et soumet ce trait à l'expert. Ce dernier donne son approbation car le trait proposé est réellement pertinent pour distinguer les deux genres (voir figure 5.11). METIS associe alors au nœud *Citharinidae* du graphe de discrimination le trait discriminant trouvé.

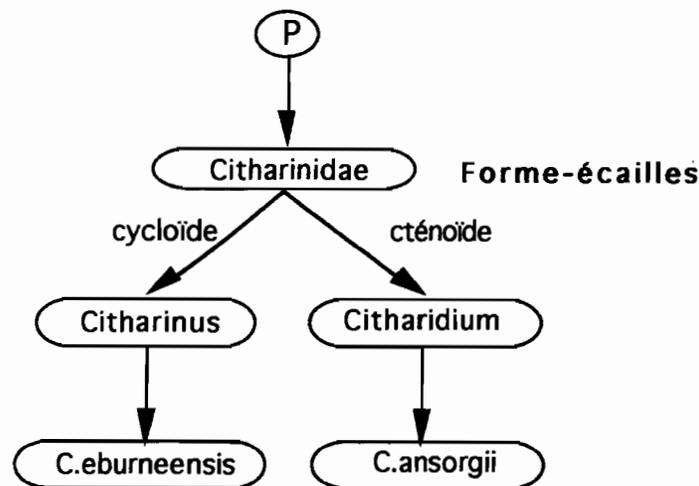


Figure 5.11. Graphe de discrimination 2

METIS met à jour la base de graphes en remontant²² la structure ([Corps] -> (Se-compose) -> [Écailles], [Écailles] -> (Caractérist) -> [Forme]) au niveau du graphe de la famille *Citharinidae*. Cette généralisation peut se faire car la structure est présente dans toutes les classes à discriminer (voir figure 5.12). Ainsi, selon la valeur de la forme des

²¹ Le seul trait intrinsèque qui reste et qui décrit le genre *Citharinus* est la forme des écailles. Étant donné qu'il y a des chances que ce trait devienne discriminant METIS pose la question pour le genre *Citharidium*.

²² La généralisation de traits se fait grâce au critère de factorisation maximale donné dans le paragraphe 5.7.3.

écailles, on évolue soit dans le graphe du genre *Citharinus* (voir figure 5.8), soit dans celui du genre *Citharidium* (voir figure 5.13). Dans ce cas précis, le trait ([Corps] -> (Se-compose) -> [Écailles]) sera hérité en tant que description intrinsèque et ([Écailles] -> (Caractérist) -> [Forme]) en tant que description discriminante, après valuation du trait (voir § 5.3.3.1). Enfin, METIS étend la base des graphes avec le graphe G₁₂ (figure 5.13).

Remarque :

Si plus tard il arrive qu'un genre ne possède pas d'écailles, ceci sera détecté par l'étape de validation de la phase de classification (voir § 5.5.2.). En effet, METIS demandera entre autres à l'expert si le trait [CORPS] -> (Se-compose) -> [Écailles] est encore valide. L'expert répond non et METIS redescend alors ce trait et toute la structure qui en dépend (dans notre cas [Écailles] -> (Caractérist) -> [Forme]) au niveau inférieur dans les classes concernées²³.

L'espèce *Citharidium.ansorgii* étant un premier descendant du genre *Citharidium* (pas de taxons frères), METIS demande à l'expert encore une fois, ce qui permet de décrire cette espèce et qui soit susceptible de la distinguer des autres espèces appartenant au même genre. Une fois que l'expert introduit la couleur du corps pour décrire l'espèce *Citharidium.ansorgii*, METIS étend la base de graphes avec G₁₂₁ (voir figure 5.14)

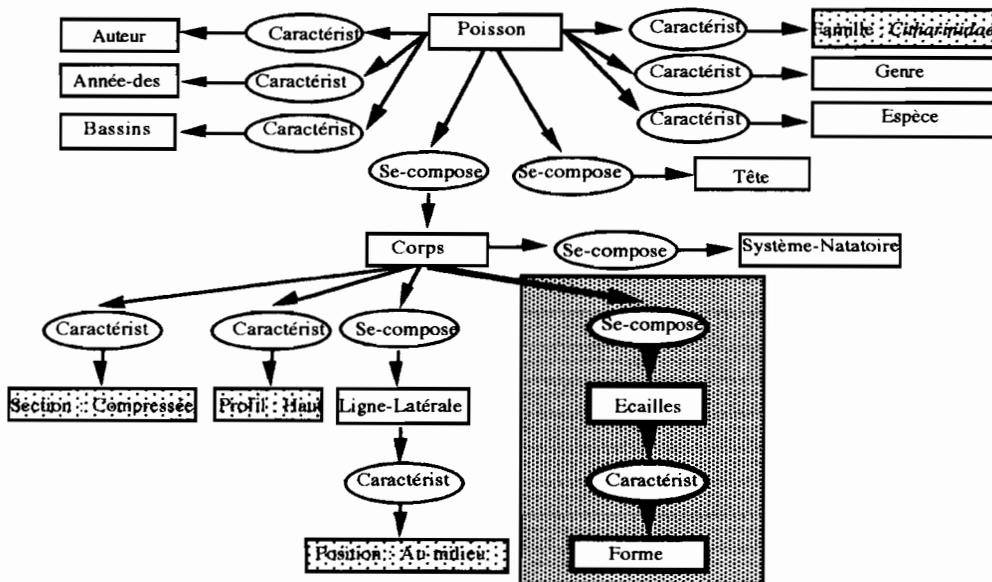


Figure 5.12. Graphe de la famille Citharinidae modifié (G₁)

²³ METIS ne traite pas les disjonctions discriminantes.

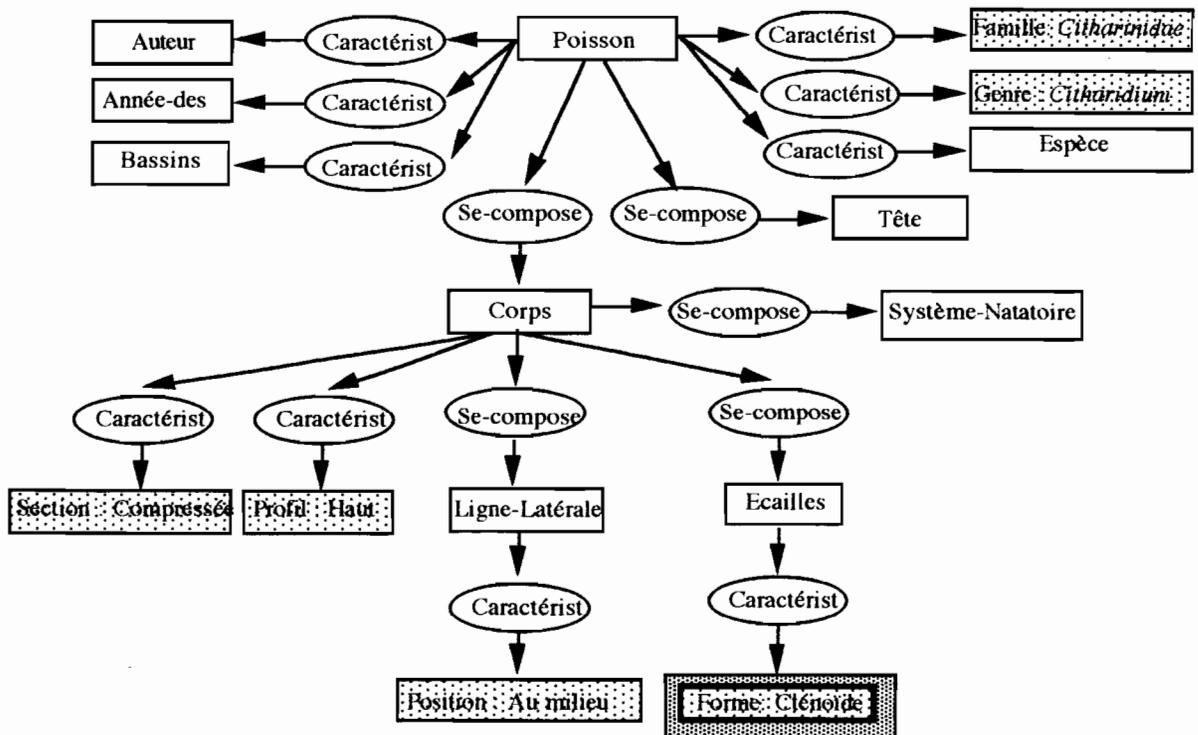


Figure 5.13. Graphe du genre Citharidium (G12)

Remarque :

A ce stade, l'expert n'a pas encore émis le souhait d'ajouter une description intrinsèque au genre Citharidium.

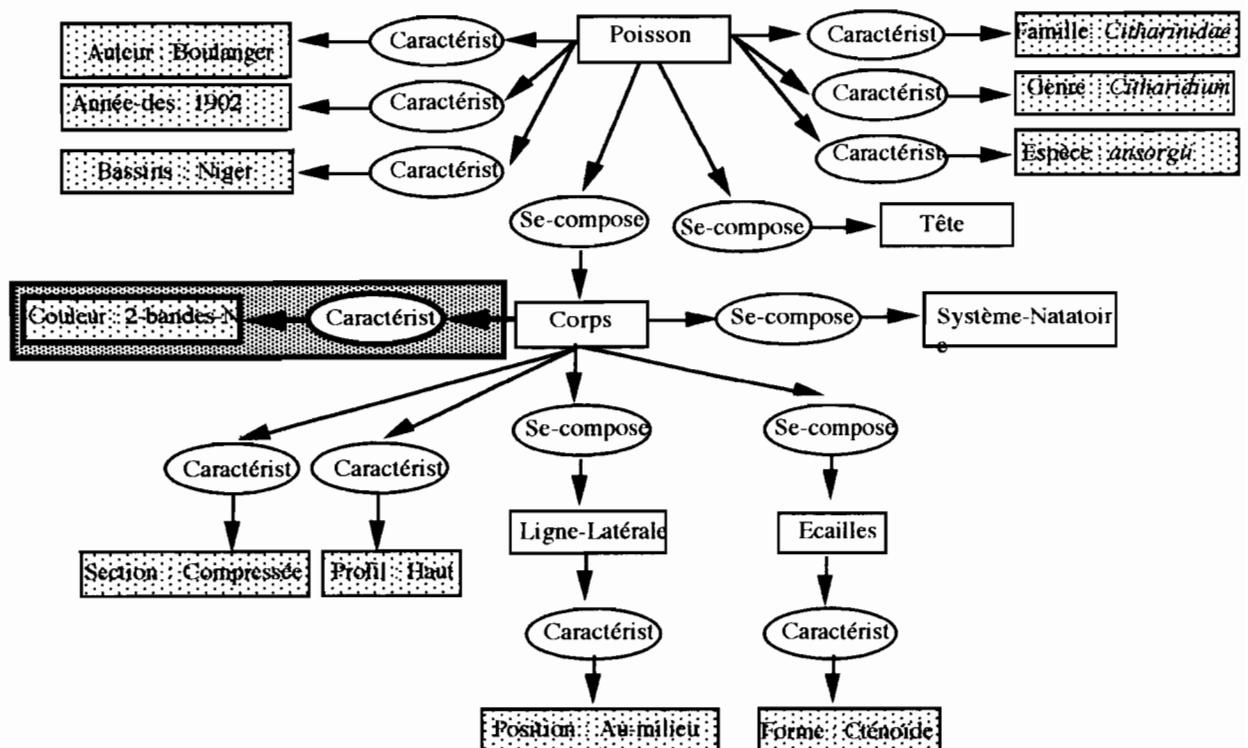


Figure 5.14. Graphe de l'espèce *Citharidium ansorgii* (G121)

5.6.2.1. Procédure d'apprentissage

Rappelons que quand il s'agit de discriminer deux ou plusieurs classes, METIS utilise un algorithme inspiré d'ID3 (voir chapitre 2) que nous appellerons IDICE (Induction et Discrimination de Classes basées sur l'Entropie).

Nous aurions pu choisir un autre algorithme d'apprentissage tel que CHARADE [GANASCIA, 87]. En effet dans [AIMEUR, 91] nous avons couplé à l'algorithme d'apprentissage CHARADE un moteur d'inférence d'ordre 0+ avec coefficients de vraisemblance afin d'évaluer la pertinence du système de règles dans le domaine de la reconnaissance de la parole. L'expérience montre que même si l'algorithme d'apprentissage est performant (nous avons obtenu un excellent taux de reconnaissance) le système de règles produit est difficile à interpréter et contient des redondances²⁴. Or, dans notre cas, nous voulons non seulement éviter les redondances, mais aussi avoir un

²⁴ La redondance est intéressante dans le cadre de l'apprentissage, car cela rend le système plus robuste pour lutter contre le bruit. En revanche, la redondance est pénalisante dans le cadre de l'acquisition des connaissances, car cela ralentit considérablement l'interaction avec l'expert.

arbre de décision qui soit le plus simple possible (plus facile à lire pour l'expert). Notre choix s'est donc porté sur ID3 pour son efficacité et la simplicité de sa mise en œuvre.

Après une étape de pré-traitement sur la description des taxons, METIS va interagir avec l'expert en lui fournissant des traits discriminants grâce à IDICE

L'algorithme (IDICE) prend en compte²⁵ aussi bien des classes non homogènes (c'est-à-dire des classes dont les descriptions ne comportent pas les mêmes attributs²⁶) que des classes dont les descriptions comprennent des attributs multivalués. De plus, il prend en compte les contraintes fournies par l'expert sur le choix et l'utilisation des attributs.

Nous allons décrire la procédure d'apprentissage avant de présenter des exemples.

a) Étape de pré-traitement

L'étape de pré-traitement va comprendre deux sous-étapes : la complétion des descriptions de classes et l'extraction de traits pertinents.

a.1.) Complétion des descriptions de classes

Quand METIS doit discriminer une classe par rapport à d'autres classes, il demande à l'expert de compléter la description des classes. Il est évident que ce n'est pas toujours possible, par exemple, l'expert ne peut donner la forme d'un composant qui n'est pas présent. Ceci illustre le cas des concepts inapplicables.

On peut définir la complétion de la façon suivante : étant donné un ensemble de classes à discriminer {C1, C2, ...Cn} METIS demande la valeur de chaque attribut (présent dans au moins une description de classe) pour toutes les classes où il n'est pas présent.

La complétion des exemples favorise la recherche de traits potentiellement discriminants et permet une certaine homogénéisation dans les descriptions de classes proches (ayant même classe ancêtre).

²⁵ Il ne s'agit pas d'un véritable traitement. En fait IDICE n'élimine pas ces attributs de la description, il les utilise dans les sous-ensembles de classes où ils ne posent pas de problème.

²⁶ Un attribut correspond dans METIS à un trait non valué. Exemples : [Corps] -> (Se-compose) -> [Écailles] , [Corps] -> (Caractérist) -> [Couleur].

a.2.) Extraction de traits pertinents

Cette seconde étape de pré-traitement consiste à récupérer les descriptions des différentes classes à discriminer (après complétion). Il suffit pour cela d'utiliser les descriptions intrinsèques stockées dans la base des graphes. Après avoir effectué cela, METIS va construire plusieurs listes, en faisant la distinction entre traits structurels et traits descriptifs.

Cas des traits structurels

Pour chaque trait structurel, on construit la liste des classes où il est présent. La liste des classes où un trait structurel est absent est obtenue, en prenant le complément de la liste des classes où il est présent.

Le résultat obtenu peut être schématisé comme suit :

{	<i>trait structurel</i>	<i>liste des classes où il est présent</i>
	<i>ts1</i>	<i>{(...)}</i>
	<i>...</i>	
	<i>tsX</i>	<i>{(...)}</i>

Cas des traits descriptifs

Étant donné un ensemble de traits descriptifs ($t_{d1v1} \dots t_{dnvn}$) où d_i est le numéro du trait et v_i la valeur associée à son concept destination. Il s'agit de calculer pour chaque trait descriptif $t_{d_i v_i}$, le généralisé de ce trait à l'aide de l'opérateur de dérérérenciation que nous définissons comme suit :

$$\mathbf{Dref}(t_{d_i v_i}) = t_{d_i}$$

Exemple :

$$\mathbf{Dref}([\text{Écailles}] \rightarrow (\text{Caractérist}) \rightarrow [\text{Forme : cycloïde}]) = [\text{Écailles}] \rightarrow (\text{Caractérist}) \rightarrow [\text{Forme}]$$

On stocke pour chaque trait ainsi que pour son généralisé la liste des classes où ces traits sont observés.

$$\begin{array}{l}
 \text{trait sans} \quad \text{liste des} \\
 \text{réfèrent} \quad \text{classes} \\
 t_{dl} \quad (\dots)
 \end{array}
 \left\{
 \begin{array}{l}
 \text{trait avec} \quad \text{liste des} \\
 \text{réfèrent} \quad \text{classes} \\
 t_{dlv1} \quad (\dots) \\
 t_{dlv2} \quad (\dots) \\
 \dots \\
 t_{dlvp} \quad (\dots)
 \end{array}
 \right.
 \text{ où } t_{dl} = \text{Dref}(t_{dlvj}) \text{ avec } v_j \in [1, p]$$

....

....

$$t_{dN} \quad (\dots)
 \left\{
 \begin{array}{l}
 t_{dNv1} \quad (\dots) \\
 t_{dNv2} \quad (\dots) \\
 \dots \\
 t_{dNvR} \quad (\dots)
 \end{array}
 \right.$$

b) Étape de Recherche d'un arbre de discrimination

A chaque étape de la construction d'un arbre de décision, ID3 sélectionne le trait qui maximise le gain d'information vis à vis des exemples à discriminer. Nous allons montrer dans notre cas que les calculs peuvent être simplifiés car nous avons un seul exemple par classe.

Soit t_{ij} le i ème trait ayant la valeur j pour son concept destination

Exemple t_i : [Corps] -> (Caractéristique) -> [Couleur]

t_{i1} : [Corps] -> (Caractéristique) -> [Couleur : rouge]

t_{i2} : [Corps] -> (Caractéristique) -> [Couleur : verte]

A chaque trait t_i on associe une probabilité d'apparition $p(t_{ij})$ telle que $0 \leq p(t_{ij}) \leq 1$

L'information propre associée à l'évènement t_{ij} est donnée par la formule suivante :

$$I(t_{ij}) = -\log_2(p(t_{ij}))$$

L'entropie $H(t_i)$ est l'information moyenne ou l'espérance mathématique de l'information propre à chacun des évènements t_{ij} .

Elle est donnée par la formule suivante :

$$H(t_i) = \sum_j p(t_{ij}) I(t_{ij})$$

$$\text{D'où } H(t_i) = - \sum_j p(t_{ij}) \log_2(p(t_{ij}))$$

De même l'information propre conditionnelle associée à l'évènement t_{ij} sachant la classe c_j est donnée par la formule suivante :

$$I(c_j / t_{ij}) = - \log_2(p(c_j / t_{ij}))$$

D'où l'entropie conditionnelle des classes connaissant le trait t_j est :

$$H(c / t_i) = - \sum_j \sum_l p(c_l \wedge t_{ij}) \log_2(p(c_l / t_{ij}))$$

Compte tenu des propriétés (associées respectivement aux probabilités et aux fonctions logarithmiques) suivantes :

$$p(a/b) = p(a \wedge b) / p(b) \text{ et } \log_2(a/b) = \log_2 a - \log_2 b$$

On peut encore écrire :

$$H(c / t_i) = H(c \wedge t_i) - H(t_i)$$

Dans notre cas il n'y a qu'un seul exemple par classe :

$$p(c_l) = \frac{1}{N}, \quad \forall l \in [1, N]$$

$$p(c_l \wedge t_{ij}) = \begin{cases} \frac{1}{N} & \text{si } t_{ij} \text{ est observé dans la classe } c_l \\ 0 & \text{sinon} \end{cases} \quad \text{on suppose qu'il y a } N \text{ classes}$$

Les traits étant exclusifs les uns des autres on a :

$$H(c \wedge t_i) = - \sum_j \sum_l p(c_l \wedge t_{ij}) \log_2(p(c_l \wedge t_{ij})) = - \sum_l \frac{1}{N} \log_2\left(\frac{1}{N}\right) = - \log_2\left(\frac{1}{N}\right)$$

Soit encore $H(c \wedge t_i) = H(c)$

D'où : $H(c) - H(c / t_i) = H(t_i)$

Finalement pour maximiser le gain d'information $G(c, t_i) = H(c) - H(c / t_i)$, il suffit de maximiser $H(t_i)$ c'est-à-dire de sélectionner le trait dont l'entropie est maximale.

b.1.) Heuristique de choix

L'heuristique de choix d'un trait discriminant repose sur l'intuition que la présence ou l'absence d'un composant est plus pertinente pour discriminer des classes que les propriétés intrinsèques relatives à ce composant²⁷. C'est-à-dire que METIS privilégie les traits structurels par rapport aux traits descriptifs.

Exemple :

soit l'ensemble de traits discriminants suivants :

[Système-Natatoire] -> (Se-compose) -> [Nageoire-Anale]

[Nageoire-Anale] -> (Caractérist) -> [Couleur]

D'après l'heuristique décrite, METIS considère en priorité le premier, puis le second. On peut constater cela lors de la discrimination des deux genres *Citharinus* et *Citharidium* (voir deuxième description), puisque METIS a d'abord demandé si les écailles étaient présentes.

b.2.) Contraintes

Il est parfois souhaitable que l'expert puisse modifier le comportement de la procédure de recherche de traits discriminants (appelée aussi procédure d'apprentissage), en interdisant par exemple l'utilisation de certains traits ou en imposant l'utilisation d'autres.

Un trait peut être interdit par l'expert bien qu'il soit discriminant (voir § 5.6.2.2.). L'expert le juge non pertinent ou difficile à observer. Par ailleurs quand l'expert impose l'utilisation prioritaire d'un trait, ce dernier peut ne pas suffire auquel cas METIS complète la recherche de nouveaux traits en **réappliquant** la procédure d'apprentissage

Par conséquent, la procédure d'apprentissage recevra deux listes de paramètres fournis par l'expert :

- la liste des traits interdits notée : Lt-interdits,

²⁷ Le nombre de composants étant forcément plus petit que le nombre de propriétés, il nous semble plus facile de nous intéresser d'abord aux composants, de plus, ils sont plus facilement observables.

- la liste des traits prioritaires notée : Lt-prioritaires,

La phase de recherche de traits discriminants étant envisagée comme une interaction, la procédure d'apprentissage peut être appelée plusieurs fois mais avec différents paramètres. Ces paramètres sont comme nous venons de le voir, des contraintes qu'impose l'expert à la procédure d'apprentissage.

c) L'algorithme IDICE

Cet algorithme fait appel à deux procédures imbriquées qui sont la recherche de traits discriminants avec contraintes et la sélection du meilleur trait à un niveau donné de l'arbre.

Recherche-traits-discriminants (Lt, Lt-interdits, Lt-prioritaires)

Lt-utilisables := Lt - Lt-interdits - Lt-prioritaires

Lt-déjà-sélectionnés := Lt-prioritaires

C := ensemble des ensembles de classes non discriminées par Lt-prioritaires

Tant qu'il reste un ensemble C non vide et que Lt-utilisables est non vide

Faire C₁ := premier élément de C ; C₁ est un ensemble de classes

Lt-sel := Sélection-trait (Lt-utilisables, C₁)

Lt-utilisables := Lt-utilisables - Lt-sel

Lt-déjà-sélectionnés := Lt-déjà-sélectionnés + Lt-sel

C := ensemble²⁸ des ensembles de classes non discriminés par Lt-déjà-sélectionnés

Fin-Tant que

Si C est vide

Alors Retourner Lt-déjà-sélectionnés,

Sinon Retourner échec.

FinSi

²⁸ Cet ensemble est produit par la procédure sélection-trait

Sélection-trait (Lt-utilisables, C_u)

Pour chaque trait appartenant à Lt-utilisables

Faire Si C_u est incluse dans la liste des classe où ce trait apparaît

Alors Si le trait n'est pas multivalué²⁹ pour toutes les classes de C_u

Alors calculer l'entropie des classes C_u

FinSi

FinSi

Fin-Pour

Si il n'y a aucun trait sélectionné³⁰

Alors Retourner Echec.

Sinon Si il n'y a qu'un trait ayant une entropie maximale

Alors Retourner ce trait

Sinon Retourner en priorité un trait structurel et dans le cas où il y a plusieurs traits structurels ayant même entropie maximale, retourner celui dont le concept origine est le plus général

FinSi

FinSi

Remarque:

Le trait ou la conjonction de traits trouvé par IDICE va remplacer le trait ou la conjonction de traits existants auparavant. Car le but d'IDICE est de minimiser le nombre de traits utilisés.

²⁹ La valeur est une disjonction de valeurs élémentaires (ne permet donc pas la discrimination)

³⁰ Le trait sélectionné est celui dont l'entropie maximale.

Exemples :

Description de la famille *Hepsetidae* (F1)

[Corps] -> (Caractérist) -> [Section : Arrondie]

[Corps] -> (Caractérist) -> [Profil : Allongé]

[Ligne-Latérale] -> (Caractérist) -> [Position : Basse]

[Écailles] -> (Caractérist) -> [Forme : Cycloïde]

[Mâchoires] -> (Se-compose) -> [Dents : présent]

[Dents] -> (Caractérist) -> [Disposition : 3R]

Description de la famille *Distichodontidae* (F2)

[Corps] -> (Caractérist) -> [Section : Comprimée ou Arrondie]

[Corps] -> (Caractérist) -> [Profil : Allongé]

[Ligne-Latérale] -> (Caractérist) -> [Position : Au-milieu]

[Écailles] -> (Caractérist) -> [Forme : Cténoïde]

[Mâchoires] -> (Se-compose) -> [Dents : Présent]

[Dents] -> (Caractérist) -> [Disposition : 1R ou 2R]

Description de la famille *Citharinidae* (F3)

[Corps] -> (Caractérist) -> [Section : Comprimée]

[Corps] -> (Caractérist) -> [Profil : Haut]

[Ligne-Latérale] -> (Caractérist) -> [Position : Au-milieu]

[Écailles] -> (Caractérist) -> [Forme : Cténoïde ou Cycloïde]

[Mâchoires] -> (Se-compose) -> [Dents]

[Dents] -> (Caractérist) -> [Disposition : 1R]

Description de la famille *Cyprinidae* (F4)

[Corps] -> (Caractérist) -> [Section : Comprimée]

[Corps] -> (Caractérist) -> [Profil : Fusiforme]

[Ligne-Latérale] -> (Caractérist) -> [Position : Au-milieu]

[Écailles] -> (Caractérist) -> [Forme : Cycloïde]

Remarque:

Le trait [Mâchoires] -> (Se-compose) -> [Dents] étant absent de la description de la famille *Cyprinidae* (F4), le trait [Dents] -> (Caractérist) -> [Disposition] est inapplicable.

Étape de pré-traitement

AC1 : [Corps] -> (Caractérist) -> [Section] (F1, F2, F3, F4) $\left\{ \begin{array}{l} \text{Arrondie (F1)} \\ \text{Comprimée (F3, F4)} \\ \text{Comprimée ou Arrondie (F2)} \end{array} \right.$

AC2 : [Corps] -> (Caractérist) -> [Profil] (F1, F2, F3, F4) $\left\{ \begin{array}{l} \text{Allongé (F1, F2)} \\ \text{Haut (F3)} \\ \text{Fusiforme (F4)} \end{array} \right.$

AC3 : [Ligne-Latérale] -> (Caractérist) -> [Position] (F1, F2, F3, F4) $\left\{ \begin{array}{l} \text{Basse (F1)} \\ \text{Au - milieu (F2, F3, F4)} \end{array} \right.$

AC4 : [Écailles] -> (Caractérist) -> [Forme] (F1, F2, F3, F4) $\left\{ \begin{array}{l} \text{Cténoïde (F2)} \\ \text{Cycloïde (F1, F4)} \\ \text{Cténoïde ou Cycloïde (F3)} \end{array} \right.$

AC5 : [Mâchoires] -> (Se-compose) -> [Dents] (F1, F2, F3, F4) $\left\{ \begin{array}{l} \text{Présent (F1, F2, F3)} \\ \text{Absent (F4)} \end{array} \right.$

$$AC6 : [Dents] \rightarrow (Caractérist) \rightarrow [Disposition] (F1, F2, F3) \begin{cases} 1R (F3) \\ 3R (F1) \\ 1R \text{ ou } 2R (F2) \end{cases}$$

1er cas : L'expert n'impose aucune contrainte

Première étape

L'expert n'ayant ni interdit ni imposé de trait, IDICE peut a priori sélectionner n'importe quel trait, afin de discriminer les classes F1, F2, F3, F4.

$$C := \{F1, F2, F3, F4\}$$

$$L\text{-utilisables} := (AC1, AC2, AC3, AC4, AC5, AC6)$$

$$C1 := \{F1, F2, F3, F4\}$$

IDICE ne sélectionne pas les traits AC1 et AC4 car ils sont respectivement multivalués pour les classes F2 et F3. Le trait AC6 ne sera pas d'avantage sélectionné car la classe F4 est absente de la liste des classes où ce trait apparaît. C'est à dire que {F1, F2, F3, F4} n'est pas inclus dans {F1, F2, F3}

$$H(AC2) = -\frac{2}{4} \log_2 \left(\frac{2}{4}\right) - \frac{1}{4} \log_2 \frac{1}{4} - \frac{1}{4} \log_2 \frac{1}{4} = 1.5 \text{ bits}^{31}$$

$$H(AC3) = -\frac{1}{4} \log_2 \left(\frac{1}{4}\right) - \frac{3}{4} \log_2 \frac{3}{4} = 0.811 \text{ bits}$$

$$H(AC5) = -\frac{3}{4} \log_2 \left(\frac{3}{4}\right) - \frac{1}{4} \log_2 \frac{1}{4} = 0.811 \text{ bits}$$

³¹l'unité est le "bit" car il s'agit d'une fonction logarithmique en base 2. Quand on est en base e, l'unité est le "nit" et quand on est en base 10 l'unité est le "dit".

A cette étape, IDICE choisit le trait AC2 car c'est celui qui maximise la mesure d'entropie. A cette étape, on a discriminé les classes F4 (profil fusiforme) et F3 (profil haut), il reste à discriminer les classes F1 et F2 car leur description comportent toutes deux la même valeur pour le trait AC2 (profil allongé).

Lt-utilisables := (AC1, AC3, AC4, AC5, AC6)

Lt-déjà-sélectionnée := (AC2)

C := {{F1, F2}}

C1 := {F1, F2}

Remarque :

Si on avait simplement pu isolé F1, F2 de F3, F4 on aurait eu C := {{F1, F2}, {F3, F4}} et C1 := {F1, F2}

Deuxième étape (possibilité d'utiliser des attributs multivalués et des attributs inapplicables)

IDICE peut sélectionner le trait AC4 car les classes F1 et F2 présentent une seule valeur chacune, de même qu'il peut sélectionner AC6 car {F1, F2} est inclus dans {F1, F2, F3}, et le trait AC6 ne prend à chaque fois qu'une valeur pour ces classes. En revanche, le trait AC1 ne peut toujours pas être sélectionné.

$$H(AC3) = -\frac{1}{2} \text{Log}_2\left(\frac{1}{2}\right) - \frac{1}{2} \text{Log}_2\left(\frac{1}{2}\right) = 1 \text{ bits}$$

$$H(AC4) = -\frac{1}{2} \text{Log}_2\left(\frac{1}{2}\right) - \frac{1}{2} \text{Log}_2\left(\frac{1}{2}\right) = 1 \text{ bits}$$

$$H(AC5) = -\frac{3}{3} \text{Log}_2\left(\frac{3}{3}\right) = 0 \text{ bits}$$

$$H(AC6) = -\frac{1}{2} \text{Log}_2\left(\frac{1}{2}\right) - \frac{1}{2} \text{Log}_2\left(\frac{1}{2}\right) = 1 \text{ bits}$$

H (AC5) est nul et cela est prévisible car cet arc présente les mêmes valeurs pour F1, F2.

H(AC3) étant égal à H[AC4] et H(AC6), IDICE choisit le trait AC3 (le premier), car il n'y pas le choix entre des traits descriptifs et des traits structurels.

A cette étape, les classes F1 et F2 sont discriminées

Lt-déjà-sélectionnée := (AC2, AC3)

L'algorithme IDICE s'arrête car $C = \{\{\emptyset\}\}$

2ème cas L'expert impose des contraintes

Supposons que l'expert impose l'utilisation du trait AC3 (Ligne-Latérale) en premier et interdit l'utilisation du trait AC2. D'emblée, la classe F1 est discriminée de F2, F3 et F4, car la ligne latérale est basse pour F1 et au-milieu pour F2, F3 et F4.

Lt-utilisables := (AC1, AC4, AC5, AC6)

Lt-déjà-sélectionnée := (AC3)

$C := \{\{F2, F3, F4\}\}$

$C1 := \{F2, F3, F4\}$

IDICE ne sélectionnera pas les traits AC1, AC4 et AC6 car ils prennent des valeurs multivaluées. le seul trait à prendre est AC5

$$H(AC5) = -\frac{2}{3} \log_2 \left(\frac{2}{3}\right) - \frac{1}{3} \log_2 \left(\frac{1}{3}\right) = 0.918 \text{ bits}$$

En choisissant le trait AC5 il reste à discriminer les classes F2 de F3.

Lt-déjà-sélectionnée := (AC3, AC5)

A cette étape METIS ne peut discriminer les classes F2 de F3 car il ne peut sélectionner aucun des traits utilisables³². Il retourne donc "Échec" afin que l'expert relâche ses contraintes ou accepte la première solution.

Évaluation de la complexité de l'algorithme

Le calcul de la complexité de l'algorithme d'apprentissage sera évalué en termes de nombre d'évaluations d'entropies de traits ($H(t_j)$) nécessaire pour la discrimination de sous-classes d'une classe c_j donné. Soit n_j , le nombre de sous-classes de la classe c_j .

Soit $T_i = \{t_{ij}\}_{j=1,m}$: l'ensemble des traits observables dans les sous-classes de c_j .

Soit T_i' l'ensemble des traits sélectionnés pour effectuer la discrimination. T_i' est un sous ensemble de T_i tel que $\text{Card}(T_i') = s_i$ (s_i étant bien sûr inférieur à m).

Étant donné que d'une étape à une autre de l'algorithme, on n'élimine qu'un seul trait discriminant, donc le nombre d'évaluations d'entropies correspondant à une étape k est de $s_i - k + 1$ (soit s_i évaluations à la 1ère étape, $s_i - 1$ à la 2ème étape, et ainsi de suite).

Sachant qu'au pire des cas, on ne discrimine qu'une classe à chaque étape, la discrimination des n_j sous-classes de c_j , est effectuée en $\min(n_j, s_i)$ étapes (c.a.d., jusqu'à épuisement des sous-classes ou des traits). Le nombre d'évaluations total est égal alors à :

$$\sum_{i=0}^{\min(n_i, s_i)} i = \min(n_i, s_i) (\min(n_i, s_i) + 1) / 2$$

Si n est le nombre moyen de sous-classes à chaque nœud de la taxinomie, et s le nombre de traits sélectionnés en moyenne au niveau de chaque nœud, la complexité moyenne est de $O([\min(n,s)]^2)$. La complexité globale de l'algorithme appliqué sur tous les nœuds de la taxinomie reste de même ordre, donc polynomial en nombre moyen de sous-classes et de traits par nœud, car le nombre de nœuds est un paramètre constant du système (il ne varie que d'un domaine à un autre ou d'un expert à un autre).

5.6.2.2. Évaluation de la pertinence des traits discriminants

Quand la procédure d'apprentissage n'a pas échoué, (voir figure 5.7) METIS apprend au moins un trait permettant de distinguer deux classes, qu'il soumet à l'expert afin que celui-ci évalue sa pertinence. Il est important de noter que l'expert peut rejeter le

³²On est dans le cas où l'on a soit valeur1 (F2) et valeur1 ou valeur2 (F3) soit valeur3 ou valeur4 (F2) et valeur3 (F3). Par conséquent il est impossible de discriminer.

trait fourni par METIS quand il est difficilement observable (exemple : le nombre de dents de la troisième rangée de la mâchoire supérieure).

Cependant, l'expert est *seul juge* quant à la pertinence des traits discriminants, d'où l'intérêt d'avoir un système interactif qui prenne en compte l'opinion de l'expert au coup par coup.

5.6.3. Traits donnés par l'expert

Quand METIS n'arrive pas à trouver des traits potentiellement discriminants, il se trouve dans l'obligation de faire appel à l'expert (voir figure 5.7).

Exemple :

L'expert souhaite enrichir sa taxinomie à l'aide des classes suivantes : le genre *Citharinops* et l'espèce *Citharinops distichodoides*. La troisième description à expliciter est celle de l'espèce *Citharinops distichodoides* appartenant au genre *Citharinops* de la famille *Citharinidae*.

Encore une fois, la description de la famille devant être commune pour les trois genres, METIS commence donc par positionner les nouvelles classes dans le graphe de discrimination, avant d'essayer de les discriminer (discrimination du genre *Citharinops* par rapport aux genres *Citharinus* et *Citharidium*).

Classification

Avant de rattacher la classe *Citharinops* à la classe *Citharinidae*, METIS doit vérifier auprès de l'expert que les traits intrinsèques de cette dernière sont toujours valides (voir § 5.5.2.), (c'est-à-dire que la section est comprimée, que le profil du corps est haut, que et que la position de la ligne latérale est au milieu). De plus, il doit vérifier que le corps se compose d'écailles (voir figure 5.12.). Après cela il rattachera la classe *Citharinops* à la classe *Citharinidae* et la classe *Citharinops distichodoides* à celle de *Citharinops*.

Discrimination

Rappelons que l'on est dans un cas où le trait discriminant existe déjà (forme des écailles). METIS demande à l'expert la forme des écailles du genre *Citharinops*. La réponse étant cycloïde le trait discriminant n'est plus suffisant puisque le genre *Citharinus* possède lui aussi le même caractère. Jusqu'ici, METIS n'a pas suffisamment d'éléments pour découvrir un nouveau trait (la description intrinsèque des genres ne contient qu'un

seul trait³³ : le corps se compose d'écaïlles). Il demande alors explicitement à l'expert de lui fournir un trait. Si l'expert indique que c'est le nombre d'écaïlles, celui-ci devient alors le nouveau trait discriminant (68 à 73 écaïlles pour le genre *Citharinus* et 50 à 56 écaïlles pour le genre *Citharinops*). Afin de vérifier si le nouveau trait discriminant suffit à lui seul pour discriminer les trois genres, METIS s'enquiert auprès de l'expert du nombre d'écaïlles pour le genre *Citharidium*. (40 à 49). Dans ce cas précis, le nombre d'écaïlles est un trait discriminant suffisant : il remplacera l'ancien (forme des écaïlles) qui appartient désormais à la description intrinsèque des trois genres et qui est supprimé du graphe de discrimination.

Après interaction avec l'expert, METIS étend la base de graphes avec G_{13} et G_{131} (Voir figure 5.3). Le nouveau graphe de discrimination est le suivant :

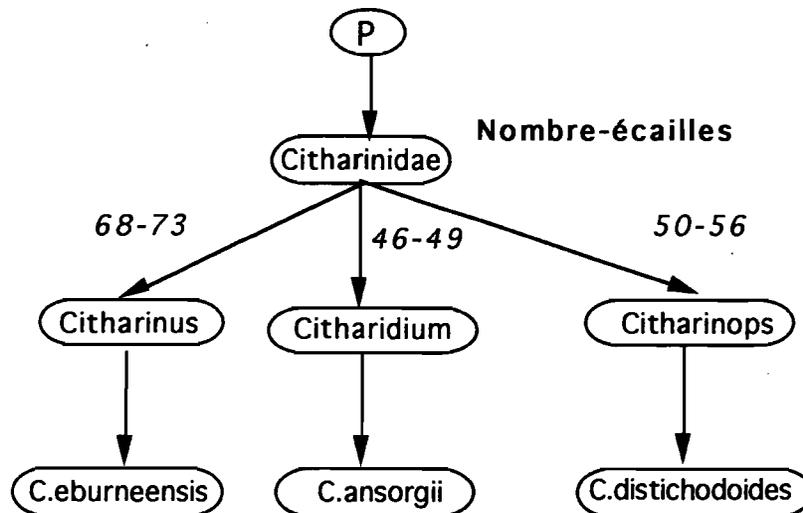


Figure 5.15. Graphe de discrimination 3

a) Parcours de l'espace de composition

L'ensemble des concepts composites définit un espace que l'on doit parcourir pour pouvoir rattacher un nouveau composant à son composite. Appelons cet espace, l'*espace de composition*.³⁴ L'explicitation d'un trait discriminant auprès de l'expert entraîne le

³³ Il ne peut y avoir de complétion des descriptions de classes (guidée par METIS), car la description disponible est trop réduite.

³⁴ C'est la partie du graphe où les relations entre concepts sont "Se-compose".

parcours de l'espace de composition dans un ordre défini au préalable par l'expert. Ainsi, si le trait discriminant donné par l'expert est [Nag-Caudale] -> (caractéristique) -> [forme] et que le concept [Nag-Caudale] est inconnu du système, le dialogue suivant est amorcé :

METIS : est ce que le concept Nag-Caudale se rattache au concept tête ?

Expert : non.

METIS : est ce que le concept Nag-Caudale se rattache au concept corps ?

Expert : oui.

METIS : est ce que le concept corps est le concept le plus spécifique auquel se rattache le concept Nag-Caudale ?

Expert : non

METIS : est ce que le concept Nag-Caudale se rattache au concept Système-Natatoire ?

Expert : oui.

METIS : est ce que le concept Système-Natatoire est le concept le plus spécifique auquel se rattache le concept Nag-Caudale ?

Expert : oui

METIS : quelle est la relation reliant Nag-Caudale à Système-Natatoire ?

Expert : la relation se-compose.

la relation étant de composition, METIS rattache au concept [Système-Natatoire] la structure ([Système-Natatoire] -> (se-compose) -> [Nag-Caudale] et [Nag-Caudale] -> (caractéristique)->[forme : Fourchue]). Le concept [Nag-Caudale] est rangé dans la liste des concepts composites associés au concept [Système-Natatoire]³⁵.

Modification de parcours de l'espace de composition

Dans le but d'optimiser le nombre de questions qu'il peut poser à l'expert, METIS peut modifier dynamiquement l'ordre de parcours de l'espace de composition (en accord avec l'expert), au fur et à mesure que de nouvelles connaissances structurelles sont

³⁵ Ce dialogue peut parfois être long ou ennuyeux pour l'expert, nous envisageons dans une version future de METIS de présenter visuellement à l'expert le graphe afin qu'il indique où connecter la nouvelle partie.

acquises. Ainsi METIS peut lui même suggérer un nouveau parcours à l'expert ou le faire sur sa demande.

METIS privilégie les concepts composites pour lesquels le nombre de concepts composants est minimum³⁶. Ainsi, dans l'exemple précédent, METIS préfère d'abord poser la question concernant le concept [tête] (nombre de concepts composites nul) avant de poser celle concernant le concept [corps] (nombre de concepts composites égal à 3). Le concept [tête] est mis avant le concept [corps] dans la liste ordonnée des concepts composites, mais si à un moment donné le nombre de concepts composants qui sont rattachés au concept [tête] est supérieur à celui du concept [corps], METIS peut suggérer à l'expert de modifier la liste des concepts composites et de placer le concept [corps] avant le concept [tête].

b) Vérification

Lorsque l'expert fournit un trait discriminant, METIS doit vérifier que ce trait n'appartient pas aux descriptions des descendants des nœuds frères du nœud considéré dans la base des graphes. Dans l'affirmative, il faut remonter le trait au niveau du nœud père du nœud considéré. Par exemple si l'expert avait donné dans l'exemple précité (voir figure 5.15) au lieu du nombre d'écailles, la couleur du corps, il aurait fallu remonter ce trait qui décrit l'espèce *Citharidium ansorgii* (voir figure 5.14) au niveau du genre auquel elle appartient, i.e. le *Citharidium*. Car le trait ([Corps] -> (Caractéristique) -> [Couleur]) étant discriminant au niveau du genre, il doit obligatoirement appartenir à la description discriminante des espèces qu'il discrimine.

Par ailleurs, METIS doit vérifier que le trait fourni par l'expert est réellement discriminant en examinant les valeurs qui doivent être exclusives. Il doit aussi vérifier que le nouveau trait suffit à lui seul, pour discriminer tous les taxons considérés. Auquel cas, METIS doit signaler à l'expert que le ou les anciens traits peuvent être supprimés au profit du nouveau.

5.6.4. Contrôle de cohérence et mise à jour des données

L'étape de contrôle de cohérence et de mise à jour des données (voir figure 5.7 étape e11) est la dernière étape à effectuer. Elle consiste à mettre à jour la base de graphes et le graphe de discrimination à chaque prise en compte d'un nouveau trait discriminant.

³⁶ Le choix de cette heuristique est purement intuitif.

Le contrôle de redondance sera assuré grâce à un critère, appelé critère de *factorisation maximale* que METIS doit respecter à chaque nouvelle mise à jour (voir § 5.7.3).

Statut des traits hérités

Si un trait discrimine un nœud de ses nœuds frères, il sera hérité par sa descendance en tant que description discriminante. Par exemple, la forme des écailles est un trait discriminant au niveau de la famille des Citharinidae et une description discriminante pour les genres Citharidium et Citharinus. Si à un moment donné ce trait est remis en cause, en tant que trait discriminant, il bascule aussitôt dans la description intrinsèque du nœud (§ 5.6.1.), c'est-à-dire qu'il ne sera présent que dans la base des graphes. Mais si le trait est remis en cause en tant que trait intrinsèque, il est immédiatement retiré du nœud et mis dans les descriptions intrinsèques des nœuds fils où il demeure valide. Ainsi, à chaque fois, seule une restructuration locale de la base des graphes et du graphe de discrimination est effectuée (voir § 5.7.3).

5.7. PHASE DE SAISIE DE DONNÉES INTRINSÈQUES

La phase de saisie de données intrinsèques a lieu après celle de classification s'il s'agit d'une classe *première descendante*³⁷, car il n'y a pas de discrimination à faire. En revanche, elle a lieu après la phase de discrimination quand METIS doit discriminer plusieurs classes, et même dans ce dernier cas, elle n'est pas obligatoire car l'expert peut décider que la description discriminante suffit pour le moment à décrire la nouvelle classe (voir deuxième description dans § 5.6.2.).

5.7.1. Acquisition de la description

Lors de la création d'une classe première descendante d'une classe donnée (voir première description dans § 5.6.2.), il faut obligatoirement laisser l'expert décrire cette première classe descendante. Ceci est dû au fait qu'à cette étape, METIS ne possède pas de classes sœurs à comparer pour guider l'acquisition. En revanche, si la classe possède d'autres classes sœurs, la saisie est guidée, car METIS posera des questions sur les traits décrivant ses classes sœurs afin de compléter la description de la classe considérée. Il faut cependant noter que, dans ce dernier cas, l'acquisition d'une description intrinsèque

³⁷ Nous appelons "classe première descendante", une classe qui n'a pas de classes sœurs

d'une classe n'est pas obligatoire car l'expert peut décider que la description discriminante et la description intrinsèque de la classe ascendante, et qui sont héritées par ses classes descendantes en tant que description intrinsèque, est suffisante (voir deuxième description dans § 5.6.2.).

5.7.2. Raffinement du modèle structurel

Le modèle structurel est enrichi à chaque nouvelle acquisition et à chaque prise en compte d'un nouveau trait discriminant. Les informations redondantes ne sont pas prises en compte, c'est à dire qu'il n'y a aucune duplication. La figure 5.16 montre l'état du modèle structurel après l'explicitation de la deuxième description (voir § 5.6.3.1.).

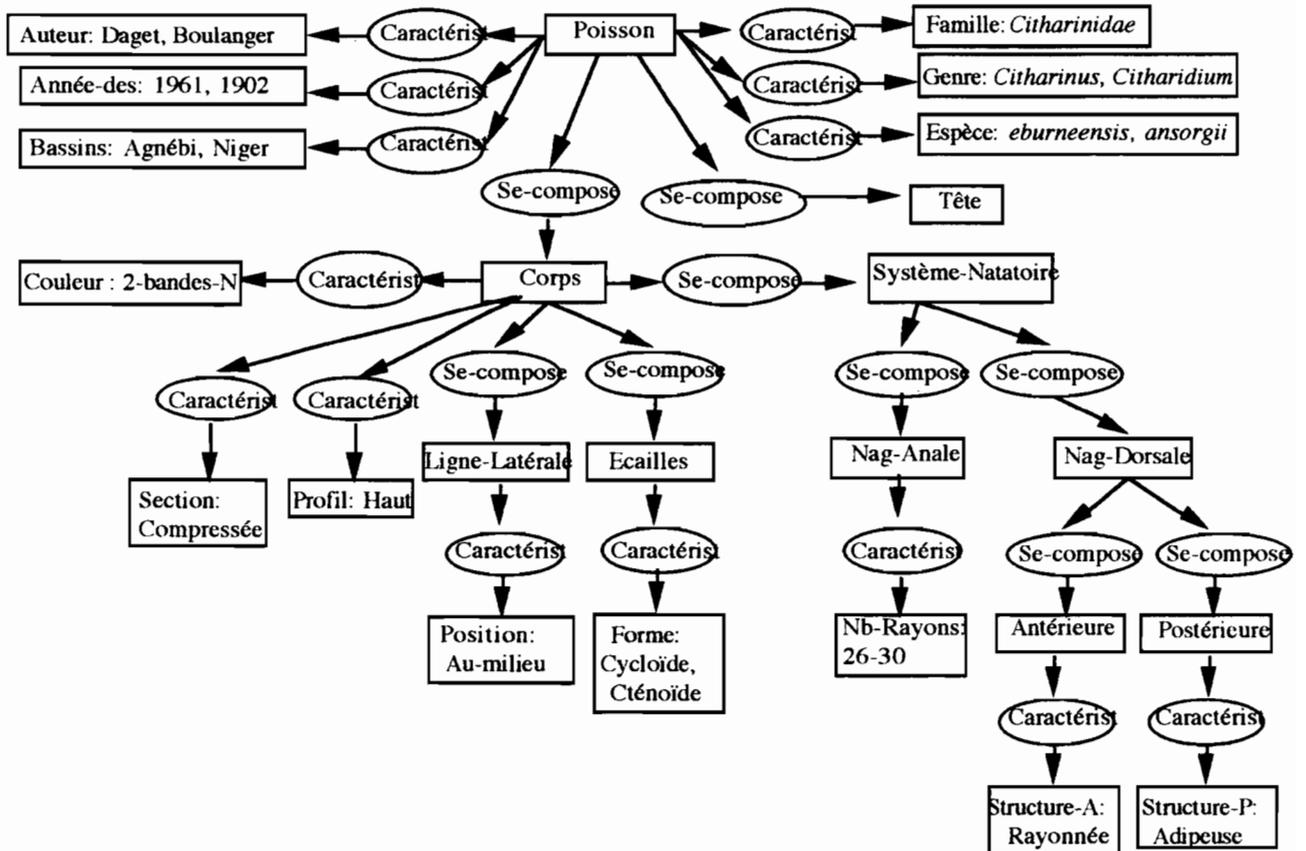


Figure 5.16. Modèle structurel

En définitive, les éléments de description s'enrichissent petit à petit et, corrélativement la base de connaissances se construit incrémentalement.

5.7.3. Cohérences des données

Selon [AYEL & ROUSSET, 90], la notion de cohérence requiert de disposer d'un modèle de référence. Deux types de cohérence sont dégagés : la *cohérence statique* qui vise à vérifier la validité des faits et des règles et la *cohérence dynamique* qui vise à vérifier la validité de l'enchaînement des déductions. Dans notre cas nous ne prendrons en considération que la cohérence statique.

Le modèle appelé C-modèle rassemble les propriétés de *cohérence statique* que doivent vérifier un fait, un ensemble de faits, une règle ou un ensemble de règles.

On définit dans le C-modèle des propriétés comme :

- l'ensemble des valeurs autorisées pour un attribut donné dans une classe donnée,
- le degré d'un attribut ou son arité maximum,
- des ensembles de valeurs exclusives,
- des attributs exclusifs les uns des autres, etc.

METIS construit incrémentalement un C-modèle auquel il se réfère pour construire de nouveaux concepts. Nous allons examiner en détail ce que contient le C-modèle dans le cas de METIS, car hormis les éléments cités il contient d'autres éléments.

Critère de factorisation maximale

Avant de définir le critère de factorisation maximale qui assure la non-redondance des traits, définissons d'abord les notions de *couverture*³⁸ de classes et de *couverture d'arcs conceptuels*.

Définition de la couverture de classes

Soit la classe Cl , on définit la couverture de Cl , notée $Couv_C(Cl)$, comme étant l'union des étiquettes³⁹ associées aux classes atomiques (classes terminales) appartenant à sa descendance. L'étiquette d'une classe atomique est son numéro d'apparition dans la

³⁸ La notion de couverture est utilisée dans divers travaux dont le système d'apprentissage CHARADE [GANASCIA, 87] où elle est appelée *couverture du système de règles*.

³⁹ La notion de couverture de classe correspond à la notion d'*étiquette de nœuds* chez MINEAU [MINEAU, 90].

base des graphes, la couverture d'une classe atomique se réduit à sa propre étiquette. La couverture d'une classe non-atomique est mise à jour en fonction d'un algorithme donné en Annexe 1.

Exemple:

Dans (figure 5.11), *C.eburneensis* recevra l'étiquette {1} tandis que *C.ansorgii* recevra l'étiquette {2}. Jusqu'ici, Citharinidae, Citharinus et Citharidium ont respectivement pour couvertures {1, 2}, {1} et {2}, car Citharinidae a deux descendants atomiques alors que Citharinus et Citharidium n'en ont qu'un.

L'illustration de l'algorithme de création et de mise à jour des étiquettes des exemples représentés dans la figure 5.17 est donné en Annexe 2.

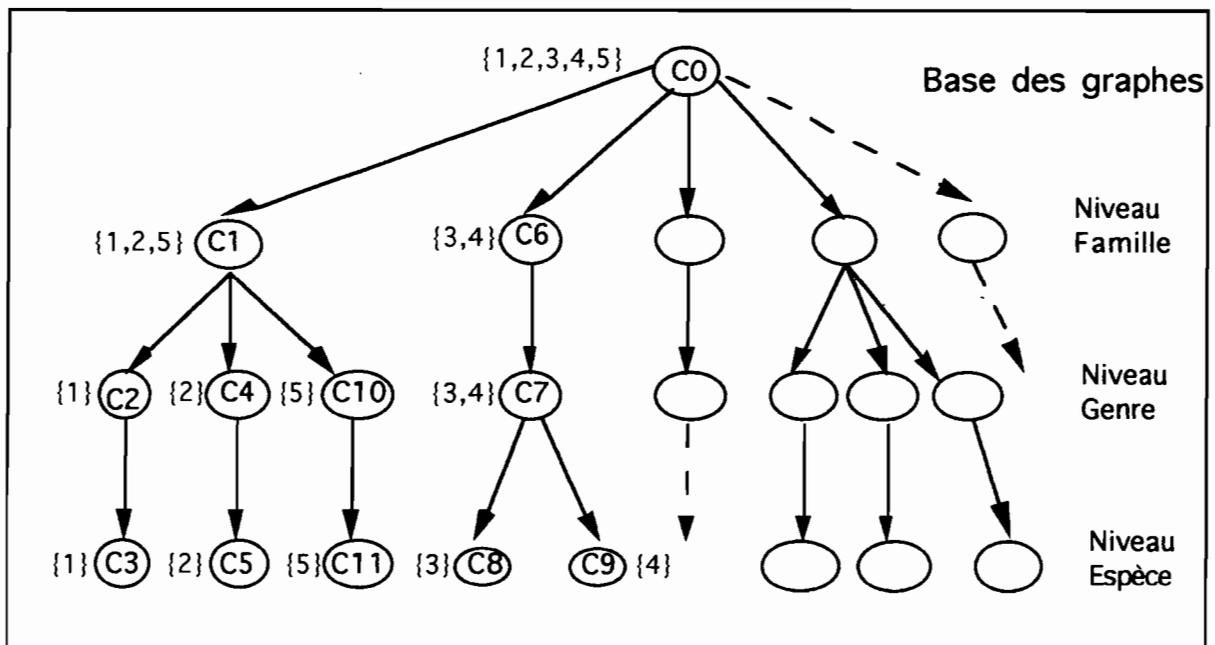


Figure 5.17. Numérotation des classes

Définition de la couverture d'un arc conceptuel

Soit un arc conceptuel ACI, par définition, la couverture de ACI, notée $Couv_a$ (ACI), est l'union des couvertures des classes les plus générales où cet arc est présent.

$$Couv_a(ACI) = \bigcup_{Classes} Cov_c(CI)$$

Exemple:

Si H. odoe (voir figure 5.4) reçoit l'étiquette {3} alors Hepsetus et Hepsetidae reçoivent l'étiquette {3}.

L'arc [Corps] -> (Se-compose) -> [Ligne-Latérale] est présent dans les familles Citharinidae et Hepsetidae qui sont les classes les plus générales.

Donc $Couv_a$ ([Corps] -> (Se-compose) -> [Ligne-Latérale]) = $\{1, 2\} \cup \{3\}$
= $\{1, 2, 3\}$

Définition du critère de factorisation maximale

Ce critère assure après chaque mise à jour, qu'un arc conceptuel (AC) est placé au "meilleur endroit", c'est-à-dire, dans toutes les classes les plus générales de la base des graphes telle que la condition suivante soit toujours vérifiée :

$$\forall C_i, Couv_c(C_i) \subseteq Couv_a(AC).$$

Remarque :

le critère de factorisation maximale n'est pas applicable aux arcs conceptuels suivants car ils sont propres au dernier niveau (niveau espèce).

[Poisson] -> (Caractérist) -> [Auteur], [Poisson] -> (Caractérist) -> [Année-des],
[Poisson] -> (Caractérist) -> [Bassins]

Exemple :

L'arc conceptuel [Ligne-Latérale] -> (Caractérist) -> [Position : Au-milieu] a pour couverture l'ensemble $\{1, 2\}$, ce qui va interdire le fait de le mettre dans la classe Hepsetidae dont la couverture $\{3\}$ n'est pas incluse dans $\{1, 2\}$.

METIS définit ainsi petit à petit avec l'aide de l'expert les contraintes de cohérence, c'est-à-dire :

- le critère de factorisation maximale,
- l'ensemble des référents autorisés pour un concept donné dans une classe donnée,
- le degré d'un concept (nombre de référents maximum que peut avoir un concept) ,

- l'ensemble des synonymes⁴⁰ d'un concept ou d'un référent,
- l'ensembles de référents simultanément contradictoires,
- les concepts respectivement exclusifs.

METIS met aussi dans le C-modèle la syntaxe à respecter pour introduire un arc conceptuel (par exemple, le trait ([Concept], [Concept], (Rel)) n'est pas admis. Par ailleurs, METIS vérifie qu'un concept n'apparaît qu'une fois afin d'éviter les circuits.

A chaque nouvelle mise à jour, METIS se réfère au C-modèle pour maintenir la cohérence dans la base de connaissances.

5.7.4. Traitement des négations

Nous distinguons deux sortes de négations : celle qui nie une propriété (la couleur des écailles n'est pas verte) et celle qui indique une absence de concept composite (le corps ne se compose pas d'écailles).

5.7.4.1. Négation de relations descriptives

Une relation descriptive s'exprime par un arc conceptuel dont le second concept n'est pas un concept composite (dans notre application il est donc descriptif). La négation d'une relation descriptive revient à mettre la négation sur le référent (voir § 6.1.8.).

Exemple :

[Écailles] -> (Caractérist) -> [Couleur : verte]

Négation : [Écailles] -> (Caractérist) -> [Couleur : ! verte]⁴¹

METIS considère que les valeurs verte et ! verte sont exclusives. Ceci sera pris en considération dans le C-modèle.

⁴⁰ La prise en compte des synonymes lexicaux vise à standardiser les termes employés pour décrire les objets. Dans [MINEAU, 93] la normalisation des descriptions conceptuelles tient compte aussi bien des synonymes lexicaux que des synonymes syntaxiques, car certaines constructions syntaxiques peuvent avoir plusieurs équivalents.

⁴¹ Si l'expert dit explicitement que la couleur n'est pas verte, n'est pas rouge et n'est pas bleue, on obtient l'arc conceptuel suivant : [Écailles] -> (Caractérist) -> [Couleur : {! verte, ! rouge, ! bleue}]

5.7.4.2. *Négation de relations structurelles*

Pour nier une relation structurelle, on peut choisir de mettre la négation sur le concept et non sur la relation. Pour cela, on peut introduire un nouveau référent, par exemple ϖ , ce qui donne le graphe conceptuel suivant : [Corps] -> (Se-compose) -> [Écailles : ϖ]. Dans notre cas on n'introduit pas de nouveau marqueur, on s'appuie sur l'heuristique suivante : "*Jusqu'à nouvel ordre toute information non donnée est considérée comme fausse*" : il s'agit de l'hypothèse du monde clos.

Toutefois, METIS permet à l'expert, s'il le désire, de marquer explicitement l'absence d'une relation structurelle (le corps ne se compose pas d'écailles). Dans le formalisme des graphes conceptuels, cela s'écrirait de la manière suivante : [Corps] -> (!Se-compose) -> [Écailles] (voir § 6.1.7.).

5.7.5. **Traitement des exceptions**

Nous distinguons deux sortes d'exceptions : celle qui concerne une propriété (la couleur des écailles est verte pour toutes les espèces, sauf pour l'espèce Ej où elle est rouge), et celle qui concerne l'absence d'un concept composite (le corps se compose d'écailles pour toutes les espèces, sauf pour l'espèce Ej).

5.7.5.1. *Exception sur les relations descriptives*

L'exception sur une relation descriptive revient à rajouter une valeur au concept non composite au niveau de la classe ascendante (dans notre cas le genre auquel appartiennent les espèces considérées) à masquer l'héritage de la relation descriptive et à redéfinir les valeurs pour les classes descendantes.

Exemple :

[Écailles] -> (Caractérist) -> [Couleur : verte] au niveau du genre

[Écailles] -> (Caractérist) -> [Couleur : verte] au niveau de toutes les espèces

[Écailles] -> (Caractérist) -> [Couleur : rouge] au niveau de l'espèce Ej (exception)

Après l'introduction de l'exception , on obtient au niveau du genre :

[Écailles] -> (Caractérist) -> [Couleur : {verte, rouge}]⁴²

5.7.5.2. Exception sur les relations structurelles

L'exception sur une relation structurelle revient à supprimer le trait comprenant cette relation au niveau de la classe ascendante et à la "redescendre" au niveau des classes filles où cette relation est présente. Puis, on ajoute explicitement la négation de la relation structurelle au niveau de la classe qui a occasionné l'exception.

Exemple :

Si [Corps] -> (Se-compose) -> [Écailles] était présent au niveau du genre, il va être supprimé à ce niveau et redescendu au niveau des espèces où ce trait est présent. On ajoute ensuite le trait [Corps] -> (!Se-compose) -> [Écailles] dans la description de l'espèce E_j.

5.8. ARCHITECTURE DU SYSTÈME

La base de connaissances est formée de trois éléments (voir figure 5.18) : la **taxinomie des classes**, la **base de graphes** et le **modèle structurel**. Elle est raffinée incrémentalement au fur et à mesure que les différents modules interagissent avec l'expert. Quant aux flèches, elles représentent les flux de données. Par exemple, si une double flèche relie un module à un élément de la base de connaissances, cela signifie que le module prend en entrée l'élément de la base de connaissances dans un état courant E₁, et moyennant un ou plusieurs processus raffine cet élément et le fait passer à un état E₂.

5.8.1. Schéma

L'algorithme de fonctionnement de ce schéma est donné au paragraphe 5.3. et est illustré par la figure 5.5.

⁴² Ceci ne veut pas dire que 50% de la population possède des écailles de couleur verte et 50 % possède des écailles de couleur rouge. Ceci veut simplement dire que la valeur rouge est désormais une valeur possible. Notons que dans METIS nous ne prenons pas en compte des critères statistiques, car on peut imaginer la notation suivante pour dire que seulement 1% de la population possède des écailles de couleur rouge : [Couleur : {verte : 99%, rouge 1%}]

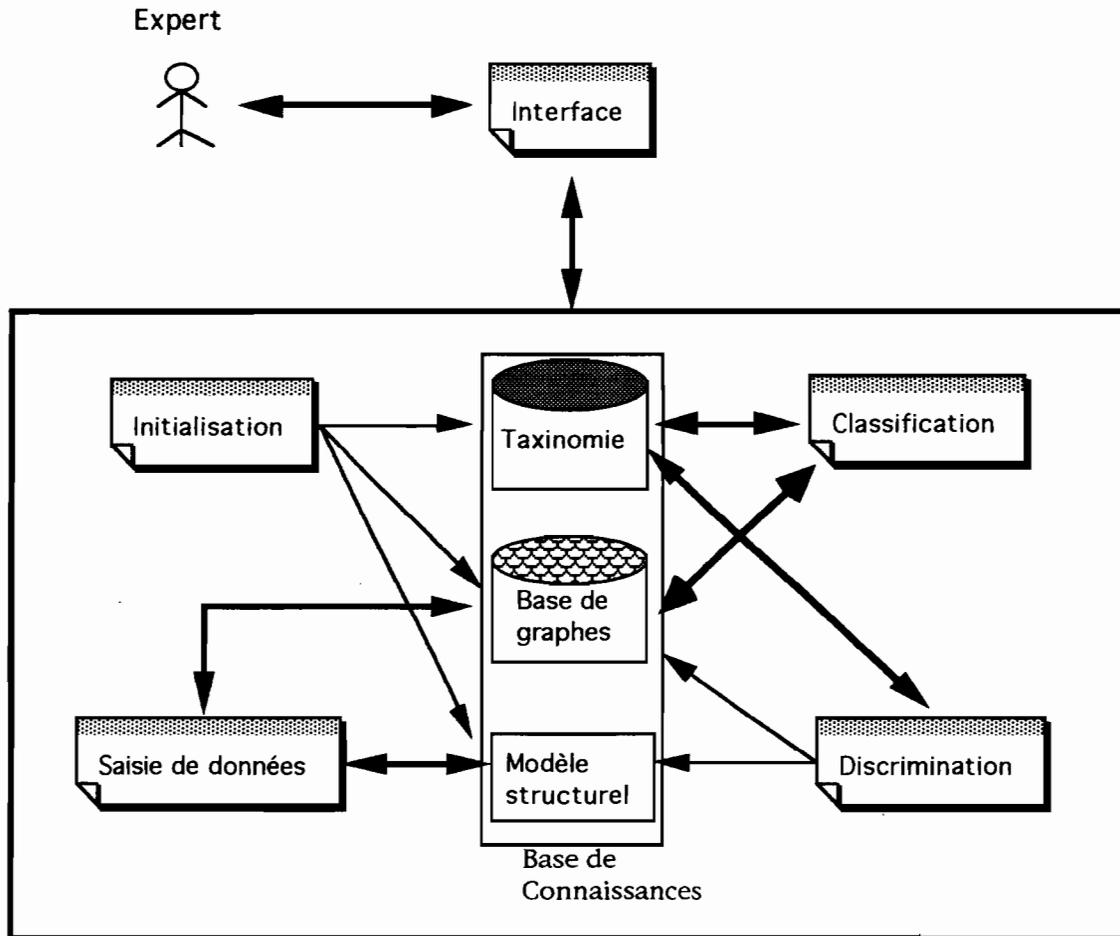


Figure 5.18. Architecture de METIS

5.8.2. Fonctionnalités

Pour atteindre les objectifs visés par notre système, nous avons dégagé un certain nombre de fonctionnalités que nous avons regroupées en modules afin d'aboutir à une architecture modulaire (voir figure 5.18.). METIS comporte ainsi quatre modules principaux qui assurent des fonctionnalités précises : le module d'initialisation, de classification, de discrimination et de saisie. Nous détaillons chacun des modules ainsi que l'interaction avec l'expert en terme d'entrées, de sorties, et de processus qui les gouvernent.

5.8.2.1. Initialisation

Le but de ce module est d'initialiser la base de connaissances, c'est à dire de doter METIS d'une connaissance minimale sur le domaine d'application (la taxinomie à considérer ainsi que certaines informations sur cette taxinomie) qui puisse lui permettre de prendre en compte les nouvelles descriptions de classes.

Entrées :

- Traits.

Sorties :

- Création d'un taxon racine pour initialiser la taxinomie,
- Création d'un noyau qui est la racine de la base de graphes et qui représente le modèle structurel initial.

Processus :

- Interrogation de l'expert au sujet des traits, concepts, noms de niveaux, et nombre de niveaux de la taxinomie,
- Construction du noyau,
- Contrôle de cohérence,
- Initialisation de la base de connaissances.

5.8.2.2. Classification

Le but de ce module est de positionner une nouvelle classe dans une taxinomie déjà existante, en préservant la cohérence de la base de connaissances. Il faut vérifier que les traits hérités, qu'ils soient discriminants ou intrinsèques, sont valides.

Entrées :

- Classe à positionner,
- Base de graphes et taxinomie courantes.

Sorties :

- Taxinomie et base de graphes enrichies (par les classes).

Processus :

- Interrogation de l'expert au sujet des noms de classes,
- Validation de traits hérités (contrôle de cohérence),
- Création de classe dans la base de graphes et dans la taxinomie,
- Initialisation de la couverture d'une nouvelle classe,
- Mise à jour des couvertures de classes existantes,
- Mise à jour des couvertures de traits.

5.8.2.3. Discrimination

Après que le module de classification ait bien positionné la classe, le module de discrimination a pour rôle de trouver des traits discriminants permettant de distinguer la nouvelle classe par rapport à ses classes sœurs, soit à l'aide d'une procédure d'apprentissage, soit en les demandant explicitement à l'expert.

Entrées :

- Taxinomie courante, nom d'une nouvelle classe, traits.

Sorties :

- Taxinomie, base de graphes et modèle structurel enrichis (par les traits discriminants).

Processus :

- Apprentissage de traits,
- Évaluation de la pertinence des traits,
- Parcours de l'espace de composition,
- Contrôle de cohérence (vérification de traits fournis par l'expert et application du critère de factorisation maximale),
- Raffinement de la base de connaissances (par les traits discriminants),
- Initialisation de la couverture d'un nouveau trait,
- Mise à jour des couvertures de traits.

5.8.2.4. *Saisie*

Le module de saisie a pour but d'acquérir la description intrinsèque des classes si l'expert désire la fournir, ainsi que les contraintes de cohérence associées à ces descriptions.

Entrées :

- Base de graphes, modèle structurel et traits intrinsèques.

Sorties :

- Base de graphes et modèle structurel enrichis (par les traits intrinsèques).

Processus :

- Saisie de traits intrinsèques,
- Raffinement de la base des graphes et du modèle structurel par les traits intrinsèques,
- Contrôle de cohérence (prise en compte de tous les éléments du C-modèle).

5.8.2.5. *Interaction avec l'expert*

Cette interaction se fait entre les quatre modules précédents et l'expert, les questions portent sur la taxinomie à considérer au départ, les classes, les traits (discriminants ou intrinsèques), les valeurs associées aux concepts formant les traits etc .

Entrées :

- Questions posées à l'expert.

Sorties :

- Réponses de l'expert.

Processus :

- Adaptation du dialogue en fonction du type de module considéré,
- Fournir les réponses aux modules.

5.9. CONCLUSION

Nous avons décrit dans ce chapitre la méthode d'explicitation incrémentale du système METIS. Ainsi, nous avons montré comment, à partir d'un noyau de description minimal, METIS explicite des descriptions homogènes, cohérentes et pertinentes, en raffinant petit à petit le langage de description du domaine et en améliorant progressivement la caractérisation des classes. La structuration des informations, renforcée par un mécanisme d'héritage, facilite leur exploitation. Ainsi, pour une description d'espèce donnée, on ne considère à chaque fois, qu'une partie du modèle structurel ; par exemple, on ne demandera pas à l'expert la couleur de la deuxième nageoire dorsale si le poisson appartient à un genre qui n'en possède qu'une.

Le choix d'un modèle générique de base (le graphe primitif) a constitué une étape importante dans le processus cognitif qui permet de passer de l'observation à la description, et qui constitue un point de départ pour le processus d'explicitation incrémental. Par ailleurs, il représente un modèle de référence qui sera interprété et instancié différemment selon les applications (voir chapitre 6), pour lesquels il constitue une modélisation du domaine (voir chapitre 3).

La procédure d'apprentissage IDICE utilise un algorithme de type ID3 afin d'apprendre de nouveaux traits discriminants. IDICE recherche, à chaque fois, l'ensemble des traits maximalement discriminants afin de le suggérer à l'expert. Contrairement à ID3, IDICE prend en compte des descriptions incomplètes et permet à l'expert, d'imposer des contraintes quant au choix et à l'utilisation de certains traits. Un traitement spécial permet de sélectionner les traits multivalués et les traits inapplicables quand ils ne posent plus de problèmes. Par ailleurs, l'étape de complétion d'exemples favorise la recherche de traits potentiellement discriminants et homogénéise dans la mesure du possible la description des taxons. ID3 a été choisi pour son efficacité et pour la simplicité de sa mise en œuvre, d'autant plus que le critère optimisé dans ce type d'algorithme se simplifie considérablement dans notre cas.

La cohérence est assurée grâce à un C-modèle auquel METIS se réfère en permanence pour valider interactivement ses connaissances. La validation se fait grâce à un dialogue permanent avec l'expert qui se trouve ainsi confronté à ses propres connaissances. Par ailleurs, le critère de factorisation maximale garantit que les informations acquises sont toujours placées au "bon niveau" de généralisation.

Contrairement à CASAIS [CASAI, 92] qui réorganise la taxinomie du domaine en remettant en cause aussi bien les traits que les classes, notre approche considère que les classes et leurs positions au sein de la taxinomie ne doivent pas être modifiées⁴³ et de ce fait, s'intéresse uniquement à la réorganisation des traits dont la hiérarchie est en perpétuel remaniement.

Par ailleurs, de façon schématique nous avons déterminé les fonctionnalités de METIS que nous avons regroupées en modules. L'expert interagit avec ces modules afin que METIS puisse enrichir la base de connaissances de manière incrémentale. Les modules ont été décrits en termes d'entrées, de sorties et de processus qui les gouvernent.

En définitive la construction de la base de connaissances se fait par évolution progressive du modèle structurel et des deux structures hiérarchiques (base de graphes et taxinomie des classes).

La partie technique de la mise en œuvre de la méthode METIS est présentée dans le chapitre suivant.

⁴³ Puisque la structure classificatoire guide le processus d'explicitation des connaissances.

6. METIS : MISE EN OEUVRE DE LA MÉTHODE

Ce chapitre traite de la représentation des connaissances et de la mise en œuvre de METIS dans le formalisme des graphes conceptuels. Nous présentons tout d'abord, les types de connaissances pris en compte par METIS et leur organisation. Nous indiquons ensuite comment sont implémentées les opérations qui manipulent ces connaissances, puis nous abordons l'expérimentation de METIS dans deux autres domaines des sciences de l'observation : la minéralogie et la pédologie.

6.1. ORGANISATION DES CONNAISSANCES

Nous allons rappeler dans ce paragraphe, toutes les connaissances manipulées par METIS en instanciant les éléments décrits dans une partie de la théorie de Sowa (introduite dans le chapitre 4 au paragraphe 4.4.). Ainsi, nous commençons par examiner la hiérarchie de généralisation et la taxinomie des classes. Nous décrivons ensuite la fonction du canon dans METIS en nous intéressant aux éléments suivants : les types concepts et relations ainsi que leurs treillis respectifs, le treillis de composition, les référents, le C-modèle et enfin, la base canonique.

6.1.1. Hiérarchie de généralisation

La Hiérarchie de Généralisation HG (voir figure 5.3.) est la structure de données dans laquelle est stocké l'ensemble des graphes associés à tous les taxons. Sa construction est guidée par la taxinomie du domaine et elle correspond à ce que SOWA nomme *hiérarchie Aristotélicienne*. Plus exactement, cela signifie qu'en pratique, on fait appel à la notion d'abstraction telle qu'elle est définie dans [SOWA, 84] (voir aussi chapitre 4, § 4.4.10) :

"A type hierarchy T is said to be *Aristotelian* if every type label t that is a proper subtype of another type label is defined by an abstraction $t = \lambda$ au."

La Hiérarchie Aristotélicienne est une hiérarchie de type dans laquelle tous les types de concepts autres que les types simples peuvent être introduits par des définitions *substantielles* spécifiant un genre commun et une ou plusieurs différences.

La base de graphes décrite dans le paragraphe 5.2.2. reflète une hiérarchie de généralisation dite Aristotélicienne [AIMEUR, 93c], [AIMEUR & GANASCIA 93b]. La racine de la hiérarchie Aristotélicienne est le type primitif auquel on associe le graphe primitif G_0 (dont la description est commune à tous les poissons). Chaque noeud de la hiérarchie est un type défini par une abstraction qui fait intervenir le type qui lui est immédiatement supérieur dans la hiérarchie des types. Ainsi, le type Genre-Citharinus (voir exemple ci après) est défini par une abstraction qui fait intervenir le type Famille-Citharinidae qui est lui même défini par une abstraction qui fait intervenir le type Primitif.

A chaque concept, on associe une définition constituée de deux parties : l'une discriminante jouant un rôle uniquement dans le processus de discrimination (voir § 5.6.1.), l'autre intrinsèque et intervenant dans la construction de la hiérarchie de généralisation. La partie intrinsèque contient la description héritée ainsi que d'autres structures qui décrivent le concept mais qui n'ont aucun rôle dans le processus de discrimination. Ces deux parties de la définition ne sont pas figées : il arrive qu'au cours du processus d'acquisition, certaines structures qui n'étaient pas discriminantes à une étape donnée, le deviennent par la suite et vice versa.

On définit sur les types de concepts définis par genre et différence (voir chapitre 4 § 4.4.7.) deux opérations : la *contraction* de type et l'*expansion* de type. Ces deux opérations vont faciliter le traitement lors du processus d'acquisition. Compte tenu des restrictions de Sowa en ce qui concerne la hiérarchie Aristotélicienne, il s'avère que dans ce cas précis, les opérations d'expansion de type et de contraction de type sont symétriques.

Exemples :

Nous allons montrer sur ces exemples, comment sont définis les différents types constituant la hiérarchie Aristotélicienne de METIS. Nous adoptons la représentation linéaire (Voir § 4.4.).

On définit un type général commun à tous les poissons en faisant une abstraction sur le type Poisson.

Le graphe correspondant à ce type est le graphe primitif G_0 (voir figure 5.2.).

type [Primitif (x)] est

[Poisson : *x] -> (Caractérist) -> [Famille]

[Poisson : *x] -> (Caractérist) -> [Genre]

[Poisson : *x] -> (Caractérist) -> [Espèce]

[Poisson : *x] -> (Caractérist) -> [Descripteur]

[Poisson : *x] -> (Caractérist) -> [Année-Description]

[Poisson : *x] -> (Caractérist) -> [Bassins]

[Poisson : *x] -> (Se-compose) -> [Tête]

[Poisson : *x] -> (Se-compose) -> [Corps]

[Corps] -> (Caractérist) -> [Section]

[Corps] -> (Caractérist) -> [Profil]

[Corps] -> (Se-compose) -> [Système-Natatoire]

[Corps] -> (Se-compose) -> [Ligne-Latérale] -> (Caractérist) -> [Position]

Après construction du graphe G_1 correspondant à la famille Citharinidae (voir figure 5.8.), on lui associe le type suivant :

type [Famille-Citharinidae (x)] est

[Poisson : *x] -> (Caractérist) -> [Famille: *Citharinidae*]

[Poisson : *x] -> (Caractérist) -> [Genre]

[Poisson : *x] -> (Caractérist) -> [Espèce]

[Poisson : *x] -> (Caractérist) -> [Descripteur]

[Poisson : *x] -> (Caractérist) -> [Année-Description]

[Poisson : *x] -> (Caractérist) -> [Bassins]

[Poisson : *x] -> (Se-compose) -> [Tête]

[Poisson : *x] -> (Se-compose) -> [Corps]

[Corps] -> (Caractérist) -> [Section : Compressée]

[Corps] -> (Caractérist) ->[Profil : Haut]

[Corps] -> (Se-compose) -> [Système-Natatoire]

[Corps] -> (Se-compose) -> [Ligne-Latérale] -> (Caractérist) -> [Position : Au-milieu]

Après abstraction sur le type Primitif, ce même type peut, après contraction, s'écrire de la manière suivante :

type [Famille-Citharinidae (x)] **est**

[Primitif : *x] -> (Caractérist) -> [Famille: *Citharinidae*]

[Ligne-Latérale] -> (Caractérist) -> [Position : Au-milieu]

[Corps] -> (Caractérist) ->[Section : Compressée]

[Corps] -> (Caractérist) ->[Profil : Haut]

De même, après contraction le type correspondant au genre *Citharinus* de la famille *Citharinidae* devient (voir figure 5.9.) :

type [Genre-Citharinus (x)] **est**

[Famille-Citharinidae : *x] -> (Caractérist) -> [Genre : *Citharinus*]

[Corps] -> (Se-compose) -> [Écailles]

[Écailles] -> (Caractérist) -> [Forme : Cycloïde]

Si on devait maintenant faire l'expansion du type [Genre-Citharinus] on devrait retrouver le graphe G_{11} qui lui est associé (voir. Figure 5.9.)

Finalement , le type correspondant au genre *Citharidium* de la famille *Citharinidae* est :

type [Genre-Citharidium (x)] **est**

[Famille-Citharinidae : *x] -> (Caractérist) -> [Genre : *Citharidium*]

[Corps] -> (Se-compose) -> [Écailles]

[Écailles] -> (Caractérist) -> [Forme : Cténoïde]

6.1.2. Taxinomie des classes

La taxinomie des classes (voir figure 6.1.) peut être extraite de la hiérarchie de généralisation. La taxinomie des classes et la hiérarchie de généralisation ont des rôles duaux. Elles sont représentées par la même structure et selon que l'on veuille accéder uniquement à la partie discriminante de la description ou à la totalité de cette dernière, on accède à la structure de discrimination ou à la hiérarchie de généralisation.

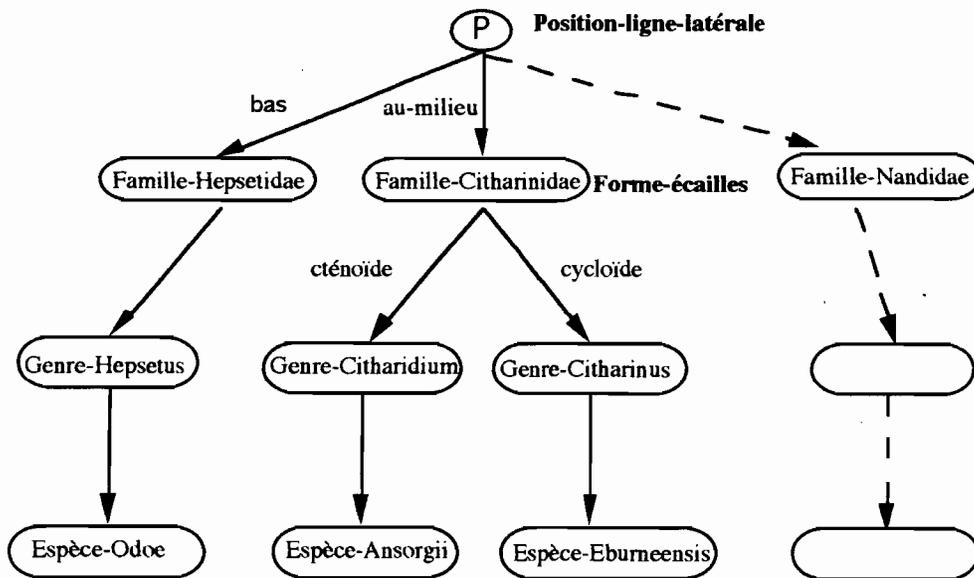


Figure 6.1. Taxinomie des classes de METIS

6.1.3. Concepts

Il existe deux types de concepts : les concepts définis par genre et différence (exemple : [Famille-Citharinidae]) et les concepts de type simple (exemple : [corps] [Tête]). Ces concepts sont organisés sous forme de treillis (voir figure 6.2.).

Conformément à ce qui a été dit dans le paragraphe 5.2.2.1, les concepts de type simples peuvent être structurels ou descriptifs. Ils appartiennent donc à deux sous-types différents : le type Élément-structurel et le type Élément-descriptif (voir figure 6.2.).

6.1.4. Treillis des types de concepts

C'est une notion très importante car elle permet de hiérarchiser les types de concepts (voir chapitre 4 § 4.4.1.). Ainsi, les concepts sont ordonnés selon leur degré de généralité par une relation d'ordre partiel noté "<". Concept1 < Concept2 signifie que Concept2 est plus général que Concept1.

Exemple :

Profil < Élément descriptif

signifie que Profil **est une sorte** d'Élément descriptif.

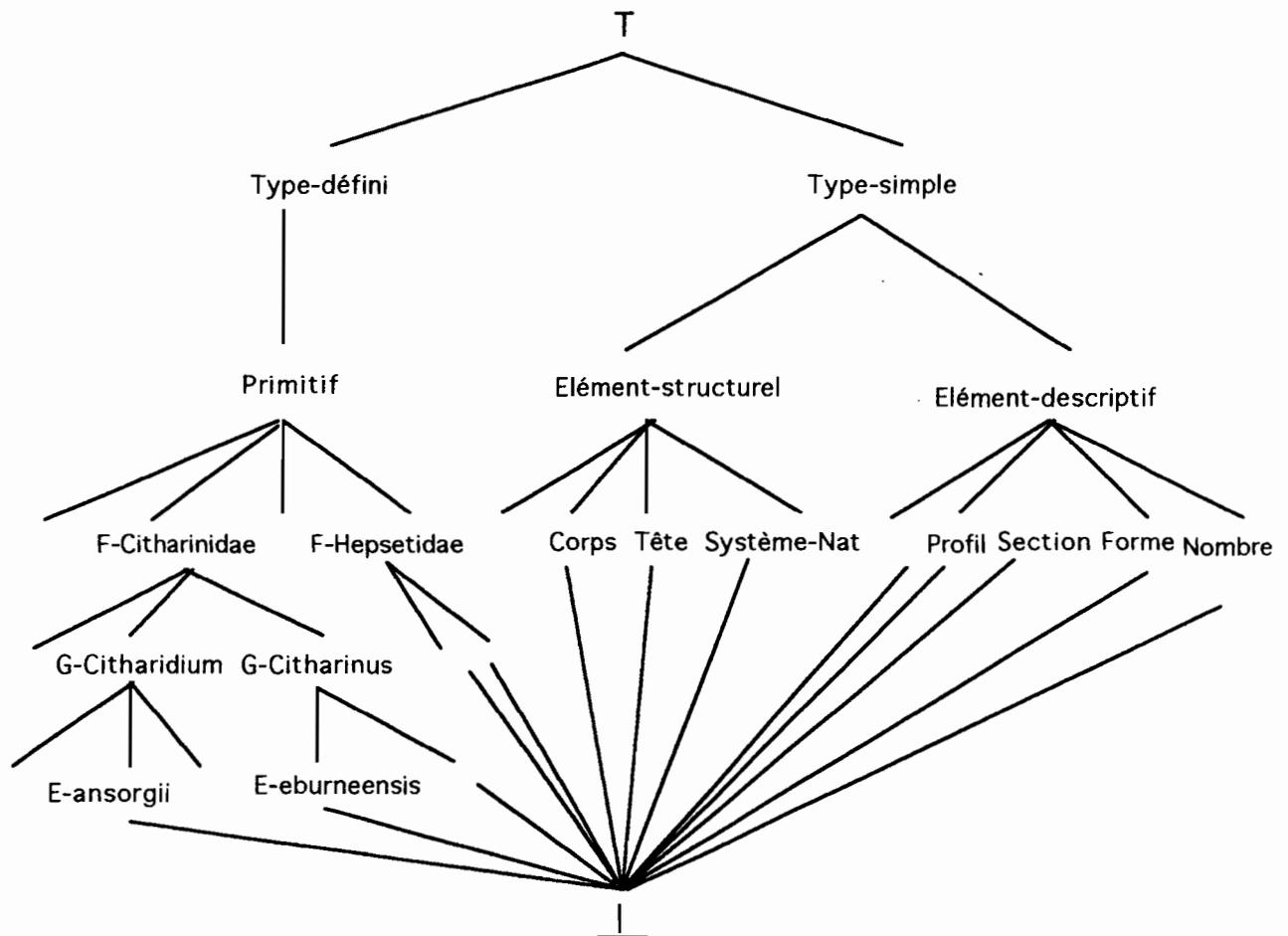


Figure 6.2. Treillis des types de concepts

Remarque :

Les concepts de types définis ont été écrits en abrégé pour alléger la présentation.

6.1.5. Treillis de composition

Les concepts composites sont hiérarchisés selon leur degré de composition (voir figure 6.3.) par une relation d'ordre partiel appelée "Composé-de" :

- Elle est réflexive : A est composé de A .
- Elle est antisymétrique : si A a pour composant B , alors B n'a pas A pour composant.
- Elle est transitive : si A a pour composant B , et B a pour composant C , alors A a pour composant C .

Remarque :

Nous adoptons le point de vue de NAPOLI [NAPOLI, 92a] qui considère que la relation est réflexive, contrairement à ce qui est avancé dans [WINSTON & al, 87] : A n'est pas composé de A .

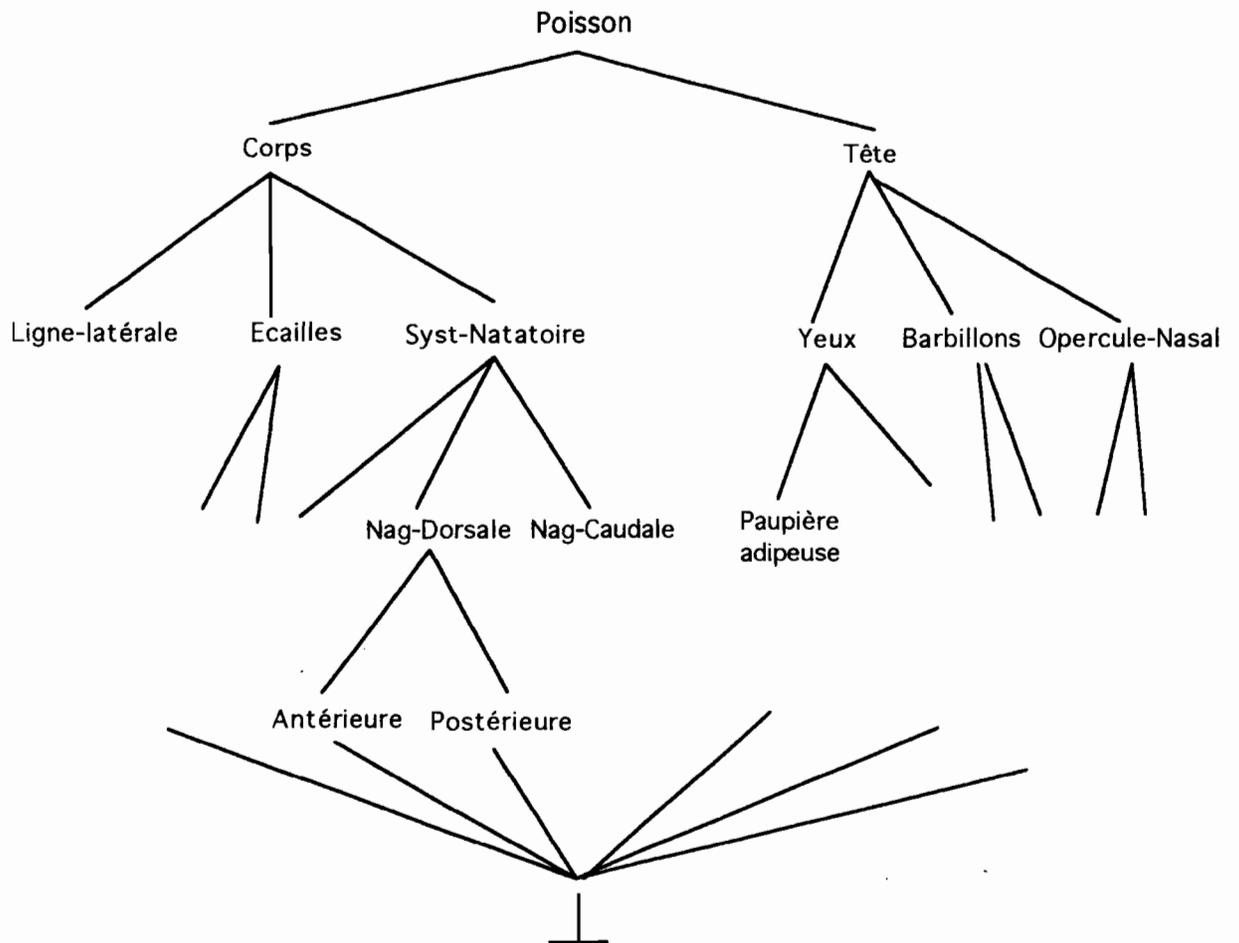


Figure 6.3. Treillis de composition

6.1.6. Relations

Il existe dans METIS trois types de relations : la relation "Caractéristique", la relation "Se-compose" et la relation "!Se-compose". Toutes ces relations sont de même niveau (voir chapitre 4 § 4.4.1.). Le treillis des relations ne contient qu'un niveau¹. Dans [SABAH & VILNAT, 93] et dans [GEY, 94], le treillis des relations est constitué de plusieurs niveaux.

La relation "Caractéristique" est une relation descriptive qui relie deux concepts dont le premier est composite et dont le second est descriptif.

¹ L'utilisation de différents niveaux ne s'est pas avérée utile dans METIS compte tenu des domaines d'applications potentiels.

Les relations "Se-compose" et sa négation "!Se-compose" sont des relations structurelles ou de composition. Elles relient deux concepts composites (présents dans le treillis de composition).

Remarque :

Il n'existe pas de négation à la relation "Caractéristique". Ceci est résolu en associant au référent du concept descriptif une négation².

6.1.7. Référents

Il existe différents types de référents dans METIS³ :

Générique : le référent peut prendre n'importe quelle valeur ; exemple : [Personne : *]

Qualitatif : le référent est désigné par une qualité ; exemple : [Couleur : verte]

Nommé : le référent est désigné par un nom ; exemple : [Bassin : Niger]

Numérique : le référent est un nombre ; exemple : [Année-description : 1923]

Intervalle : le référent est un intervalle ; exemple : [Nombre : 50-60]

Croisé : le référent est associé à un concept considéré comme paramètre formel dans une description par genre et différence (voir chapitre 4 § 4.4.1.). Exemple : [Poisson : *x] dans les exemples donnés en (§ 6.1.1.).

Remarque :

Nous rajoutons un autre type de référent que nous appelons le référent négatif ; c'est celui qui nie une qualité c'est-à-dire que c'est la négation du référent qualitatif. Exemple: [Couleur : !verte].

Que le référent soit qualitatif, négatif, nommé, numérique ou intervalle, il est soit individuel [Profil : bas], soit ensembliste⁴⁵[Forme : {cténoïde, cycloïde}].

2 Ce référent sera appelé référent négatif (voir § 6.1.7).

3 Nous ne traitons pas le référent de type hiérarchique

4 Le problème des référents ensembliste a été abordé par les auteurs suivants : [SOWA, 84], [GARDINER & al, 89], [ESCH & al, 89] et [FARGUES, 92].

5 Nous ne traitons pas les types de concepts de second ordre [NAZARENKO, 92] (où le référent peut être un graphe). Nous nous limitons aux types de concepts dits de premier ordre.

Nous rajoutons la notion de référent *ensembliste-contraint* quand il existe des contraintes entre les référents.

Remarque :

Le référent sous forme d'une conjonction de deux ou plusieurs référents est traité comme un référent individuel.

Exemple :

[Couleur : verte&rouge] pour signifier que la couleur est verte et rouge⁶

6.1.8. le C-modèle

Hormis les éléments que nous avons définis (§ 5.7.3.), le C-modèle comprend aussi le prédicat de conformité défini par Sowa [SOWA, 84].

Le prédicat de conformité noté (::) revient à vérifier si le référent associé au concept est conforme au type du concept.

Exemple :

[Couleur :: 5] ceci est faux car le type de couleur n'est pas numérique, mais qualitatif.

Le critère de factorisation maximale défini dans le paragraphe 5.7.3. utilise une matrice de correspondance appelée *matrice de traits* qui sert à mettre en correspondance les traits et les classes décrites par ceux-ci. Les traits considérés dans la matrice de traits sont les traits structurels, les traits descriptifs et leurs généralisés. La matrice de traits a aussi une grande utilité dans la phase de discrimination et plus précisément pour la procédure d'apprentissage (voir § 5.6.3.1.).

⁶ Verte&rouge est considéré comme un nouveau référent.

Remarque :

Il y a une certaine analogie entre la matrice de traits et la matrice d'intersection de MINEAU [MINEAU, 90] [MINEAU & al, 90] au sens où elles facilitent toutes deux l'apprentissage. La matrice d'intersection sert à trouver des *intersections exclusives* afin de mettre en relief des généralisations communes aux objets considérés. Dans la matrice de traits, on cherche plutôt à trouver des *différences exclusives* afin de discriminer les objets.

Exemple :

Si nous reprenons les exemples présentés dans le chapitre 5, pour expliquer l'algorithme d'apprentissage IDICE (voir § 5.6.3.1.), nous obtenons la matrice de traits suivante :

	F1	F2	F3	F4
<Corps, Caractérist, Section, *>	1	1	1	1
<Corps, Caractérist, Section, Comprimée>	0	1	1	1
<Corps, Caractérist, Section, Arrondie>	1	1	0	0
<Corps, Caractérist, Profil, *>	1	1	1	1
<Corps, Caractérist, Profil, Allongé>	1	1	0	0
<Corps, Caractérist, Profil, Haut>	0	0	1	0
<Corps, Caractérist, Profil, Fusiforme>	0	0	0	1
<Ligne-Latérale, Caractérist, Position, *>	1	1	1	1
<Ligne-Latérale, Caractérist, Position, Basse>	1	0	0	0
<Ligne-Latérale, Caractérist, Position, Au-milieu>	0	1	1	1
<Écailles, Caractérist, Forme, *>	1	1	1	1
<Écailles, Caractérist, Forme, Cténoïde>	0	1	1	0
<Écailles, Caractérist, Forme, Cycloïde>	1	0	0	1
<Mâchoires, Se-compose, Dents, présent>	1	1	1	0
<Mâchoires, Se-compose, Dents, absent>	0	0	0	1
<Dents, Caractérist, Disposition, *>	1	1	1	0
<Dents, Caractérist, Disposition, 1R>	0	1	1	0
<Dents, Caractérist, Disposition, 2R>	0	1	0	0
<Dents, Caractérist, Disposition, 3R>	1	0	0	0

Remarque1 :

Un arc conceptuel peut être représenté par un quintuplé de la forme :

<Concept1, Référent1, Relation, Concept2, Référent2>. Dans nos exemples, nous représentons l'arc conceptuel par un quadruplé car Référent1 est toujours générique⁷.

Remarque2 :

Pour améliorer la rapidité d'accès aux traits, on pourrait envisager d'avoir une matrice à deux niveaux . Le premier niveau ne contiendrait que les traits structurels et les traits ascendants dont dérivent les traits descriptifs valués. Le premier niveau servirait d'index au deuxième niveau, qui ne contiendrait que les traits descriptifs valués.

6.1.9. Base canonique

La base canonique (voir § 4.4.1.) encode les contraintes sémantiques imposées sur tout graphe valide⁸ dans le système [MINEAU, 92]. Nous stockons incrémentalement dans la base canonique , la base de graphes et le modèle structurel (voir chapitre 5 § 5.5.2.3.). Il est à noter que le graphe primitif est le graphe à partir duquel nous dérivons tous les graphes valides par spécialisation, et que le modèle structurel est le graphe à partir duquel nous dérivons tous les graphes valides par généralisation.

Remarque :

Dans METIS tous les graphes conceptuels construits sont canoniques, car nous considérons qu'un expert ne fournit que des graphes sémantiquement corrects⁹.

⁷ Concept1 est un concept composite.

⁸Un graphe conceptuel est considéré comme valide s'il peut être dérivé à partir d'un graphe canonique en utilisant les règles de formation canoniques décrites dans le chapitre 4 en 4.4.1.

⁹ Cette hypothèse s'appuie sur le fait que pour construire une base de connaissances donnée, METIS ne se réfère qu'à un seul expert.

6.1.10. Canon

Le **canon** contient tous les éléments définis dans [SOWA, 84] et qui sont nécessaires pour dériver un ensemble de graphes canoniques (appartenant à la base canonique). De plus, dans notre cas, nous rajoutons explicitement le treillis de composition, ainsi que le C-modèle qui comprend le prédicat de conformité, ce qui donne la liste suivante :

- le treillis des types de concepts,
- le treillis des types de relations,
- le treillis de composition,
- l'ensemble des référents,
- le C-modèle,
- la base canonique.

Construit incrémentalement, et constitué des éléments précédents, le canon définit la base de connaissances du domaine.

6.1.11. Opérations utilisées

Nous examinons dans ce paragraphe les différentes opérations nécessaires à la mise en œuvre de METIS et qui reposent sur le formalisme des graphes conceptuels. les algorithmes correspondant à ces opérations sont décrits an Annexe 2.

Opérations sur les arcs conceptuels

Il s'agit d'ajouter, de supprimer ou de rechercher un arc dans un graphe conceptuel. Ces opérations nécessitent d'une part l'opération de jointure maximale¹⁰ (dans le cas de l'ajout), et l'opération de projection (dans le cas de la suppression et de la recherche), d'autre part, la mise à jour du canon (voir § 6.1.10) et celle des couvertures d'arcs (voir Annexe 2).

¹⁰ En fait c'est une jointure d'ensemble que l'on fait ("Set join" [Sowa, 84] page 117), car les référents que nous manipulons sont le plus souvent des référents ensemblistes.

Remarque :

Les opérations que nous venons de citer ont été définies dans le chapitre 4 au paragraphe 4.4.

Opérations sur les treillis

Il s'agit d'ajouter ou de supprimer un concept dans le treillis des types de concepts ou dans le treillis de composition. Une autre opération sur les treillis consiste à rechercher un concept dans le treillis de composition (voir Annexe 2).

En ce qui concerne l'ajout d'un concept dans les treillis de type ou de composition, il faut vérifier que le concept n'existe pas déjà. S'il s'agit d'un nouveau concept, on l'insère à l'endroit approprié. Dans le cas de la suppression d'un concept dans les treillis de type ou de composition, il faut faire un parcours du treillis afin de localiser le concept à supprimer. S'il s'agit d'un concept composite, on supprime aussi toute sa descendance, lorsqu'elle n'intervient pas pour un autre concept composite.

Pour rechercher un concept dans le treillis de composition, le parcours se fait en profondeur en privilégiant les concepts composites ayant le moins de composants, c'est de la *profondeur ordonnée*. Afin de favoriser ce parcours, METIS posera, par exemple, des questions d'abord sur la ligne latérale puis sur les écailles et enfin sur le système natatoire (voir figure 6.3.).

Opérations sur les référents

Extension de référent¹¹

l'extension : pour tout concept c de u , si c est un concept dont le référent est individuel ou ensembliste, son référent est étendu par un référent ensembliste qui l'inclut. Ces changements sont permis seulement si le référent de c est conforme au type de c , après extension.

11 Cette opération correspond à l'opération (augmentation de l'étiquette d'un C-sommet) définie par CHEIN [CHEIN, 92], à ceci près qu'ici on ne s'intéresse qu'aux référents et non aux types de concepts.

Exemple :

Extension ([Ligne-Latérale] -> (Caractérist) -> [Position : {basse, au-milieu}]
avec le référent : "haute") => [Ligne-Latérale] -> (Caractérist) -> [Position : {basse,
au-milieu, haute}]

Définition

Si un graphe G_i est dérivé d'un graphe canonique G en utilisant la règle d'extension, alors G_i est aussi un graphe canonique généralisation de G ($G_i > G$).

Réduction de référents ensembliste¹²

la réduction : pour tout concept c de u , ayant un référent ensembliste, l'opération de réduction consiste à réduire son référent à un référent ensembliste plus petit autre que le référent générique ensembliste.

Exemple :

Réduction ([Ligne-Latérale] -> (Caractérist) -> [Position : {basse, au-milieu}]
valeur à réduire "au-milieu") => [Ligne-Latérale] -> (Caractérist) -> [Position :
{basse}]

Définition

Si un graphe G_i est dérivé d'un graphe canonique G en utilisant la règle de réduction, alors G_i est aussi un graphe canonique spécialisation de G ($G_i < G$).

Opération de déréférenciation

Nous introduisons une nouvelle opération qui concerne la hiérarchie Aristotélicienne et qui est utilisée par la procédure d'apprentissage lors du processus de discrimination (voir § 5.6.1.). Pour cela, on définit une nouvelle règle de formation canonique : la *déréférenciation*.

¹² Cette opération est l'inverse de l'extension de référents et correspond à l'opération (diminution de l'étiquette d'un C-sommet) définie par CHEIN [CHEIN, 92] , à ceci près qu'ici on ne s'intéresse qu'aux référents et non aux types de concepts.

la déréférenciation: pour tout concept c de u dans une hiérarchie Aristotélicienne¹³ si c est un concept dont le référent est individuel ou ensembliste, son référent est remplacé par le marqueur générique.

Exemple :

Déréférenciation ([Écailles] -> (Caractérist) -> [Forme : {cycloïde, cténoïde}])
=> ([Écailles] -> (Caractérist) -> [Forme])

Définition

Si un graphe G_i est dérivé d'un graphe canonique G dans une hiérarchie Aristotélicienne en utilisant la règle de Déréférenciation, alors G_i est aussi un graphe canonique généralisation de G ($G_i > G$) dans la hiérarchie Aristotélicienne.

Héritage de référents

Cette opération n'est pas définie dans Sowa et correspond à ce que l'on appelle *masquage et redéfinition*¹⁴ [MASINI & al, 89]. En ce qui nous concerne, elle permet l'héritage dans une hiérarchie Aristotélicienne. Pour cela, nous introduisons une nouvelle règle de formation canonique appelée *H-restriction* dérivée de la règle de formation canonique (restriction) et qui étend sa définition.

la H-restriction : pour tout concept c de u , $type(c)$ peut être remplacé par un sous-type de $type(c)$; si c est un concept générique, son référent peut être remplacé par un marqueur individuel. Si le référent de c est ensembliste (RE), il peut être remplacé par un sous-ensemble. Ces changements sont permis seulement si le référent de c est conforme au type de c , avant et après restriction dans une hiérarchie Aristotélicienne.

Définition

Si un graphe G_i est dérivé d'un graphe canonique G ($G_i < G$) dans une hiérarchie Aristotélicienne en utilisant la règle de H-restriction, alors G_i est aussi un graphe canonique spécialisation de G dans la hiérarchie Aristotélicienne.

¹³ Dans une hiérarchie Aristotélicienne, la règle de déréférenciation est utilisée par l'opérateur de généralisation (voir § 4.4.5).

¹⁴ Sowa n'a pas défini cette règle car il a voulu conserver à son formalisme un caractère monotone.

Exemple :

Cas de référent ensembliste

H-restriction de ([Ligne-Latérale] -> (Caractérist) -> [Position : {basse, au-milieu, haute}]) au niveau **Famille** à [Ligne-Latérale] -> (Caractérist) -> [Position : {basse, au-milieu}] au niveau **Genre**.

6.2. IMPLANTATION ET EXPÉRIMENTATION

Le langage d'implémentation choisit est CLOS (Common Lisp Orienté Objets) [KEENE, 89], [GUY & STEELE, 90]. La machine sur laquelle a été implémenté METIS est un Macintosh IIsx (9/80), son choix est dû à sa disponibilité au centre de Bondy de l'ORSTOM. Nous avons initialement implémenté une interface graphique (voir Annexe 3) à l'aide d'HyperCard (donc sur Macintosh) et l'idée était de munir METIS de cette interface graphique [AIMEUR & al, 93], [AIMEUR & KIEU, 93]. Elle s'est avérée finalement assez lourde et insuffisante pour pouvoir implémenter toutes les interactions avec les modules du système¹⁵.

Pour montrer que la méthode d'explicitation de taxonomies de METIS - développée initialement dans le domaine de l'ichtyologie - est généralisable, nous l'avons expérimenté dans deux autres domaines des sciences d'observation : le domaine de la minéralogie et celui de la pédologie. Par ailleurs, une étude critique de METIS ainsi qu'une expérimentation en archéologie a été faite par Catherine Faron de l'équipe ACASA du LAFORIA [FARON, 93].

6.2.1. Expérimentation en minéralogie

La minéralogie est la science qui s'occupe des minéraux, de leur morphologie, de leur composition, de leurs propriétés physiques et chimiques, sans oublier les conditions de leur formation. Les minéraux sont des combinaisons chimiques naturelles souvent très complexes, parfois plus rarement des éléments simples comme l'or, l'argent ou le cuivre.

Les minéraux sont donc des substances naturelles, de composition chimique et de structure cristalline déterminée. Cette dernière se présente souvent sous une forme

¹⁵ L'interface de la version actuelle de METIS n'est pas graphique.

géométrique plus ou moins régulière. La calcite, le mica et le diamant en sont des exemples.

Les minéraux sont classés d'après leur composition chimique et leur structure interne. Mais ce classement cristallographique comporte aussi quelques inconvénients. En effet, certains minéraux ont une composition chimique variable au sein d'un échantillon donné tout en conservant leur structure cristalline [DUDA & REJL, 86].

Nous avons bénéficié de l'aide du minéralogiste Alain Blot de l'ORSTOM et de celle de Hassina Mouri, chercheur à l'université Paris 7 au département de pétrologie. Les ouvrages suivants : "A course of mineralogy" [BETEKHTIN, 75] et de "La grande encyclopédie des minéraux" [DUDA & REJL, 86] ont servi de support pour l'expertise.

Exemple :

Il existe neuf classes de minéraux (1. éléments, 2. sulfures, 3. halogénures, 4. oxydes et hydroxydes, 5. nitrates, carbonates, borates, 6. sulfates, 7. phosphates, arséniates, vanadates, 8. silicates, 9. substances organiques). Ces classes se divisent ensuite en sous-classes, groupes, sous-groupes, espèces et variété. Les minéraux qui renferment une faible quantité d'un autre élément ou ceux qui, ayant même composition chimique et même structure cristalline, diffèrent extérieurement de l'espèce considérée, sont appelées des variétés.

Nous nous proposons dans cet exemple, d'appliquer METIS à la sous-classe des Oxydes afin de l'explicitier et de caractériser la taxinomie associée (voir figure 6.4.). La sous-classe des Oxydes comprend toutes les combinaisons oxygénées du règne minéral.

Initialisation :

Nombre de niveaux considérés :

En accord avec les experts nous nous sommes limités à 2 niveaux : niveau 1, "espèce", et niveau 2, "variété".

Noyau :

[Minéral] -> (Caractérist) -> [Sous-classe : Oxydes]

[Minéral] -> (Caractérist) -> [Nom-Espèce]

[Minéral] -> (Caractérist) -> [Nom-Variété]
[Minéral] -> (Caractérist) -> [Gisement]
[Minéral] -> (Caractérist) -> [Localisation]
[Minéral] -> (Caractérist) -> [Syst-cristallin]
[Minéral] -> (Se-compose) -> [Particules]
[Particules] -> (Caractérist) -> [Comp-chimique]

Informations sur le noyau :

La discrimination ne doit pas porter sur les éléments suivants : Nom-Espèce, Nom-Variété (car ce sont des noms de taxons) et la localisation (car elle appartient à la fiche signalétique de l'espèce ou de la variété¹⁶).

Type [Primitif (x)] est

[Minéral : *x] -> (Caractérist) -> [Sous-classe : Oxydes]
[Minéral : *x] -> (Caractérist) -> [Nom-Espèce]
[Minéral : *x] -> (Caractérist) -> [Nom-Variété]
[Minéral : *x] -> (Caractérist) -> [Gisement]
[Minéral : *x] -> (Caractérist) -> [Localisation]
[Minéral : *x] -> (Caractérist) -> [Syst-Cristallin]
[Minéral : *x] -> (Se-compose) -> [Particules]
[Particules] -> (Caractérist) -> [Comp-chimique]

A cette étape la taxinomie est initialisée par le type primitif, tandis que la base des graphes est initialisée par le graphe qui définit le type primitif. Le treillis des relations contient deux relations (Se-compose et Caractérist), le treillis des types contient un concept défini (Primitif) deux concepts structurels (Minéral, Particules) et plusieurs concepts descriptifs (Sous-classe, Nom-Espèce, Nom-Variété, Gisement, Localisation, ...).

Le treillis de composition contient le concept composite [Minéral] et son composant le concept [Particules].

¹⁶ Certaines espèces ne possèdent pas de variétés et il faut, donc, décrire la localisation au niveau de l'espèce.

Le graphe représentant le modèle structurel est le même que celui du graphe primitif.

Le C-modèle contient plusieurs types de référents associés aux concepts, par exemple, nommé (pour [Nom-Espèce], [Nom-Variété] et [Localisation]), et qualitatif (pour [Syst-Cristallin] [Gisement] [Comp-Chimique]).

Nous allons maintenant décrire l'espèce Opale et trois de ses variétés (l'Opale-Noble, l'Opale-de-Feu et l'Opale-Blanche-de-Lait).

Description de l'espèce Opale

Type [Espèce-Opale : (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Espèce : Opale]

[Primitif : *x] -> (Caractérist) -> [Dureté : 5,5-6,6]

[Primitif : *x] -> (Caractérist) -> [Couleur : {Blanche, Jaune, Rouge, Brune, Verte, Bleue, Noire, Opalescente}]

[Primitif : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras, Mat, Cireux}]

[Primitif : *x] -> (Caractérist) -> [Trace : blanche]

[Primitif : *x] -> (Caractérist) -> [Syst-Cristallin : Amorphe]

[Primitif : *x] -> (Caractérist) -> [Comp-chimique : SiO₂.nH₂O]

Classification :

METIS rajoute la classe Opale dans la taxinomie dans la base de graphes au niveau espèce et dans le treillis des types en tant que type défini.

La couverture de la classe est initialisée à 1 car c'est la première espèce (voir § 5.7.3.)

Saisie :

METIS demande le référent associé au concept [Nom-Espèce] car ce concept est demandable au premier niveau. Puis l'expert décrit librement ce qui caractérise cette espèce. Il y a introduction de 4 nouveaux concepts descriptifs (Dureté, Trace, Couleur, Éclat) dans le treillis des types.

Le modèle structurel est enrichi par les référents associés aux concepts [Nom-Espèce] et [Syst-Cristallin] [Comp-chimique] et par les nouveaux traits introduits. Les couvertures de traits sont initialisées à 1, car c'est la première classe.

Description de l'Opale-Noble qui est une première variété de l'Opale

Type [Variété-Opale-Noble : (x)] **est**

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-Noble]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Opalescente]

[Espèce-Opale : *x] -> (Caractérist) -> [Luminescence : {Verte, Brune}]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : rare]

[Espèce-Opale : *x] -> (Caractérist) -> [Localisation : {Tchécoslovaquie, Etats-unis, Mexique, îles Feroë, Honduras, U.R.S.S, Brésil, Australie}]

Classification :

METIS rajoute la classe Opale-Noble dans la taxinomie et dans la base de graphes au deuxième niveau (Variété) et dans le treillis des types en tant que type défini.

La couverture de la classe vaut toujours 1 car c'est la première variété de l'espèce opale.

Saisie :

METIS demande les référents des concepts [Nom-Variété], [Gisement] et [Localisation] car ils sont demandables au deuxième niveau. Puis l'expert décrit librement ce qui caractérise cette variété. Il y a introduction d'un nouveau concept descriptif [Luminescence] dans le treillis des types.

Le modèle structurel est enrichi par les référents associés aux concepts [Nom-Variété], [Gisement] et [Localisation] et par le nouveau trait [Minéral] -> (Caractérist) -> [Luminescence : {Verte, Brune}] La couverture des nouveaux traits¹⁷ introduits vaut 2 pour chacun d'entre-eux car c'est la deuxième classe décrite, sauf pour le trait [Minéral] -> (Caractérist) -> [Couleur : Opalescente] qui vaut {1,2}, car le référent Opalescente existe dans la classe 1 et 2.

¹⁷ Le calcul de couverture de traits a été donné dans le paragraphe 5.3.7.

Description de l'Opale-de-Feu qui est une deuxième variété de l'Opale

Type [Variété-Opale-de-Feu : (x)] est

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-de-Feu]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Rouge]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : rare]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Mexique, îles Feroë, Turquie, U.R.S.S. }]

Classification :

METIS rajoute la classe Opale-de-Feu dans la taxinomie et dans la base de graphes au deuxième niveau (Variété) et dans le treillis des types en tant que type défini.

La couverture de la classe vaut 2 car cette variété est la deuxième descendante de l'espèce. La couverture de l'espèce est mise à jour, elle vaut alors {1,2}.

Discrimination :

METIS valide les traits intrinsèques de l'espèce Opale, puis consulte les traits intrinsèques de la variété Opale-Noble et demande le référent associé au concept [Couleur]¹⁸. L'expert répond "Rouge". METIS propose alors à l'expert le trait discriminant [Minéral] -> (Caractérist) -> [Couleur] car la couleur étant Opalescente pour l'Opale-Noble et rouge pour l'Opale-de-Feu, elle suffit donc à discriminer les deux variétés. L'expert accepte et METIS initialise la description discriminante de l'espèce Opale par le trait [Minéral] -> (Caractérist) -> [Couleur].

Saisie :

METIS demande les référents des concepts [Nom-Variété], [Gisement] et [Localisation]. L'expert décide de ne pas continuer la description de l'Opale-de-Feu.

Le modèle structurel est enrichi par les référents associés aux concepts [Nom-Variété] [Couleur] et [Localisation]¹⁹. Les couvertures de traits sont initialisées à 3

¹⁸ METIS interroge l'expert sur la couleur plutôt que sur la luminescence, car c'est le concept couleur qui est décrit en premier.

¹⁹ METIS rajoute dans le modèle structurel le référent Turquie car il n'a jamais été rencontré.

(troisième classe décrite) sauf pour les traits [Minéral]-> (Caractérist) -> [Gisement : rare], [Minéral] -> (Caractérist) -> [localisation : Mexique], [Minéral] -> (Caractérist) -> [localisation : îles Feroë] et [Minéral] -> (Caractérist) -> [localisation : U.R.S.S], où elle vaut {2,3}.

METIS généralise le trait [Minéral] -> (Caractérist) -> [Gisement : rare] au niveau de l'espèce Opale. Si à un moment donné, l'expert décrit une variété appartenant à l'espèce Opale et que le gisement est fréquent le trait [Minéral] -> (Caractérist) -> [Gisement : rare] n'est plus valide. METIS fait alors appel à l'expert qui lui demande d'étendre le domaine de valeurs du gisement au niveau de l'espèce. METIS masquera alors l'héritage du référent afin de le redéfinir pour les variétés concernées (voir § 5.5.2.).

Description de l'Opale-Blanche-de-Lait qui est une troisième variété de l'Opale

Type [Variété-Opale-Blanche-de-Lait : (x)] **est**

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-Blanche-de-Lait]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Blanche]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : rare]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Tchécoslovaquie, Etats-unis, Mexique, îles Feroë, Honduras, U.R.S.S, Brésil, Australie}]

Classification :

METIS rajoute la classe Opale-Blanche-de-Lait dans la taxinomie et dans la base de graphes au deuxième niveau (Variété) et dans le treillis des types en tant que type défini.

La couverture vaut 3 car cette variété est la troisième descendante de l'espèce. La couverture de l'espèce est mise à jour et vaut maintenant {1,2,3}.

Discrimination :

METIS valide les traits intrinsèques de l'espèce.

METIS ne consulte aucun trait intrinsèque, le trait discriminant étant [Minéral] -> (Caractérist) -> [Couleur], il demande le référent associé au concept [Couleur], l'expert répond "Blanche".

Le trait discriminant est suffisant pour discriminer l'Opale-Blanche-de-Lait de l'Opale-Noble et de l'Opale-de-Feu.

Saisie :

METIS demande les référents des concepts [Variété], [Gisement] et [Localisation].

L'expert décide de ne pas continuer la description de l'Opale-Blanche-de-lait.

Le modèle structurel est enrichi par les référents associés aux concepts [Variété] et [Couleur]. En ce qui concerne le concept [Localisation], il n'y a pas de nouveau référent introduit.

Nous allons maintenant décrire deux nouvelles espèces qui ne possèdent pas de variétés. Nous nous focaliserons uniquement sur la discrimination.

Description de l'espèce Valentinite (espèce appartenant à la sous-classe des oxydes)

Type [Espèce-Valentinite : (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Espèce : Valentinite]

[Primitif : *x] -> (Caractérist) -> [Dureté : 2-3]

[Primitif : *x] -> (Caractérist) -> [Couleur : {Blanche, Grise, Rouge, Gris-Jaune}]

[Primitif : *x] -> (Caractérist) -> [Éclat : {Adamantin, Nacré}]

[Primitif : *x] -> (Caractérist) -> [Comp-chimique : Sb_2O_3]

[Primitif : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, France, Algérie, Bolivie}]

Discrimination :

METIS discrimine l'espèce Opale de l'espèce Valentinite grâce à la propriété de dureté (5,5-6,6 pour l'Opale et 2-3 pour la Valentinite).

Description de l'espèce Sénarmontite (espèce appartenant à la sous-classe des oxydes)

Type [Espèce-Sénarmontite : (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Espèce : Sénarmontite]

[Primitif : *x] -> (Caractérist) -> [Dureté : 2-3]

[Primitif : *x] -> (Caractérist) -> [Couleur : {Blanche, Gris-Blanche}]

[Primitif : *x] -> (Caractérist) -> [Éclat : {Adamantin, Soyeux, Gras}]

[Primitif : *x] -> (Caractérist) -> [Comp-chimique : Sb_2O_3 ²⁰]

[Primitif : *x] -> (Caractérist) -> [Localisation : {Algérie, Canada, Tchécoslovaquie, Allemagne, Roumanie, Italie}]

Discrimination :

La procédure d'apprentissage échoue car elle ne trouve aucun trait discriminant, METIS demande à l'expert de compléter les espèces Valentinite et Sénarmontite par les référents associés aux concepts [Trace], [Syst-Cristallin] et [Gisement]²¹. Ces derniers étant décrits dans l'espèce Opale.

L'expert rajoute respectivement aux descriptions des espèces Valentinite et Sénarmontite, les ensembles de traits suivants :

([Minéral] -> (Caractérist) -> [Trace : blanche],

[Minéral] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique],

[Minéral] -> (Caractérist) -> [Gisement]).

([Minéral] -> (Caractérist) -> [Trace : Blanche],

[Minéral] -> (Caractérist) -> [Syst-Cristallin : Cubique],

[Minéral] -> (Caractérist) -> [Gisement]).

METIS propose alors à l'expert le trait discriminant [Minéral] -> (Caractérist) -> [Syst-Cristallin], car il prend des valeurs exclusives pour les trois espèces décrites.

Le trait ayant été approuvé par l'expert, METIS remplace l'ancien trait [Minéral] -> (Caractérist) -> [Dureté] par le nouveau, car à lui seul il suffit à discriminer les trois espèces.

20 La formule chimique de l'espèce Sénarmontite est la même que celle de l'espèce Valentinite mais les pourcentages en Sb et O sont différents, de plus il y a inclusion d'As pour la Valentinite.

21 Le lecteur se souvient que le trait [Minéral] -> (Caractérist) -> [Gisement] a été généralisé au niveau de l'espèce Opale.

La taxinomie représentant les classes décrites ci-dessus est illustrée par la figure 6.4.

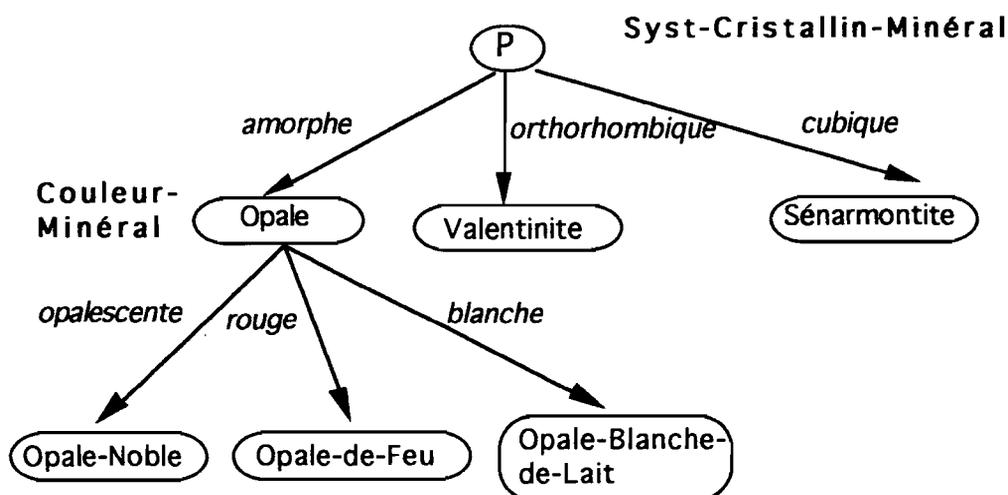


Figure 6.4. Taxinomie partielle de la sous-classe des oxydes

Remarque :

Une expérimentation sur une centaine d'exemples en minéralogie ainsi qu'un bilan critique de l'utilisation de METIS sont présentés en Annexe 4.

6.2.2. Expérimentation en pédologie

La pédologie est la branche de la géologie qui étudie les caractères physiques et biologiques, l'évolution et la répartition des sols. Nous nous sommes inspiré dans notre expérimentation des ouvrages suivants : "Keys to Soil Taxonomy" [VAN WAMBEKE & al, 83], et "Une épistémologie des sciences du sol" [CHATELIN, 79].

Nos nombreuses discussions avec les pédologues Vincent Simmoneaux et Yvon Chatelin de l'ORSTOM nous ont fait prendre conscience de la difficulté à caractériser les sols et surtout du fait qu'il n'existe actuellement aucun consensus international sur une taxinomie de sol donné²².

²² Rappelons que METIS tient compte de l'avis d'un expert unique et l'une des perspectives est justement de pouvoir faire de la "multi-explicitation des connaissances".

Exemple :

Nous nous proposons ici, d'appliquer METIS à l'ordre des oxisols qui sont, entre autres²³, des sols minéraux saturés en eau pendant une certaine période de l'année.

Nous ne donnerons pas le détail des phases, nous présentons simplement les descriptions de classes et la taxinomie associée.

Phase d'initialisation

Nombre de niveaux considérés : 3

Sous ordre, groupe, sous-groupe

Noyau

[Sol] -> (Caractérist) -> [Ordre : Oxisol]
[Sol] -> (Caractérist) -> [Sous-ordre]
[Sol] -> (Caractérist) -> [Groupe]
[Sol] -> (Caractérist) -> [Sous-groupe]
[Sol] -> (Caractérist) -> [Régime]
[Sol] -> (Se-compose) -> [Horizon]
[Sol] -> (Caractérist) -> [Granulométrie : 25-100]
[Horizon] -> (Caractérist) -> [Nom]
[Horizon] -> (Caractérist) -> [Base]
[Horizon] -> (Caractérist) -> [Profond-apparition]
[Horizon] -> (Caractérist) -> [Epaisseur : mince]

Informations sur le noyau :

La discrimination ne doit pas porter sur les éléments suivants : Ordre , Sous-ordre, Sous-groupe (car ce sont des noms de taxons), la granulométrie et l'épaisseur car elles prennent les mêmes valeurs pour tous les descendants des oxisols.

²³ Nous ne rentrons pas dans les détails des caractéristiques de ces oxisols.

Type [Primitif (x)] est

[Sol : *x] -> (Caractérist) -> [Ordre : Oxisol]
[Sol : *x] -> (Caractérist) -> [Sous-ordre]
[Sol : *x] -> (Caractérist) -> [Groupe]
[Sol : *x] -> (Caractérist) -> [Sous-groupe]
[Sol : *x] -> (Caractérist) -> [Régime]
[Sol : *x] -> (Se-compose) -> [Horizon]
[Sol : *x] -> (Caractérist) -> [Granulométrie : 25-100]
[Horizon] -> (Caractérist) -> [Type]
[Horizon] -> (Caractérist) -> [Base]
[Horizon] -> (Caractérist) -> [Profond-apparition]
[Horizon] -> (Caractérist) -> [Epaisseur : mince]

Description du sous-ordre Aquox de l'ordre des Oxisols

Type [Sous-ordre- Aquox (x)] est

[Primitif : *x] -> (Caractérist) -> [Sous-ordre : Aquox]
[Primitif : *x] -> (Caractérist) -> [Régime : Saturé-en-eau]
[Horizon] -> (Caractérist) -> [Type : oxique]

Description du groupe Plinthaquox du sous-ordre Aquox

Type [Groupe- Plinthaquox (x)] est

[Sous-ordre- Aquox : *x] -> (Caractérist) -> [Groupe : Plinthaquox]
[Sous-ordre- Aquox : *x] -> (Se-compose) -> [Plinthite]
[Sous-ordre- Aquox : *x] -> (!Se-compose) -> [Epipédon]
[Plinthite] -> (Caractérist) -> [Phase : discontinue]
[Horizon] -> (Caractérist) -> [Chroma : 2]

Description du groupe Ochraquox du sous-ordre Aquox

Type [Groupe- Ochraquox (x)] est

[Sous-ordre- Aquox : *x] -> (Caractérist) -> [Groupe : Ochraquox]

[Sous-ordre- Aquox : *x] -> (Se-compose) -> [Plinthite]

[Sous-ordre- Aquox : *x] -> (Se-compose) -> [Epipédon]

[Plinthite] -> (Caractérist) -> [Phase : discontinue]

[Epipédon] -> (Caractérist) -> [Nom : Ochrique]

[Horizon] -> (Caractérist) -> [Base : 1]

[Horizon] -> (Caractérist) -> [Texture : Argilo-sableuse]

Description du sous-ordre Ustox de l'ordre des Oxisols

Type [Sous-ordre-Ustox (x)] est

[Primitif : *x] -> (Caractérist) -> [Sous-ordre : Ustox]

[Primitif : *x] -> (Caractérist) -> [Régime : ustique]

[Horizon] -> (Caractérist) -> [Type : oxique, sombrique]

Description des groupes du sous-ordre Ustox

Type [Groupe- Sombriustox (x)] est

[Sous-ordre- Ustox : *x] -> (Caractérist) -> [Groupe : Sombriustox]

[Horizon] -> (Caractérist) -> [Type : sombrique]

Type [Groupe- Acrustox (x)] est

[Sous-ordre- Ustox : *x] -> (Caractérist) -> [Groupe : Acrustox]

[Horizon] -> (Caractérist) -> [Type : oxique]

[Horizon] -> (Caractérist) -> [Saturation-base : 0,015]

[Horizon] -> (Caractérist) -> [CRC : 1,5]

Type [Groupe- Eustrustox (x)] est

[Sous-ordre- Ustox : *x] -> (Caractérist) -> [Groupe : Eustrustox]

[Horizon] -> (Caractérist) -> [Type : oxique]

[Horizon] -> (Caractérist) -> [Saturation-base : 0,5]

Description des sous-groupes du groupe Eustrtox

Type [Sous-groupe-Eustrtox-tropéptique (x)] est

[Groupe- Eustrtox (*x)]-> (Caractérist) -> [Sous-groupe : Eustrtox-tropéptique]

[Horizon] -> (Caractérist) -> [Structure : Granulaire-fine]

Type [Sous-groupe-Eustrtox-typique (x)] est

[Groupe- Eustrtox *x] -> (Caractérist) -> [Sous-groupe : Eustrtox-typique]

[Horizon] -> (Caractérist) -> [Base : 1,25]

[Horizon] -> (Caractérist) -> [Texture : Argilo-sableuse]

[Horizon] -> (Caractérist) -> [Structure : !Granulaire-fine]

Type [Sous-ordre-Humox (x)] est

[Primitif : *x] -> (Caractérist) -> [Sous-ordre : Humox]

[Primitif : *x] -> (Caractérist) -> [Moyenne-organique : 16]

[Horizon] -> (Caractérist) -> [Type : oxique]

Type [Sous-ordre-Torrox (x)] est

[Primitif : *x] -> (Caractérist) -> [Sous-ordre : Torrox]

[Primitif : *x] -> (Caractérist) -> [Régime : Torrique]

[Horizon] -> (Caractérist) -> [Type : oxique]

La taxinomie représentant les classes décrites ci-dessus est illustrée par la figure 6.5.

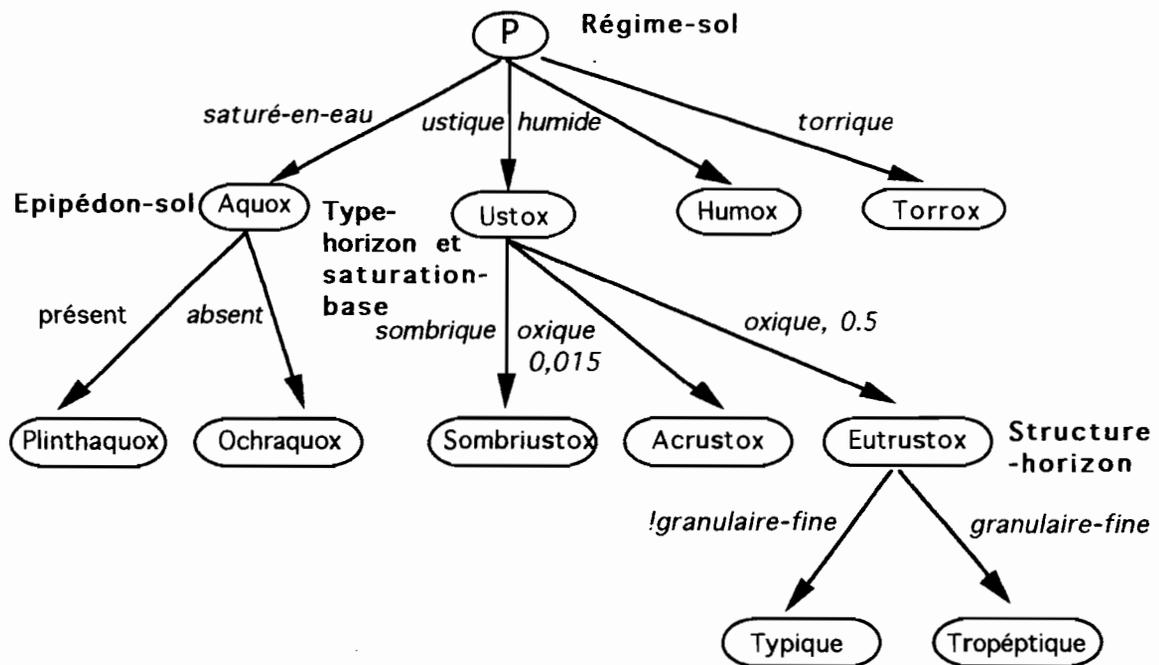


Figure 6.5. Taxinomie partielle de l'ordre des oxisols

6.3. CONCLUSION

Nous avons décrit dans ce chapitre les différents éléments de connaissance et leur organisation dans METIS. Nous avons introduit deux nouvelles composantes dans le canon : le treillis de composition et le C-modèle. En effet, METIS traite des objets composites, de plus il doit assurer la cohérence des connaissances explicitées et acquises.

La mise en œuvre des opérations manipulées a mis en évidence la nécessité d'introduire de nouvelles règles de formation canonique non définies par Sowa et qui répondent à nos besoins telles que l'extension de référents, la réduction de référents ensemblistes, la déréférenciation et la H-restriction. Il est à remarquer que nous avons remplacé par la règle de H-restriction, la règle de restriction définie par Sowa car nous l'avons jugée insuffisante dans notre cas. Par ailleurs, nous avons indiqué que la règle de jointure revient en fait à faire une jointure d'ensemble dans notre cas (voir § 6.1.11). Nous avons également mentionné le fait que la contraction et l'expansion de type sont symétriques.

Enfin, le dernier aspect concerne l'expérimentation de METIS dans deux nouveaux domaines des sciences expérimentales la minéralogie et la pédologie. Ceci montre que la

méthode METIS est applicable à d'autres domaines que l'ichtyologie et donc facilement généralisable²⁴.

²⁴ METIS n'est applicable qu'aux domaines où il existe une taxinomie préalablement connue, (en général, elle est exprimée dans des catalogues), c'est à dire que l'expert est en mesure de classer toutes les connaissances. De plus, il faut que la taxinomie comporte un nombre limité de classes et de niveaux.

7. CONCLUSION GÉNÉRALE

Nous avons présenté dans cette thèse METIS, un système interactif qui aide les taxinomistes à redécrire les taxons de manière cohérente et homogène, à construire progressivement un langage de description d'objets structurés qui soit suffisamment riche pour assurer la description et la discrimination des taxons, et enfin à représenter les connaissances qui permettent la comparaison des taxons entre eux et leur identification. De plus, le système réalise des diagnostics beaucoup plus pertinentes que celles qui existent dans les catalogues, permettant ainsi l'aide à la détermination des espèces que les taxinomistes collectent.

L'interaction avec l'expert permet l'élaboration de la base de connaissances, par évolution progressive du modèle structurel et des deux structures hiérarchiques (base de graphes et taxinomie des classes). La description et la caractérisation des taxons utilise la structure classificatoire du domaine, comme fil conducteur. Partant d'un noyau de description générique choisi avec l'expert, et constituant le point de départ pour le processus d'explicitation incrémental, METIS aide à modéliser petit à petit le domaine d'application. L'obtention du modèle descriptif est une étape fondamentale dans le processus cognitif qui permet de passer de l'observation à la description.

L'explicitation des connaissances se fait par étapes selon l'organisation hiérarchique des connaissances. Le mécanisme d'héritage permet une économie d'espace mémoire, une restriction de l'espace de recherche et donc une diminution du nombre de questions à poser à l'expert. De plus, le nombre de questions diminue quand METIS a explicité un certain nombre de taxons¹ ce qui conduit petit à petit à une stabilité du modèle structurel.

Le langage de description n'est pas figé et évolue au fur et à mesure que l'on considère un nouveau taxon ou que l'on spécialise la description de ceux qui existent déjà. Une fois construit, il est suffisamment concis (pas de redondances) pour ne contenir que les informations pertinentes et suffisamment riche pour décrire tous les taxons considérés.

¹ Le nombre de taxons doit être suffisamment grand et les descriptions suffisamment représentatives.

La représentation des connaissances à l'aide du formalisme des graphes conceptuels permet une meilleure lisibilité que celle du formalisme logique, une bonne visualisation graphique de la description des objets, une richesse d'expression sémantique, une représentation "claire" de la structuration des connaissances et enfin, de nombreuses facilités de manipulation dues à la souplesse des opérations définies sur les graphes conceptuels. La restriction imposée sur les graphes conceptuels pris en compte (concepts n'apparaissant qu'une fois et relations binaires) permet de réduire la complexité du processus de généralisation. Par ailleurs, la mise en œuvre des opérations manipulées a mis en évidence la nécessité d'introduire de nouvelles règles de formation canoniques (voir § 6.1.12) non définies par Sowa et qui répondent aux besoins du système. De même le canon a été enrichi par deux nouveaux composants : le treillis de composition et le C-modèle.

La procédure d'apprentissage IDICE (voir § 5.6.2.1.) utilise un algorithme de type ID3 afin d'apprendre de nouveaux traits discriminants. IDICE recherche à chaque fois l'ensemble des traits maximale discriminant afin de le suggérer à l'expert. Contrairement à ID3, IDICE prend en compte des descriptions incomplètes et permet à l'expert, d'imposer des contraintes quant au choix et l'utilisation de certains traits. Un traitement spécial permet de sélectionner les traits multivalués et les traits inapplicables quand ils ne posent plus de problèmes. Par ailleurs, l'étape de complétion d'exemples favorise la recherche de traits potentiellement discriminants et homogénéise dans la mesure du possible la description des taxons. L'expert est seul juge quant à la pertinence des traits : METIS suggère mais ne décide pas à la place de l'expert.

La cohérence statique des informations entrées dans le système est assurée par la prise en compte d'un C-modèle (voir § 5.7.3.) construit lui aussi de manière incrémentale. METIS s'y réfère en permanence pour valider interactivement ses connaissances. La validation se fait grâce à un dialogue avec l'expert qui se trouve ainsi confronté à ses propres connaissances. Ainsi METIS vérifie à chaque fois l'ensemble de valeurs autorisées pour un concept donné, son degré, l'ensemble de ses synonymes, l'ensemble de valeurs simultanément contradictoires, les attributs respectivement exclusifs, etc.

La non-redondance des informations de la base est assurée grâce d'une part, à la prise en compte des synonymes ce qui évite la duplication d'informations et d'autre part, au critère de factorisation maximale défini en (§ 5.7.3.) qui garantit qu'à tout moment, n'importe quel trait est placé au "bon niveau" de généralisation. Ce critère utilise une

matrice de correspondance appelée matrice de traits qui sert à mettre en correspondance les traits et les classes décrites par ces traits.

Enfin, l'expérimentation de METIS dans trois domaines : l'ichtyologie, la minéralogie et la pédologie, montre que la méthode METIS est facilement généralisable. Toutefois, l'expérimentation en minéralogie a été plus facile à réaliser que celles en ichthyologie ou en pédologie. En effet en minéralogie la plupart des concepts sont descriptifs, il y a très peu de concepts structurels. De plus, les concepts descriptifs considérés dans les descriptions, sont à peu de choses près, les mêmes. L'expérimentation en pédologie a montré qu'il est difficile de caractériser les sols surtout du fait qu'il n'existe actuellement aucun consensus international sur une taxinomie de sol donné. Quant à l'expérimentation en ichthyologie elle a été, de loin, la plus difficile ; les descriptions sont ambiguës (voire vagues), et non homogènes, les exceptions sont nombreuses, les caractères ne sont pas applicables tout le temps, et il y a un nombre important de concepts structurels. De plus, selon les familles ces concepts structurels ne sont pas décrits par les mêmes concepts descriptifs.

Discussion

L'originalité de notre approche consiste à mettre à contribution une technique d'induction (IDICE) pour l'explicitation des connaissances. En effet, nous avons adapté l'algorithme ID3 afin qu'il tienne compte d'un langage de descriptions structurées. La complexité d'IDICE est relativement faible, elle est d'ordre polynomial en nombre moyen de sous classes et de traits par nœud de la taxinomie.

L'heuristique permettant de choisir un trait discriminant dans IDICE considère que la présence ou l'absence d'un composant est plus pertinente pour discriminer des classes que les propriétés intrinsèques relatives à d'autres composants. C'est-à-dire que METIS privilégie les traits structurels par rapport aux traits descriptifs. Cette heuristique repose sur le fait qu'un trait structurel (présent ou absent) est plus facilement observable qu'un trait descriptif. De plus, METIS favorise un trait structurel de haut niveau par rapport à un trait structurel de plus bas niveau, car il nous semble naturel dans le cas de descriptions structurées de nous intéresser d'abord aux concepts composites avant d'examiner leurs composants.

Actuellement, IDICE ne prend pas en compte des traits discriminants disjonctifs. En effet, elle n'élimine pas les attributs multivalués de la description des classes, mais ne les utilise que pour discriminer des sous-ensembles de classes où ces attributs ne prennent qu'une seule valeur. C'est une utilisation en différé. Cependant, nous pensons

que l'intégration d'une procédure traitant des traits discriminants disjonctifs dans METIS reste à réaliser.

METIS met à contribution des taxinomies naturelles existantes possédant une sémantique riche, afin de construire des caractérisations de classes. Il ne s'applique donc qu'à des domaines où il existe une taxinomie préalablement connue.

Perspectives

Plusieurs directions de recherches et prolongements de nos travaux découlent de notre approche. Il serait ainsi possible :

- d'adjoindre à METIS un algorithme d'identification qui utilisera les descriptions contenues dans la base de graphes. En effet, en phase d'utilisation le système peut servir à identifier la classe d'un spécimen par interaction avec un nouvel utilisateur. Si l'utilisateur ne sait pas répondre à une question associée à un nœud de la structure de discrimination, il pourra répondre aux questions relatives à la description intrinsèque des nœuds fils dans la base de graphes. Ceci évitera le blocage (dû au déterminisme) et permettra de revenir à la structure de discrimination en ayant fait un pas d'inférence dans le processus d'identification,
- d'utiliser du raisonnement à partir de cas pour la reconnaissance de spécimens. Il s'agirait de considérer la taxinomie comme étant une mémoire de classes² de cas. Par conséquent, quand un nouveau cas (description de spécimen) arrive, le système essaie de trouver la classe de cas à laquelle pourrait appartenir le nouveau cas en parcourant la structure de discrimination,
- de proposer à l'expert de réviser la taxinomie du domaine afin de mettre en évidence de nouveaux regroupements, ou au contraire de diviser des groupes déjà formés. On pourrait par exemple, envisager de créer entre les classes *naturelles* de nouvelles classes appelées classes *artéfactuelles*.³

² La classe de cas ne contient pas des descriptions de cas distinctes mais une généralisation des descriptions de cas pouvant appartenir à cette classe. C'est une description en intension des cas.

³ En fait, ce serait des descriptions intermédiaires que l'on conviendra d'appeler classes *artéfactuelles*.

- de concevoir un algorithme d'apprentissage qui recherche toutes les corrélations que l'on peut observer entre les traits décrivant les classes. Ainsi, METIS constituerait un outil supplémentaire d'exploitation de la taxinomie,
- de doter METIS d'un module d'explication qui permettrait à l'expert de souligner ses hésitations et les causes qui l'ont conduit à opter pour une solution plutôt que pour une autre,
- d'organiser METIS selon le modèle client-serveur (architecture réseau) qui permettrait de prendre en compte les réponses de plusieurs experts. Il s'agira alors, de faire de la "*multi-explicitation*" de connaissances, afin de construire une taxinomie consensuelle. Le problème qui se posera alors sera de résoudre les conflits d'accès qui peuvent survenir.

Par ailleurs, si l'acquisition des connaissances, vise à extraire et à modéliser le savoir-faire des experts afin de construire des bases de connaissances qui assureront la pérennité d'un savoir-faire, et qui permettront aux générations futures d'en bénéficier et de le fructifier, il serait pertinent de garder une trace des sources qui ont permis leurs construction et de gérer l'historique des principales modifications subies.

En outre, il est d'usage en sciences naturelles, que les informations soient stockées dans des catalogues sous forme de textes et d'images, il serait opportun qu'à terme on puisse conserver ces catalogues à l'aide d'*hypermédia* [TREUIL & al, 94] afin que l'on puisse d'une part, allier l'image et le texte plus facilement et d'autre part, accéder plus rapidement aux informations qui nous intéressent.

Enfin, il serait possible d'étendre le formalisme de représentation des connaissances de METIS à d'autres formalismes afin de représenter des informations de types divers. Car les taxinomies que nous avons pris en compte sont basées essentiellement sur des données morphologiques⁴, mais il serait intéressant de prendre aussi en compte des taxinomies basées sur des données cytologiques, biochimiques ou génétiques afin de tenir compte de l'évolution des techniques employées par les taxinomistes...

⁴ Pendant longtemps, par nécessité puis par habitude, les systématiciens se sont limités à utiliser des caractères faciles à observer. Ils s'adressaient essentiellement à la morphologie externe et à l'anatomie osseuse. Actuellement, si les mêmes normes sont utilisées pour la description des espèces nouvelles afin de faciliter les comparaisons avec les diagnoses anciennes, les spécialistes sont d'accord pour admettre que les critères distinctifs les plus valables sont souvent de nature très différente.

RÉFÉRENCES

- [**ABBOTT & al, 85**] L.A. ABBOTT, F.A. BISBY, D.J. ROGERS
"Taxonomic Analysis in Biology"
Computers, Models and Databases. Columbia University Press, New York, 1985.
- [**AGUIRRE, 89**] J.L. AGUIRRE CERVANTES
"Construction Automatique de Taxonomies à partir d'Exemples dans un Modèle de Connaissances par Objets"
Thèse de docteur de l'Institut National Polytechnique de Grenoble, 1989.
- [**AIMEUR, 91**] E. AIMEUR.
"Application d'un Système d'Apprentissage Empirique au Décodage Acoustico-Phonétique"
Colloque de Taxinomie Informatisée, UNESCO, Paris, France, 1991.
- [**AIMEUR, 93a**] E. AIMEUR
"Une méthode d'Acquisition des Connaissances Destinée à la Réalisation de Diagnoses"
Journées VOLCAN IA, pp 338-352, Clermont-Ferrand, France, 1993.
- [**AIMEUR, 93b**] E. AIMEUR
"Acquisition d'Objets Structurés par Caractérisation"
JAVA: Journées Françaises d'Acquisition de Validation et d'Apprentissage, Saint Raphaël, France, 1993.
- [**AIMEUR, 93c**] E. AIMEUR
"Apport des Graphes Conceptuels à l'Acquisition des Connaissances"
INFORSID, pp 563-582, Lille, France, 1993.
- [**AIMEUR & GANASCIA 93a**] E. AIMEUR, J.G. GANASCIA
"Acquisition of Discriminating Features of a Classification"
Proceedings of the Thirteenth International Conference on Artificial Intelligence Expert Systems and Natural Language, pp 199-210, Avignon, France, 1993.
- [**AIMEUR & GANASCIA 93b**] E. AIMEUR, J.G. GANASCIA
"Elicitation of Taxonomies Based on the Use of Conceptual Graph Operators"
ICCS'93, Proceedings of the First International Conference on Conceptual Structures: Conceptual Graphs for Knowledge Representation. Lecture Notes in Artificial Intelligence, nb 699, pp 361-380, G.W. MINEAU, B. MOULIN, J.F. SOWA Editors, Québec City, Canada, Springer Verlag, 1993.
- [**AIMEUR & GANASCIA 93c**] E. AIMEUR, J.G. GANASCIA
"Reasoning with Classification in Interactive Knowledge Elicitation"
Proceedings of the Fourth International Workshop on Classification Research
American Society for Information Science, pp 19-38, Columbus, Ohio, USA, 1993.
- [**AIMEUR & KIEU, 93**] E. AIMEUR, Q. KIEU
"A Method of Incremental Acquisition of Structured Objects by Discrimination :
Application to Organisms' Biology"
Knowledge & Data Engineering Workshop, pp 169-182, Strasbourg, France, 1993.
- [**AIMEUR & al, 93**] E. AIMEUR, Q. KIEU, J.G. GANASCIA
"Acquisition of Structured Descriptions with Sowa's Graphs"
ICO'93: 4th International Conference on Cognitive and Computer Sciences for Organizations, pp 263-271, Montréal (Québec), Canada, 1993.

- [**ANDERSON & BOWER, 73**] J.R. ANDERSON, G. BOWER
"Human Associative Memory"
 Washington D.C., H.V. Winston and Co Publishers, 1973.
- [**ANQUETIL & VAUCHER, 93**] N. ANQUETIL, J. VAUCHER
"Acquisition et Classification de Concepts pour la Réutilisation"
 ICO: 4th International Conference on Cognitive and Computer Sciences for Organizations, pp 463-472,
 Montréal (Québec), Canada, 1993.
- [**ANTONACCI & al, 88**] F. ANTONACCI, M.T. PAZIENZA, M. RUSSO, P. VELARDI
"Representation and Control Strategies for Large Knowledge Domains: An Application to NLP"
 Applied Artificial Intelligence 2 (3/4), pp 213-249, 1988.
- [**AUROI, 92**] C. AUROI
"La Diversité Biologique, la vie en péril"
 Collection : "Dossier pour l'environnement", vol 7, Georg Editeur, 1992.
- [**AUSSENAC, 89**] N. AUSSENAC
"Conception d'une méthodologie et d'un outil d'acquisition de connaissances expertes"
 Thèse de l'université de Paul Sabatier, Toulouse, Novembre 1989.
- [**AYEL & ROUSSET, 90**] M. AYEL, M.C. ROUSSET
"La cohérence dans les bases de connaissances"
 Cepadues-Editions, 1990.
- [**BERARD-DUGOURD & al, 88**] A. BERARD-DUGOURD, J. FARGUES, M.C. LANDAU
"Natural Language Analysis Using Conceptual Graphs"
 Proceedings Of the Third Annual Workshop on Conceptual Graphs, St Paul, Minnesota, 1988.
- [**BETEKHTIN, 75**] A. BETEKHTIN
"A Course of Mineralogy"
 Moscow Peace Publishers, 1975.
- [**BIEBOW, 92**] B. BIEBOW
"Comparaison entre les Graphes Conceptuels et un Noyau KL-ONE"
 Actes des 4èmes Journées Nationales du PRC-IA *Formalisation et Extensions des Graphes Conceptuels*,
 pp 99-109, Marseille, 1992.
- [**BIEBOW & CHATY, 93**] B. BIEBOW, G. CHATY
"A comparison between Conceptual Graphs and KL-ONE"
 ICCS'93, Proceedings of the First International Conference on Conceptual Structures: Conceptual Graphs
 for Knowledge Representation. Lecture Notes in Artificial Intelligence, nb 699, pp 75-89, G.W.
 MINEAU, B. MOULIN, J.F. SOWA Editors, Québec City, Canada, Springer Verlag, 1993.
- [**BISSON, 93**] G. BISSON
"KBG: Induction de Bases de Connaissances en Logique des prédicats"
 Thèse de docteur d'université, Paris XI, 1993.
- [**BOOSE, 84**] J. BOOSE
"Personal Theory and the Transfer of Human Expertise"
 Proc 3th AAAI, 1984.
- [**BOOSE, 89**] J.H. BOOSE
"A Survey of Knowledge Acquisition Techniques and Tools"
 Knowledge Acquisition vol 1 no 1, pp 3-38, Academic Press, 1989.
- [**BOOSE & BRADSHAW, 93**] J.H. BOOSE, J.M. BRADSHAW
*"Expertise Transfer and Complex Problems: using AQUINAS as a Knowledge-Acquisition Workbench
 for Knowledge-Based Systems"*
 Readings in Knowledge Acquisition and Learning, pp 240-252, Edited by B.G. Buchanan & D.C.
 Wilkins, Morgan Kaufmann Publishers, 1993.

- [**BOOSE & al, 89**] J.H. BOOSE, J.M. BRADSHAW, C.M. KITTO, D.B. SHEMA
"From ETS to AQUINAS: Six years of Knowledge Acquisition Tool Development"
 Proceedings of the Third European Knowledge Acquisition Workshop, pp 502-515, Paris, 1989.
- [**BOOSE & al, 93**] J.H. BOOSE, J.M. BRADSHAW, J.L. KOSZAREK, D.B. SHEMA
"Knowledge Acquisition Techniques for Group Decision Support"
 Knowledge Acquisition vol 5, pp 405-448, Academic Press, 1993.
- [**BOUROCHE & SAPORTA, 87**] J.M. BOUROCHE, G. SAPORTA
"L'analyse des données"
 Presses Universitaires de France, Que-sais-je, no 1854, Paris, 1987.
- [**BOY, 91**] G. BOY
"Intelligent Assistant System"
 Knowledge Based Systems vol 6, Academic Press, London, 1991.
- [**BOY & RAPPAPORT, 87**] G. BOY, A. RAPPAPORT
"Incremental Knowledge Acquisition for Assistance to Space Robot Arm Manipulation"
 In ROBEX'87, Pittsburg, PA, USA, 1987.
- [**BOY & al, 88**] G. BOY, B. FALLER, J. SALLANTIN
"Acquisition et ratification des connaissances"
 Actes des journées nationales du PRC-GRECO, Intelligence Artificielle, pp 321-356, Toulouse, France, 1988.
- [**BRACHMAN, 77**] R.J. BRACHMAN
"What's in a Concept: Structural Foundations for Semantic Networks"
 International Journal of Man-Machine studies, vol 9, pp 127-152, 1977.
- [**BRACHMAN, 79**] R.J. BRACHMAN
"On the Epistemological Status of Semantic Networks"
 In [FINDLER, 79], pp 3-50, 1979.
- [**BRACHMAN & SCHMOLZE, 85**] R.J. BRACHMAN, J.G. SCHMOLZE
"An Overview of the KL-ONE Knowledge Representation System"
 Cognitive Science vol 9, pp 171-216, 1985.
- [**BRACHMAN & al, 91**] R.J. BRACHMAN, D.L. McGUINNESS, P.F. PATEL-SCHNIEDER,
 L.A. RESNICK, A. BORGIDA
"Living with CLASSIC: When and How to Use a KL-ONE Language"
 In J.F. SOWA Editor, pp 401-456, Morgan Kaufmann Publishers, San Mateo, California, 1991.
- [**BURTON & al, 87**] A.M. BURTON, N.R. SHADBOLT, A.P. HEDGECOCK, G. RUGG
"A Formal Evaluation of Elicitation Techniques for Expert Systems : domain 1"
 Moralee D.S. Ed, Research and Developments in Expert Systems, Cambridge University Press, 1987.
- [**CASAI, 92**] E. CASAI
"An Incremental Class Reorganization Approach"
 ECOOP'92, Lecture Notes in Computer Science no 615, pp 114-132, O. LEHRMANN MADSEN (Ed),
 Springer Verlag, 1992.
- [**CAPPONI, 93**] C. CAPPONI
"Classification des Classes par les Types"
 Actes des Journées Représentation par Objets, La Grande Motte, pp 215-224, 1993.
- [**CELEUX & al, 89**] G. CELEUX, E. DIDAY, G. GOVAERT, Y. LECHEVALLIER, H.
 RALAMBONDRAINY
"Classification automatique des données"
 Dunod informatique, 1989.

- [CHANDRASEKARAN, 87] B. CHANDRASEKARAN
"Towards a Functional Architecture for Intelligence Based on Generic Information Processing Tasks"
 In Proceedings of the 10th International Joint Conference on Artificial Intelligence, pp 1183-1192,
 Milan, Italy, 1987.
- [CHARLET & al, 94] J. CHARLET B. BACHIMONT, J. BOUAUD, P. ZWEIGENBAUM
"Ontologie et réutilisabilité : expérience et discussion"
 Actes des Sièmes Journées Françaises de l'acquisition des connaissances, Strasbourg, 1994.
- [CHATELIN, 79] Y. CHATELIN
"Une épistémologie des sciences du sol"
 Mémoire ORSTOM, No 88, France, 1979.
- [CHEIN, 92] M. CHEIN
"Un Noyau pour les Graphes Conceptuels Simples"
 Actes des 4èmes Journées Nationales du PRC-IA *Formalisation et Extensions des Graphes Conceptuels*,
 pp 60-72 , Marseille, 1992.
- [CLANCEY, 85] W. CLANCEY
"Heuristic Classification"
 Artificial Intelligence Journal, 27, pp 289-350, 1985.
- [CLANCEY, 89a] W. CLANCEY
"The knowledge Level Reinterpreted: Modeling How Systems Interact"
 Machine learning, 4, pp 285-291, 1989.
- [CLANCEY, 89b] W. CLANCEY
"Viewing Knowledge Bases as Qualitative Models"
 IEEE Expert, pp 9-23, 1989.
- [COLLINS & QUILLIAN, 69] A. COLLINS, R. QUILLIAN
"Retrieval time from semantic memory"
 Journal of Verbal Learning and Verbal Behaviour, vol IX, pp 240-247, 1969.
- [CONRUYT & al, 93] N. CONRUYT, M. MANAGO, J. LE RENARD, C. LEVI
"Une méthode d'acquisition de connaissances pour la classification et l'identification d'objets biologiques"
 Proceedings of the Thirteenth International Conference on Artificial Intelligence Expert Systems and
 Natural Language, pp 485-495, Avignon, France, 1993.
- [CORI & LASCAR, 93] R. CORI, D. LASCAR
"Logique mathématique"
 Masson, vol 1, Paris, 1993.
- [CRAUSER & al, 89] J.P. CRAUSER, Y. HARVATOPOULOS, P. SARNIN
"Guide pratique d'analyse des données"
 Les éditions d'organisation, 1989.
- [DALLWITZ, 74] M. DALLWITZ
"A Flexible Computer Program for Generating Diagnostic Keys"
 Syst. Zoology, vol 23, pp 50-57, 1974.
- [DALLWITZ, 89] M. DALLWITZ
"Diagnostic descriptions from INTKEY and CONFOR"
 DELTA News-letter, no 3, pp 8-13, 1989.
- [DECAESTECKER, 91] C. DECAESTECKER
"Apprentissage en Classification Conceptuelle Incrémentale"
 Thèse de docteur de l'université libre de Bruxelles, 1991.
- [DELAHAYE, 86] J.P. DELHAYE
"Outils Logiques pour l'Intelligence Artificielle"
 Editions Eyrolles, 1986.

- [DETIENNE & VERNANT, 74] M. DETIENNE, J.P. VERNANT
"Les ruses de l'intelligence : la MÉTIS des grecs"
 Flammarion, 1974.
- [DIDAY & al, 80] E. DIDAY et collaborateurs
"Optimisation en classification automatique"
 INRIA Editeur, 1980.
- [DIEDERICH & al, 88] J. DIEDERICH, I. RUHMANN, M. MAY
"KRITON: a Knowledge-Acquisition Tool for Expert Systems"
 Knowledge Acquisition Tools for Expert System, Knowledge-Based Systems, vol 2,
 pp 83-94, Edited by J.H. BOOSE and B. GAINES, Academic Press, 1988.
- [DIENG, 92] R. DIENG
"Méthodes et outils d'acquisition des connaissances"
 Dans J.C. Sperandio Editeur : *"L'ergonomie dans la conception des projets informatiques"*, 1992.
- [DOMINE, 88] C.H. DOMINE
"Techniques de l'intelligence artificielle, un guide structuré"
 Dunod informatique, 1988.
- [DUDA & REJL, 86] R. DUDA, L. REJL
"La grande encyclopédie des minéraux"
 Gründ, Paris, 1986.
- [ELLIS, 91] G. ELLIS
"Compiled Hierarchical Retrieval"
 Proceedings of the 6th Annual workshop on Conceptual Graphs
 pp 187-207, Binghamton, USA, 1991.
- [ERICSSON & SIMON, 84] K.A. ERICSSON, H.A. SIMON
"Protocol Analysis: Verbal Reports as Data"
 Cambridge, MA, MIT Press, 1984.
- [ERMINE, 93] J.L. ERMINE
"Génie logiciel & Génie cognitif pour les systèmes à base de connaissances"
 Volume 2 : Étude de cas. Technique et Documentation, Lavoisier, 1993.
- [ESCH & al, 89] J.W. ESCH, T.E. NAGLE, M.L. YIM, L.L. GERHLZ
"Polymorphism in the Theory of Conceptual Graphs"
 Proceedings of the 4th Annual workshop on Conceptual Graphs
 pp 223-232, Detroit, Michigan, 1989.
- [FARGUES, 92] J. FARGUES
"Les référents ensemblistes"
 Actes des 4èmes Journées Nationales du PRC-IA Formalisation et Extensions des Graphes Conceptuels,
 pp 83-89 , Marseille, 1992.
- [FARON, 93] C. FARON
"Acquisition d'une classification : la représentation et la discrimination des classes"
 Mémoire de DEA, Laforia, Université Paris VI, 1993.
- [FINDLER, 79] N.V. FINDLER
"Associative Networks: Representation and Use of Knowledge by Computers"
 Academic Press, Orlando, Florida, 1979.
- [FININ, 86] T. FININ
"Interactive Classification : A Technique for Acquiring and Maintaining Knowledge Bases"
 Proceedings of The IEEE-COMPUTER, vol 74, no 10, pp 1414-1421, 1986.

- [FIRLEJ & HELLENS, 91]** M. FIRLEJ, D. HELLENS
"Knowledge Elicitation : a Practical Book"
 Prentice Hall International (UK) Ltd, 1991.
- [FISHER, 87]** D.H. FISHER
"Knowledge Acquisition Via Conceptual Clustering"
 Machine Learning, vol 2, pp 139-172, 1987.
- [FORTUNER, 86]** R. FORTUNER
"A Better Assessment of Variability of Qualitative Characters for the Computer Identification Program Nemaïd"
 Revue Nématol, vol 9, no 3, pp 277-279, 1986.
- [FORTUNER, 93a]** R. FORTUNER
"The NEMISYS Solution to Problems in Nematode Identification"
 in [Fortuner, 93b], pp 137-164, 1993.
- [FORTUNER, 93b]** R. FORTUNER
"Advances in Computer Methods for Systematic Biology: Artificial Intelligence, Databases, Computer Vision"
 Edited by R. Fortuner, The John Hopkins University Press, 1993.
- [GAINES, 90]** B. GAINES
"Knowledge Acquisition Systems"
 Knowledge Engineering, vol 1, Fundamentals, pp 52-102, Edited by Hojjat Adeli, McGraw-Hill Publishing Company, 1990.
- [GALLOUIN, 88]** J.F. GALLOUIN. *"Transfert de Connaissances, Systèmes Experts : Techniques et Méthodes"*, Editions Eyrolles, 1988.
- [GANASCIA, 87]** J.G. GANASCIA
"AGAPE & CHARADE: deux techniques d'apprentissage symbolique appliquées à la construction de base de connaissances"
 Thèse d'état, université de Paris Sud, 1987.
- [GANASCIA, 90]** J.G. GANASCIA
"L'âme machine les enjeux de l'intelligence artificielle"
 Editions SEUIL, collection Sciences ouvertes, 1990.
- [GARDINER & al, 89]** D.A. GARDINER, B. TJAN, J.R. SLAGLE
"Extended Conceptual Structures Notation"
 Proceedings of the 4th Annual workshop on Conceptual Graphs
 pp 211-222, Detroit, Michigan, 1989.
- [GARNER & TSUI, 89]** B.J. GARNER, E. TSUI
"Knowledge Acquisition and Reasoning with a Canonical Graph Model in Personal Financial Planning"
 In Expert Systems in Economic Banking and Management, Edited by L.F. PAU, Y.H. PAOJ. MOTIWALA and H.H. TEH, North-Holland, Amsterdam, 1989.
- [GARNER & al, 89]** B.J. GARNER, K.E. LARKING, E. TSUI
"Prototypical knowledge for Case-Based Reasoning"
 Proceedings DARPA, Workshop on Case-Based Reasoning, Pensacola Beach, Florida, 1989.
- [GENNARI & al, 89]** J. GENNARI, P. LANGLEY, D. FISHER
"Models of Incremental Concept Formation"
 In Artificial Intelligence vol 40, pp 11-61, Elsevier Science, 1989.
- [GEY, 91]** O. GEY
"COCLUSH: Un Générateur de Classifications d'Objets Structurés, Suivant Différents Points de Vue"
 Actes des 6ièmes Journées Françaises de l'apprentissage, pp 165-184, Sete, 1991.

- [GEY, 94]** O. GEY
 "Saturation et généralisation de graphes conceptuels"
 Actes des 9ièmes Journées Françaises de l'apprentissage, Strasbourg, 1994.
- [GRUBER, 93]** T.R. GRUBER
 "A Translation Approach to Portable Ontology Specifications"
 Knowledge Acquisition vol 5, pp 199-220, 1993.
- [GRUBER, 90]** T.R. GRUBER
 "Acquiring strategic knowledge from experts"
 The Foundations of Knowledge Acquisition, Knowledge-Based Systems, vol 4, pp 115-133, Academic Press, 1990.
- [GUERDOUD & LEBBE, 94]** M. GUERDOUD, J. LEBBE
 "Une structuration des connaissances adaptée à leur acquisition à partir de textes en Systématique"
 Actes des Sièmes Journées Françaises de l'acquisition des connaissances, Strasbourg, 1994.
- [GUY & STEELE, 90]** L. GUY, J.R. STEELE
 "Common Lisp. The Language"
 Digital Presse, Second Edition, 1990.
- [HAMROUN & CARRE, 94]** H. HAMROUN, B. CARRE
 "Construction d'un SBC pour l'orientation juridique. Une utilisation conjointe de KADS et KOD"
 Actes des Sièmes Journées Françaises de l'acquisition des connaissances, Strasbourg, 1994.
- [HANSON & BAUER, 89]** S. HANSON, M. BAUER
 "Conceptual Clustering, Categorization and Polymorphy"
 In Machine Learning, vol 3, no 4, pp 343-372, Kluwer Academic, 1989.
- [HARMON & KING, 85]** P. HARMON, D. KING
 "Expert System: Artificial Intelligence in Business"
 A. Wiley Press Book, J. Wiley & Sons Inc., 1985.
- [HATON & al, 91]** J.P. HATON, N. BOUZID, F. CHAPILLET, M.C. HATON, B. LAASRI, H. LAASRI, P. MARQUIS, T. MONDOT, A. NAPOLI
 "Le Raisonnement en Intelligence Artificielle"
 InterEditions, Paris, 1991.
- [HENNING, 49]** W. HENNING
 "Zur Klärung einiger Begriffe der phylogenetischen Systematik"
 Forschungen und Fortschritte, pp 136-138, 1949.
- [HERNERT, 93]** P. HERNERT
 "Un système d'acquisition de définitions basé sur le modèle des graphes conceptuels"
 Thèse de docteur d'université, Paris XIII, 1993.
- [HULL, 86]** D. HULL
 "Les Fondements Epistémologiques de la Classification Biologique"
 In [TASSY, 86], pp 161-203, 1986.
- [KABBAJ & FRASSON, 93a]** A. KABBAJ, C. FRASSON
 "Un Algorithme de Formation Dynamique de la Mémoire"
 ICO'93 : 4th International Conference on Cognitive and Computer Sciences for Organizations, pp 473-484, Montréal (Québec), Canada, 1993.
- [KABBAJ & FRASSON, 93b]** A. KABBAJ, C. FRASSON
 "Représentation et Acquisition des Connaissances dans les Systèmes Tutoriels Intelligents"
 Actes des Journées ACTI' 93, Limoges, France, pp 77-99, 1993.

- [KAYSER, 84] D. KAYSER
"Examen de diverses méthodes utilisées en représentation des connaissances"
 Actes du 4ème Congrès AFCET, Reconnaissances des formes et Intelligence artificielle, RFIA '84, Paris, pp 115-144, 1984.
- [KAYSER, 91] D. KAYSER
"La Représentation des Connaissances en Intelligence Artificielle"
 Les Sciences Cognitives en Débat, pp 89-105, Editions du CNRS, Paris, 1991.
- [KEENE, 89] S.E. KEENE
"Object-Oriented Programming in Common Lisp. A Programmer's Guide to CLOS"
 Addison-Wesley Publishing Company, 1989.
- [KELLY, 55] G.A. KELLY
"The Psychology of Personal Constructs"
 New York: Norton, 1955.
- [KOLODNER & al, 85] J.L. KOLODNER, R.L. SIMPSON, K. SYCARA-CYRANSKI
"A Process of Case-Based Reasoning in Problem Solving"
 Proceedings of the 9th IJCAI, pp 284-290, Los Angeles, California, 1985.
- [KRIVINE & DAVID, 91], J.P. KRIVINE, J.M. DAVID
"L'acquisition des connaissances vue comme un processus de modélisation : méthodes et outils"
 Intellectica, 12, pp 101-138, 1991/2.
- [LAKOFF, 87] G. LAKOFF
"Women, Fire, and Dangerous Things (What Categories Reveal about the mind)"
 The University of Chicago Press, 1987.
- [LAUBLET & al, 91] P. LAUBLET, M. MEPHAM, E. MOTTA, M. ZACKLAD, M. EISENSTADT, M. GEERDERS, G. RUGG, K. O'HARA
"ID211.4 Review of Software Mediated Knowledge Acquisition"
 Rapport Interne T211.4/11, ONERA, Projet VITAL, 1991.
- [LEBBE, 84] J. LEBBE
"Manuel d'utilisation du logiciel XPER"
 Micro Application, Rueil Malmaison, 1984.
- [LEBBE, 91] J. LEBBE
"Représentation des concepts en biologie et en médecine : introduction à l'analyse des connaissances et à l'identification assistée par ordinateur"
 Thèse de docteur d'université, Paris VI, 1991.
- [LEBBE & al, 89] J. LEBBE, R. VIGNES, J.P. DEDET
"Computer-aided Identification of Insect Vectors"
 Parasitology Today, vol 5, No 9, Elsevier Science Publishers, 1989.
- [LEBBE & VIGNES, 91] J. LEBBE, R. VIGNES
"Génération de graphes d'identification à partir de descriptions de concepts"
 Induction Symbolique Numérique à Partir de Données. Kodratoff Y., Diday E. Eds, Cepadues, Toulouse, pp 193-239, 1991.
- [LEBOWITZ, 87] M. LEBOWITZ
"Experiments with Incremental Concept Formation: UNIMEM"
 Machine Learning, vol 2, pp 103-138, 1987.
- [LEROUX & LAUBLET, 93a] B. LEROUX, P. LAUBLET
"A Conceptual and Compositional Approach or Knowledge Acquisition : examples & lessons"
 Rapport interne no 93/06, Laforia, université Paris VI, 1993.
- [LEROUX & LAUBLET, 93b] B. LEROUX, P. LAUBLET
"Modèles de Méthodes de Résolution de Problèmes et Acquisition des Connaissances"
 JAVA: Journées Françaises d'Acquisition de Validation et d'Apprentissage, Saint Raphaël, France, 1993.

- [LEVEQUE & al, 91] C. LEVEQUE, D. PAUGY, G.G. TEUGELS
"Faune des poissons d'eaux douces et saumâtres de l'Afrique de l'Ouest"
 ORSTOM / MRAC Editions, 1991.
- [LEVINSON, 85] R.A. LEVINSON
"A Self-Organizing Retrieval System for Graphs"
 Proc. 4th AAAI, Austin, Texas, pp 203-206, 1985.
- [LINSTER, 92] M. LINSTER
"L'Ingénierie des Connaissances : une Symbiose de deux Perspectives sur le Développement de Modèles"
 3èmes Journées d'Acquisition des Connaissances, pp 5-17, Dourdan, France, 1992.
- [LIQUIERE, 90] M. LIQUIERE
"Apprentissage à partir d'objets structurés. Conception et Réalisation"
 Thèse de docteur d'université, Montpellier II, 1990.
- [McDERMOTT, 88] J. McDERMOTT
"Preliminary Steps Towards a Taxonomy of problem-Solving Methods"
 In S. Marcus, Ed., *Automating Knowledge Acquisition for Expert Systems*. Kluwer Academic Publishers, 1988.
- [MacGREGOR, 91] R. MacGREGOR
"The Evolving Technology of Classification-based Knowledge Representation Systems"
 In *"Principles of Semantic Networks: explorations in the representation of knowledge"*
 J.F. SOWA Editor, pp 385-400, Morgan Kaufmann Publishers, San Mateo, California, 1991.
- [MAJOR, 91] N.P. MAJOR
"CATO An Automated Card Sort Tool"
 Proceedings of the Fifth European Knowledge Acquisition for Knowledge-Based Systems
 University of Strathclyde, 1991.
- [MAJOR & REICHGELT, 90] N.P. MAJOR, H. REICHGELT
"ALTO: An Automated Laddering Tool"
 In B. Wielinga, J. Boose, B. Gaines, G. Schreiber & M. Van Someren, Eds. *Current Trends in Knowledge Acquisition*, pp 222-236, Amsterdam: IOS Press, 1990.
- [MARCUS, 93] L. F. MARCUS
"The Goals and Methods of Systematic Biology"
 in [FORTUNER, 93b], pp 31-54, 1993.
- [MARCUS & al, 85] S. MARCUS J. McDERMOTT, T. WANG
"Knowledge Acquisition for Constructive Systems"
 Proceedings of the 9th Joint conference on Artificial Intelligence, pp 637, 639, Los Angeles, 1985.
- [MARCUS & McDERMOTT, 93] S. MARCUS J. McDERMOTT
"SALT: A Knowledge Acquisition Language for Propose-and-Revise Systems"
 Readings in Knowledge Acquisition and Learning, pp 263-281, Edited by B.G. Buchanan & D.C. Wilkins, Morgan Kaufmann Publishers, 1993.
- [MARINO & al, 90] O. MARINOS, F. RECHENMANN, P. UVIETTA
"Multiple Perspectives and Classification Mechanism in Object-Oriented Representation "
 Proceedings of the 9th ECAI, pp 425-430, Stockolm, 1990.
- [MARTY & al, 89] J-C. MARTY, N.R. SHADBOLT, B. WIELINGA
"The ACKnowledge Project: Improving the Knowledge Acquisition Process"
 In Results and Progress of Esprit Projects, XIII/416/89, 1989.
- [MASINI & al, 89] G. MASINI, A. NAPOLI, D. COLNET, D. LEONARD, K. TOMBRE
"Les Langages à Objets, langages de classes, langages de frames, langages d'acteurs"
 InterEditions, Paris, 1989.

- [**MATILE & al, 87**] L. MATILE, P. TASSY, D. GOUJET
 "Introduction à la systématique zoologique"
 Biosystème 1. SFS Editions, Paris, 1987.
- [**MAYR, 86**], E. MAYR
 "La Systématique Evolutionniste et les Quatre Etapes du Processus de Classification"
 In [TASSY, 86], pp 143-160, 1986.
- [**MICHALSKI & STEPP, 83**] R.S. MICHALSKI, R. STEPP
 "Learning from Observation: Conceptual Clustering"
 In Machine Learning An Artificial Intelligence Approach, Morgan Kaufmann Edit, 1983.
- [**MINEAU, 90**] G.W. MINEAU
 "Structuration des bases de Connaissances par Généralisation"
 Ph.D. d'Informatique, Université de Montréal, 1990.
- [**MINEAU, 92**] G.W. MINEAU
 "Acquisition d'Objets Structurés destinés à la Classification Symbolique"
 1ères Journées Francophones d'Apprentissage et d'Explicitation des
 Connaissances, pp 131-146, Dourdan, France, 1992.
- [**MINEAU, 93**] G.W. MINEAU
 "Transformations Représentationnelles d'Objets destinés à la Classification Symbolique"
 JAVA : Journées Françaises d'Acquisition de Validation et d'Apprentissage, Saint Raphaël, France, 1993.
- [**MINEAU & al, 90**] G.W. MINEAU, J. GECSEI, R. GODIN
 "La Classification Symbolique Une Approche Non-Subjective"
 Actes des Sièmes Journées Françaises de l'Apprentissage, pp 169-189, Lannion, 1990.
- [**MORSE, 74**] L.E. MORSE
 "Computer programs for specimen identification, key construction and description printing"
 Publications of the Museum of the Michigan State University, Biol., vol 5, pp 1-128, 1974.
- [**NAGLE & al, 92**] T.E. NAGLE, J.A. NAGLE, L.L. GERHOLZ, P.W. EKLUND
 "Conceptual Structures: current research and practice"
 T.E. NAGLE, J.A. NAGLE, L.L. GERHOLZ, P.W. EKLUND Editors, ELLIS HORWOOD
 Publishers, 1992.
- [**NAPOLI, 92a**] A. NAPOLI
 "Représentation à Objets et Raisonnement par Classification en Intelligence Artificielle"
 Thèse d'état ès Sciences Mathématiques de l'université de Nancy 1, 1992.
- [**NAPOLI, 92b**] A. NAPOLI
 "Subsumption and Classification-Based Reasoning in Object-Based Representations"
 ECAI 92, Proceedings of the 10th European Conference on Artificial Intelligence, pp 425-429,
 B. Neumann Editor, John Wiley & Sons Publishers, 1992.
- [**NAPOLI, 94**] A. NAPOLI
 "Catégorisations, raisonnement par classification et raisonnement à partir de cas"
 Actes des 9ièmes Journées Françaises de l'apprentissage, Strasbourg, 1994.
- [**NAPOLI & RECHENMANN, 93**] A. NAPOLI, F. RECHENMANN
 "Introduction à la Journée d'Etude sur la Classification"
 Représentation par Objets, pp 211-213, La Grande Motte, 1993.
- [**NAZARENKO, 92**] A. NAZARENKO
 "Les types de concepts spéciaux"
 Actes des 4èmes Journées Nationales du PRC-IA Formalisation et Extensions des Graphes Conceptuels,
 pp 82-91, Marseille, 1992.
- [**NEBEL, 90**] B. NEBEL
 "Reasoning and Revision in Hybrid Representation Systems"
 Lecture Notes in Computer Sciences 422, Springer verlag, Berlin, 1990.

- [NEWELL & SIMON, 72] A. NEWELL, H.A. SIMON
 "Human Problem Solving"
 Prentice Hall, Inc., Englewood Cliffs, N.J., 1972.
- [NICOLAS, 87] J. NICOLAS
 "ALLY, un Système logique pour la généralisation en apprentissage automatique"
 Thèse de docteur d'université, Rennes I, 1987.
- [NIQUIL, 93] Y. NIQUIL
 "Acquisition d'exemples en discrimination : spécification des exemples par génération de scénarios"
 Thèse de docteur de l'université de Paris IX Dauphine, 1993.
- [NOGIER, 91] J.F. NOGIER
 "Génération automatique de langage et graphes conceptuels"
 Editions HERMES, Paris, 1991.
- [PAJAUD, 90a] D. PAJAUD
 "Nomenclature et Taxinomie"
 La Taxinomie Bionaturaliste, vol 1, Lachurie Editeur, 1990.
- [PAJAUD, 90b] D. PAJAUD
 "Taxinomie et Taxilogie"
 La Taxinomie Bionaturaliste, vol 2, Lachurie Editeur, 1990.
- [PANKHURST, 70] R.J. PANKHURST
 "A computer program for generating diagnostic keys"
 Computer Journal, vol 13, pp 145-151, 1970.
- [PANKHURST, 88] R.J. PANKHURST
 "An interactive program for the construction of identification keys"
 Taxon, vol 37, pp 747-755, 1988.
- [PANKHURST, 91] R.J. PANKHURST
 "Practical Taxonomic Computing"
 Cambridge University Press, 1991.
- [PANKHURST, 93] R.J. PANKHURST
 "Principles and Problems of Identification"
 in [FORTUNER, 93b], pp 125-136, 1993.
- [PANKHURST & AITCHISON, 75] R.J. PANKHURST, R.R. AITCHISON
 "An on-line identification program"
 In: Pankhurst R.J. (ed.), *Biological identification with computers*. London and New York: pp 181-185,
 Academic Press, 1975.
- [PRIMIO & BREWKA, 85] F. DI PRIMIO, G. BREWKA
 "BABYLON: Kernel System of an Integrated Environment for Expert System Development and
 Operation"
 In Proceedings of the 5th International Workshop on Expert Systems and their applications, pp 573-583,
 Avignon, France, 1985.
- [QUILLIAN, 68] M.R. QUILLIAN
 "Semantic Memory"
 In *Semantic Information Processing*, M. MINSKY (Ed), pp 216-270, MIT Press, 1968.
- [QUINLAN, 83] J.R. QUINLAN
 "Learning Efficient Classification Procedures and their Application to Chess end Games"
 In *Machine Learning I, an Artificial Intelligence Approach*, 1983.
- [QUINLAN, 86] J.R. QUINLAN
 "Induction of Decision Trees"
Machine Learning 1, vol 1, Kluwer Academic Publishers, pp 81-132, 1986.

- [QUINLAN, 90] J.R. QUINLAN
"Learning Logical Definitions from Relations"
 Machine Learning 5, vol 3, Kluwer Academic Publishers, pp 239-266, 1990.
- [RECHENMANN, 88] F. RECHENMANN
"SHIRKA: Système de Gestion de Base de connaissances Centrées-Objets"
 Manuel de référence, INRIA/ARTEMIS, Grenoble, 1988.
- [RICH, 87] E. RICH
"Intelligence Artificielle"
 Masson, Paris, 1987.
- [ROSCH & LLYOD, 78] E. ROSCH, B.B. LLYOD
"Cognition and Categorization"
 Lawrence Erlbaum Associates, Publishers, New Jersey, 1978.
- [ROUSSELOT, 93] F. ROUSSELOT
"Un Tour d'Horizon des Principales Approches en Rapport avec l'Acquisition des Connaissances à partir de Données Textuelles"
 Knowledge & Data Engineering Workshop, pp 105-122, Strasbourg, France, 1993.
- [RUGG & al, 90] G. RUGG, P. McGEORGE, N. SHADBOLT
"On the use of laddered grids in knowledge elicitation"
 Artificial Intelligence Group, Technical Report , Feb, 1, University of Nottingham, 1990.
- [RUGG & al, 92] G. RUGG, C. CORBRIDGE, N.P. MAJOR, A.M. BURTON, N.R. SHADBOLT
"A comparison of sorting techniques in knowledge acquisition"
 Knowledge Acquisition 4, pp 279-291, 1992.
- [RUMELHART & NORMAN, 73] D.E. RUMELHART, D.A. NORMAN
"Active Semantic Networks as a Model of Human Memory"
 Proc of the Third IJCAI, pp 450-457, Stanford, Cal, 1973.
- [SABAH, 88] G. SABAH
"L'Intelligence Artificielle et le Langage, Représentation des Connaissances"
 Editions Hermés, 1988.
- [SABAH & VILNAT, 93] G. SABAH, A. VILNAT
"Hierarchy of relational types in conceptual graphs to handle natural language parsing"
 ICCS'93: Proceedings of the First International Conference on Conceptual Structures: Theory and Application, pp 198-215, Québec City, Canada, 1993.
- [SHADBOLT & BURTON, 90] N. SHADBOLT, A.M. BURTON
"Knowledge Acquisition"
 In E.N. Wilson and J.R. Corlett (Eds) *Evaluation of human work: Practical ergonomics methodology*.
 Taylor and Francis, 1990.
- [SHORTLIFFE, 76] E.H. SHORTLIFFE
"Computer-based medical consultations : MYCIN"
 American Elsevier (N.Y.), 1976.
- [SIMMONS & BRUCE, 71] R.F. SIMMONS, B.C. BRUCE
"Some Relations Between Predicate Calculus and Semantic Net Representation of Discourse"
 Proc of the Second IJCAI, pp 524-530, London, 1971.
- [SNEATH & SOKAL, 73] P.H.A. SNEATH, R.R. SOKAL
"Numerical Taxonomy"
 Donald Kennedy and Roderic Park Ed., Freeman and company, San Fransisco, 1973.
- [SOKAL & SNEATH, 63] R.R. SOKAL, P.H.A. SNEATH
"Principles of Numerical Taxonomy"
 San Francisco, W.H. Freeman, 1963.

- [SOWA, 84] J.F. SOWA
 "Conceptual structures: Information Processing in Mind and Machine"
 Addison-Wesley Publishing Co, 1984.
- [SOWA, 91] J.F. SOWA
 "Principles of Semantic Networks: explorations in the representation of knowledge"
 Edited by J. F. SOWA, Morgan Kaufmann Publishers, 1991.
- [SOWA, 92] J.F. SOWA
 "Conceptual Graphs as a Universal Knowledge Representation"
 In An International Journal Computers & Mathematics with Applications : Special Issue on Semantic Networks in Artificial Intelligence, vol 23, pp 75-93 , 1992.
- [STEELS, 90] L. STEELS
 "Components of Expertise"
 The AI Magazine 11(2), 1990.
- [STEPP & MICHALSKI, 86] R.E. STEPP, R.S. MICHALSKI
 "Learning From Observation : Conceptual Clustering : Inventing Goal Oriented Classifications of Structured Objects"
 Machine Learning: An Artificial Intelligence Approach, vol 2, pp 471-498, Morgan-Kaufman Publishing Co, 1986.
- [SZULMAN, 90] S. SZULMAN
 "Enrichissement d'une Base de Connaissances à partir de textes en Langage Naturel"
 Thèse de docteur d'université, Paris XIII, 1990.
- [TASSY, 86] P. TASSY
 "L'ordre et la diversité du vivant"
 Nouvelle Encyclopédie des Sciences et des techniques, fondation Diderot, 1986.
- [THAYSE & al, 90] A. THAYSE, P. DELSARTE, J. HAGELSTEIN, G. LOUIS, A. RIFAUT, M. VAUCLAIR, E. DUBOIS, A. VAN LAMSWEERDE, F. VAN DER LINDEN.
 "Approche Logique de l'Intelligence Artificielle"
 Vol 1, Dunod, 1990.
- [TREUIL & al, 94] J.P. TREUIL, F. ACHARD, F. PARIS
 "La réalisation d'un Hypertexte sur l'Archéologie au Niger : Bilan Méthodologique et Perspectives"
 CARI 94, 2^e Colloque Africain sur la Recherche en Informatique, Yaoundé, 1994 (soumis).
- [VAN MECHELEN & al, 93] I. VAN MECHELEN, J. HAMPTON, R.S. MICHALSKI, P. THEUNS
 "Categories and Concepts, Theoretical Views and Inductive Data Analysis"
 Cognitive Sciences series, Academic Press, 1993.
- [VAN WAMBEKE & al, 83] VAN WAMBEKE et collaborateurs
 "Keys to Soil Taxonomy"
 Département d'agronomie, Université Cornell, Ithaca, N.Y. USA, 1983.
- [VIGNES, 91] R. VIGNES
 "Caractérisation automatique de groupes biologiques"
 Thèse de docteur d'université, Paris 6, France, 1991.
- [VOGEL, 88] C. VOGEL
 "Genie cognitif"
 Editions Masson, France, 1988.
- [WIELINGA & BREUKER, 86] B.J. WIELINGA, J.A. BREUKER
 "Models of Expertise"
 ECAI 86, Proceedings of the 7th European Conference on Artificial Intelligence, Elsevier Science Publishers B.V, North Holland, pp 306-318, Brighton, 1986.

[WIELINGA & al, 92] B.J. WIELINGA, W. VAN DE VELDE, G. SCHREIBER, H. AKKERMANS

"The CommonKADS Framework for Knowledge Modelling"

Rapport interne, Esprit Project KADS II P5248. Excerpt of Deliverable 1.1., 1992.

[WINSTON & al, 87] M.E. WINSTON, R. CHAFFIN, D. HERRMANN

"A Taxonomy of Part-Whole Relations"

Cognitive Science, no 11, pp 417-444, 1987.

[WOODS, 75] W. WOODS

"What's in a link"

In Representation and understanding : studies in cognitive science, Bobrow & Collins, Academic Press, New York, pp 35-82, 1975.

[WOODS & SCHMOLZE, 92] W. WOODS, J.G. SCHMOLZE

"The KL-ONE Family"

An International Journal Computers & Mathematics with Applications : Special Issue on Semantic Networks in Artificial Intelligence, vol 23, pp 133-177, 1992.

[ZACKLAD, 93] M. ZACKLAD

"Principes de modélisation qualitative pour l'aide à la décision dans les organisations : Méthode d'utilisation du logiciel d'acquisition des connaissances C-KAT"

Thèse de docteur de l'université de technologie de Compiègne, 1993.

ANNEXE 1

Algorithme de numérotation des étiquettes de classes à la création

Algorithme de numérotation des étiquettes de classes à la création

Soit Li la liste contenant les classes Ci et leur couverture $Couv_C (Ci)$ selon l'ordre d'apparition dans la base des graphes

$Li = \{(C1, Couv_C (C1)), (C2, Couv_C (C2)), \dots\}$

Soit η le compteur des numéros d'apparition des classes atomiques.

Soit la relation d'ordre partiel " $>$ ", $Ci > Ck$ signifie que Ci est une classe ascendante de la classe Ck .

Une classe Ck est une *première descendante* de la classe Ci s'il n'existe aucune classe $Cj \neq Ck$ telle que : $Ci > Cj$

Initialisation

$\eta := 1$

$Li := \{(C0, \{1\})\}$; création d'une classe racine dont la couverture vaut 1

Numérotation (Cj)

Début

Soit Ci la classe ascendante directe de la classe créée Cj

Si Cj est une classe première descendante de Ci

Alors • $Couv_C (Cj) := Couv_C (Ci)$

Sinon • $\eta := \eta + 1$

• $Couv_C (Cj) := \{ \eta \}$

Pour tout Ci telle que $Ci > Cj$

$Couv_C (Ci) := Couv_C (Ci) \cup Couv_C (Cj)$

FinPour

FinSi

$Li := Li \cup (Cj, Couv_C (Cj))$

Fin

Remarque :

en cas de suppression de classe, l'algorithme se réécrit très simplement.

Numérotation (Cj)

Début

Pour tout Ci telle que Ci > Cj

$Couv_C(Ci) := Couv_C(Ci) - Couv_C(Cj)$

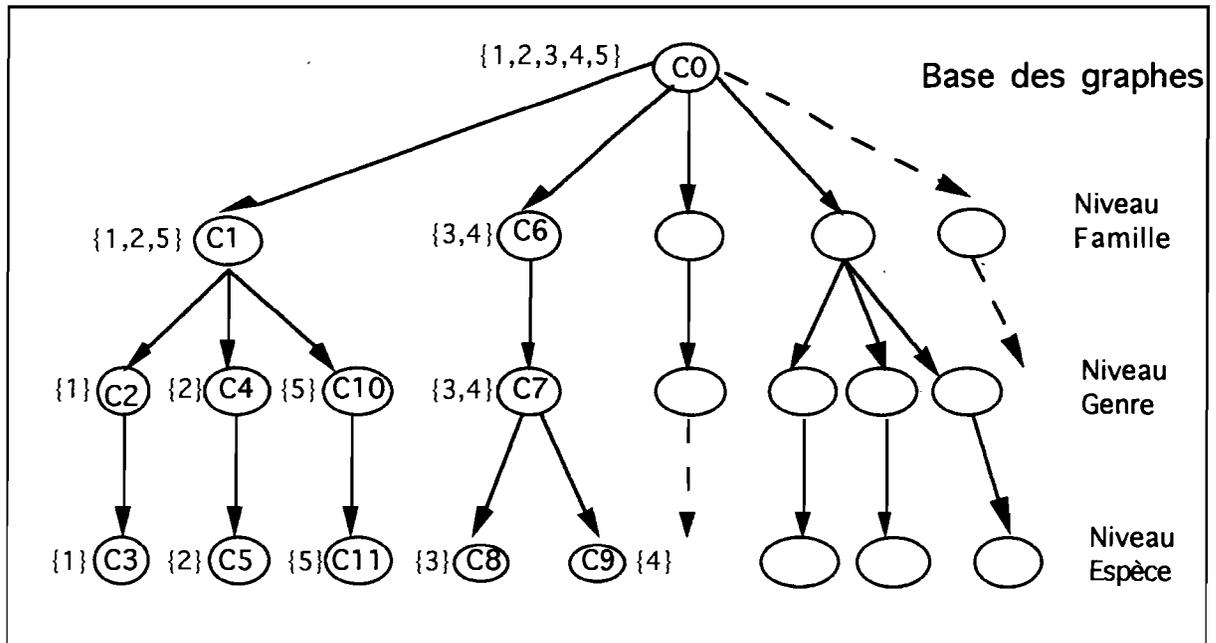
FinPour

Li := Li - (Cj, Couvc(Cj))

Fin

Exemple :

Nous allons illustrer l'algorithme de création et de mise à jour des étiquettes à l'aide de plusieurs exemples (création des classes C0 jusqu'à C11) représentés dans la figure ci-dessous



Numérotation des classes

Première étape : initialisation

$\eta := 1$

Li := {(C0, {1})};

Deuxième étape : création de classes

Numérotation (C1) $Couv_C(C1) := 1$; C1 première descendante de C0

$Li := \{(C0, \{1\}), (C1, \{1\})\}$

Numérotation (C2) $Couv_C(C2) := 1$; C2 première descendante de C1

$Li := \{(C0, \{1\}), (C1, \{1\}), (C2, \{1\})\}$

Numérotation (C3) $Couv_C(C3) := 1$; C3 première descendante de C2

$Li := \{(C0, \{1\}), (C1, \{1\}), (C2, \{1\}), (C3, \{1\})\}$

Numérotation (C4) $\eta := 2$; $Couv_C(C4) := \{2\}$; $Couv_C(C1) := \{1, 2\}$; $Couv_C(C0) := \{1, 2\}$

Mise à jour des couverture de l'ascendance de C4

$Li := \{(C0, \{1,2\}), (C1, \{1,2\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\})\}$

Numérotation (C5) $Couv_C(C5) := \{2\}$

$Li := \{(C0, \{1,2\}), (C1, \{1,2\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\})\}$

Numérotation (C6) $\eta := 3$; $Couv_C(C6) := \{3\}$; $Couv_C(C0) := \{1, 2, 3\}$

$Li := \{(C0, \{1,2,3\}), (C1, \{1,2\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\}), (C6, \{3\})\}$

Numérotation (C7) $Couv_C(C7) := \{3\}$

$Li := \{(C0, \{1,2,3\}), (C1, \{1,2\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\}), (C6, \{3\}), (C7, \{3\})\}$

Numérotation (C8) $Couv_C(C8) := \{3\}$

$Li := \{(C0, \{1,2,3\}), (C1, \{1,2\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\}), (C6, \{3\}), (C7, \{3\}), (C8, \{3\})\}$

Numérotation (C9) $\eta := 4$; $Couv_C(C9) := \{4\}$; $Couv_C(C7) := \{3, 4\}$; $Couv_C(C6) := \{3,4\}$; $Couv_C(C0) := \{1, 2, 3, 4\}$

$Li := \{(C0, \{1,2,3,4\}), (C1, \{1,2\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\}), (C6, \{3,4\}), (C7, \{3,4\}), (C8, \{3\}), (C9, \{4\})\}$

Numérotation (C10) $\eta := 5$; $Couv_C (C10) := \{5\}$; $Couv_C (C1) := \{1, 2,5\}$;

$Couv_C (C0) := \{1, 2, 3, 4, 5\}$

$Li := \{(C0, \{1,2,3,4,5\}), (C1, \{1,2,5\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\}), (C6, \{3,4\}), (C7, \{3,4\}), (C8, \{3\}), (C9, \{4\}), (C10, \{5\})\}$

Numérotation (C11) $Couv_C (C11) := \{5\}$

$Li := \{(C0, \{1,2,3,4,5\}), (C1, \{1,2,5\}), (C2, \{1\}), (C3, \{1\}), (C4, \{2\}), (C5, \{2\}), (C6, \{3,4\}), (C7, \{3,4\}), (C8, \{3\}), (C9, \{4\}), (C10, \{5\}), (C11, \{5\})\}$

Remarque :

$\eta := 5$, ce qui confirme le fait qu'il y ait 5 classes atomiques de créés (C3, C5, C8, C9, C11)

ANNEXE 2

Opérations sur les graphes conceptuels

Opérations sur les arcs conceptuels

Ajout d'un arc conceptuel (G_i, AC)

Soit G_i le graphe manipulé

Soit $AC : ([C_i : *] \rightarrow (Rel) \rightarrow [C_j : R_j^1])$ l'arc conceptuel à ajouter.

Faire une jointure² maximale de G_i et AC sur C_i

Si AC est un nouvel arc

Alors

Si C_j est un nouveau concept composite

Alors Ajouter C_j dans le treillis de composition en dessous de C_i

Rajouter C_j dans le treillis des types

Initialiser la couverture de l'arc

FinSi

Mettre à jour l'ensemble des référents si nécessaire

Mettre à jour le C-Modèle

Sinon Mettre à jour la couverture de AC

FinSi

¹ R_j peut être générique.

² En fait c'est une jointure d'ensemble que l'on fait ("Set join" [Sowa, 84] page 117), car les référents que nous manipulons sont le plus souvent des référents ensemblistes.

Suppression d'un arc conceptuel (Gi, AC)

Soit Gi le graphe manipulé

Soit AC : ([Ci : *] -> (Rel) -> [Cj : Rj]) l'arc conceptuel à supprimer.

Faire une projection de AC dans Gi

Si Cj n'apparaît que dans Gi

Alors

Si Cj est un concept composite

Alors Supprimer Cj et toute sa descendance dans le treillis de composition

FinSi

Supprimer Cj dans le treillis des types

Sinon

Si Rj n'apparaît que dans Gi

Alors supprimer Rj de l'ensemble des référents

FinSi

Mettre à jour la couverture de l'arc

FinSi

Remarque :

on ne supprime pas les relations car elles sont données dans le noyau minimal.

Recherche d'un arc conceptuel (Gi, AC)

Soit Gi le graphe manipulé

Soit AC : ([Ci : *] -> (Rel) -> [Cj : Rj]) l'arc conceptuel à rechercher.

Faire une projection de AC dans Gi

Si il existe au moins un noyau de projection³

Alors Retourner Succès

Sinon Retourner Echec

Finsi

³ Si le lecteur ne sait plus ce qu'est un noyau de projection, qu'il se réfère au paragraphe 4.4.4.

Opérations sur les treillis

Ajout d'un concept dans le treillis des types de concepts (Ci Tc)

Si Ci n'apparaît pas dans Tc

Alors Si Ci est un concept simple

Alors

Si Ci est un concept composite

Alors Insérer Ci en dessous du concept Element-structurel

Sinon Insérer Ci en dessous du concept Element descriptif

FinSi

Sinon Insérer Ci en dessous du concept par rapport auquel il est défini par genre et différence

FinSi

FinSi

Remarque : Après adjonction, on vérifie que la structure obtenue est toujours un treillis.

Suppression d'un concept dans le treillis des types de concepts (Ci Tc)

Localiser le concept Ci à supprimer dans le treillis des types (Tc)

Si Ci est un concept simple

Alors

Si Ci est un concept composite

Alors Supprimer Ci en dessous du concept Élément-structurel

Sinon Supprimer Ci en dessous du concept Élément-descriptif

FinSi

Sinon Supprimer Ci en dessous du concept par rapport auquel il est défini par genre et différence

FinSi

Remarque : Après suppression, on vérifie que la structure obtenue est toujours un treillis.

Ajout d'un composant dans le treillis des types de composition (Ci, Tcomp)

Parcourir le treillis de composition (Tcomp)

Rechercher le concept composite Cg dont Ci est l'un des composants

Si Ci n'existe pas dans Tcomp

Alors Insérer Ci au-dessous de Cg dans Tcomp

FinSi

Suppression d'un composant dans le treillis des types de composition(Ci, Tcomp)

Localiser Ci dans le treillis de composition (Tcomp)

Si Ci est un nœud feuille dans Tcomp

Alors Supprimer Ci de Tcomp

Sinon Supprimer Ci et toute sa descendance de Tcomp

FinSi

Parcours et recherche d'un concept dans le treillis de composition

Initialisation

Soit T_{comp} , le treillis des concepts composites associés au nombre de leurs composants

$T_{comp} := ((C1 (C11 n11) (C12 ((C121, n121) \dots) n12) \dots n1) (C2 n2)..)$

Racine := [Poisson]

Recherche (Ci, Racine, Tcomp)

Tant que Ci non trouvé et Tcomp non entièrement parcouru

Faire

Si N-courant := Ci

Alors Retourner Ci

Sinon Si N-courant := Feuille

Alors N-courant := Ancêtre-direct de N-courant

(Rechercher Ci, N-courant, Tcomp)

Sinon N-courant := descendant-direct de N-courant
avec $n_j = \text{minimum}$

FinSi

FinSi

Fin-Tant que

Si Ci non trouvé

Alors Retourner Echee

Sinon Retourner Ci

FinSi

ANNEXE 3

*Quelques copies d'écran de l'interface graphique en
Ichtyologie*

Espèce **C. citharus** Famille **Citharinidae**

Genre **Citharinus**

Descripteur
(Geoffroy Saint-Hilaire, 1809)

Synonymes *Sérrasalmus citharus*

Poids max

Photo Description Affinités

Longueur totale max

CLOFFA Biologie / Ecologie

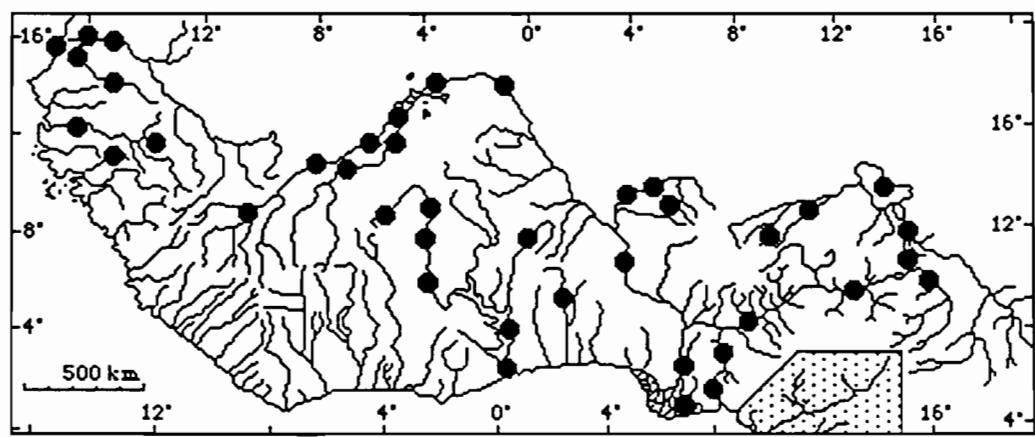
Longueur standard max 580 mm

Autres rubriques Bassins



Bassin(s) de l'espèce **C. citharus**

Liste des bassins
→ Cross
→ Gambie
→ Géba
→ Lac Mobutu



Description de l'espèce

C. citharus

Squelette

Osseux

Caractères quantitatifs

- Caractères métriques
- Caractères méristiques

Caractères qualitatifs



Caractères qualitatifs de l'espèce

C. citharus

Profil du corps

Corps élevé.



- Nageoires
- Coloration
- Description tête

Section transversale

Comprimé latéralement.



Peau

Divers

Ecailles

Ligne latérale

Ecailles cycloïdes.



Ecailles de l'espèce

C. citharus



Nombre



Forme

Ecailles cycloïdes.



Taille



Localisation

Sur tout le corps.



Tête

Pas d'écaille.



Ligne latérale de l'espèce

C. citharus

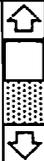


Nombre



Etat

Rectiligne et complète.



Position

Au milieu des flancs.



Nageoires de l'espèce

C. citharus

Nageoires pectorales ↵

Nageoire anale ↵

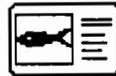
Nageoires pelviennes ↵

Nageoire caudale ↵

Nageoires dorsales ↵



nombre de dorsales 2



Nageoires dorsales de l'espèce

C. citharus

1 Dorsale

- Une dorsale adipeuse ↵
- Une dorsale molle (1ers ray. simples non épineux) ↵
- Une dorsale molle (1ers ray. simples épineux) ↵

2 Dorsales

- L'ant. molle (1ers ray. simples non épineux) et 1 adipeuse ↵
- L'ant. molle (1ers ray. simples épineux) et 1 adipeuse ↵
- Les deux dorsales molles avec ray. simples non épineux ↵
- L'ant. épineuse et 1 dorsale molle (1ers ray. simples non épineux) ↵
- L'ant. épineuse et 1 dorsale molle (1ers ray. simples épineux) ↵



Nageoires

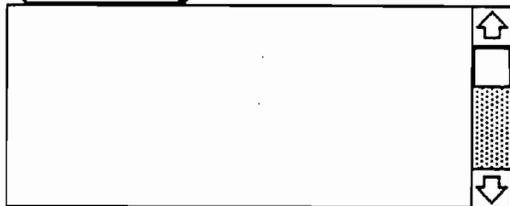


Aspect nageoires dorsales de l'espèce*C. citharus***Forme**

L'adipeuse est grande.

**Couleur**

La dorsale antérieure est grisâtre, la base de l'adipeuse est noirâtre.

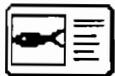
**Position****Structure**

La dorsale antérieure est rayonnée. La dorsale postérieure est adipeuse.

**Caractères métriques de l'espèce***C. latus* % Longueur standard % Longueur tête % Largeur tête % Longueur du museau % Largeur inter-orbitaire % Distance post-orbitaire % Distance prédorsale % Distance prépelvienne % Distance prépectorale % Longueur du pédoncule caudal % Longueur de la base de la dorsale

Coloration de l'espèce

C. citharus



Couleurs et localisations sur le vivant

Le corps est argenté ; les nageoires pectorales et le lobe supérieur de la caudale sont grisâtres ; le lobe inférieur de la caudale, les ventrales et l'anale (sauf sa partie basale) sont rouges. La partie basale de l'adipeuse est noirâtre.



Couleurs et localisations sur le conservé

Sensiblement le même patron de coloration que sur le vivant.



Description tête de l'espèce

C. citharus

Yeux Mâchoires Narine et appendice nasal

Barbillons ↔

Fontanelles ↔

Dents

Forme



Opercule ↔

Disposition



Forme et position bouche

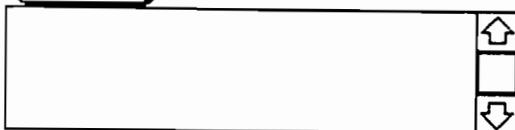
Terminale.



Aspect lèvres



Divers



Pores

Inexistants.



Formes dents de l'espèce

C. citharus

Mâchoire inférieure

Pointues.



Palatin



Mâchoire supérieure

Pointues



Ectoptérygoïde



Pharyngien



Endoptérygoïde



Yomer



Langue



Mâchoires



Disposition dents de l'espèce

C. citharus

Mâchoire inférieure

Une seule rangée de dents fines sur le bord des lèvres.



Palatin



Mâchoire supérieure

Une seule rangée de dents fines sur le bord des lèvres.



Ectoptérygoïde



Pharyngien



Endoptérygoïde



Yomer



Langue



Mâchoires



Caractères méristiques de l'espèce

C. latus



Formule scalaire

Formule dentaire

17,5-18,5/59-71/18,5-19,5 ; 13,5-15,5



Formule branchiale

Nombres de vertèbres

42



Nombre de prédorsales

Formules des nageoires



Formule des nageoires de l'espèce

C. latus



Première dorsale

Deuxième dorsale

IV-VI-16-18



Pelviennes

Pectorales

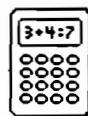
10

15-16



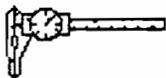
Caudale

Anale



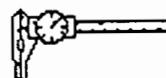
III-IV-20-23





% LONGUEUR STANDARD

Hauteur du corps.....	1,6-1,9
Longueur de la tête.....	3,3-4,3
Longueur du pédoncule caudal.....	
Longueur prédorsale.....	
Longueur prépectorale.....	
Longueur prépelvienne.....	
Distance interdorsale.....	
Longueur du plus long rayon de la dorsale.....	
Longueur de la base de la dorsale ant.....	
Longueur de la base de la dorsale post.....	
Longueur de la base de l'anale.....	

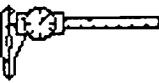


% LONGUEUR DE LA TETE

Largeur de la tête.....	
Diamètre de l'oeil.....	3,9-6,0
Longueur du museau.....	
Largeur inter-orbitaire.....	
Distance post-orbitaire.....	
Longueur de la bande vomérienne.....	
Longueur de la bande prémaxillaire.....	
Longueur de la base de la dorsale ant.....	
Longueur de la base de la dorsale post.....	
Longueur du plus long rayon de la dorsale.....	1,0-2,3

Fiche3 de l'espèce

C. latus



% LARGEUR DE LA TETE

Largeur inter-orbitaire

Diamètre de l'oeil

% DISTANCE POST-ORBITAIRE

Largeur inter-orbitaire

Diamètre de l'oeil

% LONGUEUR DU MUSEAU

Distance post-orbitaire

Largeur inter-orbitaire

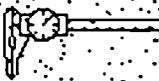
Diamètre de l'oeil

% LARGEUR INTER-ORBITAIRE

Diamètre de l'oeil

Fiche4 de l'espèce

C. latus



% LONGUEUR BASE DE LA DORSALE

Longueur du plus long rayon de la dorsale

% DISTANCE PREPECTORALE

Distance prépelvienne

Distance préanale

% DISTANCE PREDORSALE

Distance prépelvienne

Distance prépectorale

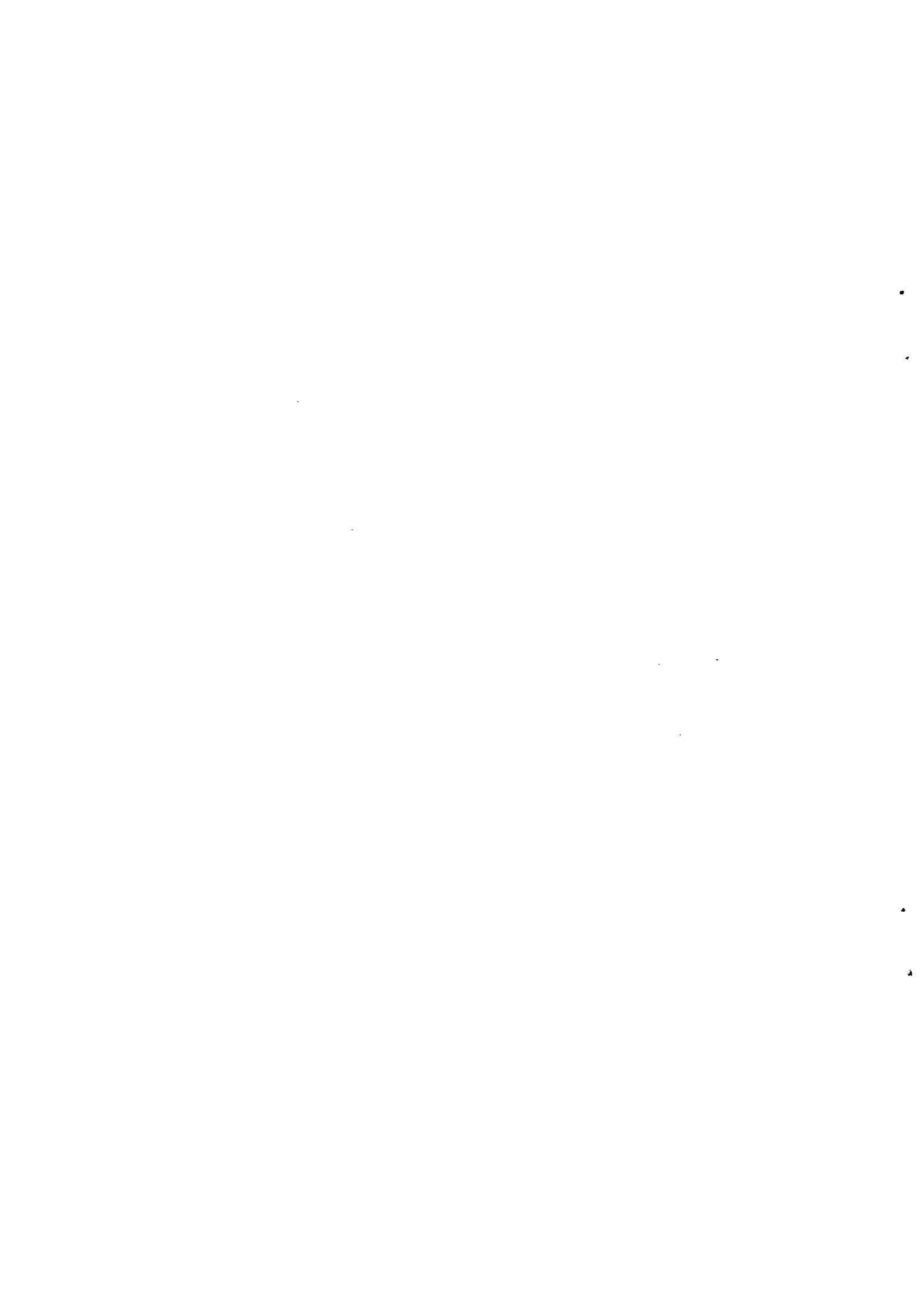
Distance préanale

% DISTANCE PREPELVIENNE

Distance préanale

% LONGUEUR DU PEDONCULE CAUDAL

Hauteur du pédoncule caudal



ANNEXE 4

Expérimentation en minéralogie

EXPÉRIMENTATION EN MINÉRALOGIE

Nous avons expérimenté METIS dans trois domaines (ichtyologie, minéralogie et pédologie). Cette annexe présente l'expérimentation de METIS en minéralogie. Elle fournit les descriptions explicitées par METIS : 100 Taxons qui comprennent 5 classes, 71 espèces, et 24 variétés. Étant donné que METIS est un système d'acquisition interactif et compte tenu du nombre d'exemples considérés, nous ne pouvons présenter toutes les étapes de l'explicitation. Ces dernières ont été largement illustrées dans le chapitre 6, sur quelques descriptions d'espèces, appartenant à la classe des Oxydes-Hydroxydes. Nous allons donc nous intéresser uniquement aux critères de discrimination retenus par la procédure d'apprentissage IDICE.

Notons toutefois, que les descriptions présentées dans cette annexe sont les descriptions obtenues après l'étape de discrimination qui comprend entre autres le traitement de complétion (voir § 5.6.2.1.). En effet, il ne faut pas perdre de vue que nous sommes dans le cas d'une acquisition incrémentale, par conséquent, par soucis de clarté et de lisibilité, nous avons omis de présenter les descriptions intermédiaires des taxons, en cours d'acquisition.

Cette annexe comprend d'une part, la phase d'initialisation dans laquelle nous donnons le nombre et les noms de niveaux de la taxinomie ainsi que le noyau du domaine d'application, d'autre part, elle comprend les descriptions de taxons et les traits discriminants qui ont permis de les distinguer. Les traits discriminants ne sont pas fournis de façon incrémentale comme l'aurait fait METIS. En fait, nous donnons d'abord les traits ayant permis de discriminer les taxons de premier niveau, ensuite ceux permettant de discriminer les taxons de troisième niveau (quand il existe), avant ceux permettant de différencier les taxons de deuxième niveau.

Enfin, nous analysons le comportement du système à travers les traits discriminants choisis puis nous donnons quelques caractéristiques de la base de connaissances obtenue.

Initialisation

Nombre de niveaux considérés : 3

Niveau 1 : Classe

Niveau 2 : Espèce

Niveau 3 : Variété

Description du noyau

Pour mémoire, rappelons que le noyau est composé essentiellement de structures constantes (présentes chez tous les minéraux), des concepts classes ([Nom-Classe], [Nom-Espèce], [Nom-Variété]) et de quelques concepts descriptifs servant à désigner le minéral.

[Minéral] -> (Caractérist) -> [Nom-Classe]
[Minéral] -> (Caractérist) -> [Nom-Espèce]
[Minéral] -> (Caractérist) -> [Nom-Variété]
[Minéral] -> (Caractérist) -> [Gisement]
[Minéral] -> (Caractérist) -> [Localisation]
[Minéral] -> (Caractérist) -> [Syst-cristallin]
[Minéral] -> (Se-compose) -> [Particules¹]
[Particules] -> (Caractérist) -> [Comp-chimique]

La discrimination ne doit pas porter sur les éléments suivants : Nom-Classe, Nom-Espèce, Nom-variété (car ce sont des noms de taxons) et la localisation (car elle appartient à la fiche signalétique de l'espèce ou de la variété).

Création du type primitif

Type [Primitif (x)] est

[Minéral : *x] -> (Caractérist) -> [Nom-Classe]
[Minéral : *x] -> (Caractérist) -> [Nom-Espèce]
[Minéral : *x] -> (Caractérist) -> [Nom-Variété]
[Minéral : *x] -> (Caractérist) -> [Syst-Cristallin]
[Minéral : *x] -> (Caractérist) -> [Gisement]
[Minéral : *x] -> (Caractérist) -> [Localisation]
[Minéral : *x] -> (Se-compose) -> [Particules]
[Particules] -> (Caractérist) -> [Comp-chimique]

¹ La matière des minéraux est le résultat d'un arrangement systématique de ses particules : atomes, ions et molécules, constituant un édifice périodique homogène, appelé matière cristalline.

Description des taxons de premier niveau : les classes.

Description de la classe des Éléments²

Type [Classe-Éléments (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Classe : Éléments]

[Primitif : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique³ : N]

Description de la classe des Sulfures⁴

Type [Classe-Sulfures (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Classe : Sulfures]

[Primitif : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (N_{i1}l ...N_{in})_x S_y]

Description de la classe des Halogénures⁵

Type [Classe-Halogénures (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Classe : Halogénures]

[Primitif : *x] -> (Se-compose) -> [Particules]

*[Particules] -> (Caractérist) -> [Comp-chimique : N*_i (N_{jl}l ...N_{jn})]*

Description de la classe des Oxydes-Hydroxydes⁶

Type [Classe-Oxydes-Hydroxyde⁵ (x)] est

[Primitif : *x] -> (Caractérist) -> [Nom-Classe : Oxydes-Hydroxydes]

[Primitif : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (N_{i1}l ...N_{in})_x O_y n(H₂O)]

Description de la classe des Substances-Organiques⁷

Type [Classe-Substances-Organiques (x)] est

² Ce sont des éléments chimiques qui se trouvent dans la nature à l'état pur. C'est par exemple le cas de l'or ou de l'argent.

³ N pouvant être n'importe quel élément du tableau périodique des éléments.

⁴ Ce sont des minéraux dont la composition chimique comprend le soufre (S).

⁵ les N* sont des halogènes tels que : le chlore (Cl), le Brome (Br), le fluor (F) et l'iode (I). L'oxygène n'est pas un halogène.

⁶ Cette classe comprend toutes les combinaisons oxygénées du règne animal (oxydes), quand il y a en plus, dans leur structure des groupements (OH), ce sont des hydroxydes.

⁷ Ce sont des minéraux dont la composition chimique comprend le carbone (C) et l'hydrogène (H).

[Primitif : *x] -> (Caractérist) -> [Nom-Classe : Substances-Organiques]

[Primitif : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : CxHy (N1i1 ...NninOz)]

Discrimination

Nous constatons que la discrimination entre les 5 classes s'est faite à l'aide du critère composition chimique. Le trait discriminant [Particules] -> (Caractérist) -> [Comp-chimique] correspondant est mis en gras italique dans la description des taxons.

Description des taxons du deuxième niveau : les espèces de la classe des Éléments.

Nous avons choisi de décrire 10 espèces de la classe des Éléments. Ce sont des éléments chimiques qui se trouvent dans la nature à l'état pur. C'est par exemple le cas de l'or ou de l'argent.

Description de l'espèce Bismuth

Type [Espèce-Bismuth : (x)] **est :**

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Bismuth]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 2-2,3]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : Blanc-argent]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 9,7-9,8]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : {Gris-de-plomb, Brillante}]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Triclinique]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement⁸ : Assez-Rare]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {Allemagne, Canada, Tchécoslovaquie, Espagne, Bolivie, Mexique, Pérou, Chili}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Bi]

Description de l'espèce Cuivre

Type [Espèce-Cuivre : (x)] **est :**

⁸ La localisation fournie n'est pas exhaustive, l'expert se limite à donner les lieux où l'on trouve les plus grands gisements.

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Cuivre]
 [Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 2,5-3]
 [Classe-Éléments : *x] -> (Caractérist) -> [Couleur : {Rouge-clair, Rouge-brun}]
 [Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]
 [Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]
[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 8,93]
 [Classe-Éléments : *x] -> (Caractérist) -> [Clivage : !Clivable]
 [Classe-Éléments : *x] -> (Caractérist) -> [Trace : Rouge-cuivre]
[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]
 [Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Abondant]
 [Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {États-Unis, Allemagne, U.R.S.S., Mexique, Australie, Chili, Namibie}]
 [Classe-Éléments : *x] -> (Se-compose) -> [Particules]
 [Particules] -> (Caractérist) -> [Comp-chimique : Cu]

Description de l'espèce Argent

Type [Espèce-Argent : (x)] est :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Argent]
 [Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 2,5-3]
 [Classe-Éléments : *x] -> (Caractérist) -> [Couleur : Blanc-argent]
 [Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]
 [Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]
[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 10,5]
 [Classe-Éléments : *x] -> (Caractérist) -> [Clivage : !Clivable]
 [Classe-Éléments : *x] -> (Caractérist) -> [Trace : {Blanche, Brillante}]
[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]
 [Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Assez-Rare]
 [Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {Allemagne, Norvège, Canada, Tchécoslovaquie, États-Unis, Bolivie, Mexique}]
 [Classe-Éléments : *x] -> (Se-compose) -> [Particules]
 [Particules] -> (Caractérist) -> [Comp-chimique : Ag]

Description de l'espèce Or

Type [Espèce-Or : (x)] est :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Or]
 [Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 2-5,3]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : {Jaune, Blanc-jaunâtre, Jaune-rouge}]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 19, 29]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : Jaune-brillant]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {Afrique du Sud, U.R.S.S., Canada, États-Unis, Mexique, Australie, îles Fidji, Ghana, Zimbabwe, Égypte, Tchécoslovaquie, Roumanie, Allemagne}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Au]

Description de l'espèce Arsenic

Type [Espèce-Arsenic : (x)] **est** :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Arsenic]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 3-4]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : Blanc-étain]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 5,7]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : Noire]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Triclinique]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {Norvège, États-Unis, France, Tchécoslovaquie, Allemagne}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : As]

Description de l'espèce Antimoine

Type [Espèce-Antimoine: (x)] **est** :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Antimoine]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 3-3,5]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : Blanc-étain]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 6,6-6,7]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Bon]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : Gris-de-plomb]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Trigonal]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {Suède, France, Tchécoslovaquie, Allemagne, Portugal, Bornéo, Australie}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : AsSb]

Description de l'espèce Fer

Type [Espèce-Fer : (x)] est :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Fer]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 4-5]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : Gris-acier]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 7,88]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : {Grise, Brillante}]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {Allemagne, Irlande, Nouvelle-Zélande, Groenland}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Fe]

Description de l'espèce Graphite⁹

Type [Espèce-Graphite : (x)] est :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Graphite]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 1-1,5]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : Gris-acier]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Mat]

⁹ Minéral constitué de carbone pur, dont il est la variété polymorphe la plus stable.

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 2,25]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : {Gris-métallique-foncé, Gris-métallique-brillant}]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Hexagonal]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {États-Unis, Allemagne, U.R.S.S., Tchecoslovaquie, Madagascar, Sri Lanka, Canada}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : C]

Description de l'espèce Soufre

Type [Espèce-Soufre : (x)] est :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Soufre]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 1,5-2]

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : {Jaune-soufre, Jaune-miel, Jaune-verdâtre, Brun-jaune}]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : {Adamantin, Mat}]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 2]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : {Blanche, Jaune-clair}]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : {Orthorhombique, Monoclinique, Amorphe}]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {États-Unis, Pologne, U.R.S.S., Espagne, Italie, Turquie, Indonésie, Mexique, Japon, Yougoslavie}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : S]

Description de l'espèce Diamant¹⁰

Type [Espèce-Diamant : (x)] est :

[Classe-Éléments : *x] -> (Caractérist) -> [Nom-Espèce : Diamant]

[Classe-Éléments : *x] -> (Caractérist) -> [Dureté : 10]

¹⁰ C'est la pierre précieuse la plus appréciée et la substance naturelle la plus dure que l'on connaisse. C'est un autre polymorphe du carbone.

[Classe-Éléments : *x] -> (Caractérist) -> [Couleur : {Incolore, Gris-bleuâtre, Verdâtre, Jaunâtre, Brune, Noire}]

[Classe-Éléments : *x] -> (Caractérist) -> [Éclat : Adamantin]

[Classe-Éléments : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Éléments : *x] -> (Caractérist) -> [Densité : 3,5]

[Classe-Éléments : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Éléments : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Éléments : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Éléments : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Éléments : *x] -> (Caractérist) -> [Localisation : {États-Unis, Pologne, U.R.S.S., Espagne, Italie, Turquie, Indonésie, Mexique, Japon, Yougoslavie}]

[Classe-Éléments : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : C]

Discrimination des espèces de la classe des Éléments

La discrimination des espèces de la classe Éléments s'est faite à l'aide d'abord du trait [Minéral] -> (Caractérist) -> [Syst-Cristallin] puis à l'aide du trait [Minéral] -> (Caractérist) -> [Densité]. L'expert a contraint IDICE de commencer par le critère système cristallin¹¹ (facilement observable) afin de discriminer les espèces, pourtant le critère densité a lui seul aurait pu suffire.

Description des taxons du deuxième niveau : les espèces de la classe des Sulfures

Nous allons maintenant fournir quelques descriptions (9 espèces et deux variétés) de la classe des sulfures.

A chaque fois qu'une espèce comprend des variétés, nous fournissons d'abord les traits discriminant les variétés de cette espèce, avant de donner les traits discriminant cette espèce de ses sœurs.

Description de l'espèce Orpiment

Type [Espèce-Orpiment : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Orpiment]

[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 1,5-2]

[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : {Jaune-or, Jaune-orange, Brune}]

[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : {Gras, Nacré}]

¹¹ Les cristaux de chaque minéral ont un aspect propre mais déterminé répondant immédiatement à sa structure interne.

[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : Transparente]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 3,48]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]
[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : Jaune-clair]
[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Assez-Rare]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {États-Unis, Pérou, U.R.S.S., Yougoslavie, Suisse}]
 [Classe-Sulfures : *x] -> (Se-compose) -> [Particules]
 [Particules] -> (Caractérist) -> [Comp-chimique : As₂S₃]

Description de l'espèce Réalgar

Type [Espèce-Réalgar : (x)] est :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Réalgar]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 1,5]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : {Rouge-brun, Rouge-orangé}]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : {Gras, Adamantin}]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 3,5]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]
[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : {Rouge-orange, Jaune-orange}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Assez-Rare]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {États-Unis, U.R.S.S., Yougoslavie, Roumanie, Suisse}]
 [Classe-Sulfures : *x] -> (Se-compose) -> [Particules]
 [Particules] -> (Caractérist) -> [Comp-chimique : As₄S₄]

Description de l'espèce Covelline

Type [Espèce-Covelline : (x)] est :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Covelline]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 1,5-2]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : {Bleu-indigo, Violet}]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : {Mat, Résineux}]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : Opaque]
 [Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 4,68]

[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : {Grise, Noire}]

[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Hexagonal]

[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {États-Unis, Yougoslavie, Roumanie, Allemagne, Chili, Bolivie}]

[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : CuS]

Description de l'espèce Tétradymite

Type [Espèce-Tétradymite : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Tétradymite]

[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 1,5-2]

[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : Gris-acier]

[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 7,2-7,9]

[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : Gris-de-plomb]

[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Triclinique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {Roumanie, États-Unis, Japon}]

[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Sb₂S₃]

Description de l'espèce Berthiérite

Type [Espèce-Berthiérite : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Berthiérite]

[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 2-3]

[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : {Brune, Gris-acier}]

[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 4,6]

[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Bon]

[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : Gris-brun]

[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {Allemagne, France, Roumanie, Tchécoslovaquie, Chili, Bolivie, Japon, Pérou}]

[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : FeSb₂S₄]

Description de l'espèce Jamesonite

Type [Espèce-Jamesonite : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Jamesonite]

[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 2,5]

[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : Gris-de-plomb]

[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 5,63]

[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : Noire]

[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {Allemagne, États-Unis, U.R.S.S., Suède, Roumanie, Tchécoslovaquie, Australie, Mexique, Bolivie, Japon}]

[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : PbFeSb₆S₁₄]

Description de l'espèce Livingstonite

Type [Espèce-Livingstonite : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Livingstonite]

[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 2]

[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : Gris-acier]

[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : {Métallique, Adamantin}]

[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : {Opaque, Translucide}]

[Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 4,9]

[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : Rouge]

[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {U.R.S.S., Mexique}]

[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]
[Particules] -> (Caractérist) -> [Comp-chimique : HgSb₄S₈]

Description de l'espèce Stannite

Type [Espèce-Stannite : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Stannite]
[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 4]
[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : {Gris-acier, Vert-olive}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : {Métallique, Mat}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : Opaque]
[Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 4,3-4,5]
[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Imparfait]
[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : Noire]
[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Quadratique]
[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : Abondant]
[Classe-Sulfures : *x] -> (Caractérist) -> [Localisation : {Allemagne, Grande-Bretagne, Australie, Tchécoslovaquie, Bolivie, U.R.S.S.}]
[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]
[Particules] -> (Caractérist) -> [Comp-chimique : Ag₃SbS₃]

Description de l'espèce Sphalérite

Type [Espèce-Sphalérite : (x)] **est** :

[Classe-Sulfures : *x] -> (Caractérist) -> [Nom-Espèce : Sphalérite]
[Classe-Sulfures : *x] -> (Caractérist) -> [Dureté : 3,5-4]
[Classe-Sulfures : *x] -> (Caractérist) -> [Couleur : {Brun-clair, Brun-foncé, Brun-rouge, Jaune, Rouge, Verte, Jaune-vert, Blanche, Noire}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Éclat : {Gras, Mat, Adamantin, Vitreux}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Transparence : {Transparente, Translucide, Opaque}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Densité : 3,9-4,2]
[Classe-Sulfures : *x] -> (Caractérist) -> [Clivage : Parfait]
[Classe-Sulfures : *x] -> (Caractérist) -> [Trace : {Blanche, Brun-clair}]
[Classe-Sulfures : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]
[Classe-Sulfures : *x] -> (Caractérist) -> [Gisement : {Abondant, Assez-Rare}]
[Classe-Sulfures : *x] -> (Se-compose) -> [Particules]
[Particules] -> (Caractérist) -> [Comp-chimique : ZnS]

Description des taxons du troisième niveau : les variétés de l'espèce Sphalérite

Description de la Marmatite qui est une première variété de la Sphalérite

Type [Variété-Marmatite : (x)] est

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Nom-Variété : Marmatite]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Couleur : {Brun-foncé, Noire}]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Trace : Brune]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Transparence : Opaque]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Gisement : Abondant]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Localisation : {Tchécoslovaquie, Italie, Roumanie, Allemagne, Yougoslavie, Grande-Bretagne}]

Description de la Cléophane qui est une deuxième variété de la Sphalérite

Type [Variété-Cléophane : (x)] est

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Nom-Variété : Cléophane]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Couleur : {Blanche, Blanc-jaunâtre, Jaune-vert, Jaune-miel}]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Trace : Blanche]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Transparence : {Transparente, Translucide}]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Sphalérite : *x] -> (Caractérist) -> [Localisation : {Pologne, Tchécoslovaquie, Roumanie, Espagne, Mexique, États-Unis}]

Discrimination des variétés de l'espèce Sphalérite

Les variétés de l'espèce Sphalérite auraient pu être discriminées grâce aux critères Couleur, Trace ou encore Transparence (même valeur d'entropie pour tous les trois). Le choix de l'expert s'est porté sur le critère de Transparence.

Discrimination des espèces des espèces de la classe des sulfures

Les espèces de la classe des Sulfures ont pu être discriminées à l'aide du critère Système-cristallin puis à l'aide du critère Trace

Description des taxons du deuxième niveau : les espèces de la classe des Halogénures

Nous allons fournir quelques descriptions (7 espèces) de la classe des Halogénures.

Description de l'espèce Calomel

Type [Espèce-Calomel : (x)] est

[Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Calomel]

[Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 1-2]

[Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Blanche, Gris-Jaune, Grise}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : Adamantin]

[Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : {Transparente, Translucide}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 6,4-6,5]

[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Halogénures : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Triclinique]

[Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Allemagne, Espagne, États-Unis, Mexique, U.R.S.S.}]

[Classe-Halogénures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : HgCl]

Description de l'espèce Chlorargyrite

Type [Espèce-Chlorargyrite : (x)] est

[Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Chlorargyrite]

[Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 1,5]

[Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Grise, Jaunâtre, Brunâtre, Noire}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : {Résineux, Adamantin, Mat}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : !Transparente]

[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 5,5-5,6]

[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Halogénures : *x] -> (Caractérist) -> [Trace : {Blanche, Brillante}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Allemagne, États-Unis, Australie, Chili}]

[Classe-Halogénures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : AgCl]

Description de l'espèce Iodargyrite

Type [Espèce-Iodargyrite : (x)] est

[Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Iodargyrite]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 1-1,5]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Jaune, Jaune-verdâtre}]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : {Gras, Adamantin}]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : Transparente]
[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 5,7]
[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : Parfait]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Trace : {Brillante, Jaune}]
[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Hexagonal]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Rare]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Allemagne, Mexique, Australie, Chili}]
 [Classe-Halogénures : *x] -> (Se-compose) -> [Particules]
 [Particules] -> (Caractérist) -> [Comp-chimique : AgI]

Description de l'espèce Carnallite

Type [Espèce-Carnallite : (x)] **est**
 [Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Carnallite]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 2,5]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Blanche, Jaunâtre, Rosâtre, Brunâtre, Vert-clair}]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : {Gras, Vitreux}]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : Translucide]
[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 1,6]
[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : !Clivable]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Trace : Blanche]
[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Abondant]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Allemagne, U.R.S.S.}]
 [Classe-Halogénures : *x] -> (Se-compose) -> [Particules]
 [Particules] -> (Caractérist) -> [Comp-chimique : KMgCl₃.6H₂O]

Description de l'espèce Sylvine

Type [Espèce-Sylvine : (x)] **est**
 [Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Sylvine]
 [Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 2]

[Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Blanche, Jaunâtre, Rougeâtre, Grise, Bleuâtre}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : {Transparente, Translucide}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 1,99]

[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Halogénures : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Allemagne, U.R.S.S., Canada, États-Unis}]

[Classe-Halogénures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : KCl]

Description de l'espèce Halite

Type [Espèce-Halite : (x)] **est**

[Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Halite]

[Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 2]

[Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Blanche, Grisâtre, Rosâtre, Bleuâtre, Violette, Orange}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 2,2]

[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Halogénures : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Allemagne, Autriche, Pologne, U.R.S.S., États-Unis}]

[Classe-Halogénures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : NaCl]

Description de l'espèce Cryolite

Type [Espèce-Cryolite : (x)] **est**

[Classe-Halogénures : *x] -> (Caractérist) -> [Nom-Espèce : Cryolite]

[Classe-Halogénures : *x] -> (Caractérist) -> [Dureté : 2,5-3]

[Classe-Halogénures : *x] -> (Caractérist) -> [Couleur : {Blanche, Grise, Rougeâtre, Brunâtre}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Éclat : {Gras, Vitreux}]

[Classe-Halogénures : *x] -> (Caractérist) -> [Transparence : Transparente]

[Classe-Halogénures : *x] -> (Caractérist) -> [Densité : 2,95]

[Classe-Halogénures : *x] -> (Caractérist) -> [Clivage : Bon]

[Classe-Halogénures : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Halogénures : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Halogénures : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Halogénures : *x] -> (Caractérist) -> [Localisation : {Groenland, Nigéria, U.R.S.S., États-Unis}]

[Classe-Halogénures : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Na₃AlF₆]

Discrimination des espèces de la classes des Halogénures

Les espèces de la classe des Halogénures ont pu être discriminées à l'aide des critères respectifs suivants : Système-cristallin, Clivage et Densité.

Description des taxons du deuxième niveau : les espèces de la classe des Oxydes-Hydroxydes.

Nous allons fournir quelques descriptions (43 espèces et 22 variétés) de la classe des Oxydes-Hydroxydes.

Description de l'espèce Opale

Type [Espèce-Opale : (x)] est

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Opale]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5,5-6,6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Jaune, Rouge, Brune, Verte, Bleue, Noire, Opalescente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras, Mat, Cireux}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 2,1-2,2]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Amorphe]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : {Abondant, Rare}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : SiO₂.nH₂O]

Description des taxons du troisième niveau : les variétés de l'espèce Opale

Description de l'Opale-Noble qui est une première variété de l'Opale

Type [Variété-Opale-Noble : (x)] est

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-Noble]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Opalescente]

[Espèce-Opale : *x] -> (Caractérist) -> [Luminescence : {Verte, Brune}]

[Espèce-Opale : *x] -> (Caractérist) -> [Structure : !Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Opale : *x] -> (Caractérist) -> [Localisation : {Tchécoslovaquie, États-unis, Mexique, îles Feroë, Honduras, U.R.S.S, Brésil, Australie}]

Description de l'Opale-de-Feu qui est une deuxième variété de l'Opale

Type [Variété-Opale-de-Feu : (x)] est

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-de-Feu]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Rouge]

[Espèce-Opale : *x] -> (Caractérist) -> [Structure : !Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Mexique, îles Feroë, Turquie, U.R.S.S, }]

Description de l'Opale-Blanche-de-Lait qui est une troisième variété de l'Opale

Type [Variété-Opale-Blanche-de-Lait : (x)] est

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-Blanche-de-Lait]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Blanche]

[Espèce-Opale : *x] -> (Caractérist) -> [Structure : !Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Tchécoslovaquie, États-unis, Mexique, îles Feroë, Honduras, U.R.S.S, Brésil, Australie}]

Description de la Prasopale qui est une quatrième variété de l'Opale

Type [Variété-Prasopale : (x)] est

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Prasopale]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : Verte]

[Espèce-Opale : *x] -> (Caractérist) -> [Structure : !Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Pologne, Nouvelle-Calédonie, Tanzanie, Tchécoslovaquie, États-unis}]

Description de l'Opale-Ligneuse qui est une cinquième variété de l'Opale

Type [Variété-Opale-Ligneuse : (x)] **est**

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Opale-Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : {Jaune-Clair, Brun-jaune, Rouge, Noire}]

[Espèce-Opale : *x] -> (Caractérist) -> [Structure : Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Cassure : Concoïdale]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : Abondant]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Tchécoslovaquie, États-unis, Roumanie, U.R.S.S., Argentine, Egypte}]

Description de l'Hyalite qui est une sixième variété de l'Opale

Type [Variété-Hyalite : (x)] **est**

[Espèce-Opale : *x] -> (Caractérist) -> [Nom-Variété : Hyalite]

[Espèce-Opale : *x] -> (Caractérist) -> [Couleur : {Bleue, Verte}]

[Espèce-Opale : *x] -> (Caractérist) -> [Structure : !Ligneuse]

[Espèce-Opale : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Espèce-Opale : *x] -> (Caractérist) -> [Gisement : Abondant]

[Espèce-Opale : *x] -> (Caractérist) -> [localisation : {Tchécoslovaquie, Allemagne, Islande, Mexique, îles Feroë, Japon, États-unis, Canada, U.R.S.S., Nouvelle-Zélande}]

Discrimination des variétés de l'espèce Opale

Les variétés de l'espèce Opale ont pu être discriminées grâce aux critères Couleur, puis Structure.

Description de l'espèce Valentinite

Type [Espèce-Valentinite : (x)] **est**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Valentinite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 2-3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Grise, Rouge, Gris-Jaune}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Nacré}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,6-5,8]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Primitif : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, France, Algérie, Bolivie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Sb₂O₃]

Description de l'espèce Sénarmontite

Type [Espèce-Sénarmontite : (x)] est

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Sénarmontite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 2-3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Gris-Blanche}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Soyeux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Transparente, Translucide}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,2-5,3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Algérie, Canada, Tchécoslovaquie, Allemagne, Roumanie, Italie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Sb₂O₃]

Description de l'espèce Calcédoine

Type [Espèce-Calcédoine : (x)] est

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Calcédoine]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Grise, Gris-Verte, Gris-Bleue, Gris-Blanche}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Mat, Soyeux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 2,59-2,61]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Romboédrique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : {Abondant, Assez-Rare, Rare}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : SiO₂]

Description des taxons du troisième niveau : les variétés de l'espèce Calcédoine

Description de la Cornaline qui est une première variété de la Calcédoine

Type [Variété-Cornaline : (x)] est

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Nom-Variété : Cornaline]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Couleur : {Rouge, Jaune-Rouge}]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Gisement : Abondant]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Localisation : {Arabic, Inde, Brésil, Tchécoslovaquie, Roumanie, Egypte}]

Description de la Sardoine qui est une deuxième variété de la Calcédoine

Type [Variété-Sardoine : (x)] est

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Nom-Variété : Sardoine]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Couleur : {Brune, Orange}]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Localisation : {Asie-Mineure, U.R.S.S.}]

Description de la Chrysoprase qui est une troisième variété de la Calcédoine

Type [Variété-Chrysoprase : (x)] est

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Nom-Variété : Chrysoprase]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Couleur : Verte]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Localisation : { Pologne, U.R.S.S., Australie, États-Unis}]

Description de l'Héliotrope qui est une quatrième variété de la Calcédoine

Type [Variété-Héliotrope : (x)] est

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Nom-Variété : Héliotrope]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Couleur : Verte-tachée-de-rouge]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Calcédoine : *x] -> (Caractérist) -> [Localisation : {Autriche, Chine, Allemagne, Grande-Bretagne, Brésil, U.R.S.S., Australie, États-Unis}]

Discrimination des variétés de l'espèce Calcédoine.

Le critère Couleur a pu à lui seul distinguer les variétés de l'espèce Calcédoine.

Description de l'espèce Tridymite

Type [Espèce-Tridymite : (x)] **est**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Tridymite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6,5-7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Blanc-jaunâtre, Grise}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 2,27]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Dimorphe]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {France, Japon, Mexique, États-Unis, Tchécoslovaquie, Allemagne}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : SiO₂]

Description de l'espèce Cristobalite

Type [Espèce-Cristobalite : (x)] **est**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Cristobalite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Blanc-laiteux}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 2,32]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Dimorphe]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {France, Inde, U.R.S.S., Mexique, États-Unis, Tchécoslovaquie, Allemagne}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : SiO₂]

Description de l'espèce Rutile

Type [Espèce-Rutile : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Rutile]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Jaune, Rouge, Brun-rouge, Brun-noir, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Semi-Métallique, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,2-4,3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Brun-jaune, Brun-rouge}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Quadratique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Suisse, Italie, Norvège, Brésil, Mexique, États-Unis, Australie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : TiO₂]

Description de l'espèce Périclase

Type [Espèce-Périclase : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Périclase]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Gris-vert}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 3,7-3,9]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Italie, Suède, États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : MgO]

Description de l'espèce Pseudobrookite

Type [Espèce-Pseudobrookite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Pseudobrookite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brun-foncé, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Semi-Métallique, Adamantin}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,4]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Brun-rouge, Jaune-ocre}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Italie, Suède, États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Fe₂TiO₅]

Description de l'espèce Bixbyite

Type [Espèce-Bixbyite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Bixbyite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brun-foncé, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,9-5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Argentine, Inde, Afrique du Sud, Espagne, Suède, États-Unis, Mexique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique¹² : (Mn, Fe)₂O₃]

Description de l'espèce Braunite

Type [Espèce-Braunite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Braunite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brun-foncé, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Métallique, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire-brunâtre]

**[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin :
Quadratique]**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Inde, Afrique du Sud, États-Unis, Brésil}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : 3Mn₂O₃ . MnSiO₃]

Description de l'espèce Franklinite

Type [Espèce-Franklinite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Franklinite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Métallique, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,0-5,2]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Brun-rouge]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {États-Unis, U.R.S.S.}]

¹² Dans l'écriture des formules en minéralogie, on convient d'utiliser une virgule (,) au sens de ou et un point (.) au sens de et. On doit lire (Mn, Fe)₂ et (Mn .Fe)₂ respectivement (Mn ou Fe)₂ et (Mn₂ et Fe₂).

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (Zn, Mn, Fe²⁺) (Fe²⁺, Mn³⁺)₂ O₄]

Description de l'espèce Euxénite

Type [Espèce-Euxénite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Euxénite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brune, Brun-noir, Brun-olive, Jaune}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Gras, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,6-5,9]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Brun-rouge]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {États-Unis, U.R.S.S.}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (Y, Ce, U, Th, Ca) (Nb, Ta, Ti)₂ O₆]

Description de l'espèce Hématite

Type [Espèce-Hématite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Hématite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brun-rouge, Gris-noire, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Mat, Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Opaque}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,2-5,3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Rouge-cerise]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Rhomboédrique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {États-Unis, U.R.S.S.}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Fe₂ O₃]

Description de l'espèce Pyrolusite

Type [Espèce-Pyrolusite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Pyrolusite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Grise, Gris-noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Mat, Métallique, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Quadratique]*

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Inde, Brésil, Grande-Bretagne, U.R.S.S.}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : MnO₂]

Description de l'espèce Magnésioferrite

Type [Espèce-Magnésioferrite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Magnésioferrite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,6-4,7]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Rouge-sombre]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]*

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {France, États-Unis, Allemagne, Italie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : MgFeO₄]

Description de l'espèce Tantalite

Type [Espèce-Tantalite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Tantalite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6-6,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brune, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Gras, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : !Transparente]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 8,1]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Brune]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Zimbabwe, Suède, Australie, Finlande, Brésil, États-Unis, Allemagne}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (Fe, Mn)Ta₂O₆]

Description de l'espèce Tapiolite

Type [Espèce-Tapiolite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Tapiolite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Fortement-Métallique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 7,3-8]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Brune]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Quadratique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Maroc, Sri Lanka, Autriche, Italie, Zaïre, Finlande}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (Fe, Mn) (Ta, Nb)₂O₆]

Description de l'espèce Thorianite

Type [Espèce-Thorianite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Thorianite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brune, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Semi-Métallique, Résineux}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 10]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Madagascar, Sri Lanka, U.R.S.S.}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : ThO₂]

Description de l'espèce Uranite

Type [Espèce-Uranite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Uranite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Noire, Grise, verte}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Gras, Mat, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 10,6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Brune, Grise, Verte}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {France, Canada, Australie, Zaïre, Tanzanie, Tchécoslovaquie, Allemagne, Suède, Norvège, Afrique du Sud}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : UO₂]

Description de l'espèce Quartz

Type [Espèce-Quartz : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Quartz]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Blanche, Grise, Gris-Vert, Brune, Brun-rouge, Noire, Violet, Verte, Bleue, Jaune, Rose}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence: {Translucide, Transparente, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 2,65]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Hexagonal]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : {Assez-Rare, Rare}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : SiO₂]

Description des taxons du troisième niveau : les variétés de l'espèce Quartz

Description du Cristal-de-Roche qui est une première variété du Quartz

Type [Variété-Cristal-de-Roche : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Cristal-de-Roche]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : Limpide]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: Transparente]

[Espèce-Quartz : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Suisse, France, Italie, Pologne, Tchécoslovaquie, Allemagne, États-Unis, U.R.S.S, Sri Lanka, Madagascar, Brésil}]

Description de l'Améthyste qui est une deuxième variété du Quartz

Type [Variété-Améthyste : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Améthyste]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : Violet]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Roumanie, Tchécoslovaquie, Allemagne, États-Unis, U.R.S.S., Madagascar, Brésil}]

Description du Quartz-Enfumé qui est une troisième variété du Quartz

Type [Variété-Quartz-Enfumé : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Quartz-Enfumé]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : {Brun-clair, Brun-foncé}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Italie, Autriche, Pologne, Tchécoslovaquie, Allemagne, États-Unis, U.R.S.S., Madagascar, Brésil, Australie}]

Description du Quartz-Rose qui est une quatrième variété du Quartz

Type [Variété-Quartz-Rose : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Quartz-Rose]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : {Rose-clair, Rose-vif}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Luminescence : Violacée]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Brésil, Finlande, Tchécoslovaquie, États-Unis, U.R.S.S., Madagascar, Namibie, Afrique du Sud, Japon}]

Description du Morion qui est une cinquième variété du Quartz

Type [Variété-Morion : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Morion]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : Noire]

[Espèce-Quartz : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {!Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Brésil, Finlande, Tchécoslovaquie, États-Unis, U.R.S.S., Madagascar, Namibie, Afrique du Sud, Japon}]

Description de la Citrine qui est une sixième variété du Quartz

Type [Variété-Citrine : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Citrine]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : {Jaune-clair, Jaune-doré¹³}]

¹³ Dans le commerce, ces citrines portent les noms de topaze mère, topaze dorée, topaze de Bahia.

[Espèce-Quartz : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Brésil, Finlande, Tchécoslovaquie, États-Unis, U.R.S.S., Madagascar, Namibie, Afrique du Sud, Japon}]

Description du Quartz-bleu qui est une septième variété du Quartz¹⁴

Type [Variété-Quartz-bleu : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Quartz-bleu]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : Bleue]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Autriche, Brésil, Scandinavie, Afrique du Sud, États-Unis}]

Description de l'Oeil-de-Chat qui est une huitième variété du Quartz

Type [Variété-Oeil-de-Chat : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Oeil-de-Chat]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : Gris-vert]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Birmanie, Inde, Brésil, Sri Lanka, Allemagne}]

Description de l'Aventurine qui est une neuvième variété du Quartz

Type [Variété-Aventurine : (x)] est

[Espèce-Quartz : *x] -> (Caractérist) -> [Nom-Variété : Aventurine]

[Espèce-Quartz : *x] -> (Caractérist) -> [Couleur : Brun-rouge]

[Espèce-Quartz : *x] -> (Caractérist) -> [Transparence: {Transparente, Translucide}]

[Espèce-Quartz : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Quartz : *x] -> (Caractérist) -> [Localisation : {Inde, Brésil, Espagne, U.R.S.S., Chine, Italie}]

Discrimination des variétés de l'espèce Quartz

Encore une fois, le critère Couleur a pu à lui seul, distinguer les variétés de l'espèce Quartz.

¹⁴ Connu aussi sous le nom de Saphir quartzeux.

Description de l'espèce Hercynite

Type [Espèce-Hercynite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Hercynite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 7,5-8]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 3,95]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Vert-foncé]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, Italie, Suède États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : FeAl_2O_4]

Description de l'espèce Cassitérite

Type [Espèce-Cassitérite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Cassitérite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brune, Noire, Jaune, Grise}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, !Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 6,8-7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Blanche, Jaune, Brun-clair}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Quadratique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, Grande-Bretagne, France, U.R.S.S., États-Unis, Malaisie, Thaïlande, Bolivie, Australie, Mexique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : SnO_2]

Description de l'espèce Spinelle

Type [Espèce-Spinelle : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Spinelle]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 8]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Incolore, Jaune, Bleue, Verte, Rouge, Brune, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Vitreux]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 3,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Italie, U.R.S.S., États-Unis, Sri Lanka, Birmanie, Thaïlande, Suède}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : $MgAl_2O_4$]

Description de l'espèce Gahnite

Type [Espèce-Gahnite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Gahnite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 8]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Vert-bleuâtre]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Grise]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, Italie, Bulgarie, Finlande, États-Unis, Brésil, Australie, Suède}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : $ZnAl_2O_4$]

Description de l'espèce Nigérite

Type [Espèce-Nigérite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Nigérite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 8,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brun-clair, Brun-rouge}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : !Transparente]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,51]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Gris-blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Hexagonal]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Nigéria, Portugal, U.R.S.S.}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (Zn, Mg, Fe²⁺)(Sn, Zn)₂(Al, Fe³⁺)₁₂O₂₂(OH)₂]

Description de l'espèce Chrysobéryl

Type [Espèce-Chrysobéryl : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Chrysobéryl]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 8,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Jaune, Vert-clair, Vert-émeraude, Vert-foncé, Brun-vert}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 3,7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Bon]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Italie, Suède, Tchécoslovaquie, Suisse, Finlande, Norvège, U.R.S.S., États-Unis, Canada, Brésil, Inde, Australie, Madagascar, Ghana, Zaïre, Zimbabwe, Tanzanie, Sri Lanka, Birmanie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Al₂BeO₄]

Description de l'espèce Corindon

Type [Espèce-Corindon : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Corindon]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 9]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Incolore, Bleue, Verte, Rouge, Rose-rouge, Brune, Grise, Jaune, Violet}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 3,9-4,1]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Bon]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Rhomboédrique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : {Assez-Rare, Rare}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Al₂O₃]

Description des taxons du troisième niveau : les variétés de l'espèce Corindon

Description du Rubis qui est une première variété du Corindon

Type [Variété-Rubis : (x)] est

[Espèce-Corindon : *x] -> (Caractérist) -> [Nom-Variété : Rubis]

[Espèce-Corindon : *x] -> (Caractérist) -> [Couleur : Rouge]

[Espèce-Corindon : *x] -> (Caractérist) -> [Transparence : Transparente]

[Espèce-Corindon : *x] -> (Caractérist) -> [Luminescence : {Rouge, Rose}]

[Espèce-Corindon : *x] -> (Caractérist) -> [Gisement : Rare]

[Espèce-Corindon : *x] -> (Caractérist) -> [Localisation : {Sri Lanka, Birmanie, Thaïlande, Inde, Chine, États-Unis, Kenya, Tanzanie, Zambie, Angola, Yougoslavie, U.R.S.S.}]

Description du Saphir qui est une deuxième variété du Corindon

Type [Variété-Saphir : (x)] est

[Espèce-Corindon : *x] -> (Caractérist) -> [Nom-Variété : Saphir]

[Espèce-Corindon : *x] -> (Caractérist) -> [Couleur : {Bleue, Bleue-clair, Jaune, verte}]

[Espèce-Corindon : *x] -> (Caractérist) -> [Transparence : Transparente]

[Espèce-Corindon : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Corindon : *x] -> (Caractérist) -> [Localisation : {Sri Lanka, Birmanie, Thaïlande, Inde, Chine, États-Unis, Cambodge, Australie, Kenya, Malawi, Angola, Tanzanie, Zambie, Angola, Tchécoslovaquie, Finlande}]

Description du Leucosaphir qui est une troisième variété du Corindon

Type [Variété-Leucosaphir : (x)] est

[Espèce-Corindon : *x] -> (Caractérist) -> [Nom-Variété : Leucosaphir]

[Espèce-Corindon : *x] -> (Caractérist) -> [Couleur : incolore]

[Espèce-Corindon : *x] -> (Caractérist) -> [Transparence : Transparente]

[Espèce-Corindon : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Espèce-Corindon : *x] -> (Caractérist) -> [Localisation : {Sri Lanka, États-Unis}]

Discrimination des variétés de l'espèce Corindon

Le système a discriminé les variétés de l'espèce Corindon à l'aide du critère Couleur.

Description de l'espèce Minium

Type [Espèce-Minium : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Minium]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 2-3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Rouge-clair, Brun-rouge}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Mat, Gras}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : !Transparente]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 8,2]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Jaune-orangé]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Tétragonal]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Yougoslavie, Grande-Bretagne, Mexique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Pb₃O₄]

Description de l'espèce Lanthinite

Type [Espèce-Lanthinite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Lanthinite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 2-3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Violet-foncé, violette, Violet-pourpre}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Vitreux, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,16]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Brun-violet]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Zaïre, France}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : $UO_2(OH)_2$]

Description de l'espèce Cuprite

Type [Espèce-Cuprite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Cuprite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 3,5-4]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Brun-rouge, Rouge, Grise, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Mat, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 6,15]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Brun-rouge]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Grande-Bretagne, Namibie, États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Cu_2O]

Description de l'espèce Ténorite

Type [Espèce-Ténorite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Ténorite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 3-4]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Gris-acier, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Gras, Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,8]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Noire, Verdâtre}]

**[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin :
Monoclinique]**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Roumanie,
Espagne, Zaïre, Namibie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : CuO]

Description de l'espèce Ramsdellite

Type [Espèce-Ramsdellite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Ramsdellite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 3]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Gris-acier, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Métallique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,37]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire-mat]

**[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin :
Orthorhombique]**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {États-Unis, Tchécoslovaquie,
Turquie, Égypte, Inde}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : MnO₂]

Description de l'espèce Zincite

Type [Espèce-Zincite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Zincite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 4,5-5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Rouge, Rouge-Jaune, Rouge-
brun}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Opaque}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,4-5,7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Jaune-orangé, Brun-jaune}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Hexagonal]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {États-Unis, Allemagne, Pologne, Italie, Yougoslavie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : ZnO]

Description de l'espèce Brannérite

Type [Espèce-Brannérite : (x)] **est** :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Brannérite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 4,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin-Gras, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Opaque}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 6,35]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Brun-vert, Brun-foncé}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Espagne, Maroc, États-Unis, Canada, Australie, Afrique du Sud}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (U, Ca, Th, Y)(Ti, Fe)₂O₆]

Description de l'espèce Bismite

Type [Espèce-Bismite : (x)] **est** :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Bismite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 4,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Jaune-paille, Jaune-vert, Gris-blanc, Gris-jaune}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Terreux}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Transparente}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 9]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Gris-jaune, Jaune-clair}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, Bolivie, U.R.S.S.}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Bi₂O₃]

Description de l'espèce Curite

Type [Espèce-Curite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Curite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 4,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Jaune, Rouge-orangé, Brun-jaune}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : Adamantin]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 7,19]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Bon]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Orange]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Zaïre, Madagascar, Australie}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : 3PbO.8UO₃.4H₂O]

Description de l'espèce Anatase

Type [Espèce-Anatase : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Anatase]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5,5-6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Bleue, Bleu-foncé, Jaune, Rouge, Brune, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Gras, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : {Translucide, Opaque}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 3,8-3,9]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Quadratique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Suisse, France, États-Unis, Afrique du Sud}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : TiO₂]

Description de l'espèce Psilomélane

Type [Espèce-Psilomélane : (x)] **est :**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Psilomélane]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Semi-Métallique, Mat}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,4-4,7]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : {Noire, Brune}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, U.R.S.S., Grande-Bretagne, Brésil, États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : (Ba, H₂O) Mn₅O₁₀]

Description de l'espèce Pyrophanite

Type [Espèce-Pyrophanite : (x)] **est :**

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Pyrophanite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : {Rouge-foncé, Rouge-framboise, Noire}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Adamantin, Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Parfait]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Jaune-ocre]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Rhomboédrique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Tchécoslovaquie, U.R.S.S., Grande-Bretagne, Brésil, États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : MnTiO₃]

Description de l'espèce Ilménite

Type [Espèce-Ilménite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Ilménite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5-6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Gras, Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,5-5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noir-brun]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Rhomboédrique]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Suisse, France, Norvège, U.R.S.S., Grande-Bretagne, Canada, États-Unis}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : FeTiO₃]

Description de l'espèce Jakobsite

Type [Espèce-Jakobsite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Jakobsite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5,5-6]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Vert-noir]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Métallique, Semi-Métallique}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 4,75]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : !Clivable]

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Rouge-Noir]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]*

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Suède, U.R.S.S., Inde, Australie, Afrique du Sud}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : $MnFe_2O_4$]

Description de l'espèce Magnétite

Type [Espèce-Magnétite : (x)] est :

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Nom-Espèce : Magnétite]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Dureté : 5,5]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Couleur : Noire]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Éclat : {Métallique, Gras, Mat}]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Transparence : Opaque]

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Densité : 5,2]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Clivage : Imparfait]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Trace : Noire]*

*[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Syst-Cristallin : Cubique]*

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Gisement : Abondant]

[Classe-Oxydes-Hydroxydes : *x] -> (Caractérist) -> [Localisation : {Allemagne, Finlande, Suisse, Suède, U.R.S.S., Roumanie, Italie, Autriche, Inde, Egypte, Australie, États-Unis, Brésil}]

[Classe-Oxydes-Hydroxydes : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : Fe_3O_4]

Discrimination des espèces de la classe des Oxydes-hydroxydes

Les espèces de la classe des Oxydes-Hydroxydes ont pu être discriminées à l'aide des critères respectifs suivants : Système-cristallin, Trace, Densité, et Clivage.

Description des taxons du deuxième niveau : les espèces de la classe des Substances-Organiques.

Nous allons fournir deux descriptions (espèces) de la classes des Substances-Organiques

Description de l'espèce Évenkite

Type [Espèce-Évenkite : (x)] est

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Nom-Espèce : Évenkite]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Dureté : 1]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Couleur : {Incolore, Blanche, Jaunâtre, Verdâtre, Jaune-vineux}]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Éclat : {Vitreux, Gras}]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Densité : 0,87]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Clivage : En-Feuillets]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Syst-Cristallin : Monoclinique]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Gisement : Rare]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Localisation : {U.R.S.S., Tchécoslovaquie}]

[Classe-Substances-Organiques : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : C₂₄H₅₀]

Description de l'espèce Idrialite

Type [Espèce-Idrialite : (x)] **est**

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Nom-Espèce : Idrialite]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Dureté : 1-1,5]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Couleur : {Jaune-verdâtre, Verte, Grise, Brun-clair, Brun-foncé}]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Éclat : {Vitreux, Adamantin}]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Transparence : Translucide]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Densité : 1,23]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Clivage : En-Feuillets]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Trace : Blanche]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Syst-Cristallin : Orthorhombique]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Gisement : Assez-Rare]

[Classe-Substances-Organiques : *x] -> (Caractérist) -> [Localisation : {Yougoslavie, Tchécoslovaquie, États-Unis}]

[Classe-Substances-Organiques : *x] -> (Se-compose) -> [Particules]

[Particules] -> (Caractérist) -> [Comp-chimique : C₂₂H₁₄]

Discrimination des espèces de la classe des Substances-Organiques

Les deux espèces de la classe des Substances-Organiques ont pu être discriminées à l'aide du critère Système-cristallin.

Conclusion

Après l'explicitation d'une centaine de description de taxons en minéralogie, nous pouvons dégager les points suivants :

La discrimination par IDICE n'a pris en compte que des concepts descriptifs car les seuls concepts structurels¹⁵ qui existent ([Minéral] et [Particules]), sont déjà présents dans le noyau. Donc il ne peuvent être candidats à la discrimination.

Pour les critères de discrimination retenus, l'expert a privilégié le système-cristallin (critère facilement observable), par rapport à la trace, à la densité et au clivage. De même, pour distinguer les variétés, la couleur semble être le critère discriminant dominant par rapport à la structure et à la transparence.

A propos de complétion, celle-ci a été faite dans la mesure du possible, afin d'homogénéiser les descriptions de taxons et de favoriser la recherche de traits discriminants. Toutefois, l'expert n'a pu fournir des informations concernant le clivage (qui est indistinct) des espèces : Euxénite, Ténorite, Ramsdellite et Brannérite appartenant à la classe des Oxydes-Hydroxydes.

Nous pouvons également observer une certaine stabilité dans l'introduction des concepts fournis par l'expert au système. En effet, le système étant incrémental, l'expert fournit à chaque fois un nouveau concept, mais à une étape donnée, l'expert aura fourni tous les concepts nécessaires à la description des taxons, et de ce fait seuls de nouveaux référents sont pris en compte. Cette stabilité s'explique par le fait que les taxons (classes et espèces) sont décrits pratiquement par les mêmes concepts, sauf pour les variétés où l'on constate l'apparition de nouveaux concepts tels que : [Luminescence], [Structure] ou encore [Cassure].

Par ailleurs, il y a eu très peu de redondance dans la base de connaissances, du fait de la diversité des référents associés aux concepts appartenant aux sous-classes d'une même classe. En effet, METIS n'a factorisé que quelques traits comme par exemple : le

¹⁵ Il n'a qu'un seul concept composite, auquel est rattaché un seul concept composant. Les descriptions fournies par l'expert sont toutes "à plat".

trait [Classe-Substances-Organiques : *x] -> (Caractérist) -> [Transparence : Translucide] qui est généralisé au niveau de la classe des Substances-Organiques¹⁶.

Par ailleurs, à la saisie des descriptions, des synonymes portant particulièrement sur des référents associés aux concepts [Couleur], [Trace] et [Clivage] ont été relevés. Par exemple le référent "très-bon" a été remplacé par le référent "parfait" en ce qui concerne le concept [Clivage].

Quelques caractéristiques de la base de connaissances obtenue

Nombre de taxons : 100

Nombre de concepts structurels : 2

Nombre de concepts descriptifs : 17

Nombre de concepts définis : 101 (nombre de taxons plus la racine)

Nombre de relations¹⁷ : 2

Types de référents : les référents (numériques, intervalle) et qualitatifs sont de loin les plus nombreux ils interviennent respectivement dans les concepts [Dureté] [Densité] et dans les concepts [Trace], [Couleur], [Transparence], [Clivage], [Éclat], [Gisement] etc. Les référents nommés interviennent dans les concepts tels que : [Nom-Classe], [Nom-Espèce], [Nom-Variété], [Localisation] etc. Quelques référents négatifs sont associés aux concepts [Clivage], [Structure] et [Transparence], tels que : "!Clivable", "!Ligneuse", "!Transparente".

Temps de construction de la base de connaissances : environ 5 heures. En effet le temps moyen de traitement d'une description est de 3 mn.

En définitive, la méthode METIS était assez inhabituelle pour les minéralogistes car leur démarche est plutôt inverse. Par exemple, ils commencent d'abord par décrire les espèces puis les classes. En effet, il est plus naturel de commencer par décrire des entités concrètes (espèces) avant d'aborder la description des entités plus abstraites (classes). Néanmoins, une approche descendante (avec possibilité de retour arrière) comme celle de METIS se justifie amplement par la prise en compte du mécanisme d'héritage inhérent à la structure hiérarchique des taxons du domaine (taxinomie). Cela optimise la quantité d'interactions avec l'expert, ainsi que le nombre de descripteurs nécessaires à la caractérisation des taxons.

¹⁶ La description du taxon Substances-Organiques a été présentée dans cette annexe avant factorisation.

¹⁷ La relation !Se-compose n'a pas été utilisée.

Cependant, force est de constater, que les classifications de minéralogie sont des édifices artificiels, qui changent d'un expert à un autre et d'une époque à une autre. Depuis quelques décennies, il y a eu évolution des techniques d'identification surtout celles utilisées en laboratoire¹⁸ telles que l'analyse aux rayons X, l'analyse thermique différentielle, l'analyse spectrale, la spectrographie à l'infra-rouge etc. Ces critères de laboratoire sont évidemment à mettre en compétition avec les critères de terrain (présentés dans cette annexe) qui restent malgré tout de bons critères d'identification.

¹⁸ Ces techniques sont très coûteuses.