

Le Schéma du modèle conceptuel, étape dans la modélisation des connaissances

Nada MATTA - Nathalie AUSSENAC-GILLES

IRIT - URA 1399 du CNRS
UPS - TOULOUSE

ARAMIHS - UMR 115 du CNRS
MATRA-ESPACE - TOULOUSE

e-mail : {matta, aussenac}@irit.irit.fr

RESUME : Nos travaux sur la méthode d'acquisition des connaissances MACAO nous ont amenés à étudier comment mieux expliciter la méthode de résolution de problème qu'appliquera le système que l'on souhaite réaliser. Nous soulignons l'intérêt de disposer d'une représentation abstraite du modèle conceptuel (MC), que nous appelons Schéma du modèle conceptuel, qui permette de le caractériser et qui rende compte en particulier de la méthode de résolution du système. Nous rapportons alors une expérience de modélisation d'une expertise de diagnostic technique dans le domaine spatial pour explorer plusieurs façons de construire ce Schéma. Nous spécifions ce que doit être le Schéma du MC dans MACAO ainsi qu'un langage adapté à sa représentation. Nous retenons également des recommandations méthodologiques pour le mettre en évidence, en particulier l'intérêt d'étudier de manière approfondie la façon de raisonner de l'expert par catégorie de problèmes.

MOTS-CLES : Acquisition des connaissances, conception de systèmes à base de connaissances, modélisation, méthode de résolution de problèmes, diagnostic.

1 Introduction

Dans le cadre de la construction d'un système à base de connaissances (SBC), un modèle conceptuel (MC) est vu comme une construction abstraite finalisée qui permet de réduire la complexité de l'expertise en se focalisant sur certains aspects des connaissances [KAR 90]. La mise en évidence de plusieurs modèles au cours de l'acquisition permet de s'intéresser successivement à différentes facettes de l'expertise : la tâche réalisée, le type de raisonnement effectué, la coopération possible entre système et utilisateur pour réaliser la tâche, etc. [WIE 92]. Depuis 1989, nos travaux sur la méthode d'acquisition des connaissances MACAO nous ont amenés à étudier comment mieux expliciter la méthode de résolution de problème

qu'appliquera le système que l'on souhaite réaliser¹. Pour cela, nous nous appuyons sur nos expériences de développement de SBC, l'une d'elles étant décrite dans cet article, une autre faisant l'objet d'un deuxième article de cet ouvrage [LEP 95]. Notre objectif n'est pas de présenter les modifications apportées à MACAO (tant à la méthode qu'à la représentation des connaissances) suite à son utilisation mais plutôt de montrer comment celle-ci nous a permis de spécifier des évolutions. Notre texte s'articule autour de quatre parties.

- Nous soulignons l'intérêt de disposer d'une représentation abstraite du MC, que nous appelons *Schéma du modèle conceptuel*, qui permette de le caractériser et qui rende compte en particulier de la méthode de résolution du système (partie 2). Nous donnons un aperçu des problèmes liés à sa représentation et à sa construction de manière générale, en s'appuyant sur un état de l'art.
- Nous présentons la méthode MACAO, la représentation des connaissances sur laquelle elle s'appuie et nous exposons ses limites pour rendre explicite une méthode de résolution de problème (partie 3).
- Nous rapportons alors une expérience de modélisation d'une expertise de diagnostic technique dans le domaine spatial (partie 4) pour explorer plusieurs façons de construire la partie du Schéma relative au raisonnement : par abstraction à partir de problèmes résolus par un expert, tel que le recommande MACAO ; et en adaptant des composants génériques, tel que le propose KADS.
- De l'analyse des modèles obtenus, nous tirons des enseignements sur les évolutions à apporter à MACAO pour représenter et mettre en évidence les méthodes de résolution adoptées dans une expertise (partie 5). Nous spécifions ce que doit être le Schéma du MC dans MACAO ainsi qu'un langage adapté à sa représentation. Nous retenons également des recommandations méthodologiques pour le mettre en évidence : nous proposons de combiner une analyse détaillée des connaissances expertes et l'adaptation de composants génériques.

2 Place d'un Schéma du modèle conceptuel dans l'acquisition des connaissances

2.1 Objectifs de l'étude

Le modèle conceptuel constitue à la fois un cadre pour comprendre et interpréter les informations provenant de l'expert et un langage pour les formaliser en vue de construire un système [AUS 92], [LIN 92]. Il intervient comme un support avec différents rôles selon les étapes du développement du système et de l'acquisition des

¹ Ce choix fait suite à notre participation au projet Sisyphus-II proposé à la communauté de recherche en acquisition des connaissances [SIS 92]. Cette expérience visait à confronter approches et langages de modélisation, en s'intéressant justement à leur capacité à mettre en évidence des méthodes de résolution de problème dans un modèle.

connaissances, qui peut alors être décomposée en quatre phases principales (fig 1) : (1) acquisition de connaissances initiales ou acquisition des connaissances conduite par les données, (2) construction d'un premier MC, parfois appelé *skeletal conceptual model* ou *framework of the conceptual model* [WIE 92], (3) définition du MC complet ou raffinement du MC et (4) validation opérationnelle de ce dernier avant le codage de la base de connaissances. Le contenu de ces étapes est détaillé dans la partie 3.

Notre étude s'est focalisée plus particulièrement sur l'étape (2) et sur le modèle produit à son issue. En effet, l'état de l'art fait ressortir une grande diversité de ce premier modèle. Suivant les approches, il est plus ou moins abstrait ; il est un sous-ensemble du modèle final ou, au contraire, il le décrit de manière plus abstraite. Il est ensuite complété ou bien plus fortement modifié. De plus, nous voulions mieux caractériser et mettre en évidence les méthodes de résolution de problème dans le modèle final. Nous avons cherché, à travers l'état de l'art et par notre expérience d'application de MACAO, à préciser ce qu'était ce premier modèle, comment le représenter, le mettre en évidence et son rôle dans la modélisation.

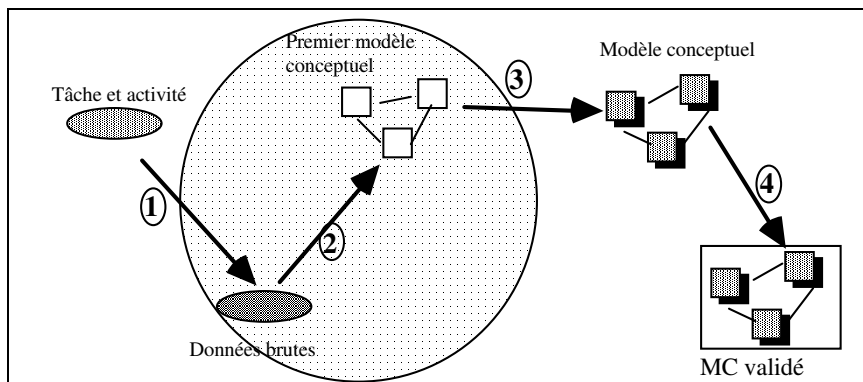


Figure 1. Parmi les différentes étapes de l'acquisition des connaissances, nous nous intéressons à l'étape (2) ainsi qu'à la nature du Schéma du modèle conceptuel (symbolisé par les carrés 'vides' sur la figure).

2.2 La notion de Schéma du modèle conceptuel

Notre hypothèse est qu'il serait pertinent que le résultat de l'étape (2) constitue une première version de ce que nous appelons *Schéma du modèle conceptuel*, dans la mesure où il peut être modifié au cours de l'étape (3). Ce Schéma caractérise les tâches décrites dans le modèle, les méthodes de résolution de problème appliquées et la structuration du domaine associée. Il définit les rôles que jouent les connaissances dans la résolution de problème et les types d'opérations effectuées. Il comporte aussi des indications sur la structuration des connaissances du domaine (modèle causal,

fonctionnel, etc.) en précisant les classes de concepts et la sémantique des relations qui les associe. Il est donc propre à une application mais défini en des termes abstraits ou indépendants du domaine [AUS 94a]. Le Schéma du MC est proche d'une modélisation générique, et peut donc être construit en adaptant un modèle générique. Il constitue un cadre précis à partir duquel le cognicien peut analyser l'expertise et rechercher les connaissances du domaine jouant les rôles qu'il définit.

Ce que nous appelons Schéma du MC correspond exactement au *framework of the conceptual model* dans CommonKADS [WIE 92]. Celui-ci est composé des couches tâches et inférence du modèle. Le raisonnement y est décrit sous forme de rôles sur lesquels s'appliquent des opérateurs, et par le contrôle sur la mise en œuvre de ces opérateurs au cours de la résolution de problèmes. Par contre, il ne comporte aucune indication ou caractérisation des connaissances du domaine, alors qu'on pourrait penser que l'ontologie du modèle de KADS puisse convenir pour compléter la caractérisation du modèle conceptuel. Cette ontologie permettrait d'énoncer clairement les présupposés sur la structuration du domaine induits par la couche inférence [LER 94].

Dans KREST [JON 92] ou OMOS [LIN 93], il n'existe pas d'équivalent direct au Schéma du MC. On pourrait considérer, par analogie avec CommonKADS, qu'il s'agit des modèles des tâches et des méthodes. Mais rien, dans la démarche préconisée par ces approches, ne garantit que ces modèles offrent une caractérisation du raisonnement abstraite ou indépendante du domaine. Toutefois, on peut penser que c'est la description des méthodes qui s'en rapproche le plus, puisqu'il est recommandé dans KREST qu'elles soient une adaptation de méthodes génériques, et dans OMOS qu'elles caractérisent le raisonnement indépendamment du domaine.

Enfin, dans des approches plus anciennes comme les Tâches Génériques (GT) [CHA 86] ou les environnements s'appuyant sur 'des méthodes délimitant des rôles' (RLM) [MAR 88], on peut associer au Schéma du MC le modèle du raisonnement, à savoir le modèle des tâches dans les GT, ou la méthode de résolution de problème dans les RLM. Dans les deux cas, ces modèles manipulent des entités de raisonnement et non les entités du domaine, le lien étant fait plus tard, au moment de détailler le modèle.

2.3 Construction du Schéma du modèle conceptuel : abstraction ou réutilisation

Le Schéma du MC peut être construit par abstraction, par réutilisation ou bien en associant les deux procédés [LIN 92], [KRI 91]. L'abstraction consiste ici à rechercher des traits caractéristiques de l'expertise, à en dégager des buts, des actions et des critères de décision, puis à les exprimer en des termes indépendants du domaine (comme dans MACAO [AUS 94b]). Au contraire, la réutilisation de composants génériques, tirés d'une bibliothèque, consiste à réduire à son minimum une analyse non guidée, à choisir le plus tôt possible un de ces composants pour

revenir analyser l'expertise et construire ainsi progressivement le Schéma du MC [AKK 93]. Selon un premier type de réutilisation, le cogniticien reprend des actions types ou des rôles de connaissances prédéfinis. Il peut aussi agencer des composants génériques non modifiables (comme pour les Tâches Génériques (GT) [CHA 86] ou les premiers modèles d'interprétation de KADS) ou adapter des composants modifiables (comme ceux de la bibliothèque de CommonKADS [BRE 94] ou les méthodes proposées dans KREST). De manière complémentaire, les travaux sur les ontologies ainsi que l'étude des liens entre méthodes de résolution et structuration du domaine fournissent des exemples de caractérisations du domaine.

Or, il semble que la plupart des modèles construits par abstraction uniquement énumèrent sans recul les étapes suivies par l'expert au cours de la résolution de problème et souffrent de la difficulté que rencontre le cogniticien, proche de la démarche et du vocabulaire de l'expert, pour abstraire des concepts et des méthodes. Au contraire, la réutilisation facilite la construction du Schéma dans la mesure où la caractérisation de la résolution de problème est un passage obligé pour sélectionner des composants réutilisables. Ces résultats généraux sont confirmés par notre expérience rapportée en partie 4 ainsi que celle décrite dans [LEP 95]. Elles nous ont permis de comparer la démarche par abstraction de MACAO avec une démarche basée sur la réutilisation comme KADS, et ensuite d'évaluer leur complémentarité pour proposer une manière de les associer [LEP 95].

2.4 Le raffinement du Schéma : une démarche descendante

Dès que l'on dispose d'éléments sur le Schéma du MC, par exemple une première caractérisation de la tâche effectuée, on peut les utiliser pour orienter le recueil de nouvelles connaissances et corriger ou 'instancier' le Schéma. Ce sont les deux facettes de l'étape (3) de l'acquisition. Le cogniticien peut alors chercher à valider ou affiner ces caractéristiques, à identifier comment la tâche se décompose en sous-tâches, etc. . Avec OMOS ou KREST, le Schéma du MC est ainsi enrichi et révisé. Ces évolutions itératives ne sont pas possibles dans les Tâches Génériques ou dans MORE ou MOLE [MAR 88], où le Schéma du MC est figé une fois pour toutes. Détailler ou 'instancier' le Schéma revient à associer aux termes génériques des entités du domaine et à organiser dans le modèle des connaissances spécifiques. Ainsi, à partir de structures d'inférence KADS ou de diagrammes de KREST, on dispose de cadres précis pour situer les connaissances du domaine intervenant dans les processus qu'ils représentent. On peut même définir des outils dédiés au recueil des connaissances comme dans les systèmes de l'approche RLM [MAR 88].

3 MACAO : présentation

Nous rappelons ici les principes de la méthode telle qu'elle était définie en 1993, version utilisée dans notre étude, ainsi que quelques définitions [AUS 89] [AUS 94b]. Nous présentons ensuite les structures de représentation des connaissances

proposées pour construire un MC. Puis nous décrivons le contenu de chacune des étapes de la méthode, afin de mieux en comprendre les points faibles et les points forts. Cet exposé nous permet de situer MACAO par rapport à d'autres approches. Enfin, nous formulons les motivations de notre étude dans le cadre de MACAO.

3.1 Un processus incrémental et ascendant

Avec MACAO, la modélisation des connaissances se fait à la fois de manière incrémentale, en s'intéressant progressivement à des sous-ensembles de plus en plus larges de l'expertise, et par abstraction, car on essaie de décrire la résolution de problème en la caractérisant de manière de plus en plus abstraite. Pratiquement, à ce processus incrémental et ascendant, correspondent trois niveaux de granularité dans l'étude de l'expertise et dans la formulation de la résolution de problème.

- Ainsi, on s'intéresse d'abord à des *problèmes*, assimilables à des cas, du type de ceux que l'expert sait traiter et que le futur système devra résoudre. Un problème est donné au moins par son énoncé, et au mieux par la modélisation de sa résolution et des concepts manipulés. Les résolutions des problèmes étudiés dans MACAO étant faites ou simulées par l'expert, elles reflètent sa démarche.
- Au niveau au dessus, on étudie les raisonnements mis en œuvre pour des problèmes dont les énoncés et/ou les résolutions sont analogues, définissant des *catégories de problèmes*. Le raisonnement valide pour une catégorie de problèmes correspond à une démarche plus générale. L'inspiration en est toujours la manière de procéder de l'expert, que l'on cherche à davantage caractériser. Le cognicien repère ces catégories en classant des problèmes selon la technique des grilles répertoires. Une catégorie est définie par son énoncé, des attributs la caractérisant, une liste de problèmes associés, et surtout par le modèle de sa résolution.
- Enfin, *l'expertise* dans son ensemble est prise en compte, rassemblant les différentes catégories. Expertise fait ici aussi bien référence aux connaissances de l'expert qu'à celles, représentées dans le MC, qui deviendront la base de connaissances du système. En unifiant selon une seule démarche les différents modes de résolutions des catégories, un plus grand nombre de connaissances est pris en compte. Le raisonnement est représenté soit à un niveau plus abstrait, éloigné des particularités contextuelles, soit en faisant apparaître les caractéristiques du contexte qui permettent de se situer dans telle ou telle catégorie.

3.2 Représentation des connaissances

La distinction entre *expertise*, *catégories* et *problèmes* est mise en évidence à l'aide de classes d'objets structurés de même nom. Le MC correspond au modèle de l'expertise. A chaque niveau (*expertise*, *catégories* et *problèmes* ainsi que dans le

Schéma du MC), on distingue les connaissances structurelles du domaine (*modèle du domaine*) des connaissances décrivant le raisonnement (*modèle du raisonnement*).

Nous avons considéré un MC avant tout comme un support pour comprendre et interpréter les informations provenant de l'expert, et beaucoup moins comme un moyen de formalisation, même si nos travaux depuis 1993 portent sur la prise en compte de ce deuxième rôle. De ce fait, nous avons défini une représentation des connaissances qui favorise des mises à jour faciles ; qui autorise des représentations provisoires, pas nécessairement cohérentes ou exactes, mais utiles pour clarifier les idées du cognitif ; qui permette des visualisations simples et faciles à comprendre par un expert. Cette représentation des connaissances offre ainsi un cadre minimal pour assurer la compréhension de l'expertise, qui nécessite de nouveaux raffinements avant d'être formalisée. Elle sert aussi bien à décrire la démarche de l'expert dans le MC, en des termes proches du domaine, qu'à caractériser les méthodes de résolution appliquées par le futur système dans le Schéma du MC. Nous présentons ce langage, écrit à l'aide de la couche objet de Le_Lisp².

3.2.1 Modélisation du domaine

Les connaissances du domaine sont représentées par des *concepts* et des relations (ou *liens*) typées entre les *concepts*. Les connaissances du domaine décrivant un exemple ou une situation spécifique sont représentées par des instances de ces *concepts* reliées par des relations de mêmes types. La liste des types de relations peut être complétée selon les besoins de la modélisation. L'ensemble des *concepts* ou de leurs instances, et des relations forme un réseau de type réseau sémantique : le *Graphe du domaine, générique ou instancié* suivant que sont représentés des *concepts* ou leurs instances. Ces graphes peuvent être vus, entièrement ou partiellement, à l'aide d'éditeurs de graphe, qui en facilitent la construction progressive.

Cette représentation ne prétend pas être un véritable langage dans la mesure où, par exemple, les relations n'ont pas de sémantique précise. Les seules vérifications faites sont syntaxiques : deux instances peuvent être reliées par un lien si celui-ci existe entre les concepts génériques correspondants ; on ne peut relier que des concepts génériques entre eux ou des instances entre elles. Notre objectif est de visualiser des associations entre concepts et de les décrire en vue de leur formalisation.

² Produit de Ilog (F)

3.2.2 Modélisation du raisonnement

La structure de *schéma*³ est utilisée pour représenter le raisonnement [AUS 94c]. Elle représente un pas de raisonnement associé à la réalisation d'un but dans un certain contexte. Un *schéma* contient donc un *contexte*, ensemble des informations retenues pour décrire la situation, le *but* visé, des *paramètres* nécessaires au raisonnement, ces trois champs étant exprimés à l'aide de concepts du domaine et de contraintes sur ceux-ci. Il contient également les actions à effectuer (*traitement*, appel de *schémas* ou de fonctions élémentaires) et la *stratégie* selon laquelle elles sont effectuées. La résolution d'un problème ou le raisonnement général s'expriment par un enchaînement de *schémas*, montré sous forme d'arbre qui constitue le *modèle de raisonnement* d'un problème ou de l'expertise.

Cette structure est comparable à la notion de *but* du langage LISA⁴ [DEL 93] ou de tâche dans KREST [JON 92]. Nous avons d'ailleurs retenu LISA comme langage pour rendre opérationnel les modèles construits avec MACAO. Mais, alors qu'en LISA et dans KREST, buts et méthodes sont distingués, un *schéma* contient à la fois l'expression d'un but et des moyens mis en œuvre pour l'atteindre dans un certain contexte. De plus, les moyens de traduire le contrôle sur l'enchaînement des *schémas* appelés sont limités à quelques opérateurs, alors qu'ils sont plus riches en LISA, puisque c'est un langage opérationnel.

3.3 Les étapes de la méthode⁵

3.3.1 Détermination des matériaux de base

Au cours de cette étape de familiarisation, le cogniticien cherche à comprendre à la fois le domaine d'expertise concerné, le vocabulaire utilisé et les types de problèmes que le système à construire devra résoudre. Il repère et exploite les sources de connaissances disponibles, commence à caractériser le domaine et la démarche générale de raisonnement puis prépare un premier modèle. Cette étape se caractérise par le fait que l'on ne dispose pas encore d'un MC, et que l'on ne sait pas comment vont être exploitées les connaissances recueillies.

³ Pour éviter la confusion entre le nom de cette structure et le Schéma du MC, nous écrivons dans l'article le nom de la structure en italique et le second avec une majuscule.

⁴ Le langage LISA a été développé au sein d'EDF et du LRI (Université d'Orsay) par I. DELOUIS pour programmer des systèmes opérationnels reflétant le MC qui a servi à les spécifier. LISA est écrit en Le_Lisp (ILog).

⁵ Nous insistons sur le fait qu'il s'agit là de la version de la méthode définie en 1993, appliquée dans l'expérience rapportée dans la partie 4, et qui a depuis évolué. Nous utilisons ici le découpage classique du processus d'acquisition des connaissances présenté en partie 2.

Avec MACAO, la familiarisation se fait à partir d'entretiens et d'observations de l'expert en situation de travail. Une autre particularité de MACAO est de proposer de consacrer du temps à cette étape, et de recueillir des connaissances (auxquelles on pourra se référer par la suite) plutôt que de se focaliser sur leur caractérisation⁶. Cette étape amorce un processus incrémental (cf. 3.1). L'étude du raisonnement se fait par l'identification de catégories de problèmes. L'expert simule la résolution de plusieurs problèmes, en décrivant à voix haute ce qu'il fait, puis, après résolution, il explique sa démarche. A partir des verbalisations, le cognicien construit le modèle du raisonnement de chaque problème, et un premier modèle du domaine de l'expertise.

3.3.2 Construction d'un premier Schéma du modèle conceptuel

De manière générale, dès que le cognicien a suffisamment de données, il commence à les structurer et à les organiser dans un premier modèle. Il y décrit la démarche de résolution de problème avec le souci de mettre en évidence les méthodes appliquées. De même, il commence à représenter des connaissances du domaine, à repérer leur organisation à travers les types de relations et de concepts mis en jeux dans le raisonnement.

Selon MACAO, on poursuit le processus d'abstraction (cf. 3.1) pour dégager un premier MC en deux étapes. Dans un premier temps, un modèle de raisonnement est établi pour chaque catégorie de problèmes en analysant les différents modèles des problèmes de cette catégorie. Dans un deuxième temps, le modèle de l'expertise est construit en comparant les modèles des catégories et en complétant le premier modèle du domaine. On continue à recueillir beaucoup de connaissances, la préoccupation essentielle étant de parvenir à une structuration du modèle qui intègre toutes les connaissances disponibles, et non de mettre en évidence des méthodes de résolution.

3.3.3 Définition du modèle conceptuel complet

Le recueil de données du domaine s'effectue ici à partir du premier modèle, qui peut être enrichi ou remis partiellement en question à cette occasion. D'une part, il s'agit d'acquérir, conceptualiser et représenter toutes les connaissances du domaine prévues par le modèle : les instances de concepts, leurs valeurs, les règles et heuristiques correspondant à des pas de raisonnement ou à des méthodes, etc. D'autre part, il s'agit d'enrichir et d'affiner la structuration du modèle, en particulier les étapes et méthodes de raisonnement. Ce cycle de recueil orienté par le modèle est répété jusqu'à ce que cognicien et expert estiment le modèle valide et considèrent que le système fait à partir du modèle vérifiera les spécifications imposées.

⁶ Nous verrons à travers l'expérience présentée dans la partie 4 que ce choix est discutable.

Dans MACAO, le premier modèle d'expertise est complété d'abord à l'aide de connaissances disponibles dans les modèles de problèmes et de catégories, puis de données recherchées auprès de l'expert ou dans des documents. Ce modèle est jugé valide lorsque l'expert a approuvé sa structure et son contenu et que le cognicien pense qu'on pourra le traduire dans un modèle opérationnel.

3.3.4 Validation du modèle conceptuel complet

Cette validation consiste à tester les capacités de résolution de problème du modèle et à vérifier la qualité des solutions qu'il produit. Dans MACAO, elle passe par l'opérationnalisation du modèle. Pour cela, nous avons choisi d'utiliser le langage LISA [DEL 93], dont les structures sont proches de celles de MACAO. La traduction en LISA se fait désormais de manière interactive, le cognicien étant guidé par la correspondance entre structures du MC et structures opérationnelles.

4 Expérience de modélisation d'une expertise de diagnostic technique

4.1 Le problème

L'intégration des différents composants d'un sous-système spatial (partie de lanceur ou de satellite) requiert une phase de test de manière à éprouver la conformité de l'ensemble du système par rapport aux spécifications de conception. Les procédures de test, organisées suivant un plan, provoquent l'envoi de commandes électriques au système, comparent la réponse reçue par rapport aux réponses prévues et signalent éventuellement des incidents à l'opérateur de test. Dès qu'un incident est signalé, l'opérateur commence une démarche de diagnostic dont l'objectif est de localiser l'équipement du système en défaut, et si possible de proposer une cause possible à cet incident [SOL 92]. Dans le cadre du test de la case à équipement du lanceur ARIANE 4 au centre toulousain de Matra Marconi Space, un système expert d'aide au diagnostic a été développé pour assister l'intégrateur dans la démarche de diagnostic et localiser l'équipement défectueux.

4.2 Méthode suivie et modèles obtenus

4.2.1 Détermination des matériaux de base

Comme le propose MACAO, le contexte de l'expertise et la démarche générale ont été étudiés au moyen d'observations de l'activité réelle et d'entretiens, complétés par l'analyse de documents (procédures de tests, documentation de la case, résultats des tests, etc.). Un modèle du domaine a été construit, et complété par la suite. Nous ne le présentons pas ici car nous nous focalisons sur le modèle du raisonnement.

Plusieurs problèmes ont été retenus par l'expert et classés à l'aide des grilles répertoires pour retrouver les catégories de problèmes qu'il a l'habitude de

rencontrer. Pour chaque catégorie, un exemple a été résolu par simulation puis modélisé. Les modèles obtenus décrivent de manière chronologique les actions effectuées pour diagnostiquer les incidents, et ceci avec le vocabulaire du domaine, les noms des *schémas* reflétant les données relatives aux problèmes de la catégorie (fig. 3).

4.2.2 Construction ascendante d'un premier Schéma du MC

L'expert et le cognicien ont cherché à rendre explicite un premier modèle du raisonnement valable pour tous les types de problèmes et décrit avec des termes du domaine. Ce modèle a été établi suivant une démarche ascendante puisqu'il est issu de l'analyse des différents exemples de problèmes [SOL 92]. Le niveau de modélisation des catégories, bien que recommandé par la méthode, n'a pas été exploité. En effet, un seul exemple a été choisi par catégorie, alors que la méthode recommande d'en simuler au moins deux. Cet exemple prototypique a été assimilé à la catégorie et complété de manière à prendre en compte tous les cas associés.

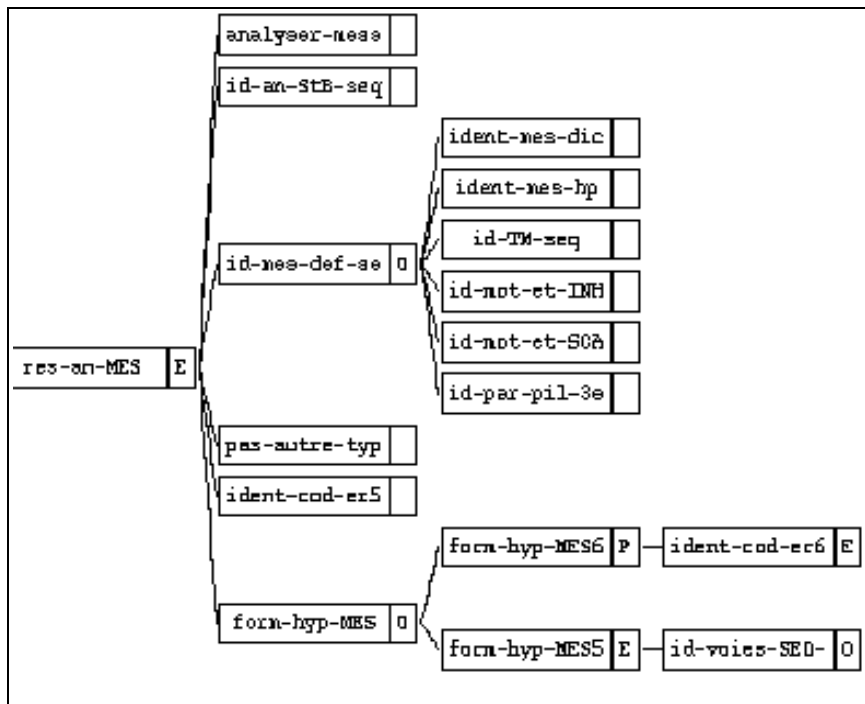


Figure 3. Une partie du modèle du raisonnement d'une catégorie de problèmes (panne indiquée par un mot d'état en mode séquentiel -

MES-). La démarche consiste à analyser le message issu du test (analyser-mess), à vérifier qu'il est apparu en mode séquentiel (id-an-StB-seq), à corrélér ce message à d'autres paramètres (id-mes-def-seq), à vérifier qu'il n'y a pas d'autre type d'anomalie présent (pas-autre-type) puis à tester les différentes hypothèses envisagées dans ce cas (form-hyp-mess). Les lettres majuscules à côté du nom des schémas rappellent l'instruction qui contrôle leur décomposition : E=ET, P=PUIS, O=OU, T=TANT QUE, C=CAS.

La figure 4 illustre le Schéma du Modèle du Raisonnement obtenu. On y retrouve des étapes analogues à celles citées dans l'exemple (fig 3) : la recherche du contexte correspond à *id-an-StB-seq*, la formulation des hypothèses à *id-mes-def-seq*, affiner la situation à *pas-d'autre-type* et *ident-cod-er5*, etc. Cependant, on a différencié la génération de l'évaluation des hypothèses, car la génération peut être complexe alors qu'elle était directe dans l'exemple. De même, l'évaluation des hypothèses était considérée comme un sous-but de leur formulation, hypothèse par hypothèse (elle correspondait aux schémas terminaux décomposant *form-hyp-MES5* et *form-hyp-MES6*), alors qu'il s'agit d'un but du même niveau que la formulation (on formule un ensemble d'hypothèses testées une à une ensuite).

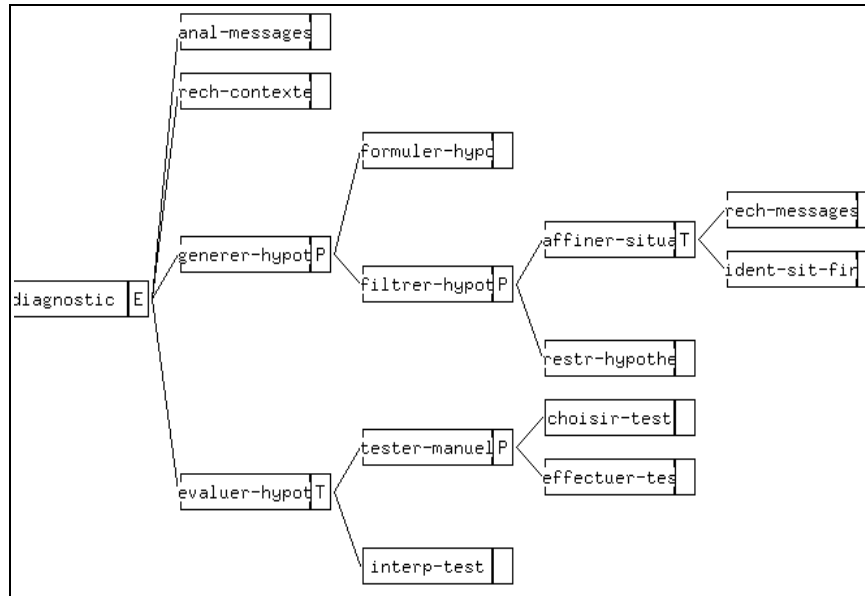


Figure 4. Un premier Schéma du modèle du raisonnement de l'expertise ARIANE 4 : Pour diagnostiquer une anomalie, (diagnostic) on analyse les messages produits à l'issue du test (anal-message), on recherche le contexte de ces messages (rech-contexte),

on génère des hypothèses (generer-hypotheses) et on évalue ces hypothèses (evaluer-hypotheses)...

4.2.3 Amélioration du Schéma du MC : Réutilisation de composants génériques

Nous avons ensuite essayé de réorganiser ce modèle à l'aide des méthodes proposées par les travaux issus de KADS sur le diagnostic [BEN 93a]. Les méthodes proposées pour le diagnostic découpent le raisonnement en trois étapes principales : génération, filtrage puis discrimination des hypothèses (fig. 5). Ces étapes se retrouvent indirectement dans les modèles des problèmes. Une étude approfondie de ces modèles permet de constater que l'expert maîtrise plus ou moins la formulation des hypothèses. Suivant la complexité et la nouveauté du problème, il utilise des méthodes différentes. Il essaie en premier de rapprocher la situation courante de cas déjà rencontrés [BEN 93b]. S'il n'arrive pas à trouver un cas similaire, il met en cause les composants prédécesseurs du composant où le symptôme est apparu suivant un modèle fonctionnel de la case. Cette analyse a permis de choisir des méthodes (comme la génération empirique *M-emp-gen* et la génération suivant le modèle fonctionnel du système *M-mod-bas-gen*) parmi les méthodes de diagnostic présentées dans [BEN 93a] pour décrire cette phase de génération des hypothèses. Elles ont été intégrées dans le modèle.

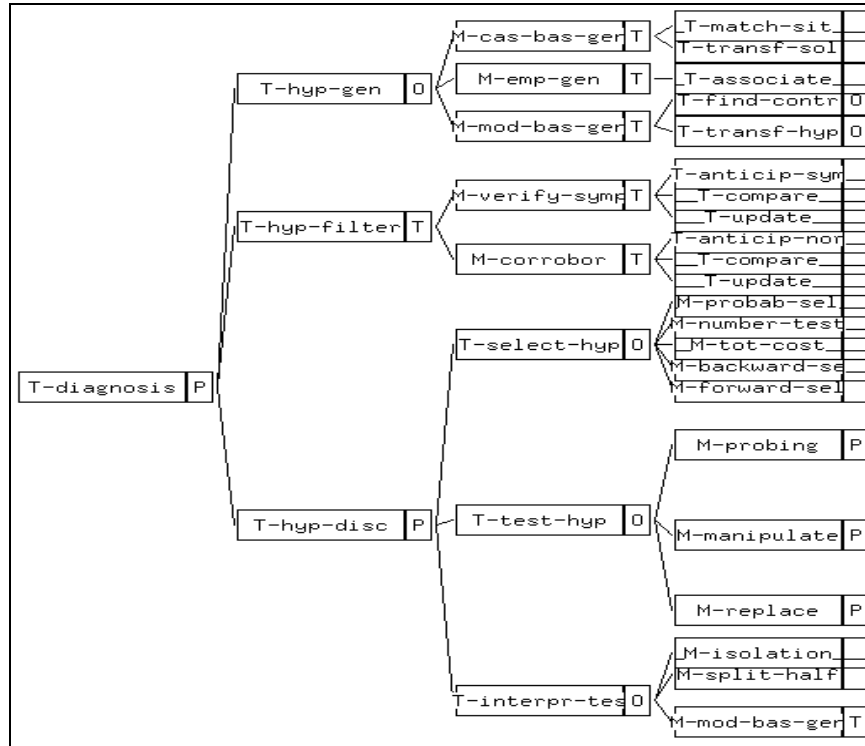


Figure 5. Une partie du Schéma du modèle du raisonnement de l'application ARIANE-4 [BEN 93b]. Les noms des méthodes commencent par un M, ceux des tâches ou buts par un T. Pour réaliser le diagnostic, on génère des hypothèses (T-hyp-gen), on filtre les hypothèses générées (T-hyp-filter) puis on les discrimine pour localiser le défaut (T-hyp-disc).

Avant de commencer à évaluer les hypothèses en testant les composants mis en cause, l'expert procède à un filtrage des hypothèses engendrées en demandant des informations complémentaires à celles prises en compte lors de la détection des symptômes. Deux méthodes choisies dans [BEN 93a] ont permis de bien structurer le processus utilisé pour ce filtrage. En effet, l'expert filtre ses hypothèses en anticipant les effets des symptômes sur l'ensemble des composants (méthode *M-verify-symptoms*) et en prédisant le fonctionnement normal de la case (*M-corroborate*). Au lieu de tester tous les composants mis en cause par les hypothèses, l'expert choisit l'hypothèse la plus prioritaire en fonction du pourcentage de panne des composants, du coût des tests et de la facilité d'accès aux composants. Ces trois critères ont été indiqués pour préciser les méthodes de choix prédéfinies dans [BEN 93a]. De la

même manière, plusieurs méthodes sont applicables pour interpréter les tests des hypothèses, suivant la nature du test et la maîtrise qu'en a l'expert. Ce second Schéma du MC est présenté figure 5.

4.2.4 Définition et validation du MC complet

Les contraintes liées au projet nous ont conduit à aménager la mise en œuvre des deux dernières étapes de MACAO. C'est en fait le premier Schéma de MC construit qui a servi à organiser les connaissances détaillées disponibles dans les modèles des problèmes, ainsi qu'à en recueillir de complémentaires. Suite à la construction du second Schéma, les connaissances détaillées ont été réorganisées. Ces modèles ont été vérifiés par l'expert puis codés dans l'environnement Nexpert pour la validation fonctionnelle.

4.3 Analyse des modèles

Pour comparer les deux modèles obtenus, reprenons le *schéma formuler-hypothèse*. Dans le modèle obtenu selon MACAO (fig. 4), ce *schéma* n'est pas décomposé. En effet, les différentes solutions décrites dans les modèles des problèmes ont été jugées trop spécifiques pour être rassemblées sous un intitulé qui les englobe et permette de reconnaître deux classes de méthodes. Si on reprend le problème modélisé sur la figure 3, on constate que deux hypothèses sont formulées (MES5 et MES6) mais le contenu des *schémas* ne dit pas comment on formule ces hypothèses, pourquoi ces deux-là et pas d'autre. Si on poursuit la décomposition de ces *schémas*, on trouve que l'on cherche à identifier dans un cas un code d'erreur (n°6) et dans l'autre que l'on teste le SED. Ces deux buts reviennent à affiner la situation, étape suivante du raisonnement dans l'arbre de la figure 4.

Ce sont les modèles des catégories qui ont servi de support privilégié au dialogue entre expert et cognitif, plus que le premier Schéma du MC, pourtant jugé satisfaisant par l'expert. Même s'il est formulé en des termes indépendants du domaine, il ne traduit donc pas *comment* se déroule la résolution de problèmes, mais seulement les *buts* que l'on cherche à atteindre. Ce résultat est renforcé par la représentation sous forme d'arbre, qui escamote complètement le contrôle sur l'enchaînement des *schémas*, ainsi que la nature des données manipulées.

Que devient la génération des hypothèses dans le deuxième modèle ? En fait, le premier *schéma* consiste à générer des hypothèses (*T-hyp-gen*). Il se décompose en trois méthodes possibles (OU est l'instruction de contrôle). La première, génération basée sur des cas déjà rencontrés (*M-cas-bas-gen*), est celle appliquée dans l'exemple (Fig. 3). La situation est ensuite reconnue (*T-match-sit*), ce qui correspond à l'analyse de messages et aux quatre *schémas* suivants (recherche du contexte) dans l'exemple. La solution connue pour cette situation est évoquée (*T-transf-sol*), ce qui correspond à la formulation des hypothèses MES5 et MES6. Ce deuxième modèle

est donc beaucoup plus précis dans sa façon de présenter les méthodes mises en œuvre, et plus fouillé. Par contre, il n'est pas forcément plus facile à lire car il est plus éloigné de l'exemple que le premier. Surtout, il faut de souligner l'utilité du premier pour assurer la transition entre l'exemple et le deuxième Schéma.

Enfin, il est intéressant de comparer le premier Schéma au code de l'application, puisque c'est ce modèle qui a été utilisé. On constate que la résolution des différentes catégories a été programmée en restant très près des modèles associés, le modèle de l'expertise ayant peu joué son rôle de 'canevas unificateur'. C'est aussi de cette façon que l'on a pu constater que la formulation des hypothèses se décomposait de plusieurs manières différentes : la programmation a donc été l'occasion de noter des connaissances détaillées qui n'apparaissaient dans aucun modèle.

5 Bilan de l'expérience et discussion

Même si cette utilisation de MACAO n'ait pas été réalisée dans des conditions expérimentales complètement définies (le degré de maîtrise et le respect de la méthode par le cognicien sont difficiles à mesurer), l'analyse des modèles obtenus fait ressortir des limites de MACAO. Nous en avons recherché les causes dans la démarche de modélisation et dans la représentation des connaissances. Nous rapportons ici nos conclusions et nos propositions pour les modifier dans l'objectif de mieux représenter et modéliser les connaissances associées à la résolution de problèmes, et de préciser ce qu'est le Schéma du MC dans MACAO.

5.1 Bilan sur la méthode

En suivant la démarche proposée par MACAO, le MC construit ne caractérise pas vraiment l'expertise et se contente de la décrire. Ce modèle ne fait référence à aucune classe de problèmes et ne présente pas explicitement les méthodes de résolution appliquées. On peut dire qu'il n'y a pas 'spontanément' de Schéma du MC. On obtient un 'meilleur' modèle avec KADS, où le raisonnement est mieux caractérisé et où le lien avec le domaine doit être assuré systématiquement. Nous fixons donc comme objectif de l'étape 2 la construction d'un premier Schéma du MC, détaillé et corrigé à l'étape 3, en insistant sur la caractérisation de l'expertise et sur la mise en évidence des méthodes de résolution utilisées.

Les modèles des catégories ont été largement utilisés, en particulier pour répondre aux questions nécessaires à la sélection de composants génériques et pour organiser un modèle du raisonnement selon KADS. Ces modèles fournissent la réponse à ces questions sans recueil supplémentaire de connaissances. Nous conseillons donc désormais de les construire le plus tôt possible et d'y consacrer plus de temps qu'aux modèles des problèmes. Mais ces modèles de catégories auraient pu servir plus. En effet, le cognicien a trouvé fastidieux d'en exploiter le contenu pour détailler le MC. Il voulait très vite voir comment le système raisonnait, ce qu'il a fait

en programmant directement les modèles des catégories. La mise en œuvre de MACAO a donc été ressentie comme fastidieuse. Pour éviter cela, nous proposons de construire incrémentalement le modèle jusqu'à son opérationnalisation, en couvrant des parties de l'expertise de plus en plus larges. Jusqu'ici, même si on analysait le raisonnement incrémentalement, à partir de l'étape 2, le modèle englobait toute l'expertise. Or, un modèle partiel permettrait de tester rapidement la validité des choix de modélisation. Ce souhait fait aussi ressortir le besoin d'opérationnaliser facilement le modèle, ce que nous avons assuré grâce à un module qui guide la traduction du MC vers le langage opérationnel LISA.

Par ailleurs, nous envisageons de proposer de construire le Schéma du MC non seulement par abstraction, mais aussi en exploitant des bibliothèques de composants génériques comme celle de CommonKADS. Il y aurait alors deux façons d'exploiter les modèles des catégories pour construire le Schéma du MC : pour choisir des composants de modèle dans une bibliothèque, ou pour identifier directement ce que peut être le Schéma du MC. Les modèles des catégories fournissent en effet des connaissances précises pour sélectionner et adapter un élément d'une bibliothèque. Le Schéma du MC contiendrait alors des méthodes et des tâches soit généralisant ceux des modèles des catégories, soit adaptés de composants génériques.

5.2 Bilan sur la représentation des connaissances

Il semble que d'autres causes de la difficulté à exprimer une méthode de résolution de problème avec MACAO soient relatives à la représentation des connaissances, autant aux structures disponibles qu'à la présentation graphique du modèle.

La structure de *schéma* se rapproche plus de la notion de but que de méthode. De plus, la présentation du modèle sous forme d'arbre de *schémas* évoque une décomposition en sous-buts mais ni la transformation des données au cours de la résolution de problème ni le contrôle. Or ces représentations s'avèrent nécessaires pour traduire une méthode de résolution de problème. L'arbre des *schémas* offre une vision hiérarchique et arborescente avec quelques opérateurs contrôlant la décomposition. Par contre, les boucles ou la logique d'enchaînement des tâches ne sont pas accessibles. Les entrées (contexte) et les résultats (buts) d'un *schéma* ne peuvent être connus qu'en consultant le contenu du schéma. De plus, ce sont systématiquement des concepts, ce qui rend difficile une traduction abstraite de la résolution de problème (soit on n'utilise que des concepts correspondant à des objets du domaine, et le modèle n'est pas abstrait ; soit on définit artificiellement des concepts abstraits qui correspondent en fait à des rôles dans le raisonnement).

La structure de *schéma* est donc abandonnée pour mieux différencier tâche et méthode, notions empruntées à LISA et KREST [DEL 93] [JON 92]. La *tâche* fait référence au *but* à atteindre et aux différentes *méthodes* accessibles pour cela dans un

contexte donné ; la *méthode* traduit une *procédure* applicable dans un certain *contexte*, qui contrôle explicitement la décomposition de la tâche en sous-tâches ou qui déroule les actions (ou opérations de base) nécessaires à la réalisation de la tâche sans en avoir une représentation explicite. On peut associer une ou plusieurs méthodes à une tâche. Nous avons aussi introduit la notion de *rôle*, empruntée à CommonKADS, pour décrire les objets du raisonnement indépendamment du domaine

Chaque structure comporte trois parties, qui permettent de la définir de manière de plus en plus précise et formelle : une description en langage naturel ; une description graphique à laquelle est associée une structure interne ; une formalisation opérationnelle à l'aide de LISA ou Le_Lisp. Enfin, d'autres présentations du modèle sont envisagées en plus de l'arbre d'enchaînement des tâches et méthodes, l'une traduisant le flot des données et l'autre le contrôle sur la résolution de problème.

6 Conclusion

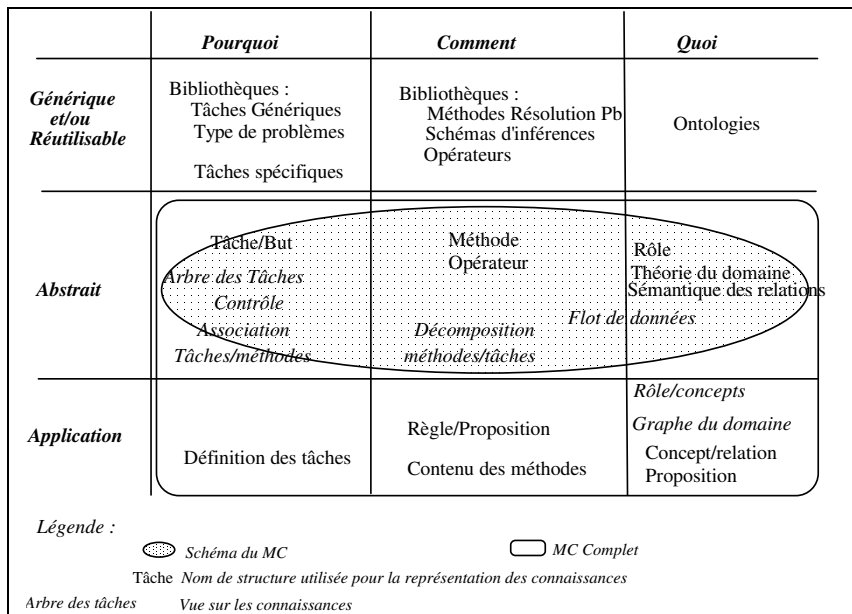


Figure 6. Les composants du MC de MACAO et leurs relations avec des composants génériques et/ou réutilisables. Verticalement, on distingue le niveau d'abstraction par rapport aux applications. Horizontalement, les structures de représentation sont regroupées en fonction des connaissances représentées : Pourquoi -> buts

poursuivis au cours de la résolution ; Comment -> manière d'obtenir ces buts, procédures ; Quoi -> objets du raisonnement.

Pour conclure, nous récapitulons sous forme de tableau (Fig. 6) nos différents résultats relatifs à ce que peut être un Schéma du MC en général, et dans MACAO en particulier, comment le représenter et en quoi il exprime les méthodes de résolutions qui seront appliquées par le futur système. Notre approche tend à rejoindre la majorité des travaux actuels. Mais elle est originale par la place accordée à l'analyse de la façon dont l'expert procède, avant de construire un MC adapté aux spécifications du système à concevoir.

Dans MACAO, le Schéma du MC est constitué des structures de représentation du raisonnement et de la théorie du domaine. Il est montré à l'aide des différentes vues qui associent des structures. Pour le mettre en évidence, nous proposons de combiner une analyse détaillée des connaissances expertes à l'adaptation de composants génériques, choisis et adaptés de bibliothèques. Les modèles de catégories de problèmes, qui, dans MACAO, représentent la façon dont l'expert résout un ensemble de problèmes analogues, s'avèrent alors particulièrement utiles. Le MC Complet regroupe, en plus du Schéma, les définitions des structures qui le composent et établit le lien avec les connaissances du domaine.

Remerciements

Nous remercions C. Soler, D.P. Vo et P. Lépine d'avoir utilisé MACAO pour modéliser leurs applications et pour les échanges fructueux que nous avons eus sur ce travail. Merci à R. Benjamins pour sa coopération à la modélisation selon KADS de l'expertise ARIANE 4, ainsi qu'aux relecteurs attentifs des premières versions de cet article.

Références

- [AKK 93] AKKERMANS H., WIELINGA B., SCHREIBER G., Steps in constructing Problem Solving Methods, *Proc. of EKAW'93*, AUSSENAC N., BOY G., GAINES B., GANASCIA J.G., KODRATOFF Y. Eds, Lecture Notes in AI 723. Bonn : Springer Verlag. 1993
- [AUS 89] AUSSENAC N., *Conception d'une méthodologie et d'un outil d'acquisition des connaissances expertes*, Thèse de Doctorat en Informatique de l'Univ. P. Sabatier, Toulouse (F), Oct. 1989.
- [AUS 92] AUSSENAC-GILLES N., KRIVINE J.P., SALLANTIN J., Editorial. *Revue d'Intelligence Artificielle*, Numéro spécial sur L'Acquisition des Connaissances. Ed. : AUSSENAC-GILLES N., KRIVINE J.P., SALLANTIN J., Paris : Hermès, 6, N°2. 1992. p 7-18
- [AUS 94a] AUSSENAC-GILLES N., MATTA N., Problèmes méthodologiques liés à la conception d'un modèle conceptuel avec MACAO, *Actes des JAC'94*, Strasbourg (F), mars 1994.
- [AUS 94b] AUSSENAC-GILLES N., PIDOUX C., *Guide méthodologique de MACAO*, Rapport IRIT 94-R-017, IRIT (Toulouse). Oct. 1994.

- [AUS 94c] AUSSENAC-GILLES N., MATTA N., Making the method of problem solving explicit with MACAO : the Sisyphus case-study, *Int. Journal of Human-Computer Studies*. 1994. **40**, 193-219.
- [BEN 93a] BENJAMINS R., Problem Solving methods of diagnosis, *Thesis Universiteit van Amsterdam*, With index ref. ISBN 90-9005877-X, Amsterdam, 1993.
- [BEN 93b] BENJAMINS R., *Report of work at Aramihs*, Toulouse, June 1993.
- [BRE 94] BREUKER J., VAN DE VELDE W., *The commonKADS library for expertise modelling : reusable problem solving components*, Amsterdam : IOS Press. 1994.
- [CAU 95] CAUSSE K., Un modèle pour les connaissances de contrôle, In *Acquisition et ingénierie des connaissances : tendances actuelles*. Eds : N. AUSSENAC-GILLES, P. LAUBLET, C. REYNAUD. Toulouse : Cépaduès. 1995.
- [CHA 86] CHANDRASEKARAN B., Generic Tasks in Knowledge based reasoning : High-level building blocks for Expert System Design, *IEEE Expert*, Autumn 86, pp. 23-30.
- [DEL 93] DELOUIS I., LISA : Un langage réflexif pour la modélisation du contrôle dans les systèmes à base de connaissances. Application à la planification des réseaux électriques, *Thèse de l'Université de Paris Sud, Centre d'Orsay*, Paris, 1993.
- [GRE 92] GREBOVAL C., SATIN : un modèle expert opérationnel, *Actes des 1^o Rencontres Nationales des Jeunes Chercheurs en IA*, Rennes, 6-9 sept 1992.
- [GRE 95] GREBOVAL C., KASSEL G., Opérationnalisation de modèles conceptuels: le générateur AIDE In *Acquisition et ingénierie des connaissances : tendances actuelles*. Eds : N. AUSSENAC-GILLES, P. LAUBLET, C. REYNAUD. Toulouse : Cépaduès. 1995.
- [JON 92] JONCKERS V., GELDOLF S., DE VROEDE K., *The Commet methodology and workbench in practice*, VUB AI memo 92-8, April 92.
- [KAR 90] KARBACH W., LINSTER M., VOSS A., Models, methods and tasks : many labels - one idea ?, *Knowledge Acquisition*, **2** (4) : pp. 279-300, déc. 1990.
- [KRI 91] KRIVINE J.P., DAVID J.M., L'acquisition des connaissances vue comme un processus de modélisation ; méthodes et outils, *Intellectica*, (12) Paris, dec. 1991.
- [LEP 95] LEPINE P., AUSSENAC-GILLES N., Modélisation de la résolution de problèmes : comparaison expérimentale de KADS et MACAO, In *Acquisition et ingénierie des connaissances : tendances actuelles*. Eds : N. AUSSENAC-GILLES, P. LAUBLET, C. REYNAUD. Toulouse : Cépaduès. 1995.
- [LER 94] LEROUX B., Eléments d'une approche constructive de la modélisation et de la réutilisation en acquisition des connaissances, *Thèse de l'Université Paris VI, Spécialité Informatique*, Déc 1994.
- [LIN 92] LINSTER M., *Knowledge Acquisition Based on Explicit Methods of Problem Solving*, Ph.D. Dissertation, D 386, Univ. of Kaiserslautern (G), Feb. 1992. 220 p.
- [MAR 88] MARCUS S. Ed., *Automated Knowledge Acquisition for Expert Systems*, Boston : Kluwer Academic Publishers, 1988, 270 p.
- [SIS 92] *Sisyphus'92 : Models of problem solving*, Ed. M. LINSTER, Rapport interne GMD n° 630 (Bonn, G), march 1992.
- [SOL 92] SOLER C., Système Expert d'aide à l'intégration Ariane, *mémoire d'ingénieur CNAM* - Toulouse, déc. 1992.
- [WIE 92] WIELINGA B., SCHREIBER G., BREUKER J., KADS : a modelling approach to knowledge acquisition, Special Issue, *Knowledge Acquisition*, **4**, (1), 1992. pp. 5-54