
Projet SHERPA

Modèles et bases de connaissances à objets

Localisation : *Grenoble*

Mots-clés : représentation de connaissances, base de connaissances, système à base de connaissances, modélisation, classification, catégorisation, révision, coopération, CO4, modèle à objets, points de vue, Tropes, similarité, typage, contrainte, modèle de tâches, résolution de problèmes, assistance à l'utilisateur, mémoire d'entreprise, mémoire technique, modélisation en biologie, génétique, génome, PowerGene, diagnostic médical, électromyographie, Myosys, World Wide Web.

1 Composition de l'équipe

Responsable scientifique

François Rechenmann DR Inria

Secrétaire

Françoise De Coninck, CDD Inria, Grenoble

Personnel Inria

Jérôme Euzenat, CR
Pierre Fontanille, IE

Personnel des établissements partenaires

Jérôme Gensel, MC, université Pierre Mendès-France
Michel Page, MC, université Pierre Mendès-France
Sylvie Robert, IR, université Pierre Mendès-France (mi-temps)
Patrice Uvietta, IE, INPG
Danielle Ziébelin, MC, université Joseph Fourier (en détachement sur un poste de CR Inria depuis le 1^{er} octobre)

Ingénieur expert

Loïc Tricand de la Goutte, (depuis le 1^{er} octobre 1996)

Collaborateur extérieur

Gilles Bisson, CR CNRS (UMR Leibniz-Imag)

Chercheurs doctorants

Isabelle Crampé, allocataire MENESR
Karine Duprez, convention Cifre
Sueli Ferreira Da Cunha, boursière CAPES, Brésil
Jean-Marc Gabriel, allocataire MENESR
Florence Lemaire, allocataire MENESR
Thibaut Parmentier, allocataire MENESR
Petko Valtchev, allocataire MENESR

Stagiaire

Jean-Yves Morneau, stagiaire CNAM, Grenoble

2 Présentation du projet

Les bases de connaissances scientifiques et techniques rassemblent et structurent des connaissances de plusieurs natures : descriptives sur les entités du domaine concerné, fonctionnelles sur le comportement de ces entités, et prescriptives sur les moyens de compléter et caractériser ces descriptions. De plus, ces connaissances formalisées doivent être associées à des connaissances terminologiques et textuelles destinées à les justifier et les commenter. Le projet SHERPA conçoit et étudie des modèles de connaissances et leurs mécanismes d'exploitation qui soient adaptés aux exigences des bases de connaissances scientifiques et techniques. Il développe un système de gestion qui autorise une construction incrémentale et coopérative de ces bases et expérimente ses résultats en participant à la construction de grandes bases de connaissances en biologie moléculaire et à des efforts d'élaboration de mémoires techniques en milieu industriel.

3 Actions de recherche

3.1 Le modèle à objets et points de vue multiples Tropes

Participant : Jérôme Euzenat

Tropes est le modèle de connaissances à objets et points de vue multiples, conçu et implémenté depuis plusieurs années par le projet SHERPA. C'est autour de ce système que sont capitalisés les différents développements logiciels menés dans le projet, eux-mêmes matérialisant des résultats de recherches fondamentales. Tropes permet ainsi d'expérimenter nos résultats à travers son utilisation dans plusieurs applications. Si la biologie moléculaire est historiquement le domaine privilégié, il n'est pas exclusif. Tropes est par exemple utilisé dans un contexte de mémoires techniques, en particulier avec SGS-Thomson dans le cadre du projet STORIA (cf. 4.1). Il est également utilisé dans le cadre du projet LIPSI de modélisation des connaissances dans un système d'enseignement des langues (participant : J. Gensel), mené en collaboration avec plusieurs équipes de recherche de l'université Pierre Mendès-France (cf. 3.9).

3.2 Serveurs de connaissances

Participant : Jérôme Euzenat

Le système Tropes a été complété par une version « serveur de connaissances » ([276],[277]). Ce serveur permet d'accéder aux objets de la base au travers du Web comme s'il s'agissait de pages HTML. Cependant la puissance du modèle à objets permet de faire des requêtes structurées sur ces objets qui dépassent la recherche usuellement disponible sur un site Web. Un tel serveur de connaissances permet

par ailleurs de structurer une base de documents (hypertextuels ou non) autour de la connaissance formalisée dans les objets. Ce travail a été fort utile dans le cadre de la conception de *FirstFly*, une base de connaissances sur les interactions géniques (cf. 3.7.3) puisque les biologistes, qui ne disposaient pas de machines sur lesquelles Tropes est porté, ont pu voir le contenu de leur base de connaissances via le réseau.

3.3 Construction incrémentale et coopérative de bases de connaissances

Participants : Isabelle Crampé, Jérôme Euzenat, Jean-Marc Gabriel, Petko Valtchev

Le développement d'une base de connaissances est un processus incrémental et coopératif, au sens où plusieurs personnes sont susceptibles d'y contribuer. Nous étudions donc les moyens qui permettent de gérer ce processus de conception, d'une part en proposant un protocole qui régit les interactions entre la base et les personnes qui y contribuent, d'autre part en étudiant les mécanismes de révision de la base de connaissances.

3.3.1 Protocole de coopération

Ce protocole, baptisé CO4, repose sur la métaphore de la soumission d'articles dans une communauté scientifique [275]. Il doit garantir que la base de connaissances est consensuelle. Ce protocole a été implémenté sous la forme d'une bibliothèque.

Un tel schéma de construction coopérative permet de développer des bases de connaissances scientifiques, mais aussi des mémoires techniques. Il fera ainsi l'objet d'expérimentations dans le cadre de STORIA (cf. 4.1).

3.3.2 Révision des bases de connaissances à objets

La mise en œuvre complète du protocole CO4 nécessite d'apporter une aide aux utilisateurs dans la gestion des conflits qui peuvent apparaître lorsqu'ils modifient la base. En effet, si leur modification entre en contradiction avec une connaissance déjà présente dans la base mais qu'ils n'ont pas introduite, il est important de pouvoir les aider à résoudre le conflit le mieux possible [271, 272]. Aussi nous sommes attachés à définir la notion de révision minimale d'une base de connaissances à partir de la sémantique de Tropes exprimée en théorie des modèles [270]. Nous avons en particulier montré la complétude entre la notion syntaxique de révision minimale et celle, sémantique, de révision la plus proche d'une base donnée. Les opérateurs de révision sont en cours d'implémentation sur un sous-ensemble du système Tropes. Les travaux, tant théoriques que pratiques, seront étendus progressivement à l'ensemble du système.

3.3.3 Explication et classification conceptuelle

Participants : Jean-Marc Gabriel, Gilles Bisson

La classification conceptuelle est un mécanisme permettant de construire automatiquement une hiérarchie de classes à partir d'un ensemble caractéristique d'instances du domaine que l'on cherche à modéliser. Cependant, la qualité de la modélisation issue d'un système de classification dépend totalement de la pertinence des données initiales. En pratique, il est donc nécessaire d'effectuer plusieurs itérations de type « essais/erreurs » au cours desquelles, en fonction des résultats obtenus, les données initiales sont révisées par l'utilisateur (ajout d'instances, modification du langage de description des instances...). Ce travail vise ainsi à coupler le système de classification avec un système explicatif capable d'aider l'utilisateur à diagnostiquer les erreurs et à formuler les corrections à effectuer. Au cours de cette année, nous avons développé un prototype d'interface explicative [279, 280], basé sur l'architecture proposée par Gilles Bisson, du laboratoire Leibniz ; elle repose sur une navigation hypertextuelle implémentée

en HTML. Cette interface sera prochainement connectée au système d'apprentissage KBG, développé antérieurement par Gilles Bisson.

3.3.4 Similarité et construction automatique de taxonomies

Participant : Petko Valtchev

Les travaux sur une similarité topologique, fondée sur les taxonomies auxquelles appartiennent les objets et les types de données, ont été poursuivis. L'objectif est de construire automatiquement des taxonomies d'objets plus complexes [269]. Un algorithme de catégorisation est ainsi en cours d'implémentation. En partant des taxonomies sur les attributs d'un concept de Tropes, il propose une taxonomie de classes.

3.4 Connaissances prescriptives et résolution coopérative de problèmes

Participants : Sueli Ferreira Da Cunha, Jean-Yves Morneau, Thibault Parmentier, François Rechenmann, Danielle Ziébelin

La résolution d'un problème revient très souvent à choisir des modules algorithmiques, puis à enchaîner leurs exécutions. Les modèles de tâches permettent de décrire, d'une part les modules en termes des objets qu'ils acceptent en entrée et produisent en sortie, d'autre part des enchaînements de ces modules (stratégies) en vue de la résolution d'une classe de problèmes. La résolution proprement dite d'un problème consiste alors à exécuter la tâche associée en la décomposant récursivement et de façon opportuniste (c'est-à-dire en tenant compte des résultats intermédiaires obtenus) en sous-tâches plus élémentaires jusqu'à l'exécution directe de procédures.

Ces modèles de tâches sont étudiés depuis plusieurs années au sein du projet. Ils ont fait l'objet de plusieurs expérimentations, la plus aboutie étant l'environnement d'aide à l'analyse de séquences génomiques appelé *PowerGene* (cf. 3.7.1). Ils continuent de faire l'objet de recherches en particulier sur l'ajout de stratégies et la résolution distribuée.

3.4.1 Ajout de stratégies de résolution

Participants : Sueli Ferreira Da Cunha, Danielle Ziébelin

Ajouter une stratégie de résolution, qui permette de résoudre un nouveau type de problème, se fait par la construction d'un nouveau schéma (ou plan) de décomposition de tâches en sous-tâches, à moins que l'on ne sache réutiliser un plan existant. Les principales difficultés rencontrées concernent alors la recherche et l'appariement des plans. Rechercher pour le plan à construire un ou plusieurs plans appropriés ne peut se faire sans une évaluation préalable des ressemblances et des dissemblances entre le problème à résoudre et les différents plans existants. Il faut donc rechercher une mesure de similarité appropriée au formalisme de résolution de problèmes utilisé, en identifiant une similarité de contenu, une similarité de structure, et une similarité de résultat.

Notre effort a surtout porté sur l'étude des similarités de structures dans le but de comparer différentes représentations d'une même tâche. Nous avons défini un modèle de représentation canonique des graphes de décomposition et de spécialisation de tâches en utilisant les structures de E-graphes (graphes orientés sans circuit ni sommet isolé, possédant un seul sommet d'entrée et un seul de sortie). Des opérations de concaténation et d'alternance ont été définies, les propriétés d'associativité et de distributivité de la concaténation par rapport à l'alternance ont été vérifiées. Ce travail s'est concrétisé par l'implémentation d'un module de recherche de formes canoniques pour les différents modèles de tâches développés au sein du projet. Ce module devra prochainement être associé à un module d'évaluation des ressemblances et des dissemblances entre hiérarchies de tâches [278].

3.4.2 Environnement distribué de résolution de problèmes

Participants : Thibaut Parmentier, François Rechenmann, Danielle Ziébelin

Dans le cadre du projet de développement d'un système d'aide à l'analyse de séquences génomiques (cf. 3.7.1), un environnement générique de résolution de problèmes par décomposition et spécialisation de tâches a été réalisé en 94 et 95. Nous avons développé une version répartie de cet environnement : les tâches et les méthodes peuvent être situées sur des sites différents, mais la résolution d'un problème se déroule comme si elles se trouvaient sur la machine de l'utilisateur. Une première maquette a été réalisée dans le cadre d'un projet de DESS « Génie Informatique ». Elle utilise le protocole *http*. Son principe est de rapatrier la description des tâches, puisqu'elles sont interprétées, mais de laisser les méthodes s'exécuter à distance dans leur contexte correct et complet. Cette maquette est actuellement réétudiée en collaboration avec le projet SIRAC afin de faire l'objet d'une version diffusable auprès de nos partenaires biologistes.

3.5 Connaissances fonctionnelles

Participants : Jérôme Gensel, Michel Page, Sylvie Robert

Dans le cadre d'une convention entre l'Inria Rhône-Alpes et l'université Pierre Mendès-France de Grenoble, le projet SHERPA accueille M. Page et S. Robert, respectivement maître de conférences et ingénieur de recherche, et aborde le problème de l'introduction de mécanismes de simulation dans un système de représentation des connaissances par objets. L'objectif applicatif est d'abord de concevoir et de réaliser un environnement pour la construction et la simulation de modèles combinant connaissances qualitatives (objets) et quantitatives (systèmes d'équations et/ou contraintes numériques). À plus long terme, ces outils et techniques pourront être utilisés pour représenter les connaissances de comportement en biologie moléculaire, en particulier sur les interactions géniques impliquées dans le développement (cf. 3.7.3).

3.6 Connaissances terminologiques et textuelles

Participants : Florence Lemaire, François Rechenmann

Le processus incrémental et coopératif de construction d'une base de connaissances scientifiques, par exemple en biologie moléculaire, soulève le problème de l'intelligibilité du contenu de la base pour un spécialiste qui n'a pas participé à son élaboration [268].

Trois types d'information nécessaires à la compréhension des connaissances représentées ont été identifiés : définition de la terminologie, justification de la validité scientifique des connaissances et explication des choix de modélisation effectués. Pour que ces trois types d'informations soient accessibles à partir des connaissances formalisées, nous proposons un environnement informatique composé de quatre bases : une base de connaissances, une base de modélisation, une base de textes et une base terminologique. Cette dernière, SiLex [267], contient la définition de l'ensemble du vocabulaire utilisé dans l'environnement et constitue l'interface entre le vocabulaire de l'utilisateur et le contenu de la base de connaissances et de la base de textes. Par ailleurs, SiLex permet une modélisation plus homogène, car elle favorise un accord entre les concepteurs d'une communauté sur l'acceptation des termes utilisés. Un prototype de SiLex a été réalisé dans un domaine de la biologie moléculaire.

3.7 Modélisation des connaissances en biologie moléculaire

Participants : Jérôme Euzenat, François Rechenmann, Danielle Ziébelin

La biologie moléculaire est le domaine d'expérimentation privilégié du projet SHERPA depuis une dizaine d'années. Sur ce thème de recherche, le projet a travaillé en collaboration étroite en 1996 avec Gilles Bisson, chargé de recherche CNRS au laboratoire Leibniz.

3.7.1 Environnement d'aide à l'analyse de séquences

Les efforts de ces trois dernières années ont porté sur la modélisation des connaissances prescriptives, destinées à aider les biologistes à analyser les séquences génomiques, et plus particulièrement les génomes entiers, tels qu'ils commencent à être obtenus depuis cette année. La collaboration se poursuit donc, à Paris avec l'équipe d'Antoine Danchin à l'institut Pasteur et le groupe de l'ABI (Atelier de Bio-Informatique), et à Lyon avec le laboratoire de génétique, biométrie et dynamique des populations de l'université Claude Bernard (Christian Gautier). L'obtention d'une version répartie (cf. 3.4.2) de l'environnement d'aide à l'analyse de séquences, réalisé avec le soutien du GIP GREG en 1994 et 1995, permettra de fédérer les efforts de conception et d'expérimentation de nouvelles méthodes et tâches d'analyse de séquences et d'envisager la création d'un « club » d'utilisateurs.

3.7.2 Visualisation de séquences génomiques

APIC est une interface graphique qui a pour but de représenter des séquences génomiques et d'en explorer le contenu. Le prototype actuel a été connecté au système PowerGene, permettant ainsi aux utilisateurs de ce système de visualiser et de sélectionner les objets biologiques qui sont produits par les tâches d'analyse de séquences. Sur la base de ce prototype, nous cherchons maintenant à développer de nouvelles collaborations, en particulier avec l'Inra. Le but est d'enrichir notre modèle d'interface en introduisant de nouvelles possibilités de représentation aptes à faciliter le travail d'interprétation des cartes par les biologistes. Ainsi, nous cherchons à modéliser d'autres types de cartes comme les cartes circulaires, ou encore de permettre l'annotation d'objets biologiques par des liens hypertextuels pointant à travers le Web sur des références bibliographiques.

3.7.3 Modélisation des interactions géniques

Grâce à une aide obtenue pour 1996 et 1997 dans le cadre de l'appel d'offres des ACC-SV (Actions Concertées et Coordonnées - Sciences du Vivant) du ministère de la recherche, nous avons débuté une nouvelle collaboration avec l'équipe de Bernard Jacq, du laboratoire de génétique et physiologie du développement de Marseille, pour la réalisation d'une base de connaissances sur les interactions géniques dans la phase du développement embryonnaire de la mouche du vinaigre (*D. melanogaster*).

La base construite cette année est stable. Outre la description des concepts mis en œuvre dans l'interaction génique (gène, protéine, motif d'expression, site de fixation...), elle contient la description à un haut niveau d'instances d'interactions. La base est visible au travers d'un client HTTP et bénéficie de divers outils de manipulation : algorithme de parcours des graphes d'interaction, affichage de ces graphes et des motifs d'expression grâce à des « applets » écrites en Java. La fonctionnalité la plus spectaculaire est un algorithme capable de déterminer l'ensemble des segments de la mouche sur laquelle agit un réseau d'interactions différent et de déterminer automatiquement si le réseau est suffisant pour expliquer les observations faites en laboratoire.

3.8 Aide au diagnostic électromyographique

Participants : Sueli Ferreira Da Cunha, Jean-Yves Morneau, Danielle Ziébelin

Le projet SHERPA poursuit sa contribution au développement du système d'aide au diagnostic électromyographique MYOSYS, initialement conçu dans le cadre d'un projet européen *Esteem* (cf. 5.2.1). Ce système met en œuvre les modèles de tâches conçus au sein du projet.

Les efforts récents ont porté sur l'amélioration de l'interface avec l'utilisateur et des fonctionnalités d'explication. Ainsi, l'intégration d'un ensemble d'images représentant des planches anatomiques a permis de faciliter l'acquisition des données, mais aussi la présentation synthétique des résultats. Ce travail a fait l'objet du stage d'ingénieur CNAM de J.-Y. Morneau [282].

Ces travaux se sont poursuivis par la création d'un serveur Web de documentation couplé à un environnement de résolution de problèmes utilisé au sein du projet. Nous avons proposé une mise à disposition, via ce serveur, de l'ensemble des planches anatomiques et des textes de documentation sur les objets et les tâches de la base de connaissances. Nous avons partiellement réalisé le mécanisme de synchronisation du navigateur *Netscape* avec l'environnement de résolution de problème.

3.9 Modélisation des connaissances dans un système d'enseignement des langues

Participant : Jérôme Gensel

L'objectif du projet LIPSI, mené en collaboration avec plusieurs équipes de recherche de l'université Pierre Mendès-France, est de permettre à un étudiant de disposer d'un environnement informatique avec lequel il puisse interagir et dont les retours appropriés lui permettront d'acquérir et d'améliorer ses connaissances dans une langue étrangère. L'idée consiste à faire s'exprimer l'étudiant, au moyen du clavier et de la souris, à propos d'un scénario animé représenté à l'écran, en l'y incitant par le biais d'une tâche de résolution de problèmes.

Les capacités de représentation du modèle Tropes sont ici mises à profit pour la description des objets et des personnages du scénario (état, position, relations entre eux). Notamment, toute modification sur un plan du scénario (déplacement d'objets ou de personnages) est propagée dans la base de connaissances Tropes et, réciproquement, toute inférence déclenchée par Tropes est reportée sur l'animation graphique. De plus, Tropes est chargé de vérifier la correction syntaxique des dialogues de l'étudiant.

4 Actions industrielles

4.1 Serveurs de connaissances et mémoires techniques

Participants : Jérôme Euzenat, François Rechenmann, Loïc Tricand de la Goutte

Le projet SHERPA intervient dans le troisième volet du projet STORIA avec SGS-Thomson. Ce projet a été labellisé par le ministère de l'industrie, dans le cadre de l'appel d'offre autour des « autoroutes de l'information » et fait l'objet d'une aide du SERICS. Sa première phase est une expérimentation d'une durée d'un an, à compter du 1^{er} octobre 1996, ciblée sur les centres de R&D de l'entreprise SGS-Thomson. L'objectif du projet STORIA est la mise en place d'un Intranet permettant le travail coopératif entre des équipes d'ingénieurs répartis sur différents sites d'une même entreprise.

L'objectif des travaux du troisième volet dans lequel intervient le projet SHERPA est de montrer l'intérêt et la faisabilité d'un système coopératif pour la construction et la consultation de mémoires techniques dans le domaine de la micro-électronique, plus précisément de la conception de circuits. L'expérimentation mettra en œuvre des techniques et outils de représentation et d'exploitation de connaissances dans un contexte réparti (Intranet), où des serveurs de connaissances seront accessibles pour la consultation et la mise à jour à travers des navigateurs Web.

Le projet SHERPA apporte des outils et méthodes pour la construction d'une mémoire d'entreprise et la mise en place d'un environnement de construction incrémental et coopératif (cf. 3.2 et 3.3) permettant à SGS-Thomson de s'approprier le système et de le faire vivre.

4.2 Gestion des connaissances et logiciels scientifiques

F. Rechenmann co-encadre avec Jean-Louis Ermine, du CEA à Saclay, plus précisément responsable du groupe « Gestion des connaissances » au sein de la section « Méthodes et Techniques de l'Information » de la DIST, une thèse sur l'association de connaissances formalisées à des logiciels de calcul en neutronique et photonique. Le doctorant, S. Picard, est inscrit à l'université d'Orsay.

4.3 Système d'aide contextuelle

Participante : Karine Duprez

Dans le cadre d'une convention CIFRE avec la société Dolphin Intégration, K. Duprez effectue sa thèse, sous la direction de F. Rechenmann, sur l'intégration d'une aide contextuelle dans une application existante, en l'occurrence un simulateur de circuits. Une première réalisation en C++ a été achevée fin octobre et totalement intégrée dans le simulateur Smash [273, 274].

5 Actions nationales et internationales

5.1 Actions nationales

J. Euzenat et F. Rechenmann sont membres du comité de programme du colloque LMO'96, « Langages et Modèles à Objets », qui s'est tenu à Leysin (CH) du 16 au 18 octobre. F. Rechenmann est membre du comité de programme de ISMB'96 (*Intelligent Systems for Molecular Biology*).

J. Euzenat a été lecteur pour la « Revue d'intelligence artificielle », TSI (Techniques et Science informatiques) et « Computational intelligence ».

F. Rechenmann a été lecteur pour les revues TSI et « Biochimie ».

F. Rechenmann a été responsable d'un dossier du bulletin de liaison de l'Afia sur « Intelligence artificielle et biologie moléculaire ». Ce dossier est paru dans le numéro de juillet.

F. Rechenmann est membre du conseil scientifique du GIS Infobiogen et du conseil scientifique du CSTB (Centre Scientifique et Technique du Bâtiment).

5.2 Actions internationales

5.2.1 Europe de l'ouest

- Continuation des travaux du consortium ESTEEM

Le projet SHERPA a été impliqué, sous la direction de D. Ziébelin, dans le projet européen AIM n° 2010 Esteem (*European Standardised Telematic tool to Evaluate EMG knowledge-based systems and Methods*). L'objectif de ce projet a été de développer, puis de mettre à disposition des centres d'examen cliniques électromyographiques, une plate-forme commune regroupant un ensemble de systèmes d'aide au diagnostic en électromyographie portant sur l'identification des pathologies neuro-musculaires. L'ensemble des médecins et informaticiens de ce projet a travaillé à l'évaluation, la validation et l'intégration des différents composants de cette plate-forme : base de données EMG, interface avec l'équipement de traitement du signal, module de télécommunication, systèmes experts, boîte à outils d'évaluation. Une plate-forme est maintenant opérationnelle sur l'ensemble des sites d'expérimentation (principalement en CHU). Le niveau de compétence des systèmes à base de connaissances, dont le système MYOSYS développé au sein

du projet SHERPA (cf. 3.8), demeure encore insuffisant pour permettre une utilisation courante du système.

Un consortium réduit (amputé de deux industriels CRI et CIPRA), mais grossi du CHU de Lille, continue donc le projet afin de compléter les bases de connaissances à partir d'une base de cas réels proposés par les partenaires hospitalo-universitaires. Chaque cas fait tout d'abord l'objet d'une discussion entre experts. Une fois « standardisé », le cas est proposé à l'ensemble des systèmes de diagnostic et ceux-ci complètent éventuellement leur base. Une base de 700 cas standardisés a ainsi été constituée. Cette activité devrait se poursuivre jusqu'à l'acceptation d'un nouveau projet européen.

- Système à base de connaissances pour la gestion des ressources en eau
Une proposition de projet européen a été déposée dans le cadre du programme INCO (DG XII). Elle a pour objectif le développement d'un système de gestion des ressources en eau. Il s'agit de concevoir et réaliser un outil de simulation à base de connaissances permettant de prévoir la demande et l'offre à long terme en eau dans différents pays méditerranéens.
- Action intégrée franco-néerlandaise Van Gogh
Une demande a été déposée dans le cadre des actions intégrées franco-néerlandaises Van Gogh. L'objectif est de collaborer avec l'équipe du professeur Nicolaas Mars, de l'université de Twente à Enschede, dans le cadre de son projet Plinius d'acquisition des connaissances à partir de textes (résumés d'articles scientifiques).

5.2.2 Europe de l'est

- Programme de recherche collaborative franco-bulgare PECO
Ce programme de collaboration franco-bulgare a pour objectif de doter MYOSYS d'une fonctionnalité de validation et de structuration automatique en exploitant un ensemble de cas. La collaboration avec le département d'intelligence artificielle de l'institut des technologies de l'informatique de l'Académie des Sciences à Sofia (Bulgarie), vise l'intégration d'un module de diagnostic et d'un module de validation et de correction par cas qui permettra d'obtenir un système d'aide au diagnostic complet et de qualité.

6 Diffusion des résultats

6.1 Enseignement

- Cours « Représentation des connaissances et des raisonnements », 3^e année ENTPE, Vaulx-en-Velin, 6h., intervenant J. Euzenat, resp. Frank Lebastard (CERMICS, Nice).
- Cours et TD « Représentation de connaissances », maîtrise d'informatique, université Joseph Fourier, Grenoble, 7h.30, intervenants : J. Euzenat, G. Bisson, D. Ziébelin, responsable D. Ziébelin.
- Cours « Représentation de connaissance », DEA Sciences cognitives, INPG, 8h., intervenants : G. Bisson, J. Euzenat, responsable J. Euzenat.
- Cours « Une brève introduction aux représentations de connaissance par objets » par J. Euzenat, école d'été « Langages et modèles à objets », CIMPA-INRIA, 3h., responsables J. Euzenat et Amedeo Napoli (cf. 6.3).
- Cours « Bases de connaissances », F. Rechenmann, 6h., DEA « Informatique : systèmes et communications », UJF - INPG.

6.2 Participation à des conférences et colloques

- Dans le cadre des journées « Modélisation informatique et bases de connaissances » organisées par l'ORSTOM à Orléans, D. Ziébelin a fait un exposé sur « Modèles à objets et tâches » (4 juin).
- Le 27 juin, à Poitiers, dans le cadre de rencontres du groupe Aurige de l'Inra, D. Ziébelin a présenté « APIC : interface cartographique pour la biologie moléculaire ».
- J. Gensel a fait un séminaire sur « Programmation par contraintes et intervalles », à l'université de Valenciennes et du Hainaut-Cambresis, les 17 et 18 avril (notes séminaires de l'UVHC 96-04).
- F. Rechenmann a fait un séminaire intitulé « Modèles et bases de connaissances en biologie moléculaire et biologie du développement », au laboratoire de génétique et physiologie du développement de Marseille, le 31 mai.
- Dans le cadre des séminaires LAFORIA-EDF, F. Rechenmann a fait un exposé sur « Construction coopérative de bases de connaissances scientifiques et techniques », le 3 juin, au LAFORIA, université de Paris 6.

6.3 Organisation de colloques et de cours

- J. Euzenat a organisé, avec Amedeo Napoli (Crin - Inria-Lorraine), en juillet à Nice, l'école d'été Cimpa-Inria « Langages et modèles à objets » (15 jours ; plus de 60h de cours ; 20 intervenants ; 20 étudiants de 8 pays différents).
- J. Euzenat a organisé une journée du groupe de travail « Raisonnement spatial et temporel » du PRC-IA, en octobre, à Grenoble.

7 Publications

Thèses

- [267] F. LEMAIRE, *Une structure terminologique pour améliorer l'intelligibilité d'une base de connaissances scientifiques*, thèse de doctorat, INPG, spécialité « Sciences Cognitives », 7 novembre 1996.

Articles et chapitres de livre

- [268] F. RECHENMANN, F. LEMAIRE, « Intégration des connaissances terminologiques dans les grandes bases d'objets. Exemples en biologie moléculaire », *La banque des mots. Revue de terminologie française Numéro spécial 7-1995 « Terminologie et intelligence artificielle »*, 1996.
- [269] P. VALTCHEV, J. EUZENAT, « Classification of concepts through products of concepts and abstract data types », in : *Ordinal and symbolic data analysis*, O. O. Edwin Diday, Yves Lechevalier (éd.), Springer Verlag, Heidelberg (DE), 1996, p. 3–12,
<ftp://ftp.inrialpes.fr/pub/sherpa/publications/valtchev96a.ps.z>.

Communications à des congrès, colloques, etc.

- [270] I. CRAMPÉ, J. EUZENAT, « Fondements de la révision dans un langage d'objets simple », in : *Actes 3^es journées « Langages et modèles à objets »*, p. 134–149, Leysin (CH), 1996,
<ftp://ftp.inrialpes.fr/pub/sherpa/publications/crampe96c.ps.gz>.
- [271] I. CRAMPÉ, J. EUZENAT, « Révision interactive dans une base de connaissance à objets », in : *Actes 10^e RFIA*, p. 615–623, Rennes (FR), 1996,
<ftp://ftp.inrialpes.fr/pub/sherpa/publications/crampe96a.ps.gz>.

- [272] I. CRAMPÉ, «Révision interactive dans une base de connaissance à objets», in : *Actes 3^{es} Rencontres Jeunes Chercheurs en Intelligence Artificielle*, p. 55–62, Nantes (FR), 1996.
- [273] K. DUPREZ, «Aide interactive pour un outil de conception», in : *Actes 3^{es} journées explication*, p. 107–120, Sophia-Antipolis (FR), 1996.
- [274] K. DUPREZ, «Une aide contextuelle pour un utilisateur actif», in : *Journées IHM'96*, Grenoble (FR), 1996. (les actes seront publiés en janvier 1997).
- [275] J. EUZENAT, «Corporate memory through cooperative creation of knowledge bases and hyper-documents», in : *Actes 10th knowledge acquisition workshop*, Banff (CA), 1996.
<ftp://ftp.inrialpes.fr/pub/sherpa/publications/euzenat96b.ps.gz>,
<http://www.inrialpes.fr/sherpa/papers/euzenat96b.html>.
- [276] J. EUZENAT, «HyTropes: a WWW front-end to an object knowledge management system», in : *Actes 10th knowledge acquisition workshop demonstration track*, Banff (CA), 1996.
<ftp://ftp.inrialpes.fr/pub/sherpa/publications/euzenat96c.ps.gz>,
<http://www.inrialpes.fr/sherpa/papers/euzenat96c.html>.
- [277] J. EUZENAT, «Knowledge bases as Web page backbones», in : *Actes 5th WWW workshop on « Artificial intelligence-based tools to help W3 users »*, Paris (FR), 1996. sans pagination,
<http://www.inrialpes.fr/sherpa/papers/euzenat96a.html>.
- [278] S. FERREIRA DA CUNHA, «Task language graphs and similarity», in : *Actes 3^{es} Rencontres Nationales des Jeunes Chercheurs en Intelligence Artificielle*, p. 260, Nantes (FR), août 1996.
- [279] J.-M. GABRIEL, «Des explications dans un système d'apprentissage», in : *Actes 3^{es} journées explication*, p. 185–200, Sophia-Antipolis (FR), 1996.
- [280] J.-M. GABRIEL, «Explications et système d'apprentissage», in : *Actes 3^{es} rencontres des jeunes chercheurs en intelligence artificielle*, p. 103–109, Nantes (FR), 1996.

Rapports de recherche et publications internes

- [281] I. CRAMPÉ, J. EUZENAT, «Fondements de la révision dans un langage d'objets simple», *Rapport de recherche n°3060*, INRIA Rhône-Alpes, décembre 1996.
- [282] J.-Y. MORNEAU, «Conception et mise en œuvre d'interfaces adaptables pour un système à base de connaissances : application au diagnostic en électromyographie», *rapport de recherche*, Cnam, mémoire d'ingénieur, Grenoble (FR), 10 juin 1996.
- [283] D. RACZY, *Esquisse de formalisation de la théorie des modèles mentaux*, Mémoire, Institut National Polytechnique, Grenoble (FR), 1996, Mémoire de DEA de Sciences Cognitives.

8 Abstract

Scientific and technical knowledge bases gather and structure different categories of knowledge: descriptive on the entities of the domain, functional on the behavior of these entities, and prescriptive on the ways to complete and characterize these descriptions. Moreover, formalized knowledge has to be associated with terminological and textual knowledge in order to be justified and annotated.

The Sherpa project designs and studies knowledge models, together with their exploitation mechanisms, which are adapted to the requirements of these scientific and technical knowledge bases. A knowledge base management system is being developed. It supports an incremental and cooperative design process. All these fundamental and technical results are experimented through the design of knowledge bases in molecular biology and of corporate memories.

