

VI FORMALISATION INFORMATIQUE DES DESCRIPTIONS : HYPERQUEST

Après avoir défini mathématiquement les objets des descriptions, nous abordons maintenant l'aspect pratique du travail de cette thèse. Nous avons tout d'abord réalisé un outil informatique baptisé HyperQuest qui permet d'acquérir des descriptions robustes pour l'apprentissage. Dans le chapitre 7, nous parlerons d'un autre outil implanté au cours de ce travail (CaseWork) dont l'objectif est de répondre à une meilleure robustesse de la phase de consultation face aux données manquantes.

6.1 Pourquoi HyperQuest ?

La description d'objets est une activité qui intervient à chaque fois qu'il est nécessaire de reconnaître ou d'identifier quelque chose. L'identification se fonde sur l'observation visuelle des objets à distinguer. La description est le processus d'abstraction de ces observations.

HyperQuest est l'outil d'acquisition de connaissances descriptives que nous avons développé pendant cette thèse pour acquérir des descriptions robustes. Cette activité mentale de description pourrait apparaître purement intuitive, mais elle se réfère en fait à un modèle implicite plus ou moins structuré. L'obtention de ce modèle descriptif, pour un domaine donné, est une étape fondamentale dans le processus cognitif qui permet de passer de l'observation à la description.

HyperQuest est l'outil qui permet à l'expert et/ou au cognicien de **construire le modèle descriptif** du domaine. Il s'agit de connaissances observables structurées avec des relations entre les objets qui composent le sujet d'étude. Ces connaissances sont représentées dans un formalisme à base d'objets. Les connaissances initiales sont écrites au format LCRC (Langage Commun de Représentation des Connaissances) ou CASUEL qui est la syntaxe commune de représentation des connaissances par le consortium européen ESPRIT sur le raisonnement par cas (INRECA).

Dans un deuxième temps, HyperQuest sert à **construire un questionnaire** interactif hypertexte. Le questionnaire est généré automatiquement à partir du modèle descriptif. L'expert ou l'utilisateur final du système utilisent celui-ci pour acquérir des descriptions individuelles et constituer une base de cas (aussi au

format LCRC et CASUEL). Les cas sont les connaissances observées du domaine qui seront ensuite traitées par les programmes KATE et CaseWork.

6.2 Particularités d'HyperQuest

HyperQuest se présente sous la forme de plusieurs piles HyperCard et est écrit en HyperTalk™ 2.1 [Apple, 1988] et en langage C. L'intérêt d'utiliser HyperQuest sous HyperCard réside dans le fait de pouvoir bénéficier à la fois des avantages de la programmation par objets et de la programmation hypertexte. Le langage C est utilisé pour accroître les performances de certaines procédures au travers de commandes et fonctions externes à HyperCard (XCMD & XFCN).

Les objets conceptuels composant la structure du modèle descriptif se traduisent directement par des entités hypertextes avec une certaine correspondance (voir § 6.4.2). Les objets, attributs et valeurs correspondent à des formes graphiques différentes (cartes, boutons et champs). Les **cartes** et les **boutons** représentent les objets et les valeurs alors que les **champs** représentent les attributs. Les boutons sont des zones sensibles de la carte que l'utilisateur du questionnaire peut choisir avec un clic de la souris pour déclencher une action. Cela permet par exemple d'aller décrire les objets qu'ils représentent dans une autre carte appelée par le message associé au bouton sélectionné (voir § 6.5.2.1). Les champs sont des zones de stockage des réponses de l'utilisateur collectées par l'intermédiaire des boutons de valeurs.

Le modèle et le questionnaire permettent une navigation aisée pour la description des objets observés qui composent un cas. La description dans le questionnaire se fait sans autre contrainte que celle de l'ordre logique apporté par les connaissances initiales sur le domaine. Les connaissances de fond révélées par l'expert et le cognicien permettent de structurer le déroulement des questions posées à l'utilisateur dans un ordre logique. Le questionnaire résultant est un outil simple et convivial : on peut aisément naviguer entre les cartes et associer des images ou dessins avec les boutons pour illustrer les objets à décrire.

L'un des objectifs d'HyperQuest est de fournir un questionnaire d'acquisition de cas : par exemple "Questionnaire *Hyalonema*" est le nom du questionnaire pour notre application d'identification d'un Genre d'éponges marines (*Hyalonema*) réalisée au MNHN.

HyperQuest génère la description du modèle descriptif et des cas selon la syntaxe LCRC ou CASUEL. La génération de descriptions peut être considérée comme un but en soi pour constituer une base de données de référence. Néanmoins, HyperQuest est intégré avec KATE qui utilise des techniques d'apprentissage automatique pour générer un système expert de détermination, ou CaseWork qui est un système de raisonnement par cas (voir chapitre 7).

HyperQuest est un module particulier du système global d'acquisition de connaissances décrit au chapitre 2 (§ 2.4). Il contient à la fois un générateur de modèle descriptif et un générateur automatique de questionnaire. Chaque questionnaire est issu d'un modèle descriptif. Pour comprendre l'architecture d'HyperQuest sous forme de piles HyperCard, le lecteur intéressé peut se référer à l'annexe 4, cette partie provenant du manuel d'utilisation d'HyperQuest [Conruyt & Dumont, 1993].

6.3 Le générateur de modèle

La première étape de fabrication d'un système expert à l'aide de l'apprentissage à partir d'exemples est de définir les objets du domaine au sein d'un modèle descriptif. Cette phase primordiale n'est pas simple à réaliser. Son objectif est d'organiser et d'explicitier le plus fidèlement possible le modèle conceptuel de l'expert pour son domaine d'étude. L'outil générateur de modèle descriptif d'HyperQuest a pour but de l'aider à modéliser ses connaissances observables, à les délimiter pour qu'elles soient réutilisées ensuite par les autres modules du système d'apprentissage.

6.3.1 Conception d'un modèle descriptif

La conception d'un modèle descriptif suit une certaine logique et quelques règles d'organisation : globalement, le modèle est structuré selon une ou plusieurs hiérarchies qui dépendent les unes des autres (dans le sens du plus général vers le plus particulier). Ces hiérarchies sont des arbres (des graphes sans cycle) définis par les informaticiens sous forme récursive [Knuth, 1968] cité par [Crémilleux, 1991] :

Un arbre est un ensemble fini d'un ou plusieurs nœuds tels que :

il existe un nœud particulier appelé la racine de l'arbre,
les autres nœuds se répartissent en m ensembles disjoints¹ ($m \geq 0$) dont
chacun constitue à son tour un arbre.

Au niveau de la modélisation, les arbres permettent de représenter la connaissance observable sous une forme structurée dont chaque nœud correspond à un objet. Cette connaissance ne doit pas être confondue avec un arbre de décision qui est le fruit du traitement des descriptions (les nœuds sont alors des critères de décision) alors qu'un arbre de modélisation est indépendant de tout traitement.

¹ Les "m ensembles disjoints" correspondent aux m branches issues de la racine.

De même, ces arbres ne doivent pas être assimilés à des taxonomies qui définissent une classification des individus, et dont chaque nœud constitue un taxon (un ensemble d'individus).

Le but des arbres est ici de structurer les descriptions. Cette structuration a lieu dans deux dimensions selon les deux relations les plus naturelles que nous avons mentionnées aux chapitres 3 et 4 : dépendances (relations de composition ou assertions composites) et spécialisations (relations d'héritage ou ordre partiel entre les objets).

L'objet (le composant) se trouve à la jonction de ces deux dimensions : d'un côté, il peut être partitionné. De l'autre, il peut être précisé. Afin de bien faire la distinction entre les deux notions, nous avons choisi de représenter graphiquement l'observable dans deux dimensions correspondant à des fenêtres graphiques différentes (figure 6.1).

Cette exigence permet de répondre à un des buts principaux du modèle descriptif qui est de servir d'ossature à un guide d'observation. Un arbre possède la propriété d'avoir un début (la racine) et une ou plusieurs fins (la feuille) et la "remontée" dans l'arbre ne s'effectue que par un seul chemin. C'est pourquoi nous avons décidé d'éviter une représentation graphique sous forme de réseau sémantique comme dans KL-ONE [Brachman, 1977], [Brachman & Schmolze, 1985] ou de graphe conceptuel [Sowa, 1984] dans laquelle les relations sont mélangées dans un même plan visuel. Cette manière de procéder (que l'on retrouve dans les applications industrielles) est inapplicable dans nos applications biologiques du fait qu'aucun ordre entre relations n'y est apparent.

Or la modélisation des descriptions d'un système naturel fait apparaître un ordre naturel dans les relations entre objets : un individu à décrire se décompose en objets et chaque composant peut être à son tour précisé (si nécessaire). L'individu représente la racine de l'arbre de (dé)composition alors que si l'objet est précisable, il se trouve être à la racine de l'arbre de spécialisation.

C'est pourquoi nous avons choisi de représenter explicitement la relation de composition (ou de dépendance) dans le plan principal. Chaque objet composant ce plan peut être ensuite précisé selon la relation de spécialisation (ou de particularisation). Cette structuration à **deux dimensions "orthogonales"** subordonne la relation de spécialisation à celle de composition (figure 6.1) :

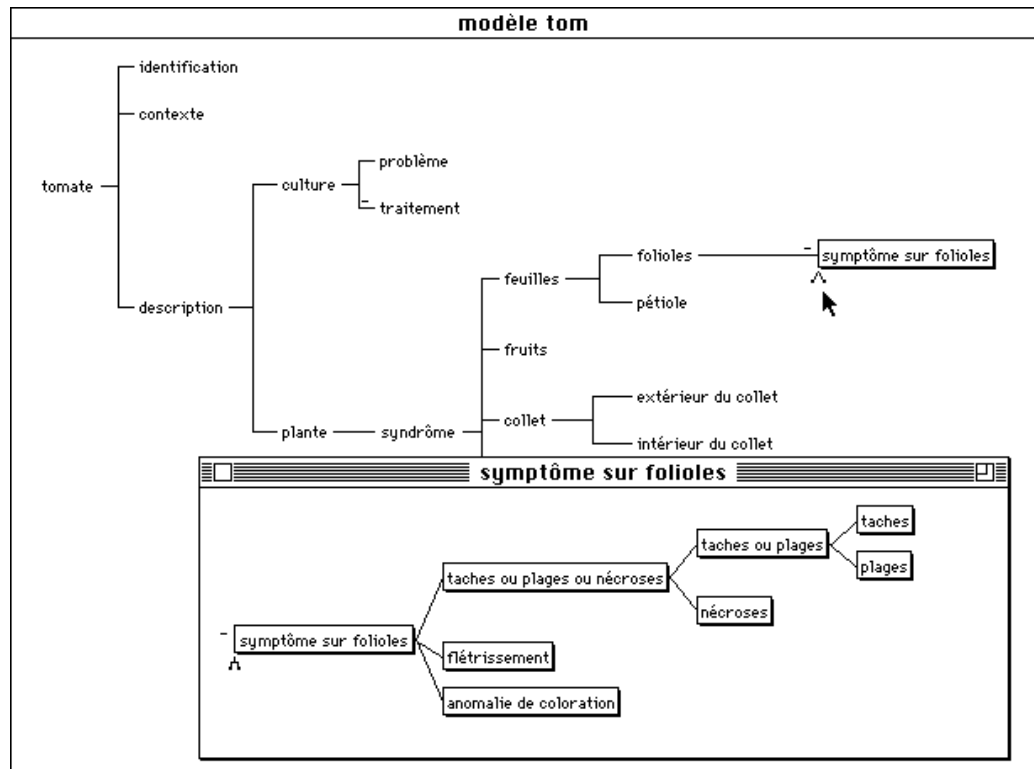


Fig. 6.1 : Vue des deux plans orthogonaux de description des objets en pathologie végétale

Dans l'exemple ci-dessus concernant le système TOM, le plan principal est celui de la décomposition d'un plan de tomate (l'individu situé à la racine de la vue globale) en sous-parties disjointes. Les objets fictifs, représentés sans encadrement, permettent de structurer le domaine, alors que les objets significatifs ont un sens vis-à-vis de l'objectif de détermination des maladies.

Pour que le schéma reste lisible, on n'a représenté que l'objet "symptôme sur folioles" comme significatif, mais tous les autres symptômes non représentés ici sur les autres organes sont aussi significatifs. De même ne sont pas représentés ici les attributs et les valeurs possibles des objets du domaine.

La sélection de l'objet "symptôme sur folioles" provoque l'affichage d'une seconde fenêtre superposée, montrant qu'il est la racine d'un arbre de spécialisation. Pour bien marquer la différence entre les relations de composition et les relations de spécialisation, nous avons symbolisé les premières par des lignes angulaires alors que les secondes le sont par des lignes directes ou obliques (voir figure 6.1).

De plus, la règle d'organisation des objets du plus général au plus particulier est respectée : les objets les plus généraux se retrouvent vers la racine (vers la gauche), quelle que soit la dimension relationnelle à expliciter (dépendance ou spécialisation). On observe ainsi que la généralité de description des symptômes sur foliole peut s'exprimer par des disjonctions de concepts ("taches ou plages" est un concept disjonctif, plus général que "taches" et que "plages").

6.3.2 Règles d'élaboration du modèle descriptif

Les règles pour établir un bon modèle descriptif sont les suivantes :

D'abord, acquérir l'observable est basé sur l'observation de la variabilité de l'observé. Le modèle descriptif n'est pas une vue de l'esprit mais repose sur l'expérience et la familiarité avec le domaine. C'est avant tout un travail de spécialiste.

Ensuite, il faut connaître les objectifs de la modélisation : faire de bonnes descriptions ne suffit pas, il faut savoir dans quel but on les fait :

- 1) classification
 - a) d'individus
 - b) d'objets

- 2) détermination
 - a) d'individus
 - b) d'objets

1a) La classification d'individus est un objectif majeur pour les biologistes au MNHN. C'est une démarche exploratoire de scientifiques qui cherchent à remettre en cause des classifications préexistantes. Les descriptions de spécimens constituent la base de ce travail de classification. Le but est de saisir le maximum de caractères sur les individus, sans *a priori* sur leur utilité pour une classification : l'**exhaustivité des descriptions** d'individus est recherchée.

1b) La classification d'objets ne met pas en œuvre le même modèle descriptif que la classification d'individus. C'est un objectif plus précis qui se focalise sur un objet particulier, et ceci indépendamment de l'individu qui le contient. Il faut pour cela étoffer la description de l'objet à classer par un nombre plus important de caractères propres. Cet objectif répond à un besoin d'**homogénéisation du vocabulaire** dans la communauté des chercheurs du domaine ainsi que de la recherche des homologues entre caractères. La logique d'itération (§ 4.4.6) permet de pallier temporairement les écarts d'interprétation sur le nom réel des objets. Dans notre modèle sur les *Hyalonema*, la classification des amphidiskues est une retombée possible de nos descriptions d'individus (par extraction des sous descriptions de cet objet), mais cela n'était pas le but recherché au départ.

2a) La détermination d'individus était le seul but recherché dans l'approche SEPV de l'INRA (§ 1.1). C'est aussi un objectif au MNHN. La détermination d'individus n'est pas que l'affaire de spécialistes. Elle peut être mise en œuvre par des bétotiens du domaine. Ces utilisateurs n'ont pas toujours la connaissance suffisante pour répondre aux questions du système expert. C'est pourquoi la **redondance des caractères** ne doit pas être écartée lors de la conception du

modèle descriptif afin que les corrélations entre caractères permettent de remplacer ceux auxquels l'utilisateur ne sait pas répondre (caractères alternatifs).

2b) La détermination d'objets n'est pas un problème que nous avons abordé jusqu'à maintenant, du fait qu'il n'existe pas de classification de ces objets reconnue par la majorité des descripteurs. Il s'agirait par exemple de reconnaître en pathologie végétale les limites et les différences d'interprétation selon les observateurs entre les concepts de tache, nécrose, plages, mosaïques, etc.. C'est un sous problème de la détermination des individus (un individu = une plante avec une certaine maladie) qui est néanmoins important car il peut influencer le diagnostic du fait que les utilisateurs n'ont pas la même vision interprétative de ces objets. En l'absence d'une véritable classification de ces objets admise par tous, la mise en œuvre de **dessins explicatifs et de photos** permet de faciliter la compréhension du vocabulaire de l'expert. La logique d'itération peut aussi être appliquée pour ne pas obliger à nommer les symptômes. Malgré son importance pratique (au niveau cognitif), la détermination d'objets est un nouvel axe de recherche dont les retombées ne sont pas directement perceptibles : les concepts visés ne sont ni des maladies, ni des taxons ; ce sont simplement des points de description.

Donc, dans un but de classification et de détermination d'individus, il ne faut pas craindre l'exhaustivité et la redondance des informations qui permettent de former un modèle complet. L'expert doit être capable de synthétiser tout ce qui a été observé sur le domaine pour fabriquer l'observable (§ 4.6.1). Il a dû auparavant analyser beaucoup de spécimens de façon à connaître toutes les situations. Les futures descriptions seront fondées sur le modèle descriptif et il est conseillé à l'expert de bien réfléchir à ce niveau. Conscient des objectifs des descriptions, l'expert peut ensuite se focaliser sur ce qu'elles devraient véhiculer.

L'expert doit résumer l'observé existant en dégagant les faits marquants d'une bonne description. Toutes les informations n'ont pas le même niveau sémantique (objets, attributs et valeurs possibles), les faits les plus marquants correspondent aux objets qui sont eux-mêmes porteurs d'une information locale (attributs et valeurs). L'objet principal est la racine de l'arbre de décomposition (§ 4.6.2). La racine correspond à une entité individuelle. Elle peut être identifiée par un nom, dont nous pensons approprié qu'il identifie le domaine.

Ensuite, le *choix des objets* dépend de la racine, de la nature des relations qu'ils entretiennent (composition ou spécialisation) et de leur niveau de généralité (du plus général au plus particulier) : il faut trouver quelles sont les relations naturelles (de bon sens) entre les objets afin de bâtir une structure de description.

6.3.3 Dépendances

Au préalable, dans toute application, on procède à une décomposition des entités du domaine en objets. Ceux-ci sont reliés par le lien de sous-partie et l'arbre de décomposition obtenu forme ce que l'on appelle une vue globale du domaine. La figure 6.2 montre la vue globale du domaine des *Hyalonema* :

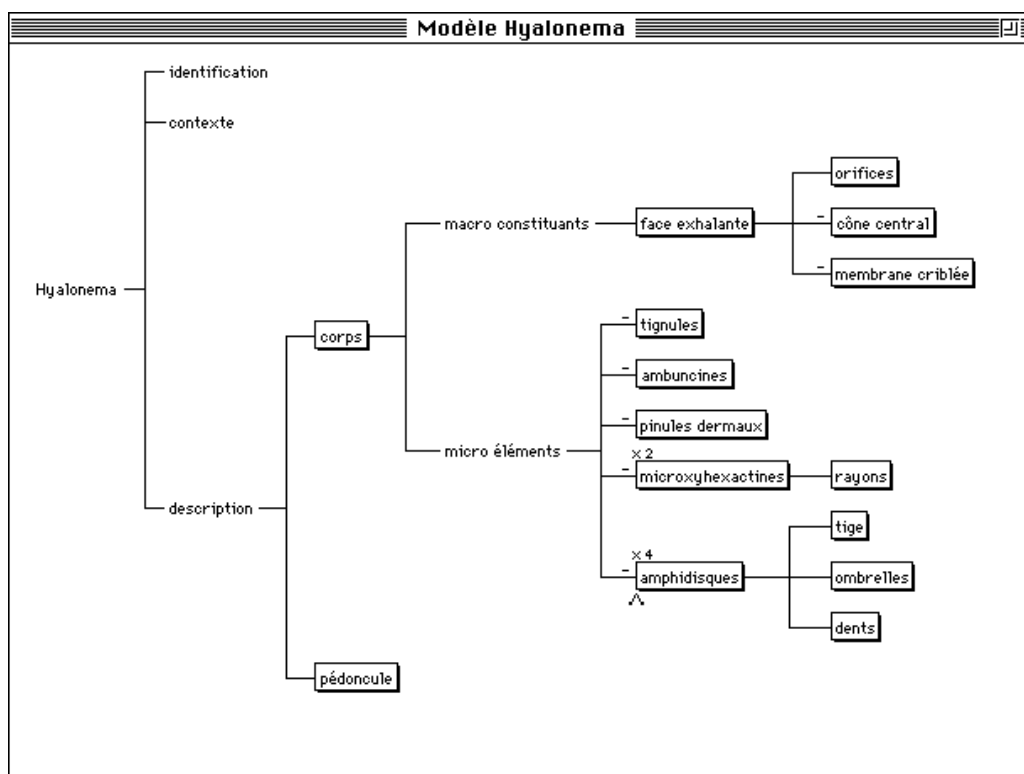


Fig. 6.2 : Vue globale du modèle descriptif des *Hyalonema*

Certains objets ne sont utiles que pour la structuration du domaine, ils sont dits fictifs (par rapport à l'objectif de classification) mais cela ne veut pas dire qu'il faille les négliger pour l'objectif de description lui-même (§ 4.6.5.1.4). Ces objets (représentés sans encadrement dans le modèle ci-dessus) ne se retrouvent que pour la vue globale du domaine et sont positionnés dans la partie gauche de l'arbre de composition : ils correspondent à la logique des points de vue (§ 4.4.3).

En se focalisant sur un objet du modèle descriptif, on peut visualiser localement toutes ses caractéristiques observables. L'exemple de l'objet "amphidisques" est illustré à la figure 6.3. Il s'agit d'un sous-arbre de la vue globale des dépendances avec comme racine l'objet "amphidisques", comme autres nœuds les objets "composants" et comme branches les liens de dépendance. Le dessin qui l'accompagne permet de mieux comprendre la nature de ce micro élément et de faire une comparaison avec l'observable reproduit dans le modèle :

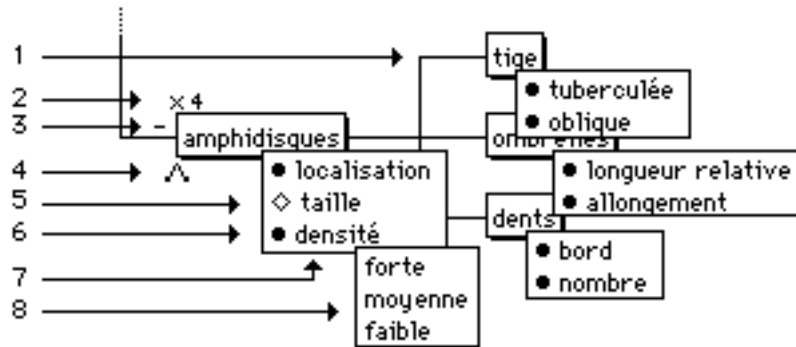
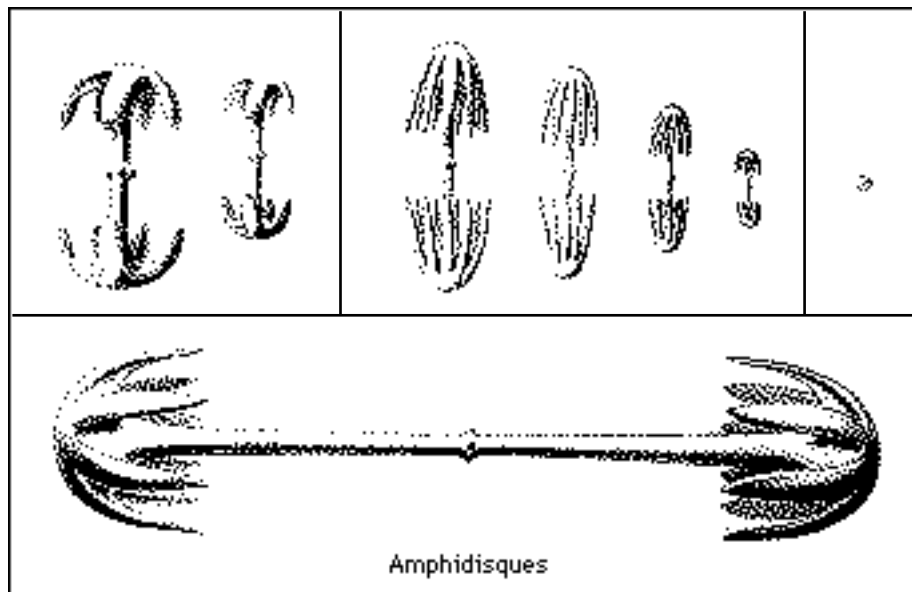


Fig. 6.3 : Vue partielle du modèle descriptif des Hyalonema



1) Cet objet contient trois sous-parties bien distinctes : la tige, les ombrelles et les dents. L'ensemble des objets avec les arcs brisés qui les relient forme le sous-arbre descriptif des amphidisques (logique de décomposition au § 4.6.2).

Remarque : on aurait pu aussi faire dépendre les dents des ombrelles.

2) Cet objet est multi-instanciable quatre fois (symbole $\times 4$). C'est un des statuts de l'objet (§ 4.6.5.1.3) qui permet d'en décrire jusqu'à quatre sortes sans être obligé de les nommer. Le dessin ci-dessus montre les quatre sortes possibles d'objets à décrire chez un spécimen particulier.

3) Les amphidisques peuvent être absents chez certains spécimens. Cette information est importante pour la classification (§ 4.6.5.1.1) et est représentée dans la vue globale par le symbole d'absence possible "-".

4) Le signe \wedge représente un arbre à arcs droits et symbolise ainsi la présence de spécialisations pour un objet. En cliquant dessus, on fait apparaître la vue globale des spécialisations d'un objet : cette action a permis par exemple d'obtenir des précisions sur la nature d'un symptôme sur folioles à la figure 6.1

ou bien encore de visualiser la classification locale des amphidisque entre les figures 6.2 et 6.4.

5) Le symbole “ ” qui précède le nom d’un attribut de l’objet permet de se rappeler que cet attribut n’a pas encore de domaine d’observation associé (phase d’élaboration du modèle descriptif).

6) Inversement, le symbole “•” indique que l’expert a déjà renseigné la liste des valeurs possibles de l’attribut d’un objet.

7) Chaque objet est porteur d’une information locale qui lui est propre comme on peut le voir sur la figure 6.3. Les attributs des objets, lorsqu’ils existent, apparaissent dans des champs accrochés aux nœuds de l’arbre. Ils sont activés par simple clic sur le nom de l’objet. Chaque attribut est indexé à l’objet avec un certain domaine d’observation. Par exemple, la taille est un attribut qui existe aussi bien pour l’objet “amphidisque” que pour l’objet “corps” (cf. figure 4.4) mais leurs domaines d’observation seront différents (le premier s’exprime en microns et le second en mm).

8) En cliquant une fois sur le nom de l’attribut, on fait apparaître son domaine d’observation (s’il a été renseigné). La liste des valeurs possibles vient se superposer à la liste des attributs observables de l’objet (sous l’attribut cliqué).

6.3.4 Spécialisations

La vue globale des dépendances est le plan principal de description des individus. L’observateur est amené à se déplacer entre les objets pour se positionner sur l’un d’entre eux : c’est pourquoi on peut parler de parcours d’un arbre de composition.

Par opposition, un arbre de spécialisation ne se parcourt pas car l’observateur est en train de décrire un objet sur place. Cet arbre donne la possibilité à l’utilisateur de choisir un nœud pour remplacer le nom du composant générique par un autre nom plus précis. C’est donc un **arbre de substitution**. La substitution concerne tout objet de l’arbre de composition que l’expert peut faire préciser. Par conséquence, les arbres de spécialisation sont “orthogonaux” à l’arbre de composition.

Les arbres de spécialisation n’étant pas de même nature que l’arbre de composition, on les visualise dans des fenêtres différentes sous la forme d’une vue globale des choix de substitution possibles.

Dans la figure 6.4, on donne l’exemple de l’objet “amphidisque” spécialisable en objets “macramphidisque”, “mesamphidisque” et “micramphidisque” selon

l'expert. Le résultat est la constitution d'un arbre de spécialisation à un seul niveau contrairement à celui de la figure 6.1.

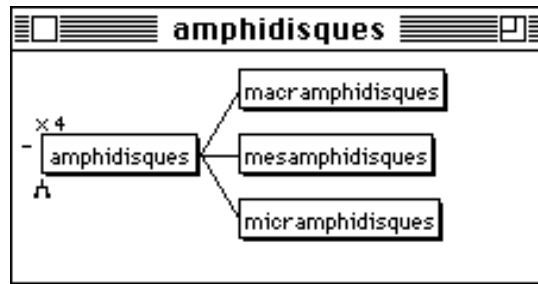


Fig. 6.4 : Vue globale des spécialisations de l'objet "amphidisques"

Les conventions utilisées dans cette vue des spécialisations sont identiques à celles de la vue des dépendances. On notera toutefois que les liens de spécialisation sont représentés par des arcs droits et obliques et on remarquera aussi la présence d'un nouveau sigle ▲ ; il représente un arbre à arcs brisés et indique pour la racine l'appartenance de celle-ci à un arbre de dépendances. Il permet aussi, par un simple clic de faire apparaître la vue des relations de dépendance de cet objet.

6.3.5 Itérations

Lorsque l'expert décèle plusieurs sortes d'un même objet à décrire sur un même spécimen et qu'il ne souhaite pas imposer sa propre classification locale des objets (comme celle qui existe pour les amphidisques ou les symptômes sur folioles sous forme d'arbres de spécialisation), il doit indiquer le nombre maximum d'itérations possibles pour la description d'un objet (objet multi-instanciable).

L'itération ne se caractérise pas par un arbre reliant plusieurs objets entre eux mais par une boucle sur un seul objet qui peut être instancié plusieurs fois (voir le schéma de la structure du modèle descriptif p. 116). Il s'agit d'une troisième dimension de description d'un objet, que nous pouvons schématiser comme un empilement de descriptions ponctuelles.

Afin de mieux comprendre l'itération, il faut se mettre dans la situation au moment où le questionnaire est utilisé pour faire une description. Au départ, on a un modèle d'objets composites ① qui constitue la 1^{ère} dimension horizontale du processus de description et que l'on va utiliser de la manière suivante :

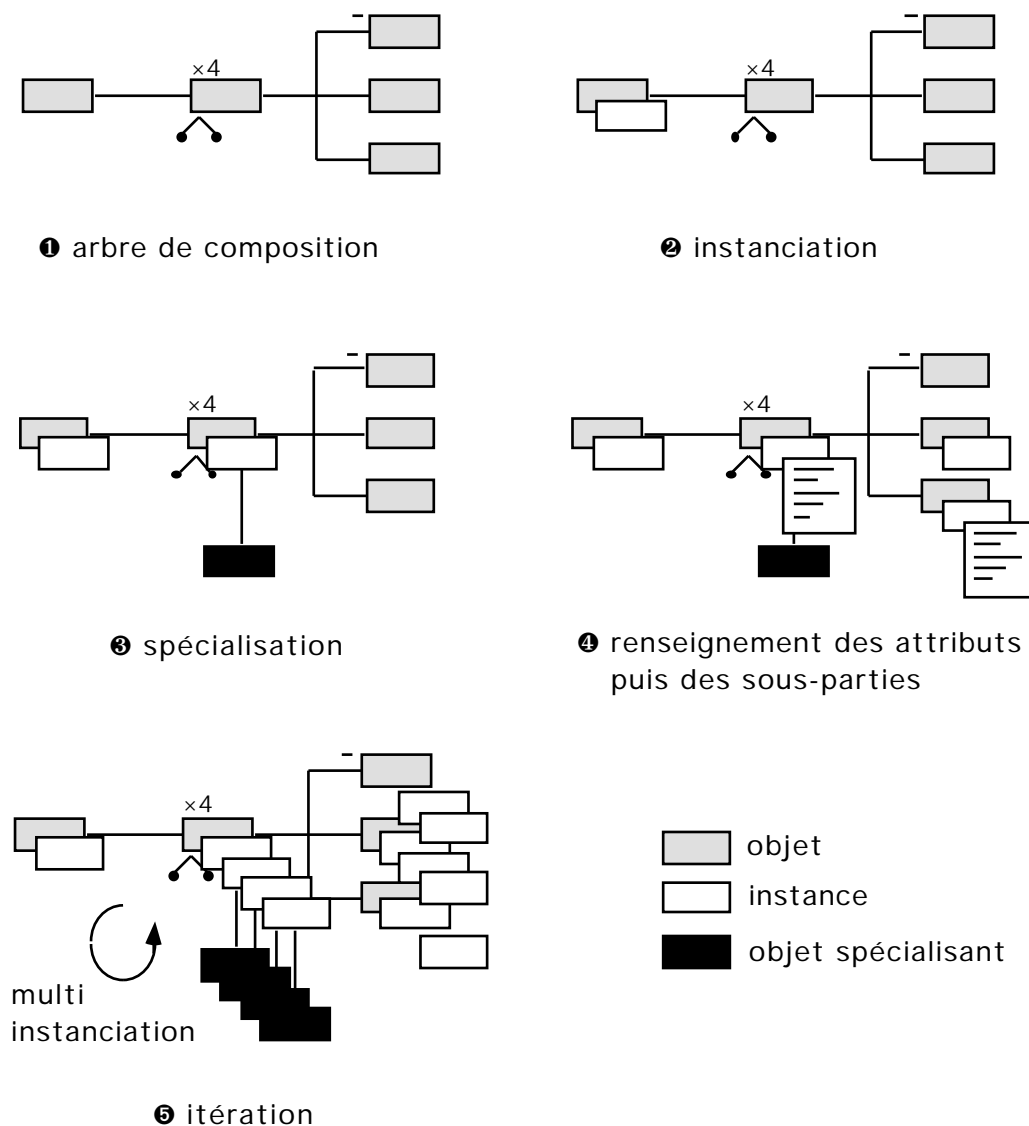


Fig. 6.5 : Processus de description d'un spécimen dans le questionnaire

Lorsque l'on veut faire une nouvelle description, on se situe à la racine. L'instanciation de cet objet est provoquée par la décision de le décrire.

La description de l'objet fait suite à l'observation du spécimen. Néanmoins, le modèle descriptif impose un certain ordre pour décrire les objets (selon un certain parcours du plus général au plus particulier). De plus, au niveau de chaque objet à décrire, le modèle donne des conditions pour ce qui est observable sur chacun d'eux. Par exemple, dans le modèle ci-dessus, l'objet racine ne peut être que présent, il ne peut pas être spécialisé, renseigné localement par des attributs ou bien encore multi-instancié comme l'objet lui succédant. C'est en confrontant l'observation à ce qui est possible d'observer que l'on produit une description d'objet.

La procédure de description d'un objet composite implique la navigation dans le questionnaire par effet de bord. Cette procédure est celle de consultation du questionnaire décrite par l'algorithme suivant :

Décrire (racine)

Fin

Procédure Décrire (objet)

si objet inconnu alors Fin procédure

si objet absent et si statut = "absence possible" alors Fin procédure

② Instancier(objet) objet_i

③ si spécialisation[objet_i] et si statut = "spécialisable" alors

Substituer(objet_i) objet_i = spécialisation[objet_i]

fin si

④ tant que attribut[objet_i] alors Renseigner (attribut[objet_i])

si objet non terminal alors Décrire (sous-partie d'objet)

⑤ tant que autre-sortie d'objet et si statut = "multi-instanciable" alors

Instancier(autre-sortie)

fin tant que

Fin Procédure

La plupart des objets de l'arbre de composition seront instanciés une seule fois lors des descriptions. Après, ils pourront ou non être remplacés par un objet spécialisant (2^{ème} dimension verticale de la description). Ensuite, s'il existe des attributs de l'objet à décrire, ceux-ci seront renseignés avant de s'intéresser à la description de ses sous-parties. Enfin, une itération permettra de boucler plusieurs fois sur la description de l'objet ; chacune des instances produites pourra à son tour être spécialisée si nécessaire.

Nous pouvons remarquer que l'itération est un processus d'instanciation (3^{ème} dimension diagonale de la description) qui est toujours mis en œuvre au moment du passage de l'observable à l'observé : quand cette itération est unique, elle correspond à une instanciation simple (l'objet se transforme en instance d'objet par copie différentielle [Bailly *et al.*, 1987]). Quand elle est multiple, nous l'appelons une multi-instanciation d'objet et le résultat correspond à la notion de horde [Diday, 1991] composite [Conruyt *et al.*, 1992].

Donc, pour constituer un modèle robuste de l'observable, il faut avoir en mémoire l'ordre prédéterminé des différentes phases du processus de description de l'observé (figure 6.5) :

- 1) instanciation d'un objet,
- 2) spécialisation d'un objet,
- 3) renseignements sur l'objet,
- 4) itération sur l'objet.

Ce fonctionnement se réfère à tout ce qui est observable sur un spécimen. Nous préconisons que ce modèle soit structuré sur papier afin de trouver les objets et les relations qu'ils entretiennent entre eux, ainsi que leurs attributs descriptifs et les valeurs possibles. En effet, la philosophie de l'éditeur de modèle n'est pas d'être un gestionnaire d'idées (les objets eux mêmes) que l'on relierait entre elles sur un brouillon pour découvrir les relations entre objets. Ces objets traduisent une réalité concrète avec un certain ordre dans le processus de description.

HyperQuest n'est pas destiné à construire un réseau sémantique avec des objets et des relations de toute nature placés un peu au hasard. La finalité est ici de construire un questionnaire "guide d'observation", ce qui implique une démarche analytique du plus général au plus particulier avec une organisation logique des objets et des relations.

Chaque nœud de l'arbre (un objet) doit être apprécié selon son homogénéité de description afin de ne décrire complètement qu'une seule chose à la fois avant de passer à la description d'un autre objet. Une fois la structure préétablie, l'expert aidé ou non du cognicien va pouvoir saisir son modèle de l'observable à l'aide des éditeurs d'objets, d'attributs et de valeurs.

6.3.6 Edition d'un objet

Les vues globales ont pour principe de représenter le modèle graphiquement sous la forme d'arbres. Elles permettent ainsi de visualiser non seulement tout ou une partie des objets mais surtout les liens entre ces objets. Les vues locales, quant à elles, permettent d'obtenir avec plus de précision des renseignements sur chacun des objets. Une fenêtre entière sous forme de carte HyperCard leur est consacrée (figure 6.6) qui est activée dès que l'utilisateur double-clique sur un des objets de la vue globale :

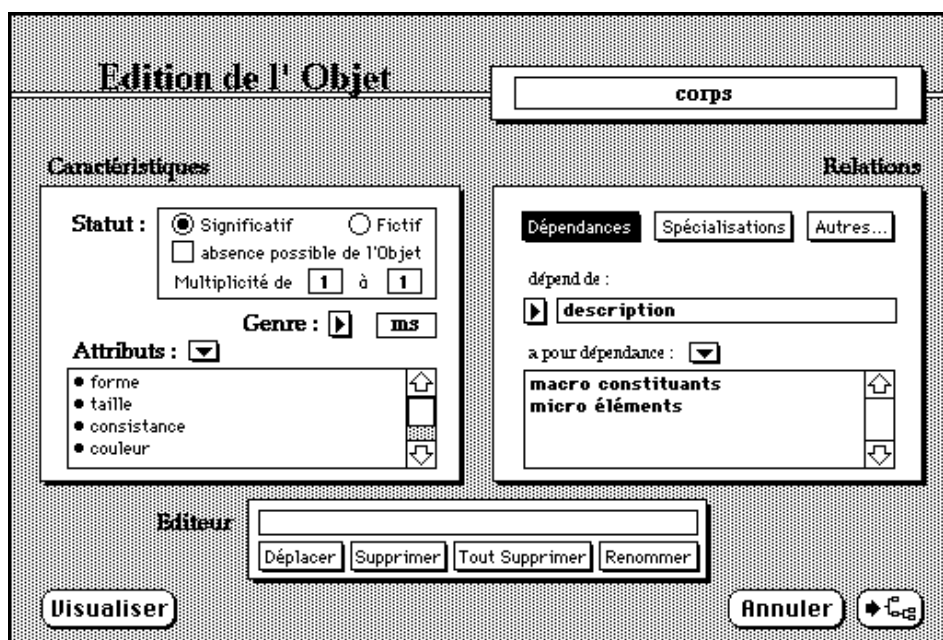


Fig. 6.6 : Fenêtre d'édition de l'objet "corps" de Hyanonema

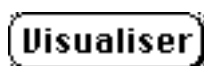
L'éditeur d'objet se décompose en quatre zones :

- ❶ Une première zone en haut à droite nous permet d'éditer le nom de l'objet.
- ❷ Une seconde zone dans le cadre à gauche donne tous les renseignements concernant les caractéristiques de l'objet observable en cours d'édition :
 - son statut (définition § 4.6.5.1),
 - son genre (définition § 4.6.5.3.1),
 - ses attributs (définition § 4.6.5.2).
- ❸ Une troisième zone à droite permet d'éditer les renseignements concernant les relations de l'objet avec d'autres objets. Les boutons "dépendances" et "spécialisations" permettent respectivement d'afficher les dépendances et les spécialisations de l'objet. Le bouton "Autres" n'est là que pour indiquer

la possibilité ultérieure de définir ses propres relations. Cette fonctionnalité n'existe pas dans cette version d'HyperQuest.

- ④ Une quatrième zone en bas de la carte permet d'éditer les éléments des listes d'attributs, de dépendances et de spécialisations et d'effectuer les corrections nécessaires qui se répercuteront dans les vues globales. Ces fonctionnalités s'intitulent "Déplacer" pour modifier la présentation des objets et des attributs, "Supprimer" pour ôter un élément d'une liste, "Tout supprimer" pour effacer la liste entière et enfin "Renommer" pour modifier le nom d'un élément d'une liste. Toutes ces fonctionnalités sont décrites dans le manuel d'utilisation d'HyperQuest [Conruyt & Dumont, 1993].

La carte d'édition d'un objet comprend aussi en plus des quatre zones précédemment décrites des boutons permettant de naviguer dans l'éditeur de modèle :



Donne la possibilité à l'expert de voir directement la transformation de la définition conceptuelle de l'objet en une carte du questionnaire telle qu'il sera généré par la suite. L'action de ce bouton est de construire en parallèle à l'édition de l'objet la carte des entités hypertextes correspondant uniquement à la vue de cet objet. Les figures 6.12 illustrent le résultat de cette action pour l'objet "corps". La carte peut alors être personnalisée directement sans attendre la phase de génération automatique de tout le questionnaire (§ 6.4.1). L'expert peut par exemple importer dans la carte un dessin schématique du corps d'une éponge pour expliquer la différence entre les macro constituants et les micro éléments ; il peut aussi indiquer les caractéristiques du microscope (grossissement) dont il faudrait disposer pour observer les micro éléments. Si l'expert possède le souci d'expliquer par un dessin et des messages les termes qu'il emploie, la robustesse du système n'en sera que meilleure. La personnalisation des cartes du questionnaire se justifie donc pleinement au niveau de l'acquisition de l'observable.



Annule toute action menée depuis l'accès à la carte. En effet, les opérations de suppression, renommage, etc. ne seront pas prises en compte. L'erreur est un droit fondamental de l'expert qui peut décider à tout moment de modifier son modèle descriptif ou de ne pas tenir compte d'informations non encore répertoriées (enregistrées dans la structure du modèle).



Par contre, le fait de cliquer ce bouton valide les informations saisies dans la carte et retourne à l'une des vues globales de dépendances ou de spécialisations.

6.3.7 Edition d'un attribut

De même que pour les objets, chaque description d'attribut nécessite la mise à jour d'informations sur une carte dont voici un exemple :

Fig. 6.7 : Fenêtre d'édition de l'attribut "forme" du corps de Hyalonema

On accède à cette carte par deux moyens : soit à partir des vues globales en double-cliquant sur le nom d'un attribut ou bien à partir de la vue locale d'un objet en double-cliquant aussi sur un attribut.

La carte "attribut" se décompose en trois zones :

- ❶ La première zone en haut à droite nous renseigne sur le nom de l'attribut.
- ❷ Une seconde zone nous indique le nom de l'objet auquel cet attribut est rattaché. Ce champ comme celui du nom de l'attribut est inaccessible depuis la carte d'édition d'un attribut (on ne peut modifier les noms de l'objet et de l'attribut qu'à partir de la vue locale de l'objet). Cette caractéristique montre que la vue locale d'un attribut est subordonnée à celle de son objet, comme cela paraît souhaitable.

③ Une troisième zone donne de multiples renseignements sur l'attribut :

Son type ... (§ 4.6.5.2.1)

Huit boutons permettent de définir le type d'un attribut. Six d'entre eux sont exclusifs : un attribut est soit entier, soit nominal, soit commentaire, soit réel, soit classifié, soit booléen.

Ses valeurs possibles ... (§ 4.6.5.2.2)

Si le type de l'attribut est nominal, commentaire, classifié ou booléen, les valeurs possibles sont éditables dans des champs défilants accessibles directement par simple clic (figure 6.7). Les espaces ne sont pas autorisés dans la saisie de valeurs nominales, ils doivent être remplacés par des tirets "-". Les valeurs "oui" "non" sont mises par défaut dans le cas d'un attribut booléen, dans les autres cas il s'agit du caractère "*" qui correspond à n'importe quelle chaîne de caractères.

Dans le cas d'un type "classifié" un triangle noir apparaît à la base de la liste des valeurs possibles ; il s'agit d'un bouton qui (figure 6.7), lorsque l'on clique dessus, donne accès à la carte d'édition des valeurs classifiées (§ 6.3.8).

Si le type de l'attribut est entier ou réel, les valeurs possibles sont saisies sous la forme d'un intervalle de valeurs dont on précise l'unité. Ces trois informations (valeur minimale, valeur maximale, unité) sont saisies dans des champs d'édition :

Valeurs possibles :

de	<input type="text" value="1"/>
à	<input type="text" value="200"/>
Unité :	<input type="text" value="mus"/>

Fig. 6.8 : Les champs d'édition d'un attribut numérique

Sa valeur par défaut ... (§ 4.6.5.2.3)

Il s'agit d'un champ dans lequel l'utilisateur est libre de saisir ce qu'il veut comme valeur par défaut de l'attribut.

Son choix ... (§ 4.6.5.2.4)

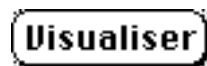
Ce champ n'apparaît que lorsque l'un des types "nominal" ou "classifié" a été choisi auparavant. Deux cases de saisie permettent de désigner le

nombre minimal et maximal de choix possibles pour un choix multiple. Les boutons “unique” et “multiple” sont exclusifs.

Sa question associée ... (§ 4.6.5.3.2)

La question associée à un attribut est saisie dans un champ défilant où l'utilisateur est libre d'inscrire ce qu'il désire.

La carte d'édition d'un attribut comprend aussi en plus des trois zones précédemment décrites des boutons permettant de naviguer dans l'éditeur de modèle :



Comme pour un objet, cette action donne la possibilité à l'expert de voir directement la transformation de la définition conceptuelle de l'attribut en une carte du questionnaire. La figure 6.17 montre le résultat de cette action pour l'attribut “forme” du corps de la Hyalema. La carte peut alors être personnalisée en y ajoutant des dessins explicatifs (figure 6.26).



Permet l'annulation de toute action menée depuis l'accès à la carte.



Accède à l'édition de l'objet auquel est rattaché cet attribut.



Valide les informations saisies dans la carte et retourne à l'une des vues globales de dépendances ou de spécialisations.

6.3.8 Edition des valeurs classifiées

Dans le cas d'un attribut classifié, la description des valeurs possibles nécessite la mise à jour d'informations sur une carte dont voici un exemple dans une classification à deux niveaux :

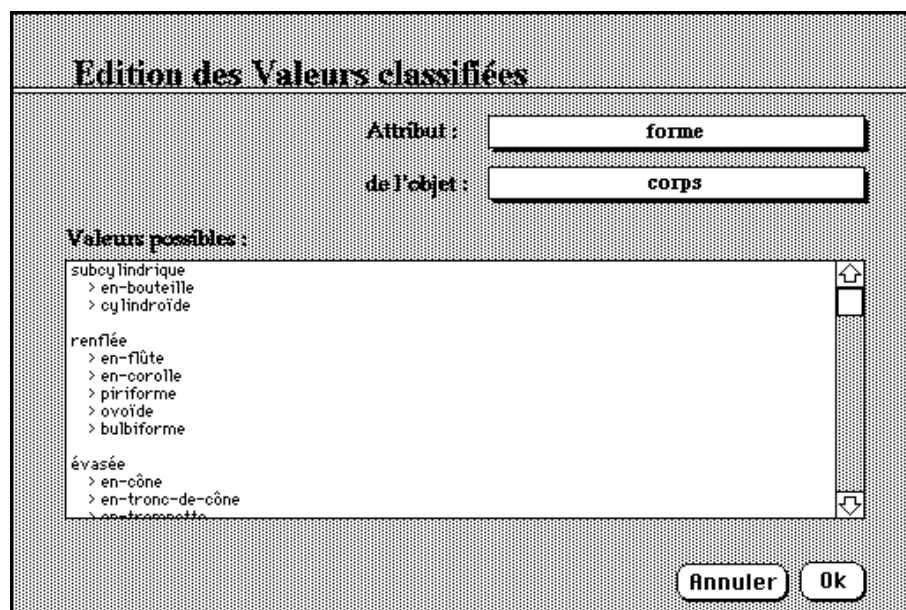


Fig. 6.9 : Fenêtre d'édition des valeurs classifiées de la "forme" du corps de Hyalonema

La carte "valeurs classifiées" se décompose en trois zones :

- ❶❷ une première et seconde zone indiquent respectivement le nom de l'attribut auquel les valeurs se réfèrent et le nom de l'objet possédant l'attribut. Ces champs sont inaccessibles en écriture depuis la carte des valeurs classifiées.
- ❸ une troisième zone sous la forme d'un champ défilant permet la saisie des valeurs classifiées selon un format d'indentation en cascade utilisant les tabulations particulières ">".

6.4 Le générateur automatique de questionnaire

Le générateur de questionnaire est un module HyperQuest qui permet de passer de l'éditeur de modèle à l'éditeur de cas (le questionnaire). Il est un élément charnière qui transforme un modèle de l'observable en un questionnaire permettant d'éditer l'observé (cf. figure 2.4).

6.4.1 Principe de construction d'un questionnaire

Pour fabriquer le questionnaire, le générateur exploite la correspondance existant entre les concepts de description (objets, attributs, valeurs) et les objets hypertextes : cette correspondance est expliquée au § 6.4.2.

Le générateur de questionnaire est formé d'une seule pile HyperCard :



Fig. 6.10 : La pile de génération du questionnaire

Cette pile sert de moule à la fabrication d'un questionnaire <Questionnaire>. Elle ne contient en principe qu'une seule carte au départ (celle de la figure 6.10). La pile est d'abord dupliquée et prend le nom du domaine d'application (par exemple <Questionnaire *Hyalonema*>). La copie est ensuite développée automatiquement par des procédures hypertextes (programmées en HyperTalk) qui font s'accroître le nombre de cartes d'autant qu'il y a d'objets, d'attributs et de valeurs hiérarchisés décrits dans le domaine descriptif.

Les cartes générées du questionnaire portent les noms <Objet> pour les objets, <Attribut(Objet)> pour les attributs, <Valeur(Attribut)> pour les valeurs classifiées (1^{er} niveau), <SousValeur(Valeur)> pour les valeurs classifiées (2^{ème} niveau), etc..

La durée de création de la pile <Questionnaire> dépend du nombre d'objets, d'attributs et de valeurs à traiter et du type d'ordinateur. A titre d'indication, il faut 6 mn à HyperQuest sur un Macintosh PowerBook 170 avec 8 Mo de mémoire vive pour créer un questionnaire de 74 cartes comprenant 25 objets, 40 attributs et 98 valeurs. Ce temps correspond à la création de toutes les entités hypertextes (cartes, boutons, champs) de la pile du questionnaire.

Néanmoins, si des objets ou des attributs ont été visualisés auparavant dans l'éditeur de modèle, le générateur de questionnaire a déjà construit la carte correspondante du questionnaire dans sa propre pile. De ce fait, la pile possède déjà un certain nombre de cartes prêtes pour le questionnaire. Au moment de la création en série de toutes les cartes par le choix "Créer le questionnaire", le générateur de questionnaire possède donc déjà certaines cartes d'objets, d'attributs ou de valeurs classifiées qu'il n'a pas besoin de fabriquer.

Une fois le questionnaire du domaine engendré, le générateur de questionnaire reprend sa taille initiale (une carte) au moment de la fermeture de la pile.

6.4.2 Entités conceptuelles et hypertextes

Pour pouvoir construire automatiquement un questionnaire à partir du modèle descriptif, il faut définir les transformations possibles entre les entités conceptuelles et les entités hypertextes.

Comme nous l'avons déjà dit au chapitre 3, les connaissances observables sont représentées dans un formalisme à base de frames (§ 4.5). Le formalisme informatique que nous utilisons pour représenter le modèle descriptif est différent du formalisme mathématique défini au chapitre 5. Ce dernier s'intéresse plus aux données elles-mêmes dont il veut formaliser la complexité (objets booléens, modaux, probabilistes, possibilistes, etc.). Or, si la complexité des données est une chose à ne pas négliger, nous souhaitons néanmoins bien distinguer la représentation de l'observable (le modèle descriptif) de la représentation de l'observé (les cas) afin d'assurer la cohérence du dernier par rapport au premier.

Par exemple, les règles chez les objets munis de méthodes et de propriétés (§ 5.4) ne font pas partie de l'observé mais de l'observable. Elles interviennent lors de l'établissement des descriptions pour maintenir leur cohérence, mais n'ont plus lieu d'être indiquées une fois celle-ci assurée.

D'autre part, nous avons choisi de nous placer dans un cadre d'objets booléens² plutôt que dans celui des objets modaux. Ce choix nous semble plus proche de la pratique concrète pour représenter des descriptions de spécimens : en les observant, l'expert sait décider si oui ou non ils appartiennent à telle classe, ou encore si telle instance d'un composant de l'individu appartient à tel objet. L'expert n'a pas d'état d'âme en décrivant des individus et ne perçoit pas bien l'intérêt de pondérer ses choix à ce moment par des incertitudes. Cette manière de décrire avec des modalités telles que couleur(corolle) = *souvent* rouge, *rarement* blanche correspond plus à l'approche par "descriptions" de concepts [Vignes, 1991], [Lebbe, 1991]. Cette approche est à nos yeux une manière de faire des descriptions potentielles (ce n'est pas de l'observé qui est décrit mais de l'observable qui est défini).

Pour représenter l'observable, nous nous baserons donc sur les langages de frames dont l'opérationnalité a déjà été largement éprouvée avec KRL [Bobrow & Winograd, 1977], FRL [Roberts & Goldstein, 1977], SHIRKA [Rechenmann, 1985], [Aguirre, 1989] ou KATE [Manago, 1988].

² Objets dont la valeur de vérité est soit vraie soit fausse.

Le formalisme informatique est écrit en LCRC ou CASUEL. Il utilise des mots clé tels que “defObject”, “defSlot” et “defValue” qui permettent de repérer les entités conceptuelles et les mots “subpart” et “spec” pour définir les relations entre objets. Cette connaissance modulaire est écrite selon une syntaxe BNF (voir annexe 4), puis est interprétée dans HyperQuest pour être convertie en entités hypertextes au moment de la création du questionnaire. Les correspondances définies sont les suivantes :

modèle descriptif ----->	pile <Questionnaire>
objet ----->	carte de la pile <Questionnaire>
attributs d'un objet----->	champs de la carte <Objet>
attribut de l'objet----->	item (ligne) dans le champs de carte <Objet>
composants d'un objet ----->	boutons de la carte <Objet>
spécialisations d'un objet ----->	boutons de la carte <Spécialisations(Objet)>
valeurs d'attribut qualitatif ----->	boutons de la carte <Attribut(Objet)>
valeurs d'attribut numérique ---->	boutons curseur de la carte <Attribut(Objet)>
valeurs d'attribut commentaire -->	boite de dialogue de carte <Attribut(Objet)>
règles et contraintes ----->	scripts liés aux champs et boutons de carte.

Les transpositions définies ci-dessus sont valables pour n'importe quel type d'application, ce qui rend le générateur de questionnaire indépendant du domaine étudié (généricité inter-applications).

De plus, chaque entité hypertexte (cartes, boutons et champs) est porteuse d'un message prédéfini qui conservera un même type d'action dans le questionnaire (par exemple une procédure de navigation). Ces message sont constitués une fois pour toute et sont reproduits à chaque création d'une instance d'un objet hypertexte (généricité intra-application).

Nous allons maintenant illustrer ces correspondances à l'aide du questionnaire généré pour le domaine des *Hyalonema*.

6.5 L'éditeur de cas : le questionnaire³

Dans le paragraphe précédent, nous avons défini les liens entre entités conceptuelles et entités hypertextes. Ce choix a été fait à la suite de nombreuses études sur le terrain pour proposer un questionnaire convivial et ergonomique aux utilisateurs dans le domaine de la biologie. Nous sommes convaincus que ces qualités ne peuvent qu'améliorer la robustesse des descriptions à acquérir. Nous exposons donc dans cette partie de la thèse les caractéristiques et fonctionnalités attendues d'un bon questionnaire d'acquisition d'exemples pour l'apprentissage.

Caractéristiques

6.5.1 Représentation de l'individu

L'individu observable est représenté dans le modèle descriptif à la racine de l'arbre de composition. Il est censé identifier le domaine. Le modèle descriptif est donc assimilable à la pile du questionnaire dont la première carte représente l'individu. On accède directement à cette carte en entrant dans le questionnaire :

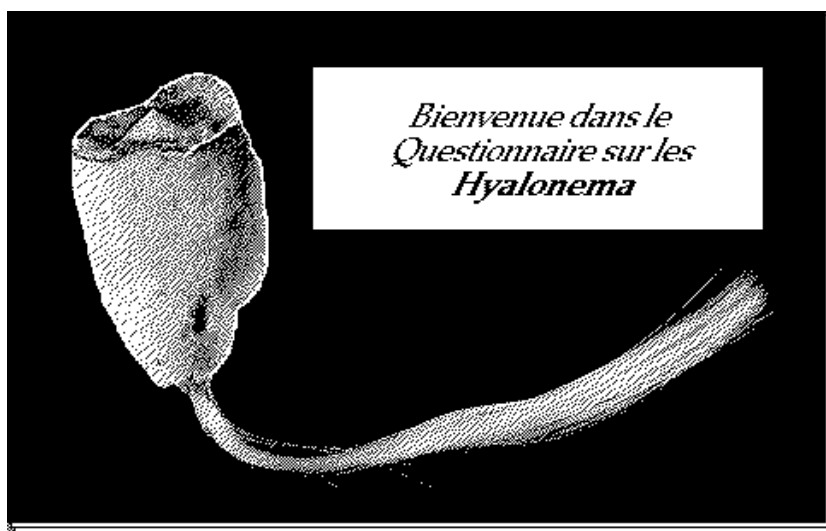


Fig. 6.11 : La première carte de la pile du Questionnaire Hyalonema

³ Cette partie reprend certaines informations contenues dans [Conruyt & Dumont, 1993].

6.5.2 Représentation des objets de l'individu

6.5.2.1 Un objet = une carte

Il y a autant de cartes dans la pile que d'objets à décrire. Chaque carte (objet) possède ses propres caractéristiques et des relations avec d'autres cartes de la pile. Les caractéristiques d'un objet donné apparaissent dans un champs de la carte alors que les liens relationnels sont au niveau des boutons de cette carte.

L'exemple suivant montre la correspondance entre la description au format LCRC de l'objet "corps" de l'éponge marine et la carte de la pile du questionnaire (figure 6.12) :

```
(defobject corps
  (gender ms)
  (subpart macro-constituants micro-elements)
  (slot forme taille consistance couleur))
```

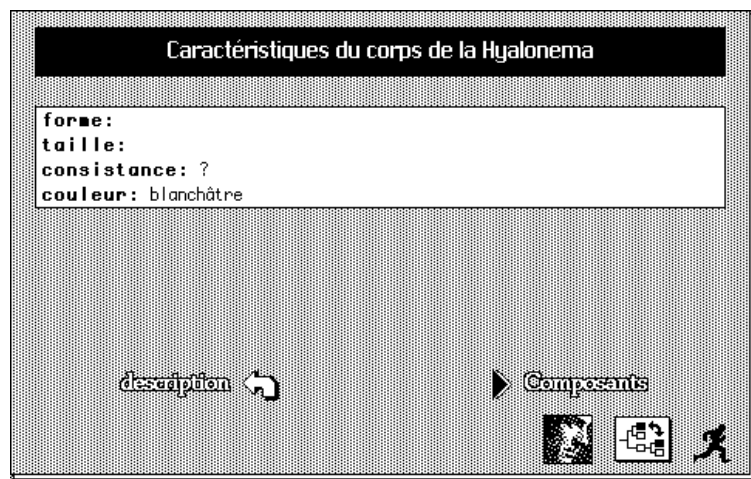


Fig. 6.12 : Cartes des caractéristiques et des composants de l'objet "corps" de Hyalonema

Cet exemple met en évidence les points suivants : pour renseigner quel est l'objet courant, un titre est construit automatiquement à partir de l'information sur l'objet, son genre (ms pour "masculin-singulier") et sur le type de description locale (caractéristiques, composants ou spécialisations).

Ici, le corps est une dépendance de l'objet "description". Par navigation, on peut l'atteindre en cliquant sur la flèche incurvée de gauche (l'incurvation indique une remontée vers la racine à gauche de l'arbre des dépendances). Le message suivant est déclenché et on aboutit à la carte <Description> :

```
on mouseUp
  get card field "Surpartie"
  visual effect scroll down fast
  go to cd it
end mouseUp
```

Les caractéristiques (slot) et les dépendances de l'objet (subpart) sont représentées par des entités hypertextes (champs et boutons). Elles sont placées sur la même carte du fait qu'il s'agit de la description locale d'un objet, et on y accède en cliquant sur la flèche triangulaire de droite. On bascule ainsi des caractéristiques de l'objet à ses composants sur la même carte et réciproquement. Cette présentation des objets avec deux volets (caractéristiques et relations) est à rapprocher avec celle de la carte d'édition des objets dans le modèle descriptif (figure 6.6).

Le personnage en bas à droite est un **coursier** qui emmène sous son bras la description du cas courant afin de procéder à un enregistrement dans la base de cas (pour la classification) ou bien à une consultation du système expert (pour l'identification d'une nouvelle observation).

Les deux autres boutons sont des fonctionnalités de navigation et de personnalisation des descriptions (voir § 6.5.10).

6.5.2.2 Représentation des statuts possibles d'un objet

Les statuts possibles sont les propriétés d'un objet observable telles que la cardinalité et la multiplicité de l'objet (voir § 4.6.5.1).

Pour la cardinalité, considérons l'objet "cone-central" de l'éponge :

```
(defobject cone-central
  (part-of face-exhalante)
  (cardinal 0 1)
  (...))
```

La cardinalité de cet objet indique qu'il peut ne pas y avoir présence de l'objet (cardinal = 0) et que s'il est présent, alors il y en a un seul (cardinal = 1). Par défaut, on n'écrit pas dans les fichiers au format LCRC la possibilité pour un objet d'être inconnu car c'est le statut qui est défini pour tous les objets du modèle descriptif. Le premier chiffre suivant le mot clé "cardinal" indique la cardinalité minimale, le second chiffre étant la cardinalité maximale de l'objet.

Au niveau du questionnaire, la représentation des statuts possibles d'un objet se traduit par des boutons sélecteurs d'état comme on peut le voir sur la figure 6.13 suivante :

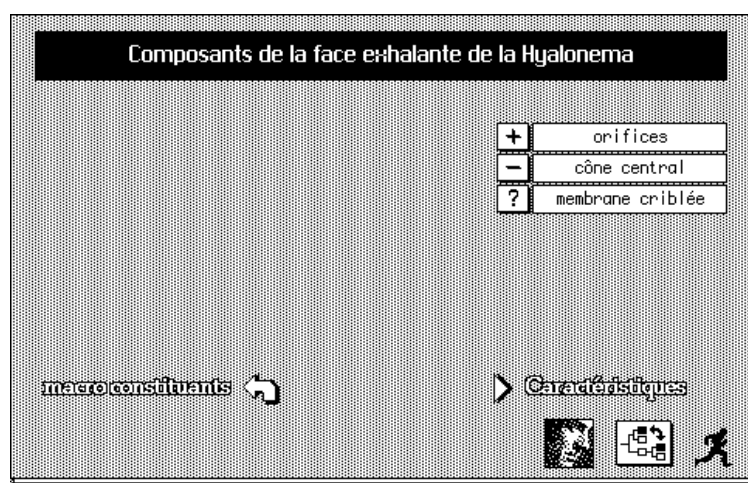


Fig. 6.13 : Représentation des différents états d'un objet dans le questionnaire

L'utilisateur doit cliquer sur le bouton précédant le nom de l'objet à décrire pour indiquer le statut de cet objet :

“+” pour la présence, “-” pour l'absence, “?” pour l'inconnu.

En ce qui concerne la multiplicité d'un objet, on peut se reporter au § 6.5.5 sur la représentation des multi-instanciations d'objets.

6.5.3 Représentation des attributs d'un objet

Pour donner une valeur à l'attribut d'un objet, on clique sur le nom de l'attribut dans le champ des caractéristiques de l'objet (figure 6.12), ce qui a pour effet d'accéder à une autre carte intitulée <Attribut(Objet)>. Celle-ci donne la liste des valeurs possibles pour cet attribut. Les attributs sont définis en LCRC dans un fichier séparé des objets et possèdent des types différents (qualitatifs, numériques, commentaire, classifié).

6.5.3.1 Attributs qualitatifs

Prenons la couleur du corps de l'éponge dont voici la définition en LCRC :

```
(defslot couleur corps nominal
  (range blanchâtre grisâtre autre)
  (default blanchâtre)
  (question "Quelle est la couleur du corps?"))
```

Cet attribut sur le corps de l'éponge est de type nominal. Les valeurs possibles sont indiquées dans le "range" ou domaine d'observation (§ 4.6.5.2.2).

L'utilisateur sera amené à faire un choix unique (§ 4.6.5.2.4). Le nombre de choix possibles est unique par défaut, sinon il serait précisé dans la définition de l'objet. La question posée pour la consultation apparaît dans le champ "question" (§ 4.6.5.3.2).

On remarque aussi qu'il existe une valeur par défaut indiquée dans le modèle descriptif pour la couleur (§ 4.6.5.2.3). Cette valeur mentionnée par l'expert est la plus fréquemment observée mais l'utilisateur peut néanmoins spécifier une autre couleur pour un spécimen donné. Lors de la construction du questionnaire, cette valeur par défaut vient s'ajouter automatiquement comme valeur de la couleur (figure 6.12).

La transformation de cette définition donne la carte de la figure 6.14 :

Quelle est la couleur du corps ?

blanchâtre
grisâtre
autre
?

Annuler

Fig. 6.14 : Correspondance entre un attribut qualitatif et la carte <couleur(corps)>

On remarque dans cette carte l'apparition du bouton "?". Il s'ajoute de manière automatique lors de la génération du questionnaire et indique la possibilité pour l'attribut d'avoir comme réponse de l'utilisateur : "je ne sais pas" ou "inconnu".

Remarque : le choix de la réponse "autre" appelle un commentaire en texte libre (voir § 6.5.3.3).

6.5.3.2 Attributs numériques

On peut prendre l'exemple de la longueur d'un micro élément de l'éponge (les "tignules") dont voici la description au format LCRC et la correspondance dans le questionnaire (figure 6.15) :

```
(defslot longueur tignules real Interval
  (range 1.0 20.0)
  (unit mm)
  (question "Quelle est la longueur des tignules ?"))
```

The image shows a graphical user interface for a questionnaire. The title is "Quelle est la longueur des tignules ?". There is a text input field containing "11.14 mm". Below the input field is a horizontal logarithmic scale with markers at 1.0 and 20. A small box below the scale contains the number "2" and the text "Chiffres significatifs". At the bottom of the window, there are three buttons: "Annuler", "?", and "Ok". To the right of these buttons are three small icons: a globe, a window, and a person.

Fig. 6.15 : Correspondance entre un attribut numérique et la carte <longueur(tignules)>

Ici, l'attribut numérique est un réel mais il pourrait aussi s'agir d'un entier naturel. Dans le cas d'une valeur comprise dans un intervalle, l'utilisateur dispose d'une échelle munie d'un minimum et d'un maximum. Cette échelle est logarithmique. L'utilisateur peut aussi indiquer la précision de ses mesures par le nombre de chiffres significatifs (1, 2 ou 3 chiffres significatifs). On recopie l'unité de mesure de la valeur numérique dans le champ "unit".

6.5.3.3 Attributs commentaires

Un attribut commentaire (ou chaîne de caractères) est une information supplémentaire sur le cas à décrire qui ne sera pas prise en compte lors de l'induction ou du raisonnement par cas.

En effet, l'apprentissage de règles de décisions se fait "en monde fermé" c'est-à-dire lorsque l'on a prédéfini tous les objets, attributs et valeurs observables. Les attributs commentaires sont des caractéristiques d'un questionnaire "ouvert" sur lesquels on ne peut fixer à l'avance les réponses possibles. C'est une information utile pour replacer les cas dans leur contexte. L'expert dispose ainsi d'un moyen de contrôle des réponses fournies. Il pourra par la suite redéfinir de nouveaux descripteurs intégrant la variation des réponses des utilisateurs.

Prenons comme exemple la localité à laquelle appartient un individu :

```
(defslot localite contexte comment
  (default ?)
  (string "localité")
  (question "Origine du Specimen?"))
```

La transformation de cet attribut donne la carte suivante (figure 6.16) avec une boîte de dialogue dans laquelle l'utilisateur peut rentrer sa valeur :



Fig. 6.16 : Correspondance entre un attribut commentaire et la carte <localité(contexte)>

6.5.3.4 Attributs classifiés

Certaines valeurs peuvent être structurées (voir § 4.6.5.2.2) selon une hiérarchie de classification (taxonomie de valeurs). Prenons l'exemple de la forme du corps de l'éponge dont les valeurs sont représentées sous la forme d'une hiérarchie à deux niveaux. Voici une description au format CASUEL d'un tel attribut :

```
defslot forme of corps
  type corps_forme_type;
  cardinal [1:2];
  french_question "Quelle est la forme du corps de l'éponge ?".

deftype corps_forme_type
  a_kind_of taxonomy;
  range [forme [subcylindrique [en-bouteille cylindroïde]]
        [renflée [en-flute en-corolle piriforme ovoïde bulbiforme]]
        [evasee [en-tronc-de-cone en-cloche en-crateres en-cone en-
trompette en-cornet]]
        [aplatie [en-lame-repliée en-galette]]
        [mal-definie [irrégulier amyboïde]]].
```

A un premier niveau, la forme du corps de l'éponge peut prendre comme valeurs "subcylindrique", "renflée", "évasée", "aplatie" et "mal-définie". Le second niveau est constitué des spécialisations de chacune de ces valeurs. La valeur "renflée" peut être précisée pour sa part en 5 sous-valeurs : "en-flûte", "en-corolle", "piriforme", "ovoïde" et "bulbiforme".

La cardinalité indique le nombre de choix possibles, l'utilisateur est limité ici à 2 valeurs au maximum. Voici la carte du questionnaire correspondant à la définition de la valeur "renflée" (figure 6.17) :

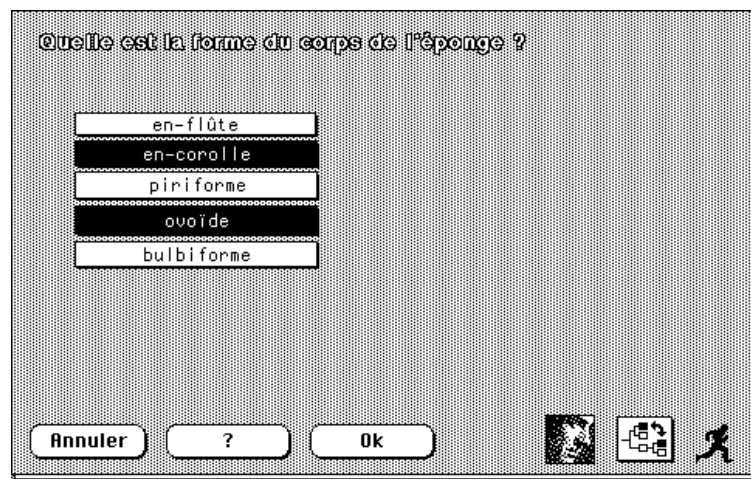


Fig. 6.17 : Les sous-valeurs de la carte <renflée(forme)>

Lorsque la cardinalité d'un attribut est supérieure à 1, on construit le bouton "Ok" pour permettre la validation des choix de l'utilisateur. Sur la figure ci-dessus, le descripteur a choisi les valeurs "en-corolle" et "ovoïde" pour indiquer le doute sur la forme renflée du corps de son spécimen. Il lui faut maintenant cliquer sur le bouton "Ok" pour retourner au niveau supérieur de la taxonomie des valeurs.

Rappel : un choix multiple traduit une imprécision et non pas une présence simultanée (voir § 4.6.5.2.4).

6.5.4 Représentation des spécialisations d'un objet

Outre la relation de dépendance entre objets, il y a les relations de parenté ou de recouvrement qui indiquent une hiérarchie entre deux objets dont l'un est plus général que l'autre. Cela se représente par une relation de spécialisation : un objet se spécialise en un "sous-objet". Par exemple, dans le système TOM, les objets "tache" et "nécrose" sont des spécialisations de l'objet "symptôme sur folioles".

L'intérêt de représenter cette connaissance des liens entre certains objets est double :

1) il permet de faire hériter les propriétés des objets les plus généraux *vers* les objets les plus spécifiques. Ainsi, on ne définit les propriétés spécifiques à chacun des objets qu'à leur niveau, les propriétés communes étant *mises en facteur* au niveau d'un objet plus général. Une procédure se charge d'enrichir les spécialisations. Prenons l'exemple de la figure 6.1 ou la description du symptôme sur folioles peut être précisée par un objet de l'arbre de spécialisation :

```
(defobject symptome-sur-folioles
  (part-of folioles)
  (gender fp)
  (subobject taches-ou-plages-ou-necroses fletrissement anomalie-de-coloration)
  (string "symptôme sur folioles")
  (slots repartition-sur-la-plante repartition-sur-la-feuille))
```



Fig. 6.18 : Représentation du lien de spécialisation de l'objet "symptôme sur folioles"

Le lien de spécialisation est symbolisé par une flèche verticale descendante dans le questionnaire (figure 6.18). En cliquant sur cette flèche, on accède à une carte intitulée <Specialisations(Objet)> (figure 6.19) qui permet de préciser l'objet générique.

Cette carte est différente de la carte parente, ce qui donne la possibilité à l'utilisateur d'associer des dessins illustrant les sous objets sans interférer avec les illustrations des composants de l'objet.

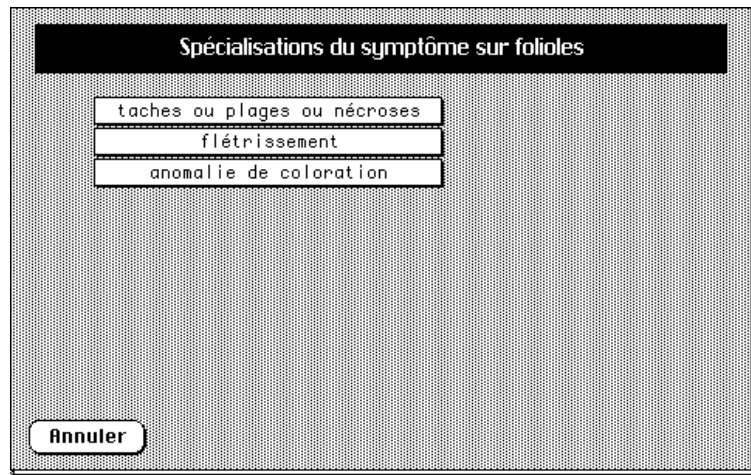


Fig. 6.19 : Vue de la carte <spécialisations(symptome-sur-foliole)>

L'utilisateur fait son choix parmi la liste des sous objets. Prenons l'objet "taches ou plages ou nécroses" décrit de la manière suivante au format LCRC :

```
(defobject taches-ou-plages-ou-necroses
  (part-of folioles)
  (gender fp)
  (superobject symptome-sur-folioles)
  (subobject taches-ou-plages necroses)
  (string "taches ou plages ou nécroses")
  (slots localisation-sur-foliole proportion-sur-foliole repartition-sur-foliole
    jaunissement-en-pourtour toucher nombre variabilite forme relief
    nombre confluence zonations mesure limites couleur))
```

On remarque dans cette définition que l'on n'a introduit que les attributs propres à l'objet. La transformation de cette description du modèle donne la carte suivante du questionnaire (figure 6.20) :

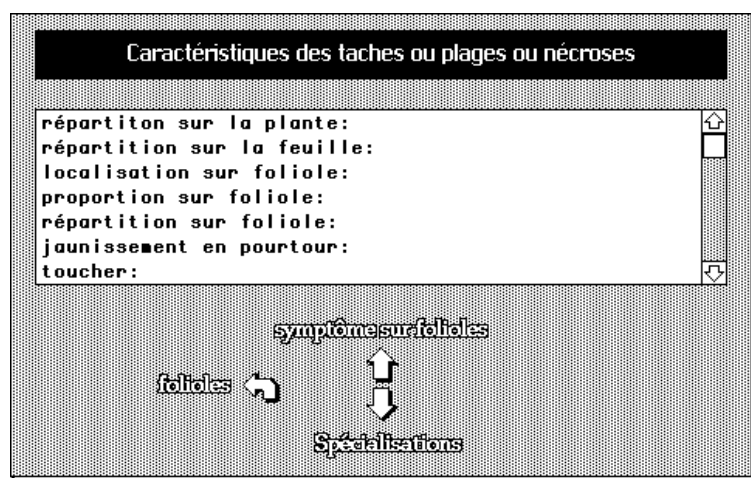


Fig. 6.20 : Représentation de la carte de l'objet "taches ou plages ou nécroses"

L'objet **héríte** des attributs du symptôme sur folioles grâce au lien de spécialisation : il en est ainsi pour les attributs “répartition sur la plante” et “répartition sur la feuille”. En continuant à spécialiser le symptôme, on en viendrait à préciser plus encore sa nature en choisissant un des concepts de la disjonction (figure 6.21). Ce faisant, on est amené à restreindre le domaine d'observation de certains attributs ou à introduire de nouveaux attributs plus spécifiques à l'objet spécialisant.

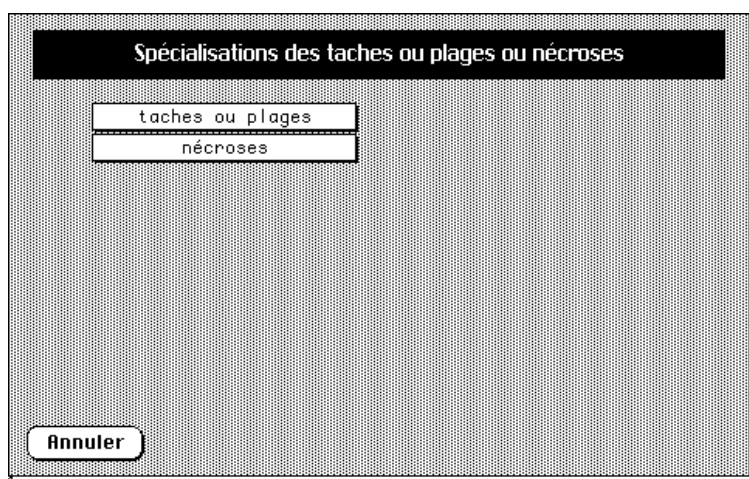


Fig. 6.21 : Vue de la carte <spécialisations(taches-ou-plages-ou-necroses)>

Ce premier intérêt met l'accent sur l'action elle même d'héritage qui permet de transférer une information plus générale à un objet particulier. C'est une vision descendante de la spécialisation, telle qu'elle est comprise par les informaticiens [Manago, 1988].

2) A l'inverse, pour le biologiste, la spécialisation se traduit par le remplacement de caractéristiques plus générales par des caractéristiques plus particulières, c'est-à-dire par une substitution (vision ascendante). Ainsi, on passe de quelque chose de plus général à quelque chose de plus particulier. L'héritage n'apparaît que comme une conséquence de la spécialisation, qui consiste essentiellement en une détermination de l'objet selon une classification conceptuelle (on reste dans l'observable).

6.5.5 Représentation des multi-instanciations d'un objet

La possibilité de décrire plusieurs sortes d'un même objet a été décrite au § 4.6.7. Prenons l'exemple de l'objet “microxyhexactines” du domaine des *Hyalonema* qui a été défini avec une multiplicité de deux. Cela signifie qu'il est possible de décrire jusqu'à deux fois cet objet pour un même cas. Lorsque l'on clique sur l'objet “microxyhexactines” dans la fenêtre “Composants des micro éléments de la Hyalonema”, la carte suivante apparaît (figure 6.22) :

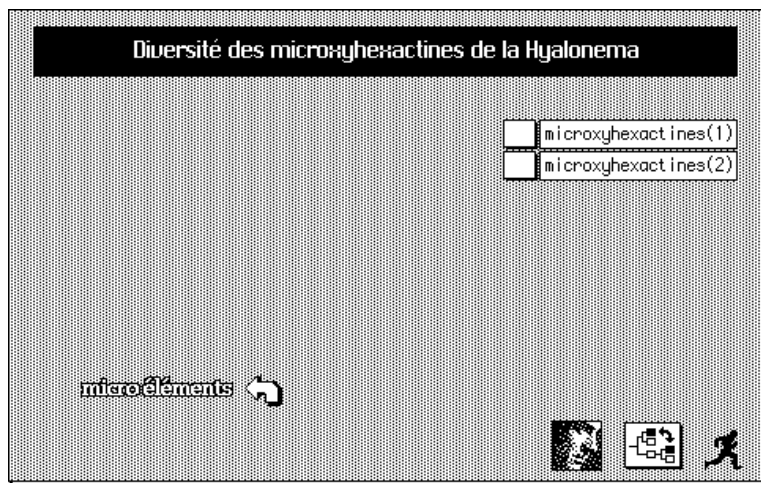


Fig. 6.22 : Vue de la carte des instances multiples de l'objet "microxyhexactines"

Elle a autant de boutons que l'on peut décrire de sortes de microxyhexactines (ici deux boutons). Chaque label de bouton est suivi d'un numéro :

- (1) pour la première description de l'objet,
- (2) pour la seconde.

Décrire l'objet microxyhexactine(2) avant microxyhexactine(1) n'a aucun sens pour la description de différentes sortes d'un même objet : on décrit toujours une première sorte avant d'en décrire une seconde ! L'ordre que l'on choisit de suivre n'a aucune signification propre.

Quel que soit l'objet décrit, microxyhexactine(1) ou microxyhexactine(2), la carte à renseigner est identique. L'utilisateur peut itérer des descriptions différentes du même objet en donnant des valeurs différentes aux attributs. Le passage à la description d'autres sortes d'un même objet s'effectue en cliquant sur la flèche incurvée vers le bas à gauche (figure 6.23) :



Fig. 6.23 : Description de la première sorte d'objet "microxyhexactines"

6.5.6 Représentation des règles entre objets

Nous avons déjà mentionné les dépendances possibles entre les différents caractères composant une description de spécimen (§ 4.6.8).

Nous prendrons un exemple qui permet de visualiser la représentation d'une règle localement sur un même objet : soit la description des orifices situés sur la face exhalante d'une éponge, dans laquelle le slot nombre a pour valeurs observables {unique, quadruple, multiple} :

```
(defobject orifices
  (part-of      face-exhalante)
  (gender      mp)
  (slots nombre (if (nombre unique quadruple) then
                    (HideSlots dimensions repartition)
                    else? (SetDefaultValues (dimensions ?) (repartition ?)))
         dimensions
         repartition))
```

Cette définition d'objet comporte une règle écrite au format LCRC. C'est une connaissance de bon sens indiquée par l'expert lors de la création de l'objet :

Si le nombre d'orifices est unique ou quadruple, alors il ne faut pas s'intéresser à la description des attributs "dimensions" et "répartition" des orifices (mot clé "HideSlots").

Cette règle entre une valeur d'attribut et d'autres attributs du même objet est représentée de la manière suivante dans le questionnaire (figure 6.24) :



Fig. 6.24 : Dépendance entre une valeur d'attribut et d'autres attributs

En effet, la dimension des orifices n'est plus pertinente dans ce contexte (information fournie par l'expert) et la répartition n'est pas observable

(information de bon sens). L'utilisateur qui répond "unique" au nombre d'orifices ne peut donc logiquement pas renseigner ces deux attributs (en italique).

La règle énonce aussi que si la valeur du nombre des orifices est "inconnu" alors les valeurs des attributs "dimensions" et "répartition" sont aussi inconnues (mot clé "SetDefaultValues"). Cette règle est activée automatiquement dans le questionnaire par un changement d'état des attributs dépendants (passage de la forme cachée en italique à la forme active en gras) et la mise à jour automatique des valeurs "inconnu" pour les attributs dépendants (figure 6.25) :

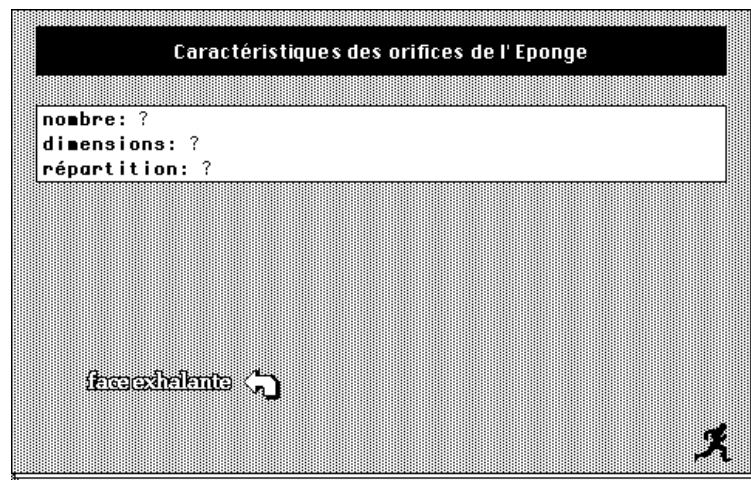


Fig. 6.25 : Dépendance entre une valeur d'attribut et d'autres valeurs d'attributs

Toutes ces règles font partie de l'observable, elle permettent d'assurer la cohérence de la procédure de description et participent donc ainsi à la robustesse globale du système.

6.5.7 Personnalisation des objets

L'intérêt d'utiliser un questionnaire sous Hypercard est que chaque carte associée à un objet ou à un attribut peut être personnalisée en utilisant les fonctionnalités associées de dessin. En effet, il est très facile d'enrichir la carte par des images dessinées ou collées, et dont certaines zones représentent d'autres objets à décrire (des composants par exemple). On peut aussi illustrer des valeurs d'attribut par des schémas explicatifs. Ensuite, on peut recouvrir ces zones de l'image par des boutons multi-formes et transparents qui permettront de sélectionner l'objet ou la valeur visée de manière équivalente à la sélection du bouton nommé.

La figure 6.26 suivante illustre la personnalisation de la carte <renflée(forme)> du corps de la *Hyalonema* :

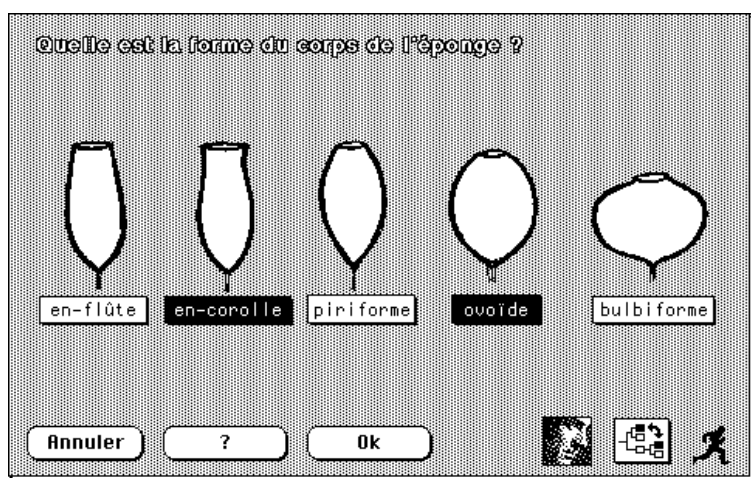


Fig. 6.26 : Association de dessins pour illustrer des valeurs d'attributs

La personnalisation des cartes du questionnaire est un facteur ergonomique important pour l'aide à la détermination des objets. Le rôle de l'expert est ici d'illustrer l'observable, c'est-à-dire son propre vocabulaire de description pour le rendre accessible aux autres utilisateurs sans interprétation lors des consultations futures du questionnaire. C'est un facteur non négligeable de la robustesse puisqu'il dissocie l'exactitude des choix de la compréhension d'un vocabulaire.

Les changements de personnalisation sont sauvés dans une autre pile (voir § 6.5.9.11). La sauvegarde est utile pour remettre à jour une nouvelle pile avec les anciennes images ou dessins. En effet, dans le cas où l'utilisateur désire modifier le modèle descriptif (rajouter un nouvel objet observable par exemple), il faudra générer un nouveau questionnaire du fait que sa structure est modifiée.

Cette fonctionnalité de sauvegarde est importante dans la mesure où il pourrait y avoir plusieurs versions de modèles descriptifs élaborés par l'expert, et par conséquent plusieurs questionnaires intermédiaires à construire. Néanmoins, il est préférable de ne fabriquer qu'un seul modèle descriptif complet dès le départ pour éviter toutes les remises à jour (le questionnaire et les cas).

Remarque : la limite actuelle est qu'HyperCard ne connaît que le noir et le blanc (mode Paint). Un autre moyen de personnaliser la pile du questionnaire peut être d'associer une photo ou une séquence vidéo en couleur à certaines cartes. Cette fonctionnalité peut être intégrée à HyperCard au travers de commandes externes capables d'afficher dans une fenêtre spéciale une image au format PICT ou une séquence filmée QuickTime.

Fonctionnalités

Le questionnaire est une pile HyperCard générée à partir d'HyperQuest. Il permet d'éditer des descriptions pour les trois objectifs suivants :

- ❶ Rentrer de nouveaux cas dans la base de cas,
- ❷ Consulter le système expert engendré par KATE,
- ❸ Consulter le système de raisonnement par cas avec CaseWork.

Outre le fait que le questionnaire est un éditeur de cas, des fonctionnalités de personnalisation et de navigation lui sont attachées pour en faire un véritable guide d'observation.

Deux menus “Cas” et “Personnaliser” traitent respectivement de la gestion des cas et de la personnalisation des cartes d'édition des cas dans le questionnaire (figure 6.27) :

Cas	Personnaliser
Nouveau	Préférences...
Dernier	Scanner une image
Ancien...	Tout importer
Enregistrer	Importer une Image ⌘G
Consulter l'Expert	Déplacer l'Image
Voir la liste	Associer un Bouton ⌘A
Convertir en Casuel	Changer le titre ⌘Y
	Déplacer les Boutons ⌘D
	Eliminer un Bouton
	Déplacer les Champs
	Sauver la Carte ⌘S

Fig. 6.27 : Les menus "Cas" et "Personnaliser"

De plus, chaque carte du questionnaire possède les trois boutons suivants : le bouton “voir une image”, le bouton “naviguer” et le bouton “coursier”.

6.5.8 Le menu “Cas”

Nous voici maintenant dans la phase d'utilisation du questionnaire généré pour éditer des descriptions observées. Deux natures de descriptions d'individus sont produites :

- ❶ Ce sont des **cas** à enregistrer dans la base de cas si elles contiennent une détermination associée, c'est-à-dire une valeur au niveau de la classe de détermination qui n'est pas inconnue.

- ② Ce sont des **observations** si la classe est inconnue. Ces observations peuvent être stockées dans la base de cas en attendant d'être classifiées mais ne peuvent pas être traitées par KATE et CaseWork sous cette forme.

6.5.8.1 Le choix “Nouveau”

Il permet de saisir un nouveau cas en initialisant le modèle descriptif : en effet, l'idée est de partir de l'objet observable situé à la racine du modèle et représenté par une carte (de nom <Hyalonema> dans notre application).

“Nouveau” instancie cette carte, c'est-à-dire en effectue une copie (cf. § 6.4.1). La nouvelle carte instanciée possède le même nom plus un numéro d'instance accolé (par exemple “Hyalonema1”). Elle devient ainsi un objet observé⁴ du domaine que l'on peut décrire, ce n'est plus un concept abstrait.

Une fois le modèle initialisé à la racine, le cas existe et l'utilisateur n'a plus qu'à compléter sa description en naviguant pour instancier ses composants. Eux aussi sont des objets observables représentés par un état nul (le carré à gauche du bouton <composant> est vide). Dès que l'utilisateur clique dans un des boutons des composants, l'objet se met à exister : la carte correspondante est instanciée et le carré représentant son statut prend la valeur “+” (voir section 6.5.2.2).

Au fur et à mesure de la navigation, la pile du questionnaire s'accroît donc en nombre de cartes, relativement au nombre d'objets observés pour l'individu décrit. Ce processus est valable aussi bien pour les objets “composites” que pour les objets “spécialisants”.

Outre la navigation entre les objets, on peut s'arrêter au niveau de la carte de l'objet observé et le décrire avec ses caractéristiques locales. A ce moment, on va renseigner les attributs situés dans le champ de la carte de l'objet observé. Il suffit pour cela de basculer au niveau des caractéristiques de l'objet courant et de cliquer l'un des items du champs associé. On accède ainsi à la carte <Attribut(Objet)> qui permet de renseigner la valeur que doit prendre l'attribut sélectionné.

Le schéma de la figure 6.28 illustre la description locale de l'objet observé “Ombrelles1” :

⁴ Très incomplètement pour l'instant !

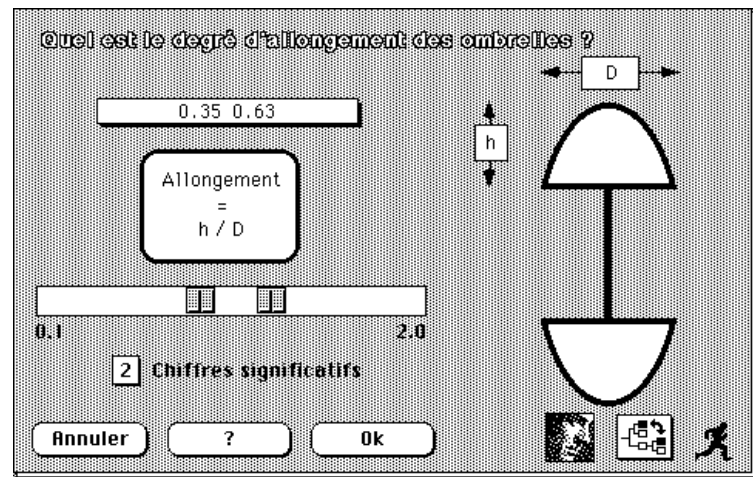


Fig. 6.28 : Le processus de description locale d'un objet⁵

⁵ On peut noter ici l'intérêt d'illustrer l'allongement des ombrelles par un dessin explicatif !

Notons qu'il n'y a pas de copie d'instanciation des cartes associées aux attributs d'un objet : ces cartes sont mises à jour à chaque fois que l'on y accède selon que l'attribut possède ou non une valeur dans le champ. Elles sont “nettoyées” de la valeur choisie à chaque sortie.

Lorsque la description du nouveau cas est terminée, il faut alors le sauvegarder soit pour la base de cas (voir § 6.5.8.4), soit pour la consultation du système expert (voir § 6.5.8.5).

6.5.8.2 Le choix “Dernier”

Après enregistrement et fermeture du questionnaire, la pile conserve néanmoins la représentation des cartes instanciées du dernier cas.

Ce choix d'implantation est justifié par la pratique d'acquisition des descriptions : en effet, on s'aperçoit que le fait de renseigner un questionnaire d'acquisition d'exemples est en général une opération répétitive, longue et fastidieuse. L'utilisateur doit s'organiser pour rentrer les cas dans un ordre où les spécimens à décrire sont similaires au niveau des objets à renseigner. Il est alors plus efficace en terme de temps de repartir de la description du dernier cas pour en renseigner un nouveau puis de modifier les quelques caractéristiques divergentes. Cette pratique est celle qui a été suivie lors de l'acquisition des descriptions de *Hyalonema* : on peut observer à l'annexe 3 les séries de cas décrits les uns à la suite des autres et concernant un même Sous-Genre (*Cyliconema* par exemple).

Il faut faire attention néanmoins à cette facilité qui peut conduire à des descriptions inappropriées par omission de la mise à jour de certains caractères du dernier cas.

Remarque : les cas dérivés les uns des autres peuvent être rapprochés des nuances critiques [Winston, 1977] qui sont très importantes dans l'isolement des caractéristiques définissant un concept lors de la discrimination à partir d'exemples.

6.5.8.3 Le choix “Ancien ...”

Le questionnaire est un éditeur de cas. Cette possibilité d'édition d'un ancien cas est nécessaire pour la gestion de la base de cas. Elle est utilisée lorsque l'utilisateur désire modifier un ancien cas ou bien se servir de lui comme modèle pour en construire un nouveau (voir paragraphe précédent). L'opération est néanmoins plus coûteuse lorsqu'il s'agit d'importer un ancien cas, du fait que la pile du questionnaire doit d'abord détruire toutes les cartes instanciées correspondant aux derniers objets observés, puis reconstruire l'ancien cas à

partir de la description stockée dans la base. Le questionnaire demande quel est l'ancien cas à importer puis reconstruit automatiquement les cartes correspondant aux objets observés dans l'ancien cas.

6.5.8.4 Le choix “Enregistrer”

Lorsque l'utilisateur a terminé sa description, l'une des possibilités est d'enregistrer ce cas dans la base, les deux autres choix étant de quitter sans enregistrer ou bien de consulter l'expert. Dans le cas d'un enregistrement, le message “Attendez quelques instants..., je récupère la description de votre cas” s'affiche dans la boîte d'information⁶, puis le cas courant vient s'afficher dans un champ (figure 6.29) pour que l'utilisateur puisse vérifier lui même l'exactitude de sa description :

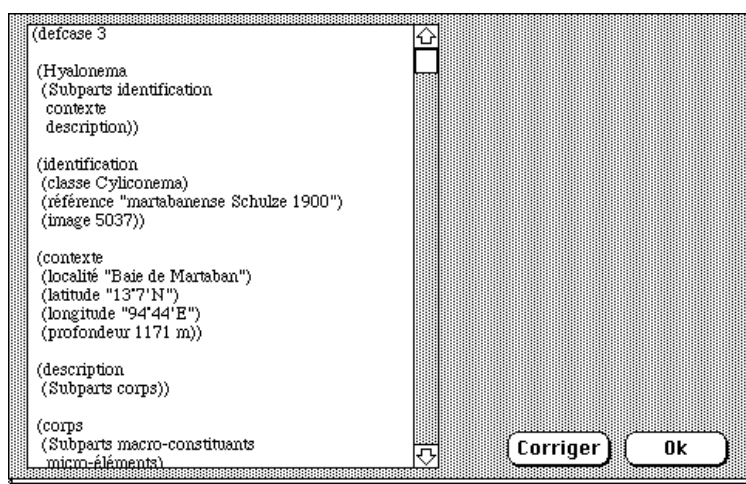


Fig. 6.29 : La description du cas en LCRC à enregistrer

Nous pensons que la visualisation du cas selon une formule lisible est important avant son enregistrement effectif. Si l'utilisateur s'aperçoit d'une erreur de description, il peut encore la corriger en cliquant sur le bouton “Corriger” qui le fait revenir à la dernière carte qu'il avait quittée pour enregistrer son cas. Sinon, l'utilisateur satisfait de sa description clique sur le bouton “Ok” pour permettre l'enregistrement de son cas dans la base.

Il serait bon aussi de présenter une description sous une forme proche du langage naturel comme cela est préconisé dans le schéma global de notre méthodologie (§ 2.4). Il s'agit là d'un module non développé actuellement mais qui pourrait servir à l'édition de descriptions “naturelles” utiles aux biologistes pour leurs monographies. Pour l'heure, l'utilisateur peut néanmoins imprimer la

⁶ Ces messages ne sont pas superflus car ils instaurent un dialogue avec l'utilisateur, ce qui apporte une certaine convivialité à la procédure et évite les fausses manoeuvres de l'utilisateur (il sait qu'il doit patienter !).

description du cas au format LCRC ci-dessus en utilisant la fonctionnalité "Imprimer un champ..." d'HyperCard.

Il est recommandé d'enregistrer dans la base de cas un cas complet, c'est-à-dire possédant le nom de la classe attachée à la description. Néanmoins, l'utilisateur peut enregistrer une observation puis la compléter plus tard : il s'agit de la possibilité de remplacer un ancien cas de la base par la description courante (figure 6.30) :



Fig. 6.30 : Le dialogue permettant de remplacer un ancien cas de la base de cas

6.5.8.5 Le choix "Consulter l'expert"

Le questionnaire est aussi un outil permettant de déterminer la classe à laquelle appartient la description courante appelée ici observation. La procédure de consultation est la suivante :

- ❶ L'utilisateur rentre sa description de l'individu observé dans l'ordre qu'il désire sans être contraint par la procédure de décision du système expert,
- ❷ Il choisit "Consulter l'expert" dans le menu "Cas"
- ❸ Comme pour l'enregistrement d'un cas, l'utilisateur peut corriger sa description en la visualisant dans le champ "Exemple" (figure 6.31) :

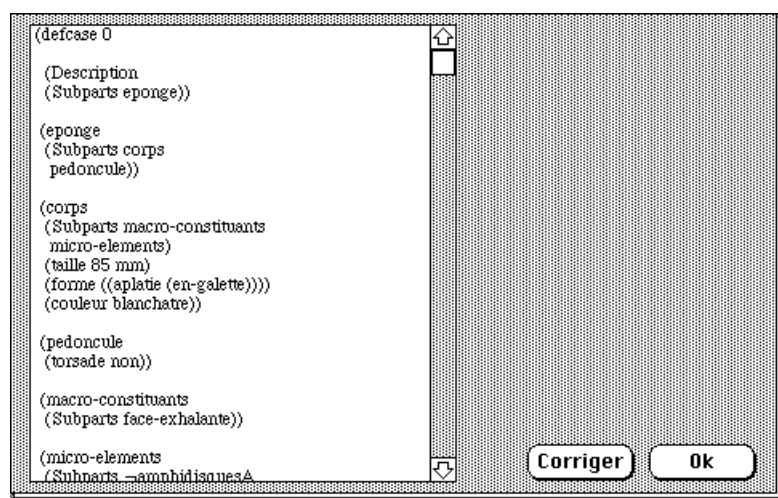


Fig. 6.31 : La description du cas à consulter

- ④ Le numéro associé à cette description est 0 car il ne comporte pas le nom de la classe. En validant la description, le cas est stocké dans un fichier texte de consultation nommé “consult”, prêt pour la procédure de détermination.
- ⑤ Que ce soit avec KATE (l'arbre de décision) ou CaseWork (raisonnement par cas), la consultation s'effectue automatiquement jusqu'à l'affichage d'un résultat (figure 6.32) :

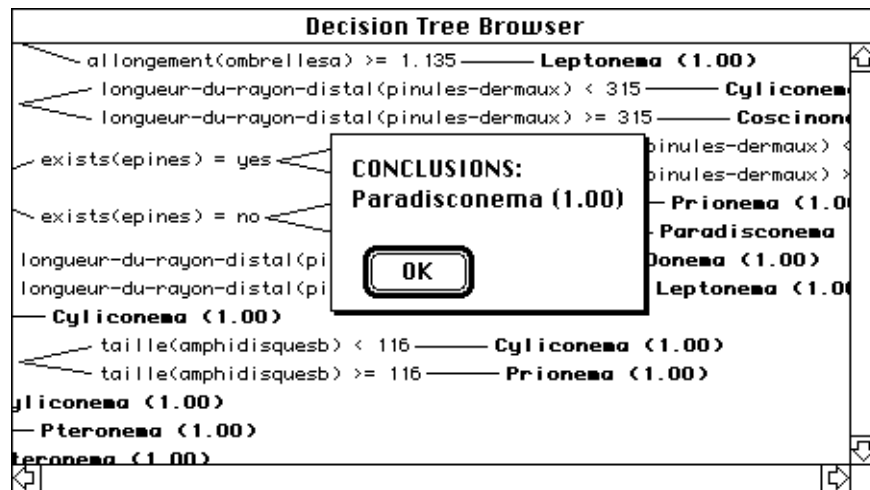


Fig. 6.32 : La consultation du système expert à partir du questionnaire

6.5.8.6 Le choix “Voir la liste”

Si l'utilisateur désire visualiser une description déjà enregistrée dans la base, il peut afficher la liste des cas par le choix “Voir la liste”. Il a alors la possibilité de voir la description de l'un des exemples en cliquant sur son numéro ou bien encore d'imprimer la liste des cas déjà saisis par le choix “Imprimer un champ...” d'HyperCard. Cette possibilité est importante pour que l'expert puisse avoir toujours le contrôle de la base de cas : un des objectifs est par exemple de ne pas renseigner deux fois le même cas !

6.5.8.7 Le choix “Convertir en Casuel”

CASUEL est une syntaxe de description des connaissances observables (le modèle descriptif) et des cas qui a été adoptée par le consortium INRECA (projet Esprit n° 6322) pour permettre l'intégration des technologies du raisonnement par cas et d'induction. C'est un langage commun de représentation des connaissances comme le LCRC. Afin de pouvoir utiliser les différents outils du consortium écrits dans différents langages (SmallTalk, C et Lisp), il est nécessaire de convertir ces connaissances au format CASUEL afin que chacun des programmes puisse les traiter.

6.5.9 Le menu “Personnaliser”

6.5.9.1 Le choix “Préférences...”

Ce choix permet de configurer l'application qui a été générée à partir d'HyperQuest. Cet aspect n'est pas développé dans cette thèse, on peut se reporter au manuel d'utilisation d'HyperQuest [Conruyt & Dumont, 1993].

6.5.9.2 Le choix “Scanner une image”

En activant cette option, l'utilisateur peut associer une image pour illustrer les objets ou attributs observables du domaine. Ce choix permet de "scanner" (scannériser) directement avec HyperScan™ des images ou dessins en noir et blanc qui viendront se superposer à la carte courante dans le questionnaire (§ 6.5.10.1). Ce choix est plutôt dédié à l'expert s'il désire illustrer le vocabulaire qu'il utilise par des dessins appropriés, ceci afin de faciliter sa compréhension par l'utilisateur final du système.

Inversement, l'utilisateur final peut illustrer sa description observée à l'aide du même outil. La souplesse d'utilisation d'HyperScan est un atout dans le choix d'un outil hypertexte pour illustrer les objets du modèle descriptif. C'est un élément clé de la constitution d'une base de connaissances multi-média. La couleur qui sera intégrée prochainement à HyperCard (version 3.0) dans le cadre de machines plus puissantes (Power PC) ouvrira des perspectives encore plus attrayantes à ces outils hypertextes.

6.5.9.3 Le choix “Tout importer”

Ce choix permet de mettre à jour les images d'une nouvelle pile de questionnaire concernant un domaine déjà modélisé auparavant et ayant eu un questionnaire personnalisé. A chaque fois qu'un utilisateur personnalise des cartes dans un questionnaire, il doit “sauver la carte” (§ 6.5.9.11). La carte est alors recopiée dans une autre pile indexée au nom du domaine. S'il s'agit d'une personnalisation de l'observable, le nom de cette pile est “Observable Nom_du_domaine”. S'il s'agit d'une personnalisation de l'observé, le nom de cette pile est “Observé Nom_du_domaine”.

La sauvegarde tient compte non seulement des images, mais aussi du nombre, de la nature et de la position de tous les champs et boutons de la carte personnalisée : il en est ainsi des boutons transparents utilisés lors de la procédure “Associer un bouton” (§ 6.5.9.6). En choisissant “Tout importer”, on n'a plus besoin de remodifier des cartes que l'on avait déjà personnalisées dans une précédente version du questionnaire.

6.5.9.4 Le choix “Importer une image”

Ce choix est identique au précédent “Tout importer” mais est spécifique à la carte courante du questionnaire.

6.5.9.5 Le choix “Déplacer l’image”

Une fois l’image importée, celle-ci peut ne pas être positionnée correctement à l’endroit voulu par rapport aux boutons de la carte par exemple. “Déplacer l’image” permet de pointer l’image sélectionnée en pointillés et de la déplacer avec la flèche (figure 6.33). La validation se fait en tapant sur la touche “Retour” comme indiqué par le message en bas de la carte.

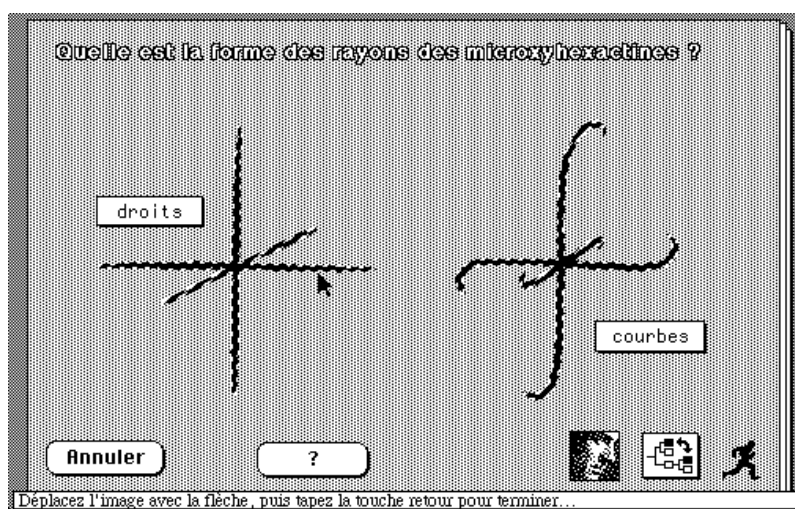


Fig. 6.33 : Déplacer une image dans le questionnaire

6.5.9.6 Le choix “Associer un bouton”

Associer un bouton consiste à rendre une zone de l’image sensible à un clic de la souris pour une action déterminée (par exemple pointer vers une sous-partie de l’image d’un objet). Les zones à délimiter pouvant avoir des formes variées, elles sont donc recouvertes de boutons transparents soit rectangulaires soit polygonaux. L'utilisateur doit choisir la forme du bouton qu'il veut associer à la partie de l'image (figure 6.34) :



Fig. 6.34 : Associer des boutons multi-formes aux images dans le questionnaire

Les boutons rectangulaires

Par exemple, à partir du dessin de dents dont on cherche à déterminer le type de bordure, l'utilisateur va cliquer sur le bouton "lisse" à associer à la partie de l'image correspondante. En fonction de la forme du dessin, il peut choisir entre un bouton polygonal ou un bouton rectangulaire afin de le recouvrir. S'il choisit le bouton rectangulaire, puis clique sur le bouton "lisse" de la carte <bord(dents)>, une copie transparente de ce bouton vient se placer au milieu de la carte.

L'utilisateur peut alors déplacer cette copie sur la partie désirée de l'image, agrandir ou diminuer les limites du bouton avec la flèche (figure 6.35). La boîte de messages en bas de la carte indique ce qu'il faut faire pour terminer :

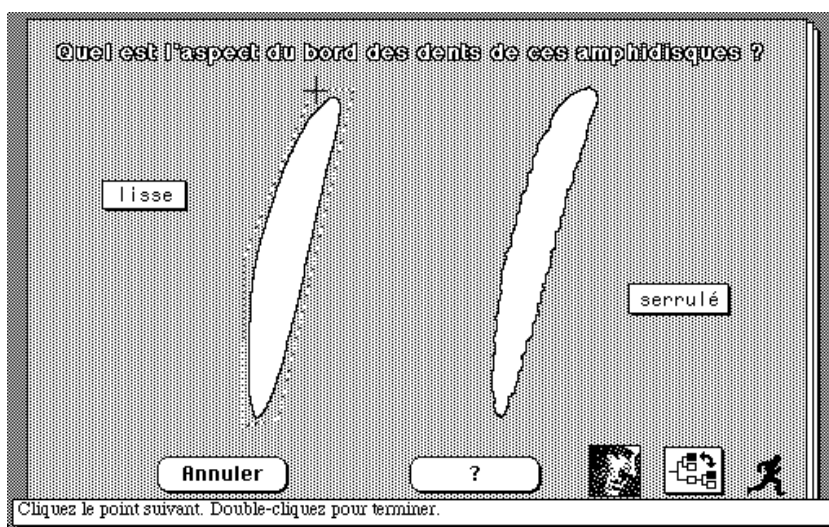
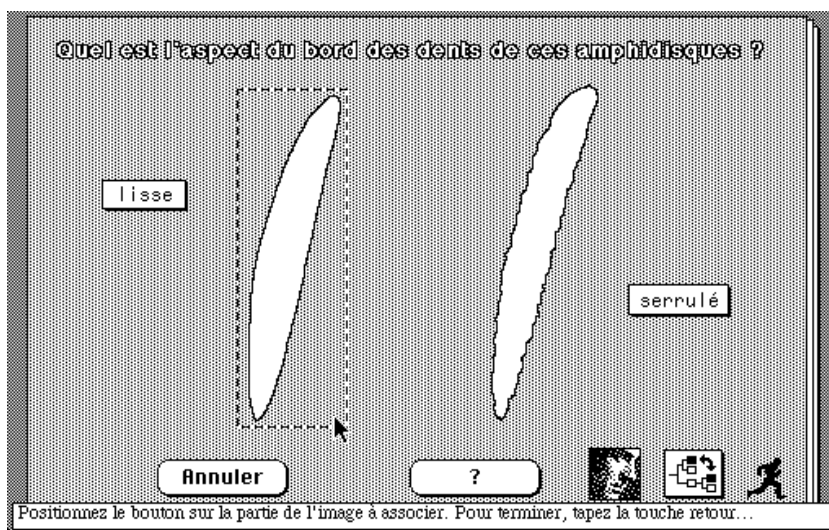


Fig. 6.35 : Associer un bouton rectangulaire ou polygonal à une partie d'image

Les boutons polygonaux

Lorsque les formes de l'image à recouvrir sont complexes ou se superposent, il est préférable de recourir à ce choix pour "Associer un bouton" (figure 6.35). Cette option permet de prendre en compte n'importe quelle forme du dessin. Un curseur en forme de croix apparaît à l'écran et l'utilisateur doit entourer la partie de l'image à délimiter. Il devra double-cliquer avec la souris pour fermer le polygone ainsi défini.

6.5.9.7 Le choix "Changer le titre"

Il est possible de changer le titre d'une carte quelconque du questionnaire (figure 6.36) pour poser une question différente de celle qui a été générée automatiquement (pour les cartes des objets) ou dans le modèle descriptif (pour les attributs) :

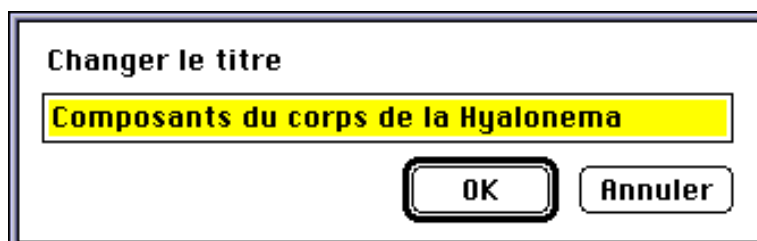


Fig. 6.36 : Changer le titre d'une question dans le questionnaire

6.5.9.8 Le choix "Déplacer les boutons"

Ce choix permet de visualiser tous les boutons de la carte, dont ceux qui recouvrent des parties d'image afin de les déplacer. Il suffit pour cela de maintenir la souris enfoncée avec la flèche sur l'objet pendant le déplacement. La touche "Retour" termine cette action.

6.5.9.9 Le choix "Éliminer un bouton"

Il est possible de détruire les boutons créés : il faut d'abord les sélectionner (en les cliquant une fois) avant de les éliminer. Tous les boutons de la carte sont concernés par ce choix.

6.5.9.10 Le choix "Déplacer les champs"

Ce choix est similaire à celui de "Déplacer les boutons". Il permet de déplacer les textes des champs de la carte.

6.5.9.11 Le choix “Sauver la carte”

Une fois que l'on a associé de nouveaux boutons sur les parties de l'image d'une carte du questionnaire et que l'on a positionné les boutons des valeurs ou objets correspondants, il faut conserver les modifications du questionnaire afin qu'elles puissent être réutilisées par la suite. Elles sont enregistrées sur la pile d'images observables. Cela permet de stocker la configuration de chaque carte personnalisée par l'utilisateur, ce qui s'avère utile lors de la génération d'un nouveau questionnaire.

Pratiquement, le fait de sauver une image et les boutons associés dans la pile d'images génère automatiquement les scripts permettant de créer les boutons de forme quelconque, de les déplacer et de commander leur action. Ils sont stockés dans le champ de scripts de la carte correspondante dans la pile d'images observables.

6.5.10 Les boutons

6.5.10.1 Le bouton “Voir une image”

Ce bouton (figure 6.37) permet de numériser à l'aide d'un scanner et du programme HyperScan™ des images sous forme de cartes HyperCard. Il est destiné à l'utilisateur final du questionnaire qui désire illustrer son cas par des images des différents objets ou attributs qu'il est en train de décrire. Il s'agit pour lui de “prouver” sa description en l'illustrant, ce qui est une information très intéressante pour l'expert afin de valider les cas des autres utilisateurs. C'est aussi la procédure inverse de la personnalisation des objets (§ 6.5.7), puisqu'elle permet d'illustrer l'observé et d'opérer un retour d'information vers l'expert.



Fig. 6.37 : Le bouton “Voir une image” de la pile du questionnaire

Les images observées sont stockées dans une pile à part et référencées dans la description au format LCRC ou CASUEL : celle d'un objet est indexée au slot “image” par le numéro d'identification de la carte de la pile contenant l'image numérisée (voir figure 6.29).

HyperScan (figure 6.38) est un logiciel de numérisation d'images ou dessins en noir et blanc très souple d'emploi :

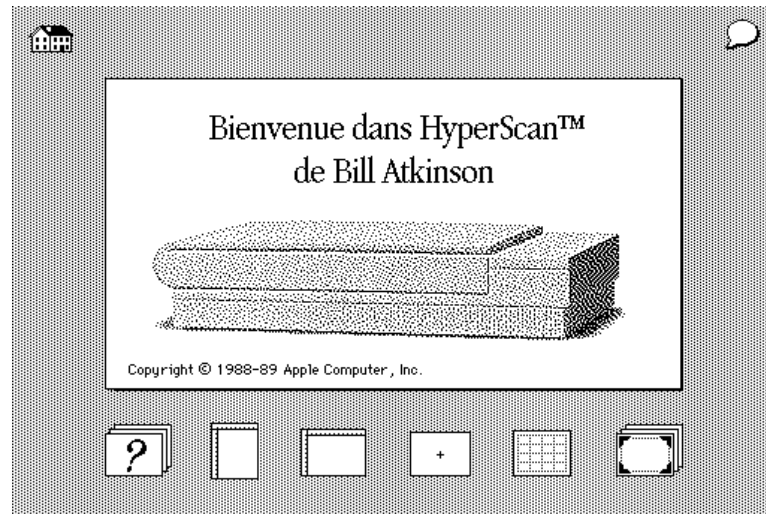


Fig. 6.38 : La pile HyperScan

Pour une documentation en ligne sur son utilisation, il suffit de cliquer sur le bouton d'aide "?". Une fois l'image numérisée, le bouton "Enregistrer" permet de coller dans la pile des images observées du questionnaire ("Observé Hyalonema" par exemple) le contenu imagé de la carte HyperScan. Chaque image observée de la pile contient l'image numérisée ainsi qu'un bouton "Ok" de retour au questionnaire. Une fois revenu, cliquer sur le bouton "Voir une image" permet directement l'accès (en plein écran) à l'image numérisée associée à l'objet ou l'attribut dans le questionnaire (figure 6.39) :

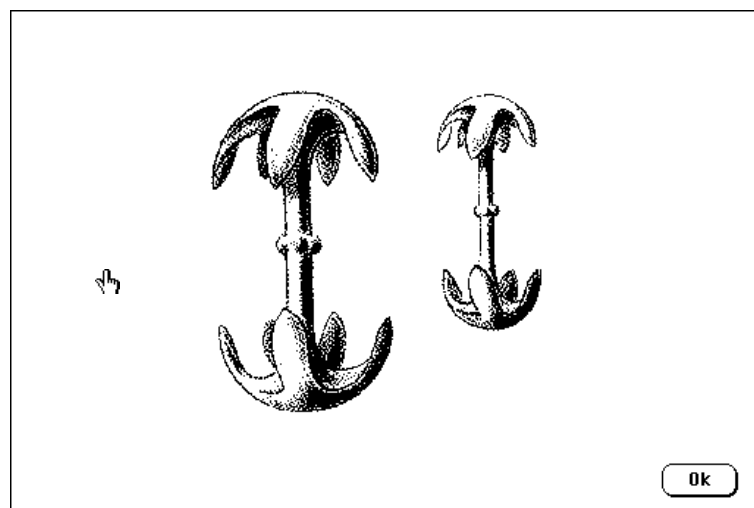


Fig. 6.39 : Illustration par l'utilisateur d'un objet observé : les amphidisks(2) du cas n° 3

6.5.10.2 Le bouton “Naviguer”

Lors de la saisie d'un cas, pour se repérer, l'utilisateur doit pouvoir accéder rapidement à la structure du domaine sur lequel il travaille et ceci à plus forte raison lorsqu'il ne s'agit pas du concepteur du modèle.



Fig. 6.40 : Le bouton “Naviguer” du questionnaire

Par simple clic sur ce bouton (figure 6.40), la vue globale du domaine s'affiche (figure 6.2) et le mode navigation est activé. Grâce à cette vue d'ensemble, l'utilisateur peut situer à nouveau les objets les uns par rapport aux autres et imaginer facilement le chemin qu'il a déjà parcouru, c'est-à-dire repérer les objets qu'il a déjà décrits. La vue d'ensemble permet de passer de la description d'un objet à un autre sans être obligé de parcourir les cartes du questionnaire correspondant aux objets intermédiaires. Lorsque l'objet visé n'a pas déjà été instancié, il n'est pas présent et ne peut donc pas être atteint par ce mode de navigation : ce saut est interdit.

6.5.10.3 Le bouton “Coursier”

Le bouton “Coursier”, situé en bas à droite de chaque carte, permet à tout moment de quitter l'application (figure 6.41) :



Fig. 6.41 : Le bouton “Coursier” du questionnaire

Pour décider si le coursier emporte la description du cas, c'est-à-dire si le cas est enregistré dans la base de cas ou pas, il faut alors répondre à la boîte de dialogue qui est déclenchée après sa course (figure 6.42) :

Vous voulez enregistrer ce cas ou quitter sans enregistrer ?

Fig. 6.42 : Le dialogue permettant de sortir du questionnaire

En choisissant “Enregistrer”, le coursier va stocker la description dans la base de cas à la suite des descriptions déjà enregistrées, sinon “Quitter” n'enregistre pas

le cas dans la base. Il est néanmoins conservé dans le questionnaire et pourra être retrouvé par la procédure “Dernier” (§ 6.5.8.2) lors de la réouverture de cette pile.

6.6 Conclusion

Toutes les caractéristiques et fonctionnalités d’HyperQuest décrites dans ce chapitre ont été implantées pour répondre à l’objectif de robustesse des descriptions en biologie.

Le générateur de modèle descriptif d’HyperQuest permet d’éditer les descriptions observables. Elles doivent répondre à des règles d’élaboration précises sur la cohérence des relations entre objets à observer. Nous avons explicité ces règles et nous les avons illustrées à l’aide de notre application sur les *Hyalonema*. Les définitions des objets du modèle sont ensuite stockées dans des fichiers texte sous forme de *frames*.

Le générateur de questionnaire permet d’automatiser le passage de l’observable à l’observé. Il analyse les frames décrits dans les fichiers et fait correspondre les objets conceptuels structurés sous forme arborescente dans le modèle avec les entités hypertextes du questionnaire. Nous pouvons donc à partir de là construire un questionnaire automatiquement par programme. Nous avons illustré ces différents modes de représentation à l’aide de notre application.

Le questionnaire se présente alors comme un guide d’observation interactif et simple d’emploi. Son utilisation pour la constitution d’une base de cas ou la consultation du système expert doit obéir aux objectifs de mise à jour, de redondance et d’ergonomie afin d’obtenir finalement des descriptions observées les plus robustes possible. C’est à partir de ces cas bien décrits (représentés dans un fichier texte) et des connaissances initiales du modèle que l’on va pouvoir améliorer le processus de classification et de détermination.

