

## Liste des sujets de thèse proposés par l'Irisa pour la rentrée 2000

Version du 9 mars 2000

Projet ADP : algorithmes répartis et protocole . . . . .	2
Projets ADP et Pampa . . . . .	5
Projet Aïda : modélisation et apprentissage pour l'interprétation de données et l'aide à la décision . . . . .	6
Projet Armor : architectures et modèles de réseaux . . . . .	11
Projet Badins : bases de données et interrogation souple . . . . .	15
Projet Cordial: dialogue oral homme-machine . . . . .	17
Projets Cordial et Aïda . . . . .	22
Projets Cordial et Badins . . . . .	23
Projet Cosi : conception de systèmes sur silicium . . . . .	24
Projet EP-ATR : environnement de programmation pour application temps réel . . . . .	26
Projet Imadoc : interprétation et reconnaissance d'images et de documents . . . . .	33
Projets Imadoc et Aïda . . . . .	38
Projet Lande : langages déclaratifs . . . . .	39
Projets Pampa et Paragraphe . . . . .	40
Projet Pampa et Ep-Atr . . . . .	41
Projet Pampa : modèles et outils pour la programmation des architectures parallèles réparties . . . . .	43
Projet Paris : programmation des systèmes parallèles et distribués pour la simulation numérique à grande échelle . . . . .	47
Projet Siames : synthèse d'image, animation, modélisation et simulation . . . . .	49
Projet Sigma2 : signaux, modèles et algorithmes . . . . .	52
Projet Temics : traitement, modélisation et communication d'images numériques . . . . .	53
Projets Vista et Temics . . . . .	57
Projet Vista : vision spatio-temporelle et active . . . . .	58

Sujet de thèse proposé pour la rentrée 2000-2001

## Algorithmique distribuée pour simulations coopératives temps réel

**Localisation :** Irisa, Rennes

**Equipe :** ADP

**Responsable :** Emmanuelle ANCEAUME, tél. 02 99 84 75 96, Emmanuelle.Anceaume@irisa.fr

**Mots-clés :** consensus, gestion de groupe, diffusion fiable, temps réel

**Sujet :** Certaines activités industrielles à risques (raffineries, centrales nucléaires, ...) demandent que leur personnel suive périodiquement un entraînement intensif et de haute qualité pour les préparer à faire face aux situations de crises les plus diverses (explosions, ...). Pour des raisons évidentes de sécurité et de coût, une grande partie de cet entraînement doit se faire sur des simulateurs. Pour que les simulations soient efficaces, les systèmes actuels utilisent des architectures dédiées pour garantir les performances et emploient des environnements de réalité virtuelle pour offrir un graphisme d'excellente qualité. Néanmoins, ces simulations ne permettent pas à un groupe de techniciens de coopérer les uns avec les autres pour mener à bien une tâche complexe qui demanderait, dans la réalité, des interactions nombreuses et fines. En effet, le niveau de coopération possible entre acteurs d'une simulation reste aujourd'hui grossier.

Algorithmes de consensus, de gestion de groupe, de diffusion fiable, de cohérence distribuée, sont parmi les briques logicielles indispensables pour construire un environnement de simulation où des techniciens, éventuellement physiquement dispersés, pourraient coopérer et interagir en temps réel.

La recherche que nous proposons consistera essentiellement à spécifier, concevoir et prouver de tels algorithmes, qui s'inscrivent dans la catégorie des algorithmes distribués temps réel. A terme, ils seront intégrés dans la plate-forme de réalité virtuelle immersive de l'Irisa. Cette recherche s'appuiera sur des travaux antérieurs de l'équipe ADP. Une très bonne maîtrise de la programmation en C++ et une habitude du développement (sur station Unix de préférence) sont requises. De plus, des notions de CORBA seront appréciées.

Sujet de thèse proposé pour la rentrée 2000-2001

## Étude de la tolérance à différents types de défaillances

**Localisation :** Irisa, Rennes

**Equipe :** ADP

**Responsables :** M. Hurfin (tél. direct : 02 99 84 75 12, email : hurfin@irisa.fr)

A. Mostéfaoui (tél. direct : 02 99 84 71 96, email : mostefao@irisa.fr)

**Mot-clés :** Système réparti asynchrone, tolérance aux défaillances, communication de groupe, panne franche, recouvrement, faute byzantine.

**Financement :** A définir.

**Sujet :** Pour de nombreuses applications critiques, la sûreté de fonctionnement est un critère de qualité de service essentiel. Depuis deux ans, un des axes de recherche du projet ADP est consacré à la conception et au développement de services permettant de garantir cette contrainte non-fonctionnelle dans un environnement réparti asynchrone. La tolérance aux défaillances peut être obtenue en dupliquant des composants du système. Le maintien de la cohérence entre les copies d'un même composant nécessite alors des prises de décision unanimes au sein du groupe de copies. L'absence d'horloges synchronisées et l'absence de bornes sur les temps de transfert des messages rendent ces prises de décision complexes.

La sûreté de fonctionnement ne peut être garantie que pour des types de défaillances préalablement identifiés. Jusqu'ici, nous n'avons considéré que des défaillances de type pannes franches : dans ce contexte, chaque copie peut, soit s'exécuter correctement, soit s'interrompre brutalement (et définitivement) suite à une panne. L'objectif de cette thèse est d'étudier de nouveaux types de défaillances (pannes avec recouvrement, fautes byzantines, ...) et de proposer, afin de les prendre en compte, des extensions au prototype de gestion d'objets dupliqués développé (en JAVA) au sein du projet ADP. Des études préliminaires montrent que la stratégie actuellement adoptée pour développer les services de communication de groupe (utilisation d'un service de consensus générique) est compatible avec la prise en compte de ces nouveaux types de défaillances.

Sujet de thèse proposé pour la rentrée 2000-2001

## Sûreté de fonctionnement et contraintes temporelles

**Localisation :** Irisa, Rennes

**Equipe :** ADP

**Responsables :** M. Hurfin (tél. direct : 02 99 84 75 12, email : hurfin@irisa.fr) et  
A. Mostéfaoui (tél. direct : 02 99 84 71 96, email : mostefao@irisa.fr)

**Mot-clés :** Système réparti asynchrone, tolérance aux défaillances, communication de groupe, panne franche, contraintes temporelles, terminaison, approche probabiliste.

**Financement :** A définir.

**Sujet :** Pour de nombreuses applications critiques, la sûreté de fonctionnement est un critère de qualité de service essentiel. Depuis deux ans, un des axes de recherche du projet ADP est consacré à la conception et au développement de services permettant de garantir cette contrainte non-fonctionnelle dans un environnement réparti asynchrone. La tolérance aux défaillances peut être obtenue en dupliquant des composants du système. Le maintien de la cohérence entre les copies d'un même composant nécessite alors des prises de décision unanimes au sein du groupe de copies. L'absence d'horloges synchronisées et l'absence de bornes sur les temps de transfert des messages rend ces prises de décision complexes.

Dans l'approche que nous avons adoptée, tous les services destinés à gérer un groupe de copies (diffusion ordonnée, gestion de la composition du groupe) sont conçus comme des instances d'un service d'accord générique qui ne requiert pas de détecteurs de défaillances fiables. Dans le cadre de cette thèse, nous souhaitons enrichir la solution proposée afin de pouvoir prendre en compte des contraintes temporelles. Or, sans hypothèses supplémentaires, l'approche adoptée pour résoudre les problèmes d'accord ne permet pas de borner le délai nécessaire à une prise de décision (en période de fort asynchronisme). L'objectif de cette thèse est donc de maîtriser la terminaison des protocoles exécutés en permettant de gérer au mieux une durée maximale d'exécution via une approche probabiliste.

Sujet de thèse proposé pour la rentrée 2000-2001

## Modélisation d'aspects qualité de service en UML application aux composants logiciels

**Localisation :** Irisa, Rennes

**Equipe :** Projets ADP et Pampa

**Responsable :** Noël Plouzeau (tél. direct : 02 99 84 71 89, email : Noel.Plouzeau@irisa.fr)

**Mot-clés :** Metamodèle, UML, conception par objets, qualité de service, composants logiciels

**Sujet :** L'emploi de composants devient de plus en plus fréquent dans l'industrie du logiciel. De tels composants s'appuient sur un modèle d'assemblage, lui-même fondé sur la spécification précise des responsabilités réciproques des composants et de de leur environnement. La notion de *contrat* joue un rôle fondamental pour spécifier les liens entre les composants qui fournissent des services et ceux qui les consomment. L'utilisation de contrats force les concepteurs à spécifier avant de mettre en oeuvre et fournit une documentation plus formelle. Dans la pratique, cette notion classique de contrat ne suffit pas pour les applications réparties ou interactives. Il est possible d'étendre cette notion pour prendre en compte des aspects de synchronisation ou de qualité de service (durée d'exécution d'un service, délai, bande passante). Une telle extension a été commencée dans le cadre de l'action Rusken de l'Irisa.

Le travail de thèse proposé reprendra les résultats de Rusken et les étendra pour définir une famille de modèles de contrats, en utilisant les possibilités de metamodélisation de la notation UML. En d'autres termes, il s'agira d'étendre cette notation pour offrir aux concepteurs de logiciels à objets une palette d'outils de spécification et conception traitant les aspects de synchronisation et de qualité de service. Les modèles définis seront mis en oeuvre dans la plate-forme de modélisation Umlaut développée par le projet Pampa. Les modèles de contrats développés pour alors s'intégrer au processus d'assemblage de mise en oeuvre développé dans l'environnement Umlaut. Enfin, ils devront être utilisables dans une architecture à composants standard (COM, Beans, OMG components,...).

Sujet de thèse proposé pour la rentrée 2000-2001

## Apprentissage et reconnaissance de scénarios pour le monitoring et les prothèses implantables en cardiologie

**Localisation :** Irisa, Rennes

**Equipe :** Aïda

**Responsables :** Marie-Odile Cordier (tél. direct : 02 99 84 73 14, email : cordier@irisa.fr) et René Quiniou (tél. direct : 02 99 84 73 19, email : quiniou@irisa.fr)

**Mot-clés :** IA, diagnostic, reconnaissance de scénarios, apprentissage automatique, cardiaque.

**Sujet :** Le domaine des prothèses implantables ainsi que les systèmes de monitoring en cardiologie connaissent un développement rapide tant sur le plan technologique que dans les indications médicales. Leurs caractéristiques nouvelles, multisites et multifonctionnelles, nécessitent des programmes embarqués de plus en plus sophistiqués, afin de traiter ces données de type nouveau, et hautement paramétrables, afin d'adapter le traitement au patient. Nous proposons d'utiliser des techniques alliant traitement du signal et intelligence artificielle qui permettent d'adapter au mieux le traitement des données à la pathologie du patient sous surveillance ou portant la prothèse.

Précisément, il s'agit d'étudier les apports de la reconnaissance de scénarios pour la détection et la reconnaissance d'arythmies cardiaques. La reconnaissance de scénarios est une technique efficace consistant à reconnaître des motifs d'événements remarquables contraints temporellement et pouvant être associés à des dysfonctionnements ou pathologies répertoriés. Nous proposons d'étudier l'utilisation d'un couplage fort entre traitement de signal et reconnaissance de scénarios exploitant, par exemple, les hypothèses émises par le reconnaisseur de scénarios afin de focaliser l'étiquetage sur les événements futurs des scénarios actifs ou, encore, focalisant la reconnaissance sur les événements étiquetés avec sûreté par le module de traitement de signal. L'acquisition de modèles de scénarios reste une tâche difficile pour laquelle nous proposons d'étudier les apports de l'apprentissage automatique à partir de séquences d'événements représentatives des pathologies présentées par le patient. Dans les deux cas il faudra proposer des solutions qui prennent en compte le caractère incertain des données et des connaissances dû au bruit sur les signaux et aux erreurs provenant des mesures.

Sujet de thèse proposé pour la rentrée 2000-2001

## Inférence de transducteurs rationnels pour la prédiction de structures

**Localisation :** Irisa, Rennes

**Equipe :** Aïda

**Responsable :** Jacques NICOLAS (tél. direct : 02 99 84 73 12, email : jnicolas@irisa.fr)

**Mot-clés :** Apprentissage; Inférence Grammaticale; Bio-informatique ; Transducteurs.

**Sujet :** La thèse étudiera l'apprentissage automatique de transducteurs dans le contexte particulier de l'étude des séquences génomiques. Les transducteurs rationnels sont des machines à états finies qui peuvent être considérées comme une extension des automates classiques, où les transitions peuvent être étiquetées par des productions sur un deuxième alphabet. Les transducteurs permettent ainsi d'établir une traduction d'un langage vers un autre. Le domaine d'application considéré ici est celui de la reconnaissance de structures de protéines. Les protéines peuvent être décrites par une séquence sur un alphabet à 20 lettres. La structure que prennent celles-ci dans l'espace correspondent à des relations entre des sous-séquences. Il s'agit de construire automatiquement les transducteurs les plus petits correspondant à des ensembles de couples connus puis de tenter de rassembler les résultats en un prédicteur de la structure globale de la protéine. La résolution de ce problème d'apprentissage suppose la définition d'une sous-classe de transducteurs suffisante pour le problème posé et la mise au point d'un algorithme capable d'identifier tout élément de cette classe. Il existe de nombreux travaux sur la prédiction de structures et relativement peu sur l'inférence de transducteurs. La méthodologie suggérée est de proposer un cadre d'inférence permettant de profiter des connaissances apportées par les autres méthodes de prédiction mais aussi des connaissances biologiques sur le domaine (relation physico-chimiques sur les lettres de l'alphabet).

Sujet de thèse proposé pour la rentrée 2000-2001

## Apprentissage des grammaires catégorielles

**Localisation** : Irisa, Rennes

**Equipe** : Aïda

**Responsables** : Jacques NICOLAS (tél. direct : 02 99 84 73 12, email : jnicolas@irisa.fr)

Christian RETORÉ (tél. direct : 02 99 84 74 48, email : retore@irisa.fr)

**Mot-clés** : Apprentissage; Langages Formels; Traitement Automatique des Langues; Logique.

**Sujet** : Un algorithme d'inférence grammaticale à partir d'exemples positifs est un algorithme qui, étant donnés une classe de grammaires et un nombre fini d'énoncés (éventuellement structurés), produit une grammaire de la classe qui soit la plus générale possible à accepter ces énoncés. La classe de grammaires considérée est intéressante si elle est vaste mais autorise des algorithmes efficaces qui peuvent atteindre tout langage de la classe dès qu'ils disposent de suffisamment d'énoncés. Parmi de nombreuses applications (notamment, étude des génomes), on s'intéresse à la linguistique informatique: engendrer une grammaire à partir d'un corpus est non seulement un module utile de traitement automatique des langues, mais aussi un test pour la plausibilité cognitive des théories et formalismes linguistiques.

Les grammaires catégorielles [1,2] sont un formalisme logique qui décrit les langages algébriques. Cette classe de grammaires, trop vaste pour être apprenable, admet néanmoins une sous-classe dite AB-rigide qui l'est et contient des langages algébriques. [1] Il s'agit de proposer, de programmer et d'étudier des algorithmes d'apprentissage pour des classes plus riches [2]: grammaires de Lambek, grammaires en réseaux, grammaires multimodales, ... Les grammaires catégorielles ayant une interface aisée avec la sémantique de Montague, les algorithmes d'apprentissage pourront utiliser la sémantique des énoncés.

[1] M. Kanazawa, *Learnable classes of categorial grammars*, CSLI, Cambridge University Press, 1998.

[2] van Benthem and ter Meulen (eds), *Handbook of Logic and Language*. Elsevier, 1997. Chapitre 12 et 2.



Sujet de thèse proposé pour la rentrée 2000-2001

## Apprentissage de lexiques sémantiques pour l'extension de requêtes au sein d'un système de recherche d'information

**Localisation :** Irisa, Rennes

**Equipe :** Aïda

**Responsable :** P. Sébillot (tél. direct : 02 99 84 73 17, email : [sebillot@irisa.fr](mailto:sebillot@irisa.fr))

**Mot-clés :** acquisition d'informations lexicales en corpus, recherche d'information, variations sémantiques de termes, apprentissage automatique.

**Sujet :** De nombreuses applications du traitement automatique du langage naturel (recherche documentaire, traduction automatique, etc.) nécessitent l'utilisation d'un lexique décrivant la sémantique des mots qu'elles manipulent ; pour être pertinents, ces dictionnaires doivent être suffisamment précis pour que les sens des mots répertoriés correspondent effectivement à ceux du domaine particulier de l'application. Puisqu'il n'est pas possible de disposer a priori de tels lexiques pour tous les domaines, il faut être capable d'acquérir automatiquement une représentation sémantique riche des mots. J. Pustejovsky a proposé un modèle théorique de description des mots, le lexique génératif, dans lequel un nom possède une structure dite *des qualia* qui le lie à différents prédicats correspondant à ses diverses facettes sémantiques (son origine, sa fonction, etc.), c'est-à-dire aux différents événements auquel il peut participer. Ainsi, un livre y est décrit comme étant un objet qui a été *écrit* et dont la fonction est d'être *lu*. Un premier travail d'acquisition partielle de la structure des qualia a été développé au sein de l'équipe Aïda, dans lequel les prédicats verbaux qui peuvent être liés à un nom sont appris par Programmation Logique Inductive.

Le but de cette thèse sera de compléter l'acquisition de cette structure des qualia, en s'intéressant d'une part à la structure argumentale des prédicats verbaux associés à un nom, et, d'autre part, à l'acquisition d'éléments nominaux complétant sa description. L'apport de cette connaissance associée aux noms pour étendre les requêtes d'utilisateurs d'un système de recherche d'information devra être testé, et une caractérisation des variations sémantiques des termes des requêtes sera effectuée. La faisabilité de l'extension de ce travail à des moteurs de recherche de type web sera étudiée.

Sujet de thèse proposé pour la rentrée 2000-2001

## Systeme de reconnaissance automatique de concepts fondé sur la classification hiérarchique. Application en agriculture.

**Localisation :** Irisa, Rennes

**Equipe :** Aïda

**Responsable :** I.C. LERMAN (tél. direct : 02 99 84 72 86, email : lerman@irisa.fr)

**Mot-clés :** Algorithmes de Classification, Coefficients d'association, Apprentissage.

**Sujet :** Le but de la thèse consiste en la mise au point d'une méthode générale de construction d'un système de reconnaissance automatique de concepts sur un ensemble d'objets structurés. La recherche sera directement nourrie par le problème important de la détermination de l'occupation agricole du sol d'un bassin versant de Rennes, à partir d'images satellitaires. L'objet élémentaire, une parcelle agricole dans notre cas, sera représenté par des distributions statistiques discrètes. Il s'agira d'abord, à partir de la description d'un échantillon d'apprendre un arbre condensé significativement, de classification dont les noeuds définiront la représentation des concepts pertinents. Cet arbre conceptuel doit permettre de prédire l'occupation du sol d'une parcelle dont la vérité terrain est inconnue. Dans cette thèse en Classification et Informatique les aspects suivants seront considérés: coefficients d'association entre objets, relativement à une échelle métrique ou une échelle probabiliste de vraisemblance du lien; algorithmes de classification de "gros" ensembles; comparaison d'arbres de classification sur un "gros" ensemble; gestion de la "grande" ressemblance; apport de la Classification dans un processus intelligent de désambiguïsation. Ce dernier point concerne une interaction directe avec le travail de thèse de Christine Largouët qui est en cours et qui relève plus spécifiquement de techniques utilisées en Intelligence Artificielle. Une réflexion importante et des travaux significatifs ont déjà été effectués et préparent le terrain. On peut ainsi signaler le rapport de DESS de Morel DAVID: "Comparaison d'arbres de classification". Cette thèse sera menée en collaboration étroite avec le laboratoire d'Informatique de l'Ensar (Ecole Nationale Supérieure d'Agronomie de Rennes). La recherche sera effectuée dans l'environnement d'une équipe réunissant Informaticiens ou Statisticiens de l'Irisa et de l'Ensar.

Sujet de thèse proposé pour la rentrée 2000-2001

## Modèles économiques de l'Internet

**Localisation :** Irisa, Rennes

**Equipe :** Armor

**Responsables :** Gerardo Rubino (tél. direct : 02 99 84 72 96, email : rubino@irisa.fr),  
Bruno Tuffin (tél. direct : 02 99 84 74 94, email : btuffin@irisa.fr)

**Mot-clés :** Tarification, Internet, DiffServ, évaluation de performance.

**Financement envisagé :** Bourse du MENRT.

**Sujet :** Une méthode pour évaluer les coûts réels ainsi que les prix à instaurer pour les utilisateurs est d'un intérêt essentiel dans un réseau tel que l'Internet. Déterminer les prix à instaurer n'est pas seulement utile pour faire des bénéfices, mais peut aussi fournir une méthode de contrôle du trafic lui-même. En effet, l'encombrement qui touche l'Internet, où les prix sont largement basés sur des taux constants, démontre que sans méthode basée sur notamment l'*utilisation*, il est difficile de contrôler cet encombrement et de partager les ressources du réseau avec les autres utilisateurs de manière stable et appréciable. D'autre part, dans un environnement compétitif, les opérateurs auront besoin de déterminer le prix des services en tenant compte de l'utilisation des ressources du réseau. Un important paramètre vient du fait que le réseau de communication doit acheminer simultanément une grande variété de types différents de trafic (comme dans *DiffServ*) et fournir des performances différentes. L'utilisateur et l'opérateur doivent négocier un contrat de trafic. Sous ce contrat, l'utilisateur est d'accord pour que son trafic soit conforme à certains paramètres (taux de transfert maximal...) alors que l'opérateur garantit une qualité de service particulière (par exemple en terme de délai ou de taux de perte).

Dans cette thèse, il s'agira de déterminer un modèle mathématique qui, entre différents contrats de services et différents types de trafic, permettra d'évaluer les prix à instaurer *optimisant*, simultanément si possible, les performances du réseau (à la satisfaction de l'utilisateur) et les gains de l'opérateur. Il faudra aussi surveiller si la tarification est *équitable*, c'est à dire si les coûts reflètent l'utilisation relative des ressources afin d'obtenir une allocation efficace des ressources qui ne privilégie pas certains flux de manière arbitraire.

Sujet de thèse proposé pour la rentrée 2000-2001

## Le multicast et les très grands réseaux à haut débit

**Localisation :** Irisa, Rennes

**Equipe :** Armor

**Responsable :** Bernard Cousin (tél. direct : 02 99 84 73 33, email : bcousin@irisa.fr)

**Mot-clés :** transmission multicast de données, QoS, réseaux à haut débit, commutation par labels, agrégation, administration de réseaux, IPv6, DiffServ

**Sujet :** La transmission multicast de données est un service de base d'Internet. Toutefois dans le cadre d'un réseau en croissance constante, et à travers le développement d'applications réparties multimédias (notamment de type VoIP), l'utilisation extrêmement fréquente d'un tel service n'est pas sans conséquence.

Une activité importante autour des protocoles de routage (par ex, PIM ou MOSPF) et des protocoles chargés s'assurer la transmission fiable des données multicast (par ex. RMRG, RMTP ou RMP) existe. Cependant parmi les points encore à explorer on peut citer : la gestion et la constitution d'arbres de diffusion au sein des réseaux de transit, l'interconnexion d'arbres appartenant à des réseaux de transit différents, l'utilisation des techniques de commutation de labels pour l'optimisation de ces arbres, leur agrégation et leur reroutage, et finalement l'intégration d'IPv6 et de DiffServ dans un tel cadre.

L'adaptation de la technique proposée par MPLS est une piste prometteuse qui semble apte à apporter une réponse aux différents points évoquer ci-dessus. Nous vous proposons plus spécifiquement d'étudier vos propositions vis-à-vis des deux critères essentiels suivants : l'adaptation au facteur d'échelle ("scalability") et la garantie de service (QoS).

Sujet de thèse proposé pour la rentrée 2000-2001

## Allocation optimale des VP dans les réseaux

**Localisation :** Irisa, Rennes

**Equipe :** Armor

**Responsable :** Raymond MARIE (tél. direct : 02 99 84 72 95, email : marie@irisa.fr) Miklós MOLNAR (tél. direct : 02 99 84 74 42, email : molnar@irisa.fr)

**Mot-clés :** Réseaux, qualité de services, routage de VP, optimisation multicritère et combinatoire

**Sujet :** Un élément efficace du contrôle de trafic et de gestion des ressources dans les réseaux à haut débit orientés connexion (ou dans tous ceux qui définissent une garantie de qualité) consiste à préconfigurer des chemins virtuels (VP). Pour une topologie donnée, l'application d'un système de VP diminue les coûts, protège la qualité et influence beaucoup la gestion du réseau (l'admission des appels individuels, le taux d'utilisation des dispositifs physiques, ...). Les réseaux d'aujourd'hui doivent satisfaire des classes de trafic très différentes. Les normes ATM prévoient plusieurs classes de services avec des exigences différentes et la technologie DiffServ nécessite également l'allocation préalable de certaines ressources du réseau Internet. Les critères de qualité sont négociés et fixés dans un contrat caractérisé par la qualité de services (QoS). Le gestionnaire du réseau doit respecter ces contrats, en utilisant - entre autres - des VP garantis.

L'optimisation d'un système de VP correspond à un cas particulier de l'optimisation des multiflots dans un graphe. Même si l'on suppose des demandes homogènes, reposant sur un seul critère, le problème est NP-difficile. Plusieurs algorithmes heuristiques nécessitant un temps de calcul polynomial sont connus pour approcher ce problème. Dans le cadre de cette thèse, nous proposons d'examiner la possibilité d'une optimisation efficace, complexe et réalisable des VP. L'optimisation doit tenir compte des importants critères de qualité du trafic futur : les coûts, les capacités nécessaires ainsi que des contraintes sur le délai et/ou sur la variation du délai. Le projet nécessite l'analyse de la gestion de trafic et vise la recherche d'algorithmes d'optimisation combinatoire efficaces et implémentables.

Sujet de thèse proposé pour la rentrée 2000-2001

## Etude d'optimisation des communications multicast

**Localisation :** Irisa, Rennes

**Equipe :** Armor

**Responsable :** Raymond MARIE (tél. direct : 02 99 84 72 95, email : marie@irisa.fr) Miklós MOLNAR (tél. direct : 02 99 84 74 42, email : molnar@irisa.fr)

**Mot-clés :** Réseaux, qualité de services, multicast, problème de Steiner

**Sujet :** Dans les réseaux à haut débit, l'importance des applications multipoint (vidéoconférences, diffusions des émissions et des messages) augmente. L'organisation et la maintenance de ces communications soulèvent plusieurs problèmes : trouver le support optimal, résoudre la coexistence des communications unicast et multicast, gérer le trafic multicast, etc. Dans le cas optimal, la communication entre les membres d'un groupe multicast se résout à l'aide d'un arbre couvrant partiel minimal (dit arbre minimum de Steiner). La construction d'un tel arbre étant NP-complète, plusieurs heuristiques existent pour approcher l'optimum. Le temps de calcul et la longueur de l'arbre varient beaucoup d'un algorithme à l'autre. Pour certaines heuristiques, la complexité de l'algorithme et la qualité du résultat dépendent de la profondeur paramétrable de la recherche ou de la topologie paramétrable des connexions. Les paramètres sont des compromis entre temps d'exécution et qualité du résultat. Pour trouver le bon compromis, le développement et l'analyse précise des algorithmes dans l'espace complexité-performance nous semblent nécessaires. Pour un groupe multicast donné, plusieurs connexions quasi-équivalentes peuvent exister. Le choix de gestionnaire doit assurer les services garantis et équilibrer l'utilisation des ressources : ce qui peut influencer la création des canaux de communications.

Dans ce cadre, il est envisagé que la thèse :

- étudie les heuristiques pour créer les liaisons des communications multicast,
- analyse la faisabilité des heuristiques dans les réseaux qui prévoient des services multicast (du point de vue des besoins d'informations et de possibilité d'incorporation dans les protocoles),
- propose des solutions validées pour l'allocation de ressources en tenant compte de la charge du réseau.

Sujet de thèse proposé pour la rentrée 2000-2001

## Propriétés flexibles dans les bases de données

**Localisation** : Enssat, Lannion

**Equipe** : Badins (projet Upresa 6074)

**Responsable** : Patrick BOSC (tél direct : 02 96 46 66 45, email : bosc@enssat.fr)

**Mot-clés** : bases de données, ensembles flous

**Sujet** : Ce sujet est proposé par le projet Badins qui vise à permettre la prise en compte de l'imprécision dans les bases de données. Deux aspects principaux sont potentiellement concernés par l'imprécision, l'incertitude ou la flexibilité, à savoir : d'une part les données que l'on veut stocker dans la base et les propriétés qu'elles vérifient, d'autre part les requêtes qui sont adressées au système de gestion de bases de données (SGBD). On a donc besoin de représenter des données imprécisément connues (capacités de réservoirs, niveau de maîtrise des langues, etc) et des conditions souples pour exprimer des contraintes (requêtes) ou propriétés souples.

Le sujet proposé concerne l'expression de propriétés souples sur les données dans le but d'étendre la notion de contrainte d'intégrité usuelle. On peut penser a priori à diverses extensions telles que la généralisation des dépendances (fonctionnelles et multivaluées en particulier) et la prise en compte des exceptions dans des expressions du genre «habituellement P». On sera amené à distinguer le cas où les données sont précisément connues de celui où les données peuvent être entachées d'imprécision ou d'incertitude. Une première phase du travail consiste à identifier la nature des contraintes exprimables et à en effectuer une typologie. Dans cette optique, il y aura lieu d'attacher une importance particulière à diverses propositions ayant été faites dans la littérature pour étendre la notion de dépendance fonctionnelle. Un second volet a trait à l'usage qui peut être fait des propriétés souples, par exemple, dans le cadre de réponses coopératives ou intentionnelles.

Sujet de thèse proposé pour la rentrée 2000-2001

## Langages de requête pour données imprécises

**Localisation :** Enssat, Lannion

**Equipe :** Badins (projet Upresa 6074)

**Responsable :** Patrick BOSC (tél. direct : 02 96 46 66 45, email : bosc@enssat.fr)

**Mot-clés :** bases de données, ensembles flous

**Sujet :** Ce sujet est proposé par le projet Badins qui vise à permettre la prise en compte de l'imprécision dans les bases de données. Deux aspects principaux sont potentiellement concernés par l'imprécision, l'incertitude ou la flexibilité, à savoir : d'une part les données que l'on veut stocker dans la base et les propriétés qu'elles vérifient, d'autre part les requêtes qui sont adressées au système de gestion de bases de données (SGBD). On a donc besoin de représenter des données imprécisément connues (capacités de réservoirs, niveau de maîtrise des langues, etc) et des conditions souples pour exprimer des contraintes (requêtes) ou propriétés souples.

Le sujet proposé concerne les bases de données dans lesquelles les données sont mal connues. Dans le cadre du modèle de données relationnel (au moins au départ), on veut définir des langages de requêtes appropriés à deux situations. Dans la première, on interroge les données sur les valeurs elles-mêmes à l'aide de requêtes booléennes ou graduelles et on obtient des réponses incertaines. L'un des problèmes est de définir des opérations de base (par analogie avec l'algèbre relationnelle). Dans la seconde, on interroge les données en introduisant des conditions graduelles portant sur l'imprécision des données. On manipule donc les mêmes concepts que ceux servant à décrire les données et on obtient des réponses graduelles. L'un des problèmes est de définir un ensemble aussi exhaustif que possible de conditions élémentaires du langage de requêtes.



Sujet de thèse proposé pour la rentrée 2000-2001

## Méthodologie d'évaluation de systèmes oraux de dialogue

**Localisation :** IRISA, Lannion

**Equipe :** Cordial

**Responsable :** Jacques SIROUX (tél. : 02 96 46 50 30, email : siroux@enssat.fr)

**Mots-clés :** évaluation, communication personne-machine, génération de test.

**Financement :** bourse MENRT (selon classement)

**Sujet :** Actuellement la mise au point de nombreux systèmes oraux de dialogue fondés sur des approches très diverses nécessite des évaluations afin d'une part de comparer les approches sous différents aspects et d'autre part de mesurer la qualité globale de l'interaction. La complexité des systèmes, le contexte oral ainsi que le rôle essentiel des utilisateurs font qu'il n'existe pas encore de méthodologie complète d'évaluation. Le sujet de la thèse consiste : -1) à mener une étude sur les propositions émises dans le cadre du traitement de la langue naturelle écrite et dans celui de l'action Dialogue oral de l'Aupelf-Uref (méthode DQR d'autoévaluation) ainsi que sur les travaux du LIMSI et d'AT&T (ces derniers sont fondés sur une approche analyse de données). Cette étude devra prendre en compte notamment la complexité du dialogue et celle de l'application ; -2) à concevoir et mettre en oeuvre une méthodologie d'évaluation suffisamment souple pour pouvoir s'adapter à des systèmes, des applications et des critères différents et qui doit permettre la possibilité de générer automatiquement des séquences de test.

**Bibliographie.** [WALK97] M. Walker, D. Litman, C. Kamm et A. Abella. PARADISE: A Framework for Evaluating Spoken Dialogue Agents. *Proceedings of 8th Conference of the European Chapter of the Association for Computational Linguistics*, Madrid, 7-12 July 1997, p. 271-280. [ANT99] J.-Y Antoine, J. Caelen, Pour une évaluation objective, prédictive et générique de la compréhension en CHM orale : le paradigme DCR (Demande, Contrôle, Résultat). *Langues*, Vol. 2, 2, juin 1999

Sujet de thèse proposé pour la rentrée 2000-2001

## Modélisations pour le traitement robuste des références multimodales

**Localisation :** IRISA, Lannion

**Equipe :** Cordial

**Responsable :** Jacques Siroux (tél. : 02 96 46 50 30, email : siroux@enssat.fr)

**Mots-clés :** multimodalité, référence, pragmatique, dialogue oral, acte de langage.

**Financement :** bourse MNERT (selon classement)

**Sujet :** Le système de dialogue multimodal GEORAL a pour but de répondre à des questions qui portent sur des informations touristiques et géographiques. L'utilisateur dispose de deux moyens pour formuler ses questions : le geste (écran tactile) et la parole. En sortie, l'information est restituée par l'affichage d'une carte et par synthèse de la parole. Dans un tel environnement, les activités de désignation multimodale (référence) sont nombreuses et essentielles. Elles doivent être prises en compte en fonction de leur caractère contextuel (état de la carte), de la complémentarité et de la redondance présentes sur les médias et enfin de l'incertitude due soit à leur traitement automatique (ex. reconnaissance de la parole) soit à leur réalisation par l'utilisateur. Le sujet de la thèse porte sur les développements à apporter à une ébauche du traitement des références qui est fondé sur des aspects linguistiques et sur une architecture particulière. Il s'agit : 1) de finaliser la modélisation des actes de référence du point de vue linguistique et pragmatique (théorie des actes de langage, aspects perceptifs) ; 2) de mettre en œuvre et évaluer différents modes de traitement et architectures au sein du système.

**Bibliographie.** [Bret98] G. Breton. Modélisation d'un contexte cartographique et dialogique. Rapport de DEA, IFSIC, 1998. [Scha97] D. Schang. Représentation et interprétation de connaissances spatiales dans un système de dialogue homme-machine. Thèse, université Nancy 1, 1997.

Sujet de thèse proposé pour la rentrée 2000-2001

## Négociation d'expressions référentielles

**Localisation :** Irisa-Enssat, Lannion

**Equipe :** Projet Cordial

**Responsable :** Marc Guyomard (tél. direct : 02 96 46 66 29 ; mèl : guyomard@enssat.fr)

**Mot-clés :** Dialogue multimodal, référence, désignation, négociation, actes de langages.

**Sujet :** Des observations ont montré que dans le dialogue à composante orale les interlocuteurs ont tendance à s'assurer du bon déroulement du processus d'identification des objets désignés (notamment linguistiquement). Une première conséquence de ceci est que les requêtes, qui dans la langue écrite sont formulées d'un seul tenant, sont dans le dialogue oral morcelées en plusieurs fragments dont la plupart visent à identifier des objets du discours. Une seconde conséquence concerne le fait que les énoncés de référence sont sujets à négociation afin que chacun ait l'assurance d'une identification correcte des objets. Une mise en oeuvre récente, celle de Heeman et Hirst, présente à nos yeux deux défauts. Le premier est que les actes élémentaires de référence sont des actes ad hoc alors qu'une intégration dans un formalisme plus proche de la théorie des actes de langage telle que [Lit,85] serait plus raisonnable. En second lieu Heeman et Hirst n'intègrent pas de considérations multimodales. L'objectif du sujet présenté est, en partant des travaux de Clarke et Wilkes-Gibbs, d'effectuer une proposition de formalisation surmontant ces difficultés. Le cadre de la mise en oeuvre sera celui du système Géoral (système de dialogue pour l'obtention d'informations géographiques et touristiques sur le Trégor) développé dans l'équipe.

Références :

[Cla,90] Clark H.H., Wilkes-Gibbs D. Referring as a Collaborative Process. *Intentions in Communication*. Edited by P.R. Cohen, J. Morgan and M.E. Pollack. The MIT Press. 1990, p. 463-493.

[Hee,95] Heeman P.A., Hirst G. Collaborating on Referring Expressions. To appear in *Computational Linguistics*, Vol. 21-3, 1995.

[Lit,85] Litman D.J. *Plan Recognition and Discourse Analysis: An integrated Approach for Understanding Dialogues*. TR 170, University of Rochester, 1985.

Sujet de thèse proposé pour la rentrée 2000-2001

## Coopération dans le dialogue homme-machine

**Localisation** : Irisa-Enssat, Lannion

**Equipe** : Projet Cordial

**Responsable** : Marc Guyomard (tél. direct : 02 96 46 66 29 ; mèl : guyomard@enssat.fr)

**Mot-clés** : Dialogue verbal, négociation, réponses coopératives, collaboration.

**Sujet** : L'importance du phénomène de coopération pour l'efficacité d'activités collaboratrices n'est plus à démontrer. C'est en particulier le cas dans le dialogue verbal où la coopération apporte une contribution essentielle au naturel du dialogue. De nombreuses variétés de sous-dialogues coopératifs ont été identifiés : énoncés suggestifs, énoncés correctifs, etc. Cependant, bien qu'une base théorique informelle à la coopération dialogique existe (le principe de coopération et les quatre maximes de Grice), la plupart des systèmes de dialogue existants se limitent à la mise en oeuvre d'une forme de coopération sur la base d'une méthode le plus souvent ad-hoc. L'objectif du sujet proposé consiste, après avoir étudié les différentes formes de coopération dialogique, de rechercher et proposer un cadre formel unificateur. Les retombées attendues sont une meilleure compréhension de la nature de la coopération, une intégration naturelle dans les théories de l'interaction et donc une plus grande facilité dans la réalisation de systèmes. Les propositions seront illustrées par une mise en oeuvre intégrée à un système de dialogue existant.

Références :

[Jok,97] Jokinen K. et al. Collaboration, Cooperation and Conflict in Dialogue Systems. IJCAI-97 Workshop, Nagoya, Japan.

[Sad,91] Sadek M.D. *Attitudes mentales et interaction rationnelle : vers une théorie formelle de la communication*. Thèse de l'université de Rennes I. Juin 1991.

[Guy,89] Guyomard M. et Siroux J. Suggestive and Corrective Answers : a Single Mechanism, in *The structure of multimodal dialogue*, Taylor M. M., Néel F. and Bouwhuis D.G. Éditeurs, North Holland, 1989, p. 361-374.

[Lev,90] Levesque H.J. et al. On Acting Together. AAI-90.

Sujet de thèse proposé pour la rentrée 2000-2001

## Modélisation de l'interaction entre un moteur de dialogue et un analyseur linguistique

**Localisation** : Irisa-Enssat, Lannion

**Equipe** : Projet Cordial

**Responsable** : J-C. PETTIER (tél. direct : 02 96 46 66 12 ; mèl : pettier@enssat.fr)

**Mot-clés** : Langue naturelle, actes de langage, dialogue, coopération.

**Sujet** : Classiquement, les systèmes de dialogue personne-machine utilisent en entrée une représentation sémantique des énoncés associée à une forme de surface indiquant la nature de l'énoncé (informatif, directif, ...). Cette représentation est ensuite mise en contexte (résolution des références, d'anaphores, d'ellipses) pour fournir une représentation pragmatique permettant ensuite d'inférer le contenu propositionnel réel de l'énoncé débarrassé des tournures de langage. Cette combinaison amont d'unifications contextuelles et d'inférences sur la force illocutoire de l'énoncé pose un certain nombre de problèmes : 1) il existe des énoncés multi-fonctionnels ou à actes multiples ainsi que des énoncés fragmentés ; 2) le domaine de définition des prédicats de la base doit être considéré pour évaluer la validité pragmatique de l'énoncé ; 3) il n'existe pas de taxinomie complète reconnue des différentes formes de surface ; 4) le vocabulaire linguistique doit trouver une correspondance dans le vocabulaire logique ; 5) certains énoncés nécessitent de connaître l'état mental de l'interlocuteur pour inférer leur contenu propositionnel réel. Afin de gérer cette problématique, il s'agit de définir un modèle d'interaction avec une chaîne linguistique du domaine public et un moteur de dialogue en cours de développement afin de lui fournir des entrées adaptées aux différents cas.

Références :

[Alle,95] J. Allen, Natural Language Understanding, Benjamin-Cummings Eds., 2<sup>nd</sup>e édition, 1995.

[Anto,94] J-Y. Antoine, Coopération syntaxe-sémantique pour la compréhension automatique de la parole spontanée, Thèse INPG, 1994.

Sujet de thèse proposé pour la rentrée 2000-2001

## Inférence approchée de grammaires régulières

**Localisation :** ENSSAT, Lannion

**Equipe :** Cordial et Aïda

**Responsable :** Laurent MICLET (tél. direct : 02 96 46 66 28, email : miclet@enssat.fr)

**Mot-clés :** Apprentissage à partir d'exemples, Grammaires régulières.

**Sujet :** L'apprentissage d'une grammaire à partir d'exemples (ou *inférence grammaticale*) consiste classiquement à produire une grammaire, si possible de taille minimale, acceptant un échantillon (ensemble d'exemples) positif  $I_+$  de phrases et n'acceptant pas un échantillon négatif  $I_-$ . Dans le cas des grammaires régulières, un théorème assure qu'il existe un espace fini et structuré de solutions possibles (sous forme d'automates finis), respectant certaines contraintes naturelles et construit à partir de  $I_+$ . Le problème de la recherche de l'automate minimal étant NP-difficile, des techniques heuristiques sont mises en oeuvre pour explorer cet espace sous le contrôle de  $I_-$ .

Ce sujet porte sur la recherche d'algorithmes d'inférence sous une contrainte un peu différente : on s'intéresse ici à trouver une grammaire aussi simple que possible, mais en tolérant qu'elle ne soit pas en accord total avec les échantillons positifs et négatifs. On cherche donc à apprendre une grammaire qui réalise un compromis entre simplicité maximale et adéquation aux données d'apprentissage. Une extension naturelle de ce problème conduira à s'intéresser à l'apprentissage combiné de  $C$  grammaires à partir d'un échantillon supervisé par  $C$  classes.

Le sujet s'insère dans un thème de recherche très actif dans les équipes CORDIAL et AÏDA. Il vise à approfondir les acquis récents et à explorer de nouvelles propriétés de l'espace de recherche pour proposer des algorithmes originaux. Ceux-ci seront testés sur les *benchmarks* mis à disposition par la communauté scientifique.

Référence : [DM,98] Dupont P. et Miclet L. *Inférence grammaticale régulière : fondements théoriques et principaux algorithmes* Rapport de recherche INRIA 3449. Juillet 1998.

Sujet de thèse proposé pour la rentrée 2000-2001

## Transformation de questions écrites en requêtes SQLf

**Localisation :** ENSSAT, Lannion

**Equipe :** Cordial et Badins

**Responsable :** Laurent MICLET (tél. direct : 02 96 46 66 28, email : miclet@enssat.fr) Patrick BOSC (tél. direct : 02 96 46 66 45, email : bosc@enssat.fr)

**Mot-clés :** Apprentissage de transducteurs, traitement de la langue naturelle, interrogation flexible de bases de données.

**Sujet :** Le projet BADINS a développé (sous le nom de SQLf) une extension du langage SQL d'interrogation des Bases de Données, autorisant la requête à intégrer des paramètres d'imprécision. Par exemple, si ">>" représente le "comparateur graduel" *beaucoup plus grand* (dont la sémantique est donnée sous la forme de fonction d'appartenance floue), la phrase : *"trouver les jeunes employés dont le salaire est beaucoup plus élevé que celui des autres employés du même département"* s'écrit en SQLf :

**select** #emp, nom **from** EMP E **where** age = jeune **and** salaire >> **all** (**select** salaire **from** EMP **where** #dep = E.#dep **and** #emp ≠ E.#emp)

Le but de la thèse est l'apprentissage automatique d'un outil de traduction, à partir de corpus de couples (phrase en français, requête SQLf) de même sémantique. Une fois la phase d'apprentissage terminée, une phrase en français pourrait donc être traduite automatiquement en requête SQLf pour interroger la base de données. L'outil de traduction qu'il est envisagé d'apprendre est une forme d'automate fini avec sortie, appelé *transducteur*. Le travail consistera donc à se familiariser avec la théorie et la pratique de l'apprentissage (ou *inférence*) des transducteurs, ainsi qu'avec les langages SQL et SQLf. La bibliographie sur le passage du texte écrit aux requêtes SQL devra également être faite. Il faudra ensuite construire un corpus d'apprentissage et tester un algorithme existant. Enfin, il s'agira de développer des heuristiques propres à ce problème et de concevoir un algorithme d'apprentissage spécifique. Une question ouverte est si la totalité de SQLf pourra être traitée ainsi, ou s'il faudra définir des sous-ensembles syntaxiques où l'apprentissage sera efficace.

Sujet de thèse proposé pour la rentrée 2000-2001

## Architecture virtuelle dynamiquement reconfigurable pour le code mobile

**Localisation** : Irisa, Rennes

**Equipe** : Cosi

**Responsable** : D. LAVENIER (tél. direct : 2 99 84 72 17, email : lavenier@irisa.fr)

**Mot-clés** : machine virtuelle, JAVA, logique reconfigurable, code mobile, programmation objet.

**Sujet** : L'exécution du code mobile doit répondre à plusieurs exigences difficiles à concilier : la performance, la portabilité et la sécurité. La solution adoptée actuellement, par exemple dans les navigateurs Internet, est de recourir à une machine virtuelle capable d'interpréter le code à la volée tout en l'isolant du système hôte. Cette solution, de part sa mise en oeuvre, a déjà atteint ses limites: d'une part elle repose sur l'universalité du langage de programmation du code mobile, principalement Java, et de son interprétation par une machine virtuelle monolithique et spécialisée, la JVM; d'autre part, l'utilisation de cette machine virtuelle oblige le processeur hôte à émuler un système coûteux en ressources et relativement lent pour l'exécution du code mobile.

Une solution technique aux problèmes de la performance, de la disponibilité des ressources et de l'évolutivité du support d'exécution peut être atteinte par l'utilisation d'une technologie différente de celle des microprocesseurs: les architectures dynamiquement reconfigurables (ADR). Associée à un processeur, sous la forme d'un co-processeur par exemple, une ADR configurée pour réaliser les tâches d'une machine virtuelle apporte à la fois la performance et la flexibilité.

La réalisation pratique suppose tout à la fois de disposer d'un moyen de description d'une machine virtuelle, de maîtriser la compilation pour une telle machine et, enfin, de parvenir à configurer rapidement une architecture dynamiquement reconfigurable pour une machine virtuelle donnée.

Ce sont les trois objectifs du travail de thèse qui s'effectuera en coopération avec le l'équipe de recherche ORCADE de l'Université de Bretagne Sud à Vannes.



Sujet de thèse proposé pour la rentrée 2000-2001

## Technologie Objet et Calcul Reconfigurable

**Localisation :** Irisa, Rennes

**Equipe :** Cosi

**Responsable :** D. LAVENIER (tél. direct : 2 99 84 72 17, email : [lavenier@irisa.fr](mailto:lavenier@irisa.fr))

**Mot-clés :** modélisation objet, parallélisme, logique reconfigurable

**Sujet :** Des microprocesseurs dont la puissance ne cesse de croître, des réseaux omniprésents, la percée des applications distribuées et multimedia, internet, etc. Telle est l'image publique de l'informatique contemporaine. En profondeur, il s'agit d'une révolution plus que d'une évolution, révolution dont l'ampleur surpasse peut-être celle provoquée par l'apparition des premiers microprocesseurs :

- les méthodes de programmation évoluent du style impératif vers le style objet, autorisant la prise en charge de nouveaux modèles de calcul ou des démarches descendantes concernant des domaines applicatifs (frameworks);

- les circuits reconfigurables (FPGA) changent d'échelle, ils peuvent être modifiés à la volée, et permettent d'implanter un parallélisme massif: 50 millions de transistors (l'équivalent de 250000 cellules logiques) peuvent être mis au service des applications multimedia, du calcul intensif, de la modélisation physique, etc.

Le sujet de thèse vise l'élaboration de modèles objets capables de tirer parti des ressources offertes par les FPGA. La démarche de recherche s'appuiera sur trois points: (1) l'expertise dans la conception des architectures régulières, leur mise en oeuvre et leur contrôle, (2) la modélisation objet et les méthodes de conception d'outils génériques, (3) un examen critique des modèles de calcul actuels et la prise en charge des opportunités des FPGA (arithmétique dédiée ou standard; contrôle statique ou reconfiguration dynamique; description structurelle ou paramétrage, typage et recompilation de modèles, etc).

Cette thèse s'effectuera en coopération avec le laboratoire d'informatique de Brest.

Sujet de thèse proposé pour la rentrée 2000-2001

## Synthèse de contrôleur sur des systèmes à événements discrets hiérarchiques

**Localisation :** Irisa, Rennes

**Equipe :** Ep-Atr

**Responsable :** Hervé MARCHAND (tél. direct : 02 99 84 75 09, email : Herve.Marchand@irisa.fr)

**Mot-clés :** synthèse de contrôleurs, système hiérarchique, méthodes symbolique

**Financement :** À définir

**Sujet :** Le problème de la synthèse de contrôleur sur des systèmes à événements discrets consiste, à partir d'une spécification du système et d'un ensemble de propriétés attendues de celui-ci, à dériver/synthétiser un contrôleur qui, une fois placé dans son environnement, va contraindre le comportement du système de manière à garantir l'ensemble des propriétés. À l'heure actuelle, l'un des problèmes majeurs de la synthèse de contrôleurs réside dans la méthodologie appliquée: alors même que les systèmes à contrôler ont été à l'origine spécifiés de manière hiérarchique (analyse descendante en génie logiciel), la synthèse s'applique sur le système mis à plat (toute la modularité verticale a été oubliée). Ceci combiné au fait que la complexité des algorithmes de calcul croit exponentiellement avec la taille des systèmes, les techniques actuelles constituent un goulot d'étranglement à la diffusion de ces méthodes dans les milieux industriels.

L'objectif de thèse portera sur l'étude de techniques générales de simplification portant à la fois sur le système à contrôler et sur les algorithmes de synthèse. Dans ce contexte, la prise en compte de la modularité verticale semble être une voie prometteuse: il est possible de spécifier des objectifs de contrôle à différents niveaux selon la "vue" que l'on a du système. (e.g. objectif d'ordonnancement de tâches à haut niveau, objectifs de sécurité propres, plus bas dans la hiérarchie).

Le travail commencera par une étude de systèmes hiérarchiques et de leur critère de contrôlabilité. Il se poursuivra par une adaptation des algorithmes de synthèse par décomposition vertical des objectifs de contrôle en fonction de la hiérarchie du système. Ce travail pourra se mener de front avec l'étude d'autres techniques telles que l'utilisation d'algorithmes de bisimulation pour la génération d'un contrôleur, ainsi que l'étude d'algorithmes de synthèse symboliques "à la volée". Finalement ces différentes techniques pourront être expérimentées dans le cadre d'applications portant sur la sécurité des systèmes informatiques.

Sujet de thèse proposé pour la rentrée 2000-2001

## Logiques pour le contrôle d'automatismes discrets

**Localisation** : Irisa, Rennes

**Equipe** : EP-ATR

**Responsable** : Sophie Pinchinat (tél: 02-99-84-72-54, email: pinchina@irisa.fr)

**Mot-clés** : systèmes à évènements discrets, synthèse de contrôleur, logiques modales.

**Financement** : A déterminer.

**Sujet** : De nos jours, des systèmes de plus en plus complexes sont mis en oeuvre dans de nombreux domaines (systèmes de production, systèmes informatiques, systèmes de transport, systèmes administratifs,...). La complexité de ces systèmes peut être liée d'une part à leur taille (nombre d'éléments du système), et d'autre part à leur fonctionnement (type et nombre des interactions entre les éléments du système). Les principaux problèmes que posent les systèmes complexes sont la compréhension et/ou la validation de leur fonctionnement, leur dimensionnement, etc. Le cas particulier des problèmes de validation des systèmes trop complexes pour être traités "manuellement" est à l'origine depuis une vingtaine d'années d'un développement spectaculaire de méthodes formelles souvent automatisées (e.g. model-checking, démonstration automatique, synthèse, simulation, test, etc). L'approche par **synthèse d'automatismes discrets** consiste à construire automatiquement un contrôleur du système qui garantit *a priori* les bonnes propriétés du système contrôlé. La méthode de construction est guidée par un objectif de contrôle qui exprime les propriétés attendues du système contrôlé.

théorique dans un premier temps, des modèles de systèmes et des logiques bien adaptées à la spécification d'objectifs de contrôle. En particulier, la décidabilité de ces logiques, ainsi que la complexité du problème de décision devra être traitée. Ensuite, si de telles logiques peuvent permettre d'énoncer (et de vérifier automatiquement) l'existence d'une solution au problème du contrôle, il n'en reste pas moins la question de synthétiser un contrôle. Dans ce deuxième volet, un enrichissement des algorithmes de décision des logiques pourront par exemple servir de point de départ. Ce sujet s'inscrit dans une grande famille de travaux pour lesquels divers résultats ont déjà été proposés. Toutefois, la problématique énoncée offre encore un très large spectre d'études à mener (e.g. modèles de théorie des jeux et logiques modales, etc.).

Sujet de thèse proposé pour la rentrée 2000-2001

## Analyse sémantique des systèmes de transitions pré-ordre

**Localisation :** Irisa, Rennes

**Equipe :** EP-ATR

**Responsable :** Jean-Pierre Talpin (tél.: 0299847436, mail: [Jean-Pierre.Talpin@irisa.fr](mailto:Jean-Pierre.Talpin@irisa.fr))

**Mot-clés :** Systèmes mobiles synchrones et asynchrones, interprétation abstraite.

**Sujet :** La spécification et la mise en oeuvre de systèmes réactifs nécessite une analyse comportementale précise afin d'assurer la sûreté de fonctionnement et de garantir l'exécution efficace. L'utilisation de langages de programmation synchrone facilitent la réalisation de systèmes critiques remplissant ces critères.

L'objectif de ce stage sera d'étudier la décision de propriétés de sûreté (endochronie, accessibilité d'états, etc) dans les systèmes de transitions hiérarchiques [1] en ayant recours à des techniques d'interprétation abstraites [2].

**Bibliographie** (disponible en <http://www.irisa.fr/prive/talpin>)

[1] J.-P. Talpin, A. Benveniste. and P. Le Guernic. Asynchronous deployment of synchronous transition systems. Research Report No. 1269. Irisa, October 1999.

[2] J.-P. Talpin, A. Benveniste., B. Caillaud and P. Le Guernic. Hierarchic normal forms for desynchronization. Research Report No. 3822. Inria, December 1999.

[3] J.-P. Talpin, A. Benveniste. Synchronous design and asynchronous deployment of mobile processes Research Report, in preparation. Inria, January 2000.

[4] "Timed polyhedra analysis for synchronous languages". Besson, F., Jensen, T., Talpin, J.-P. Static Analysis Symposium (Sas'99). Springer Verlag Lectures Notes in Computer Science, September 1999.

Sujet de thèse proposé pour la rentrée 2000-2001

## Axiomatisation des systèmes de transitions hiérarchiques

**Localisation** : Irisa, Rennes

**Equipe** : EP-ATR

**Responsable** : Jean-Pierre Talpin (tél.: 0299847436, mail: [Jean-Pierre.Talpin@irisa.fr](mailto:Jean-Pierre.Talpin@irisa.fr))

**Mot-clés** : Systèmes mobiles synchrones et asynchrones, preuve de théorème.

**Sujet** : La spécification et la mise en oeuvre de systèmes réactifs distribués nécessite la vérification de propriétés comportementales précises afin d'en assurer le fonctionnement correct et d'en garantir l'exécution efficace. Dans ce but, une théorie des systèmes de transitions synchrones a été élaborée. Cette structure possède une représentation canonique qui permet de mettre en évidence les propriétés d'endochronie et d'isochronie. Nous avons montré que cette structure pouvait être étendue afin de rendre compte de la mobilité et de comportements dynamiques [1]. L'objectif de la thèse sera d'étudier cette structure afin d'en proposer une axiomatisation à l'aide d'un assistant de preuve (e.g. Coq, [2]). En utilisant cette axiomatisation, nous pourrions étudier la formalisation et la vérification de propriétés de programmes, comme par exemple: l'endochronie (équivalence entre observations synchrones et asynchrones), l'isochronie (équivalence entre compositions synchrones et asynchrones), la sécurité (l'accès légitime aux données d'un programme), le typage (l'accès correct aux données d'un programme). Mettant ce travail en perspective, nous pourrions enfin proposer une méthode générale mettant en oeuvre les techniques de démonstration automatique dans le cadre du déploiement distribué de composants applicatifs synchrones.

**Bibliographie** (disponible en <http://www.irisa.fr/prive/talpin>)

[1] J.-P. Talpin, A. Benveniste. Synchronous design and asynchronous deployment of mobile processes Research Report, in preparation. Inria, January 2000.

[2] Nowak, D., Beauvais, J.-R., Talpin, J.-P. Co-inductive axiomatization of a synchronous language International Conference on Theorem Proving in Higher-Order Logics. Springer Verlag, October 1998

Sujet de thèse proposé pour la rentrée 2000-2001

## Hierarchisation des systèmes de transitions pré-ordre

**Localisation** : Irisa, Rennes

**Equipe** : EP-ATR

**Responsable** : Jean-Pierre Talpin (tél.: 0299847436, mail: [Jean-Pierre.Talpin@irisa.fr](mailto:Jean-Pierre.Talpin@irisa.fr))

**Mot-clés** : Systèmes synchrones et asynchrones, compilation.

**Sujet** : La spécification et la mise en oeuvre de systèmes réactifs nécessite une analyse comportementale précise afin d'assurer la sûreté de fonctionnement et de garantir l'exécution efficace. L'utilisation de langages de programmation synchrone facilitent la réalisation de systèmes critiques remplissant ces critères.

Dans ce processus d'analyse, une représentation symbolique du système, par exemple sous forme d'arbres de décision binaires [1], ou bien sous forme de systèmes de transitions hiérarchiques [2], joue un rôle primordial afin d'en analyser, d'en vérifier et d'en optimiser la structure.

L'objectif de ce stage sera de définir un ensemble d'algorithmes efficaces pour la représentation hiérarchiques de systèmes de transition symboliques.

**Bibliographie** (disponible en <http://www.irisa.fr/prive/talpin>)

[1] T. P. Amagbègnon, L. Besnard and P. Le Guernic. Implementation of the data-flow synchronous language Signal. In *Conference on Programming Language Design and Implementation*. Acm, 1995.

[1] J.-P. Talpin, A. Benveniste. and P. Le Guernic. Asynchronous deployment of synchronous transition systems. Research Report No. 1269. Irisa, October 1999.

[2] J.-P. Talpin, A. Benveniste., B. Caillaud and P. Le Guernic. Hierarchic normal forms for desynchronization. Research Report No. 3822. Inria, December 1999.

Sujet de thèse proposé pour la rentrée 2000-2001

## Description d'architectures de logiciels en Signal

**Localisation :** Irisa/Inria Rennes

**Equipe :** EP-ATR

**Responsables :** Thierry Gautier (projet EP-ATR, e-mail : [Thierry.Gautier@irisa.fr](mailto:Thierry.Gautier@irisa.fr),  
site web : <http://www.irisa.fr/ep-atr/>)

**Mots-clés :** langage de description d'architecture logicielle, langage Signal, graphe, abstraction.

**Sujet :** Les langages de description d'architectures de logiciels fournissent des notations pour décrire des composants et connecteurs et permettent de spécifier comment ces éléments se combinent pour former l'architecture d'un système. À la possibilité d'associer, dans un modèle statique, des propriétés aux composants d'un système, se rajoute parfois un modèle dynamique permettant la prise en compte de propriétés comportementales du système ou éventuellement sa reconfiguration.

Le langage Signal, conçu initialement pour décrire des applications temps réel, se base sur un modèle sémantique « à flots de données synchronisés » : les composants de base d'une description communiquent par l'intermédiaire de signaux dont les relations temporelles sont spécifiées par la description. Une description Signal peut être représentée sous la forme d'un graphe statique de composants, dont certains composants sont eux-mêmes des graphes. Un composant Signal peut se voir à différents niveaux : spécification de propriété, description comportementale, etc., niveaux entre lesquels il existe des relations d'abstraction ou, à l'inverse, de raffinement.

Le sujet de thèse proposé concerne l'utilisation de Signal comme langage de description d'architecture logicielle sous ses différents aspects : la description proprement dite à différents niveaux et le passage d'un niveau à un autre, la possibilité de vues multiples d'une architecture et le maintien de leur cohérence, l'étude de diverses propriétés d'une architecture. Une des orientations possibles de l'étude pourrait être l'extension du modèle statique sous-jacent à Signal vers un modèle de graphes dynamique.

Sujet de thèse proposé pour la rentrée 2000-2001

## Accélération du temps en Signal

**Localisation :** Irisa/Inria Rennes

**Equipe :** EP-ATR

**Responsable :** Thierry Gautier (projet EP-ATR, e-mail : [Thierry.Gautier@irisa.fr](mailto:Thierry.Gautier@irisa.fr),

**Mots-clés :** systèmes réactifs, langages synchrones, sur-échantillonnage, interface synchrone-asynchrone, compilation.

**Sujet :** Parmi les langages synchrones, une caractéristique propre au langage Signal consiste en la possibilité de créer des instants entre deux instants d'une horloge existante. On a ainsi un mécanisme, appelé *sur-échantillonnage*, de création d'horloges « plus grandes » (plus *fréquentes*) que les horloges d'origine.

Les techniques de synthèse du contrôle implémentées dans le compilateur du langage Signal reposent sur la construction d'une arborescence d'horloges dont les racines sont des bornes supérieures des horloges du programme. Lorsqu'il y a une seule racine (une borne supérieure), il est possible de générer du code indépendant du contexte du programme, l'horloge de base pour cette génération étant cette racine. S'il y a plusieurs racines, on peut se ramener au cas précédent en ajoutant une racine artificielle ; l'utilisateur devra alors en général fournir des informations supplémentaires sur la présence des signaux d'interface (ajout de booléens).

Lorsque plusieurs sur-échantillonnages sont utilisés dans un programme, les horloges créées par ces sur-échantillonnages ne sont pas toujours comparables ; si de plus le nombre d'instantés créés dépend de conditions internes au programme, il n'est pas forcément souhaitable de faire « remonter » ces informations au niveau de l'interface.

La thèse proposée s'attachera à étudier les divers aspects du sur-échantillonnage, tant du point de vue du modèle que de son implémentation. On déterminera par exemple des conditions dans lesquelles des horloges sur-échantillonnées peuvent être comparables, ou au contraire sont suffisamment indépendantes pour qu'il soit possible de générer du code selon des schémas à préciser.

Les études s'appuieront sur l'utilisation du sur-échantillonnage dans le cadre de l'interfaçage synchrone-asynchrone, mais aussi pour la simulation de boucles « while » parallèles, chaque itération correspondant à un instant dans une horloge sur-échantillonnée.



Sujet de thèse proposé pour la rentrée 2000-2001

## Modélisation de système automatique numérique de perception et d'interprétation

**Localisation :** Irisa, Rennes

**Equipe :** Imadoc

**Responsable :** Guy Lorette (tél. direct : 02 99 84 73 58, email : [lorette@irisa.fr](mailto:lorette@irisa.fr))

**Mot-clés :** Modélisation, système numérique, fusion de données, perception.

**Sujet :** La recherche sur la perception artificielle vise à doter l'ordinateur de capacités perceptives qui donnent des résultats comparables à ceux de la perception humaine. La perception automatique, par un ordinateur, de signaux (sons, parole, etc.) ou d'images (contours virtuels induits, etc.) est un problème encore non résolu, bien que certains modèles particuliers existent déjà (réseaux bayesiens, réseaux neuromimétiques, etc.). Un Système de Perception et d'Interprétation (SPI) à plusieurs couches a été développé, au sein de l'équipe Imadoc. Son fonctionnement repose sur un cadre commun de représentation de différents niveaux de connaissances, la possibilité d'introduire des connaissances a priori à chaque niveau, une triple stratégie de gestion de ces connaissances (stratégie ascendante ou inductive, stratégie descendante ou déductive, inhibition latérale), un ensemble de mécanismes de fusion de connaissances et un fonctionnement dynamique (cycles perceptifs). Ce système a été réalisé sous forme d'un réseau neuronal hiérarchique (graphe valué) qui comporte plusieurs rebouclages interne et qui fonctionne par cycles successifs de mécanismes synchrones d'activation/inhibition des neurones.

L'objectif de cette thèse est de développer une modélisation mathématique du fonctionnement de ce système numérique hiérarchique (convergence, stabilité, sensibilité au réglage des paramètres, optimisation, ...) ainsi que d'étudier sa généralisation à N couches pour aboutir à un système très générique.

Sujet de thèse proposé pour la rentrée 2000-2001

## Modélisation par arbre de décision flou et système d'inférence floue pour la reconnaissance de symboles graphiques manuscrits

**Localisation :** Irisa, Rennes

**Equipe :** Imadoc

**Responsable :** Eric Anquetil et Guy Lorette (tél. : 02 99 84 72 38, email : eric.anquetil@irisa.fr)

**Mot-clés :** logique floue, arbre de décision, système d'inférence floue, reconnaissance, classification, interprétation de symboles manuscrits.

**Sujet :** Ce sujet se situe dans le domaine de recherche associé à l'extraction et à la modélisation automatique de connaissances. Les approches étudiées pendant cette thèse s'appuieront sur les modélisations à base de Système d'Inférence Floue (SIF), sur les arbres de décision flous et sur la classification non supervisée floue ou possibiliste.

La mise en oeuvre de ces méthodes de modélisation aura pour objectif d'extraire et de modéliser de manière explicite des connaissances, à partir de bases d'apprentissage, pour concevoir des systèmes de reconnaissance et d'interprétation de formes graphiques en intégrant notamment des informations de nature contextuelle.

Ces travaux s'intégreront à plusieurs projets sur lesquels l'équipe IMADOC travaille. Ainsi, la reconnaissance portera notamment : sur l'identification des gestes graphiques effectués à l'aide d'un stylo sur une tablette à digitaliser et sur la reconnaissance en-ligne de formules mathématiques ou de symboles musicaux.

Sujet de thèse proposé pour la rentrée 2000-2001

## Introduction de connaissances lexicales et syntaxiques dans un système de reconnaissance d'écriture manuscrite

**Localisation :** Irisa, Rennes

**Equipe :** Imadoc

**Responsable :** Guy Lorette et Eric Anquetil (tél. : 02 99 84 73 58, email : [guy.lorette@irisa.fr](mailto:guy.lorette@irisa.fr))

**Mot-clés :** organisation de dictionnaire, recherche dans un dictionnaire, analyse et modélisation syntaxique, classification, reconnaissance d'écriture.

**Sujet :** Le projet IMADOC a développé un système de reconnaissance d'écriture capable de reconnaître des lettres et des mots manuscrits écrits sur une tablette à digitaliser. Cette thèse se situe dans la prolongation directe de la conception de ce système de reconnaissance d'écriture. Elle a pour objectif d'étendre le cadre d'utilisation du système à la reconnaissance de mots en contexte de phrase. Pour ce faire, il est nécessaire d'enrichir la modélisation en intégrant des connaissances linguistiques, notamment de nature lexicale et syntaxique.

L'organisation et la réduction automatique de dictionnaire seront à la base de l'étude lexicale. Cette étude aura pour objectif d'augmenter les performances et les temps de réponse du système de reconnaissance. L'intégration de connaissances syntaxiques devra, quant à elle, permettre d'exploiter le système de reconnaissance de mots manuscrits dans le contexte d'une phrase.

Sujet de thèse proposé pour la rentrée 2000-2001

## Interface stylo de correction interactive de documents électroniques

**Localisation :** Irisa, Rennes

**Equipe :** Imadoc

**Responsable :** Guy Lorette (tél. direct : 02 99 84 73 58, mél : guy.lorette@irisa.fr) et  
Hélène Richy (tél. direct : 02 99 84 73 71, mél : helene.richy@irisa.fr)

**Mot-clés :** interface homme-machine, édition automatique, correction de documents, reconnaissance de symboles.

**Sujet :** Une des tâches les plus longues en édition électronique est la correction des documents à l'aide du clavier et de la souris. Les améliorations récentes de la qualité, de la lisibilité et de la précision des tablettes graphiques, des écrans tactiles et des stylos électroniques permettent maintenant d'envisager d'utiliser ces dispositifs pour annoter directement des documents électroniques structurés.

Le sujet de la thèse proposé ici concerne la correction manuscrite de documents électroniques. L'objectif de la thèse est d'identifier, de formaliser, de reconnaître et de localiser des symboles de correction adaptés à l'édition de documents électroniques ainsi que les zones de texte sur lesquelles elles vont s'appliquer. L'étude proposée permettra de concevoir, de réaliser et de tester un système de correction automatique de documents électroniques à partir du déplacement du stylo sur une tablette. Les résultats de ce travail pourront être validés et intégrés dans un éditeur interactif tel que Thot ou Amaya, développés à l'Inria par le projet Opéra. L'application d'une telle méthode ouvre la voie à de nouvelles interfaces d'édition appliquées aussi bien à la correction de textes, qu'à celle de tableaux ou de graphiques, ou à des corrections de mise en page.

La thèse aura pour cadre le projet Imadoc (IMAgés et DOCuments) dont les activités de recherche concernent les relations entre écrit et documents. Ce projet regroupe des compétences en reconnaissance, en analyse, en synthèse des documents ainsi qu'en interfaces homme-machine orientées stylo.

Sujet de thèse proposé pour la rentrée 2000-2001

## Introduction de la connaissance *a priori* en reconnaissance de formules mathématiques

**Localisation** : Irisa, Rennes

**Equipe** : Imadoc (IMAgés et DOCuments)

**Responsable** : Jean CAMILLERAPP, Bertrand COÛASNON (tél. direct : 02 99 28 66 40, email : couasnon@irisa.fr)

**Mot-clés** : Reconnaissance de documents, analyse structurelle, gestion des connaissances *a priori*, segmentation, formules mathématiques.

**Sujet** : Dans le domaine de la reconnaissance de documents, la reconnaissance de formules mathématiques a déjà fait l'objet de nombreuses études. Cependant, selon la littérature, certaines difficultés ne sont toujours pas résolues : caractères se touchant, petits symboles...

Les résultats actuels sont insuffisants dans le cadre d'une utilisation industrielle. Il est en effet nécessaire, d'une part d'obtenir de meilleurs taux de reconnaissance, et d'autre part d'éviter une relecture fastidieuse et coûteuse de l'ensemble des formules. L'objet de cette thèse sera donc d'introduire la connaissance *a priori* déduite des règles de constitution des formules afin d'améliorer la segmentation (*i.e.* la reconnaissance) et de permettre au système de détecter avec, une grande certitude, les zones susceptibles d'être mal reconnues.

Ce travail sur l'introduction de la connaissance *a priori* reprendra et étendra les méthodes développées au sein de l'équipe Imadoc. Il sera nécessaire de définir une formalisation grammaticale des formules mathématiques, ainsi que de déterminer les connaissances sémantiques utiles à une meilleure reconnaissance. Il faudra également définir une méthode de segmentation contextuelle des symboles mathématiques. Cette méthode pourra utiliser un classifieur, développé dans l'équipe Imadoc, capable de rejeter des formes qui ne correspondent pas à des symboles.

Ce travail s'intègre dans la réalisation d'un système complet et générique de reconnaissance de documents structurés pouvant s'adapter à d'autres types de documents.

Sujet de thèse proposé pour la rentrée 2000-2001

## Modèle logique spatio-temporel & hypertexte de brouillons d'écrivains

**Localisation :** Irisa, Rennes et ITEM (ENS, Paris)

**Equipe :** IMADOC + AÏDA et ITEM

**Responsable :** Jacques André (tél. 02 99 84 73 50 ; jandre@iris.fr) et  
Jean-Louis Lebrave (Jean-Louis.Lebrave@ens.fr)

**Mot-clés :** Modèle logique, contraintes, TEI, XML, RCC, génétique, brouillons, IA, hypertexte.

**Financement :** Ce travail devrait entrer dans le cadre du projet cognitif du MRT et disposer ainsi d'une bourse de thèse.

**Sujet :** La « critique génétique » a pour objet l'analyse des brouillons d'écrivains. Ceux-ci sont caractérisés par des bouts de textes (ou de dessins) localisables (à droite de ceci, au dessus, dans la marge, etc.) produits par diverses actions (ajouter, raturer, remplacer, etc.) à des instants donnés ou supposés (tel bout de texte peut avoir été écrit juste à la suite de tel autre, ou probablement entre tel et tel autre, etc.). On propose de décrire ces brouillons par une version « diplomatique » basée sur la TEI/XML en utilisant un modèle logique spatio-temporel (utilisant des logiques du type Allen ou RCC) et des résolutions de contraintes permettant, par exemple, d'afficher le scénario de l'écriture de ce manuscrit.

Ce travail devra s'appuyer sur des textes modernes, par exemple les manuscrits de Proust et nécessitera donc une bilocalisation du travail à Paris et à Rennes.

Sujet de thèse proposé pour la rentrée 2000-2001

## Conception d'un système de détection d'intrusions opérationnel de haut niveau

**Localisation :** Irisa, Rennes

**Equipe :** Lande

**Responsable :** Mireille Ducassé (tél.: 02 99 28 64 85, email : ducasse@irisa.fr)

**Mot-clés :** Détection d'intrusions, sécurité, langages concurrents,

**Sujet :** La détection d'intrusions dans les systèmes d'information est indispensable dans un environnement sécurisé non sécuritaire. De la même manière qu'il est impensable de mettre des barreaux à toutes les fenêtres des bâtiments, on ne peut multiplier les barrières de sécurité dans les systèmes. Il existe de nombreux systèmes opérationnels mais qui sont en général de très bas niveau. D'un autre côté, il existe des prototypes de recherche de plus haut niveau, mais ils ont souvent des performances qui ne peuvent permettre une utilisation en grandeur réelle.

Le sujet de la thèse consistera, entre autres, à explorer les possibilités offertes par les langages de haut niveau, en particulier concurrents et coopératifs comme CLF [2]. Ce type de langages a atteint une maturité certaine et des techniques de compilation efficaces existent. Le résultat escompté est, d'une part, un langage de haut niveau pour spécifier les signatures d'attaques sans que l'officier de sécurité ait à se soucier des détails d'implémentation du système bas niveau (quel qu'il soit) et, d'autre part, un prototype de recherche opérationnel validé par un fonctionnement continu sur un vrai réseau de machines.

[1] *Languages and tools for rule-based distributed intrusion detection*. A. Mounji, Thèse de l'université Notre Dame de la Paix, Namur, Belgique. Sept. 1997.

[2] *The Coordination Language Facility: coordination of distributed objects*. J.-M. Andreoli, S. Freeman and R. Pareschi In: *Journal of Theory and Practice of Object Systems (TAPOS)*, 2(2), pp 77-94, 1996

Sujet de thèse proposé pour la rentrée 2000-2001

## Langages de scénarios et synthèse de contrôleurs

**Localisation :** Irisa, Rennes

**Equipe :** Pampa, Paragraphe

**Responsables :** Benoît CAILLAUD (tél. direct : 02 99 84 74 07, email : Benoit.Caillaud@irisa.fr) et Philippe DARONDEAU (tél. direct : 02 99 84 72 80, email : Philippe.Darondeau@irisa.fr)

**Sujet :** Les langages de scénarios ont été introduits pour permettre à des non-spécialistes de donner des spécifications incomplètes de systèmes à réaliser : protocoles de télécommunications, systèmes d'informations répartis, systèmes de production, etc.

Le contrôle des systèmes à événements discrets (à la Ramadge & Wonham) consiste à construire, pour un système donné, un superviseur (le plus permissif possible) garantissant la sûreté et la vivacité de ce système.

Le sujet proposé est de combiner ces deux thèmes en s'inspirant des avancées récentes en synthèse de réseaux de Petri pour automatiser à la fois la synthèse de systèmes et la synthèse de contrôleurs. L'intérêt de la démarche proposée est d'aboutir naturellement à des systèmes répartis.

Le travail de thèse s'équilibrera entre des travaux théoriques, méthodologiques et de prototypage sur l'outil SYNETH.

### **Bibliographie :**

[1] P. Darondeau, *Deriving Unbounded Petri-Nets from Formal Languages*, Rapport de recherche INRIA RR-3365, 1998, <http://www.inria.fr/RRRT/RR-3365.html>

[2] B. Caillaud, *SYNET: A Synthesizer of Distributable Bounded Petri-Nets from Finite Automata, Version 2.0a*, <http://www.irisa.fr/pampa/logiciels/synet/synet.html>

[3] L. Hélouët, *MSC-Related Online Publications*, <http://www.irisa.fr/pampa/perso/helouet/compil.html>

[4] P.J.G. Ramadge, W.M. Wonham, *The control of discrete event systems*, Proceedings of the IEEE, vol.77, no.1 p. 81-98, 1988



Sujet de thèse proposé pour la rentrée 2000-2001

## Synthèse de contrôleurs et génération de tests

**Localisation :** Irisa, Rennes

**Equipe :** Ep-Atr, Pampa

**Responsables :** Thierry JÉRON (tél. direct : 02 99 84 74 64, email : Thierry.Jeron@irisa.fr) et Hervé MARCHAND (tél. direct : 02 99 84 75 09, email : Herve.Marchand@irisa.fr)

**Mots-clés :** contrôle, synthèse, test, conformité, spécification formelle, objectif de test, objectif de contrôle, système réactif.

**Financement :** à définir.

**Sujet :** Le problème de la synthèse de contrôleur consiste, à partir d'une spécification d'un système réactif et d'une propriété attendue de celui-ci, à synthétiser un contrôleur qui, placé dans l'environnement du système et disposant de la spécification, va garantir que les exécutions possibles du système valident cette propriété. La génération de tests consiste à dériver à partir d'une spécification d'un système et éventuellement d'objectifs de tests, une suite de tests. Ces tests sont ensuite exécutés sur une implémentation du système afin de déterminer sa conformité vis à vis de la spécification. Il est clair que des similitudes existent entre les deux problèmes, mais il subsiste quelques différences. Par exemple, le problème de la synthèse n'a pas toujours de solution car le système n'est pas forcément contrôlable, l'interaction est plutôt synchrone, les objectifs de contrôle sont divers (atteignabilité, inévitabilité, invariance, etc). D'un autre côté, en génération de test on cherche à maximiser le contrôle sur le système sans l'avoir complètement, l'interaction est souvent asynchrone, on se limite pour l'instant à des objectifs de tests de type atteignabilité.

L'objectif de la thèse est de contribuer à l'amélioration des techniques de synthèse de contrôleur et de génération de tests en s'appuyant sur les résultats de chacun des problèmes et en les étendant. Une piste prometteuse est de plonger ces deux problèmes dans le contexte plus général de la théorie des jeux en voyant le problème de la synthèse (respectivement de la génération de tests) comme une recherche de stratégie gagnante entre le système contrôlé (resp. l'implémentation) modélisé par la spécification et le contrôleur (resp. le testeur). On pourra aussi faire le lien avec la théorie du contrôle des systèmes à événements discrets de Ramadge et Wonham. Enfin, ce travail se fera en lien avec d'autres études sur la génération de tests symboliques (manipulant des données) et l'expressivité des objectifs de contrôle dans le *mu*-calcul.

**Pour en savoir plus :** <http://www.irisa.fr/pampa/>, <http://www.irisa.fr/epatr/>

Sujet de thèse proposé pour la rentrée 2000-2001

## Méthodes de réalisation répartie de spécifications partielles de systèmes réactifs dans le langage BDL

**Localisation :** Irisa, Rennes

**Equipe :** Pampa, Ep-Atr

**Responsables :** Benoît CAILLAUD (tél. direct : 02 99 84 74 07, email : Benoit.Caillaud@irisa.fr)

Albert BENVENISTE (tél. direct : 02 99 84 72 35, email : Albert.Benveniste@irisa.fr)

**Sujet :** Les phases initiales de la conception de logiciels répartis reposent à l'heure actuelle sur l'utilisation de spécifications partielles exprimées dans des notations telles que les diagrammes de séquences (abréviation MSC, pour *message sequence charts*), ou les contraintes OCL de la notation UML. Le langage BDL [1] permet également l'expression et la manipulation formelle de spécifications partielles. Chaque classe ou composant d'un modèle UML peut être muni d'une spécification BDL prenant la forme d'une famille de graphes orientés dont les sommets sont des événements valués pouvant être gardés par des conditions. Ces graphes spécifient des contraintes d'ordonnancement entre leurs événements sommets. BDL récupère en fait la technologie développée pour les langages synchrones (Esterel, Lustre, et en l'occurrence plus particulièrement Signal), tout en offrant une notion d'abstraction appropriée aux spécifications incomplètes. Le point nouveau, qui repose sur des résultats fondamentaux récents, est la possibilité d'interpréter un terme BDL indifféremment selon un mode synchrone ou asynchrone, en fonction de l'architecture du système modélisé et de la sémantique de communication entre objets.

Le travail de thèse portera sur l'élaboration de méthodes de synthèse de programmes répartis à partir de spécifications partielles BDL, qui prennent en compte l'environnement d'exécution des programmes. Il pourra s'inspirer à la fois des résultats récents sur la désynchronisation de systèmes de transitions synchrones [2] ainsi que de techniques classiques de synthèse de superviseurs de systèmes à événements discrets [3].

### **Bibliographie :**

[1] *Home Page of the BDL Project*, <http://www.irisa.fr/pampa/bdl/>

[2] A. Benveniste, B. Caillaud, P. Le Guernic, FROM SYNCHRONY TO ASYNCHRONY, CONCUR'99, Lecture Notes in Computer Science, vol. 1664, pp. 162–177, August 1999.

[3] P.J.G. Ramadge, W.M. Wonham, *The control of discrete event systems*, Proceedings of the IEEE, vol.77, no.1 p. 81-98, 1988

Sujet de thèse proposé pour la rentrée 2000-2001

## Génération de tests avec des modèles de vrai parallélisme

**Localisation :** Irisa, Rennes

**Equipe :** Pampa

**Responsable :** Claude Jard (tél. direct : 02 99 84 71 93, email : Claude.Jard@irisa.fr)

**Mot-clés :** Génie logiciel, Test, Méthodes formelles, Parallélisme, Distribution, Telecom.

**Financement :** à définir.

**Sujet :** L'équipe Pampa a développé depuis plusieurs années une méthode de génération de tests pour les systèmes réactifs (TGV). La théorie et l'algorithmique sont fondées sur une représentation des spécifications (objet à tester, objectif de test, test lui-même) sous forme de graphes états-transitions (automates). Ce modèle est simple et est largement utilisé pour tout ce qui ressemble à de l'analyse formelle de programmes répartis. Il a pourtant un défaut majeur, c'est qu'il ne contient aucune information sur le parallélisme possible entre les actions des programmes. En pratique, on fait comme si l'objet à tester pouvait être abstrait en une machine séquentielle. Cette vision trouve ses limites lorsque l'on veut s'intéresser à de nouvelles formes de test dans les systèmes répartis. Dans le contexte réparti, non seulement le système à tester est parallèle, mais on souhaite aussi que le test soit aussi parallèle : le testeur est lui-même un réseau parallèle de testeurs. Par ailleurs, il existe des modèles d'automates incluant du parallélisme comme celui des structures d'événements (ou des dépliages de réseaux de Petri), décrivant non plus des séquences mais des ordres partiels. La tentation est donc grande d'essayer de reprendre notre démarche de test en utilisant ce nouveau type de modèle et utiliser l'information de parallélisme pour générer du parallélisme dans les tests. Des travaux récents montrent la possibilité pratique de générer des structures d'événements à partir de spécifications de type automates communicants. L'objectif de la thèse est de revoir complètement la chaîne de génération TGV (depuis des spécifications dans le langage SDL vers des cas de test dans le langage TTCN) en utilisant le modèle intermédiaire des structures d'événements. Il s'agit d'un sujet comportant à la fois des aspects fondamentaux (utilisation d'un nouveau modèle resté confiné jusqu'alors dans le milieu de l'informatique théorique) et des aspects appliqués (méthodologie pour la génération de tests, validation expérimentale souhaitée). Il bénéficie aussi d'un intérêt réel de la part du monde industriel.

**Pour en savoir plus :** <http://www.irisa.fr/pampa/>

Sujet de thèse proposé pour la rentrée 2000-2001

## Test d'inter-opérabilité et de QoS de protocoles IP nouvelle génération

**Localisation :** Irisa, Rennes

**Equipe :** Pampa

**Responsable :** C. VIHO (tél. direct : 02 99 84 74 16, email : viho@irisa.fr)

C. JARD (tél. direct : 02 99 84 71 93, email : Claude.Jard@irisa.fr)

**Mot-clés :** test d'inter-opérabilité, protocole, télécommunication, QoS, mobilité, systèmes répartis.

**Sujet :** Le développement croissant des réseaux informatiques en général et d'Internet en particulier pose de sérieux défis pour les concepteurs de protocoles de communication. Ceci se complexifie avec l'explosion des télécommunications mobiles. Il s'agit donc pour la recherche, d'anticiper les problèmes liés à l'interconnexion de ces réseaux et de s'assurer qu'ils interopèrent correctement. Ceci passe par le développement de méthodes et outils appropriés pour le test d'interopérabilité des nouveaux protocoles mis en oeuvre dans ces réseaux. Du point de vue des utilisateurs de ces réseaux, il faudra également garantir une qualité de service (QoS) de plus en plus forte : garantie de débit minimum, de performance, d'accès aux ressources, etc. Au test d'interopérabilité, il faudra donc rajouter des tests de QoS.

Dans l'équipe Pampa de l'IRISA Rennes, nous menons depuis quelques années une activité de recherche soutenue sur le test de conformité de protocoles. Certains aspects échappent au test de conformité (incomplétude des spécifications, des objectifs de test) et des comportements inattendus surviennent lors de l'interconnexion des protocoles. C'est notamment le cas lorsque les fonctionnalités ne sont plus seulement définies par les comportements logiques admissibles, mais aussi par le respect et la cohérence des paramètres de qualité de service tels que le temps d'acheminement, le débit minimum garanti, etc.

Dans cette thèse, il s'agit d'étendre notre savoir-faire au contexte du test d'interopérabilité des nouveaux protocoles des réseaux informatiques, avec la prise en compte des aspects quantitatifs liés à la QoS. Il consistera à développer de nouvelles méthodes, de nouveaux algorithmes et outils appropriés, et à les implémenter dans des plate-formes d'expérimentation.

Sujet de thèse proposé pour la rentrée 2000-2001

## Méthodes symboliques pour la génération de tests de systèmes réactifs comportant des données

**Localisation :** Irisa, Rennes

**Equipe :** Pampa

**Responsable :** Vlad RUSU (tél. direct : 02 99 84 75 17, email : Vlad.Rusu@irisa.fr)

Thierry JERON (tél. direct : 02 99 84 74 64, email : Thierry.Jeron@irisa.fr)

**Mot-clés :** système réactif, test boîte noire, spécification formelle, objectifs de test, génération de tests symboliques.

**Sujet :** Le test de conformité est un test de type boîte noire. On se donne une spécification d'un système, qui sert de modèle de référence, et une implémentation réelle de ce système (implementation under test ou IUT) dont on ne connaît le comportement que par ses interactions avec l'environnement. Un test consiste à stimuler l'IUT par des événements, à observer les réactions de l'IUT, et à en déduire sa conformité ou non par rapport à la spécification. Le test de conformité permet de montrer la présence d'erreurs (non-conformité), mais ne permet pas en général de montrer leur absence (conformité). En pratique, les cas de test sont des programmes comportant des paramètres et des variables symboliques.

En nous basant sur notre expérience précédente avec l'outil de génération de tests TGV (qui repose sur des explorations énumératives de graphes d'états et produit des cas de test complètement instanciés), nous avons défini des modèles et proposé une approche pour générer des cas de test symboliques, qui sont corrects par construction. Cependant, les cas de test obtenus sont complexes en terme de nombre de variables et d'états: certaines variables sont redondantes, et certains états sont inaccessibles.

Le travail de thèse abordera ces problèmes. Il s'agit de définir et d'implémenter une véritable méthodologie de génération de tests symboliques, en combinant des techniques d'exploration de graphes, des techniques de propagation et de résolution de contraintes sur les variables symboliques et des techniques d'abstraction, pour produire des cas de test symboliques compatibles avec la demande industrielle.

Sujet de thèse proposé pour la rentrée 2000-2001

## Assemblage testable et validation de composants

**Localisation :** Irisa, Rennes

**Equipe :** Pampa

**Responsable :** Yves LE TRAON (tél. direct : 02 99 84 25 68, email : yletraon@irisa.fr)

**Mot-clés :** conception-testable, composant, test, UML, qualité.

**Sujet :** Cette thèse a pour objectif général la construction fiable et économiquement efficace d'applications par assemblage de composants logiciels.

Un premier angle d'attaque concerne l'amélioration des techniques de validation du logiciel à tous les niveaux (composant, intégration, système, maintenance). Mais dans le cas d'un assemblage de composants, il semble impératif d'intervenir dès en amont, au niveau du processus de développement, en cherchant à améliorer la démarche de spécification et de conception pour élaborer par construction un produit logiciel plus testable, plus « facile » à tester. On parle alors de conception testable. Cet axe de recherche vise ainsi à établir les règles et modèles applicables au plus tôt permettant:

- d'estimer l'effort et la qualité du test,
- de déterminer les faiblesses d'une architecture logicielle,
- de planifier les étapes de test et le contrôle qualité.

Les travaux déjà engagés dans le projet Pampa font apparaître des résultats quant à la validation de composants individuels vis-à-vis de leurs propriétés fonctionnelles dans un cadre séquentiel (auto-test et évaluation de sa qualité par analyse de mutations). Si ils ont conduit à l'élaboration de modèles formels dérivables depuis UML, il reste à les adapter pour prendre encore en compte les problèmes liés à la qualité de service et aux relations inter-composants (intégration). A terme, l'élaboration de contrats pour la testabilité pourra guider la reconfiguration d'un modèle de conception UML initial en une nouvelle conception testable.

Sujet de thèse proposé pour la rentrée 2000-2001

## Gestion globale des processeurs dans une grappe de calculateurs

**Localisation :** Irisa, Rennes

**Equipe :** Paris

**Responsable :** Christine MORIN (tél. direct : 02 99 84 72 90, email : cmorin@irisa.fr)

**Mot-clés :** système distribué, grappe de calculateurs, migration de processus, répartition de charge, disponibilité, calcul haute performance, simulation numérique distribuée.

**Sujet :** Les grappes de calculateurs interconnectés par un réseau à très haut débit constituent une architecture de choix pour l'exécution des applications de simulation numérique distribuée utilisant un ensemble de codes séquentiels ou parallèles qui coopèrent. Pour allier haute performance et haute disponibilité, nous nous proposons de mettre en oeuvre une gestion globale de l'ensemble des ressources (processeur, mémoire, disque) au sein du système d'exploitation de la grappe de calculateurs. Dans le système considéré, une gestion globale de la mémoire permet aux processus d'accéder aux données de la mémoire d'un noeud quelconque. Cette gestion mémoire est couplée à un système de gestion de fichiers parallèle offrant une gestion globale des disques de l'architecture. Un mécanisme de projection de fichiers en mémoire permet de rendre transparent aux applications les transferts de données entre la mémoire et les disques. L'objet du travail de thèse proposé est de concevoir et mettre en oeuvre dans ce système une gestion globale de la ressource processeur permettant aux processus d'être ordonnancés sur des processeurs quelconques de la grappe de calculateurs au cours de leur exécution. Le travail de thèse portera sur les trois points suivants :

- (1) A partir de l'étude des politiques d'ordonnancement ayant été proposées tant dans le domaine des machines parallèles que dans celui des systèmes distribués, élaborer une politique d'ordonnancement permettant d'exécuter efficacement des applications de simulation numérique distribuée sur une grappe de calculateurs ;
- (2) Concevoir un mécanisme de migration de processus tirant profit de la gestion globale des autres ressources du système ;
- (3) Evaluer les propositions des points (1) et (2) par des expérimentations sur une grappe de PCs sous Linux pour des applications réalistes de simulation numérique distribuée.

Sujet de thèse proposé pour la rentrée 2000-2001

## Conception d'un serveur de données sur une grappe de calculateurs

**Localisation :** Irisa, Rennes

**Equipe :** Paris

**Responsable :** Christine MORIN (tél. direct : 02 99 84 72 90, email cmorin@irisa.fr)

**Mot-clés :** système distribué, grappe de calculateurs, mémoire virtuelle répartie, système de gestion de fichiers parallèle, disponibilité, base de données, serveur de données.

**Sujet :** Les grappes de calculateurs sont attrayantes pour les applications haute performance de par leur excellent rapport performance/coût, leur extensibilité et l'intégration rapide des dernières avancées technologiques. De nombreux domaines applicatifs peuvent tirer profit d'une telle architecture parmi lesquels non seulement le calcul scientifique mais aussi les serveurs de données sur Internet. Dans le projet Paris, nous travaillons à la conception et réalisation d'un système d'exploitation distribué, Gobelins, pour l'exécution d'applications haute performance sur des grappes de calculateurs. Gobelins se caractérise par une gestion globale des ressources mémoire, disque et processeur pour garantir aux applications visées à la fois la haute performance et la haute disponibilité (ie la capacité de résister à la défaillance d'un des éléments du système). Ainsi, Gobelins met en oeuvre un système de stockage de données uniforme bâti à partir d'une mémoire virtuelle partagée (MVP) à pagination en mémoire distante et d'un système de gestion de fichiers parallèle (SGFP). La MVP assure une gestion globale de la ressource mémoire. Le SGFP assure une gestion globale de la ressource disque. L'interface du SGFP est la projection de fichiers en mémoire ce qui rend les transferts de données entre la mémoire et les disques transparents aux applications. L'objectif de cette thèse est de concevoir une infrastructure logicielle fondée sur le système Gobelins en vue de l'exécution efficace et fiable de serveurs de données de type base de données sur une grappe de calculateurs. Le travail de thèse consistera (1) à étudier l'architecture logicielle des serveurs de données fondés sur une base de données, (2) à caractériser les accès aux données dans le contexte des applications de base de données (3) à proposer une architecture logicielle de serveur de données tirant profit du système Gobelins. Des adaptations du système Gobelins pourront être proposées afin de satisfaire les exigences des serveurs de données fondés sur une base de données. En vue de son évaluation, le système proposé fera l'objet d'expérimentations sur une grappe de PCs bi-processeurs interconnectés par un réseau à haut débit.



Sujet de thèse proposé pour la rentrée 2000-2001

## Animation par ordinateur en synthèse d'images incluant fonctions robotiques et interactions

**Localisation** : Irisa, INRIA Rennes

**Equipe** : Siames

**Responsable** : Guy ANDRE (tél. direct : 99 84 74 65, email : [guy.andre@irisa.fr](mailto:guy.andre@irisa.fr))

Le sujet de thèse se déroulera dans l'équipe SIAMES (Synthèse d'images, Animation, Modélisation et Simulation). L'axe de recherche se propose de développer une synergie entre Informatique Graphique d'une part et Automatique-Robotique d'autre part. Les techniques actuelles en Animation sont basées sur des modèles de mouvement (système en boucle ouverte) produisant des trajectoires prédéfinies. Les recherches actuelles ont pour objectif l'animation de scènes complètes multi-acteurs où chaque entité est dotée d'une certaine adaptativité par rapport au contexte et à l'environnement (système en boucle fermée). L'"animation comportementale" s'intéresse à des entités dotées d'attributs et de comportements. A chaque entité est associée une cellule perception-décision-action-communication. Selon une approche pluridisciplinaire, la tendance consiste à transposer certains modèles et méthodes issus de la robotique, dans le domaine de l'informatique graphique. Une autre évolution importante est l'amélioration des interactions à différents niveaux du système d'Animation. Le sujet se focalise sur l'exploration de nouvelles fonctions liées à Animation et Interactions. Les études concernent l'identification de fonctions proches de la robotique, la définition de nouveaux schémas d'animation, les interactions en animation-simulation (par exemple incluant un périphérique à retour d'effort). Deux cas pourront être abordés: soit l'animation interactive de scènes génériques en images de synthèse enrichie par la simulation robotique, soit l'animation-simulation relativement à des robots réels. Divers aspects pourront être considérés: chaînes cinématiques, contraintes, simulation de perception et boucles perception-action, contrôleur associé, modes interactifs de commande, scénario, coopération (groupe)...Les travaux impliqueront l'étude et le prototypage de ces fonctions sur un simulateur générique. Plusieurs applications variées sont visées: la première porte sur le contrôle de personnages du type "humanoïde" en image de synthèse; la deuxième est le développement de simulateurs pour la téléopération ou la robotique. D'autres applications pourraient concerner les systèmes multi-agents (par exemple simulation de trafic de véhicules, groupe de micro-robots etc...). Les travaux donneront lieu à démonstrations sur stations Silicon Graphics et le centre de Réalité Virtuelle incluant salle immersive et écran géant.

Sujet de thèse proposé pour la rentrée 2000-2001

## Problème inverse d'illumination appliqué aux tunnels

**Localisation :** Irisa, Rennes

**Equipe :** Siames

**Responsable :** Kadi Bouatouch (tél. direct : 02 99 84 72 58, email : kadi@iris.fr), Eric Maisel (tél. direct : 02 99 84 72 59, email : maisel@iris.fr)

**Mot-clés :** Synthèse d'image-Illumination globale-Perception visuelle-Rendu temps réel.

Il est nécessaire, pour des raisons de sécurité et de confort visuel, d'assurer des niveaux d'éclairage suffisants à l'intérieur de la zone d'entrée des tunnels. Il faut quantifier le minimum d'éclairage assurant une sécurité suffisante sans toutefois négliger les aspects de confort visuel pour un conducteur pénétrant de jour dans un tunnel. Il est possible de déterminer la courbe de luminance sur la chaussée au niveau de la zone d'entrée. Afin de respecter cette contrainte de luminance, il faut déterminer les réflectances et BRDFs de la chaussée, des murs et du plafond ainsi que le type de source lumineuse, leur nombre et leur emplacement. Il s'agit d'un problème inverse. La chaussée peut être mouillée donc spéculaire.

Le travail demandé comportera plusieurs étapes. La première concerne le problème d'illumination inverse qui tiendra compte de la spécularité de la chaussée, les murs étant Lambertiens. Il faudra tenir compte des aspects de perception visuelle : contraste de luminance, niveau de visibilité, luminance d'adaptation, etc. Ensuite un logiciel de simulation d'éclairage global sera développé, on choisira un modèle trois points car il est indépendant du point de vue. Ce logiciel sera utilisé dans la troisième étape pour concevoir un nouvel algorithme permettant de se déplacer en temps réel dans une scène (un tunnel dans notre cas) comportant des objets spéculaires (la chaussée dans notre cas) afin de vérifier si les luminances au sol correspondent bien aux luminances requises, que le contraste de visibilité est au dessus d'un certain seuil minimal, etc.

Ce sujet aborde plusieurs domaines de la synthèse d'image : perception visuelle, problème inverse d'éclairage (optimisation), illumination globale et rendu temps réel en présence d'objets spéculaires (carte d'environnement, champs de lumière, etc.).

Sujet de thèse proposé pour la rentrée 2000-2001

## Modèles d'interactions coopératives en univers 3D

**Localisation :** Irisa, Rennes

**Equipe :** Siames

**Reponsable :** Thierry Duval (tél. direct : 02 99 84 72 56, email : thierry.duval@irisa.fr)

**Mot-clés :** Interfaces Homme-Machine, Interactions Coopératives, Immersion 3D, Réalité Virtuelle Distribuée.

**Sujet :** Le travail demandé consiste à proposer de nouveaux paradigmes d'interaction en environnements virtuels 3D, ces environnements pouvant être immersifs ou non, ainsi que des méthodes d'implémentation de ces nouveaux concepts d'interaction.

Ces paradigmes pourront s'appliquer à un cadre classique mono-utilisateur, mais également à un cadre d'interactions coopératives multi-utilisateurs. Dans le cas de coopérations multi-utilisateurs, celles ci pourront avoir lieu sur un même site (plusieurs utilisateurs se trouvant physiquement à proximité les uns des autres, devant un même dispositif de visualisation ou devant plusieurs) ou sur des sites distants (plusieurs machines connectées via un réseau).

L'approche *Génie Logiciel* de ce projet consistera à proposer une méthode générale pour interagir avec les objets de l'univers : comment rendre ces objets interactifs, et ce le plus facilement possible pour les personnes qui les conçoivent et les implémentent ? Quels services offrir à un développeur pour qu'il puisse créer des objets interactifs *a priori* ? Quel *design pattern* appliquer *a posteriori* pour rendre interactifs des objets qui n'ont pas été pensés pour l'interaction ?

Les développements se feront autour de la plate-forme de simulation GASP de l'équipe Siames, en utilisant C++ et Performer.

Sujet de thèse proposé pour la rentrée 2000-2001

## Détection et estimation robuste

**Localisation :** Irisa, Rennes

**Equipe :** Sigma2

**Responsable :** Jean Jacques Fuchs (tél. direct : 02 99 84 72 23, email : fuchs@irisa.fr)

**Mot-clés :** estimation, détection, optimisation.

**Sujet :** L'estimation robuste a pour objet de limiter l'effet de la présence de données aberrantes sur la qualité de l'estimation. Elle peut être précédée par une phase de détection des observations entachées de perturbations anormalement grandes mais en général la détection et l'estimation se font simultanément.

De nombreuses extensions sont envisageables. Au lieu de limiter la présence de perturbations anormales aux seules sorties, on peut considérer le cas où elles sont également présentes sur les entrées. Au lieu de se limiter à des perturbations ponctuelles, on peut considérer que des perturbations s'étendant sur plusieurs échantillons, des interférences possédant une structure interne mais par ailleurs inconnues, sont à détecter et à éliminer de la même façon avant de passer à l'estimation.

On se propose dans ce travail de recherche de développer de nouveaux algorithmes et d'analyser leurs performances. Une application à des données réelles (signal radar ou télécom mobile) complètera ce travail.

Sujet de thèse proposé pour la rentrée 2000-2001

## Prise en compte des liens sémantiques entre objets vidéo dans un contexte d'édition et de post-production vidéo

**Localisation** : Irisa, Rennes

**Equipe** : Temics

**Responsables** : Henri NICOLAS, tél. 02 99 84 74 25, Henri.Nicolas@irisa.fr,  
Laurent AMSALEG, tél. 02 99 84 74 44, Laurent.Amsaleg@irisa.fr

**Mots-clés** : Post-production, objets vidéo, réalisme, description de scènes, liens sémantiques

### **Sujet:**

Dans le monde de la post-production vidéo, de nombreuses applications nécessitent l'édition de vidéo originales. Il s'agit par exemple d'ajouter ou d'enlever des objets, ou encore de modifier leurs trajectoires. Pour cela, on utilise des méthodes semi-automatiques permettant d'obtenir des séquences modifiées réalistes.

Les méthodes actuelles se basent sur une notion de réalisme liée aux relations mécaniques entre objets, c'est à dire principalement la cohérence des trajectoires et des conditions d'illumination. Dans la réalité, il existe des relations plus générales entre les objets liées à la sémantique de la scène qui doivent être pris en compte pour garantir une cohérence globale de la séquence modifiée. Actuellement, ces aspects ne peuvent être pris en compte que manuellement par un opérateur humain.

L'objectif de cette thèse consistera tout d'abord à identifier et définir les relations sémantiques entre objets de manière à enrichir la description des scènes. Dans un second temps, il s'agira de développer et d'automatiser au maximum des outils d'édition vidéo sophistiqués prenant en compte cette description pour générer des scènes modifiées encore plus réalistes.

Cette thèse s'appuiera sur des travaux antérieurs menés dans l'équipe TEMICS.

Sujet de thèse proposé pour la rentrée 2000-2001

## Tatouage robuste d'images et de séquences vidéo

**Localisation :** Irisa, Rennes

**Equipe :** Temics

**Responsable :** Stéphane Pateux et Christine Guillemot (tél. direct : 02 99 84 73 60 / 74 29, email : [Stephane.Pateux@irisa.fr](mailto:Stephane.Pateux@irisa.fr) / [Christine.Guillemot@irisa.fr](mailto:Christine.Guillemot@irisa.fr))

**Mot-clés :** Tatouage ou 'watermarking' des signaux vidéos, tatouage seconde génération, Multimédia, analyse spatio-temporelle.

**Sujet :** Les technologies numériques, qui permettent l'apport de fonctionnalités nouvelles dans les applications de communication vidéo et multimédia, ainsi que l'apparition de réseaux ouverts comme l'Internet ne sont pas sans susciter de nouveaux dangers pour la propriété intellectuelle. Ces dangers conduisent à la réticence des possesseurs de contenus envers l'Internet, ou d'autres environnements ouverts de distribution. Grâce - ou à cause - des technologies numériques, il est en effet possible de faire des copies parfaites des contenus, la re-distribution est devenue aisée et le traçage des re-distributions est difficile. Ceci a motivé récemment les études de marquage (aussi appelé tatouage ou 'watermarking' en anglais) et d'authentification des images numériques.

Cette thèse s'intéressera à l'insertion de marques à des positions temporellement stables afin d'éviter des attaques par filtrage temporel, et dans des domaines de représentation de haut niveau de l'image afin d'être plus robuste aux différentes attaques susceptibles d'être utilisées (compression vidéo, altération de la vidéo, déformations géométriques, ...). Elle s'intéressera également à l'insertion de marques avec niveaux de protection différenciés, à la robustification de la marque par utilisation de codage à descriptions multiples, par utilisation de codes correcteurs et à l'enrichissement du décodage par une mesure de confiance dans les observations. Elle s'intéressera enfin à l'utilisation de techniques dites de seconde génération afin de fournir à la fois robustesse et souplesse au tatouage.

Sujet de thèse proposé pour la rentrée 2000-2001

## Codage et décodage conjoint source-canal pour transmission d'images sur des liens sans fil

**Localisation :** Irisa, Rennes

**Equipe :** Temics

**Responsable :** Christine Guillemot (tél. direct : 02 99 84 74 29, email : Christine.Guillemot@irisa.fr)

**Mot-clés :** Codage robuste des signaux vidéo, décodage pondéré, compression, dualité protection/compression, Théorie de l'information, Multimédia sur réseaux mobiles.

**Sujet :** Le modèle traditionnel adopté en communication repose sur un principe de séparabilité du codage de source et du codage de canal. Cependant, une optimisation séparée des codeurs de source et de canal, dans des environnements de communications hétérogènes, multipoints et/ou à caractéristiques variant dans le temps, s'avère largement sous-optimale. On a vu alors apparaître de nombreux travaux visant la prise en compte des caractéristiques du canal dans la conception des algorithmes de compression. Le premier réflexe a été d'abandonner les quantificateurs uniformes suivis de codes à longueurs variables, et de concevoir des quantificateurs non uniformes et des algorithmes d'indexation permettant de minimiser la distorsion globale (source + canal). Une autre approche consiste à exploiter la redondance résiduelle dans l'information compressée, ou encore une information "a priori" sur le signal compressé, dans des processus de décodage conjoint canal-source. Enfin, on a aussi vu apparaître récemment des transformations et des quantificateurs, dites à descriptions multiples, maintenant une quantité "contrôlée" de corrélation ou d'informations redondantes.

Dans ce contexte, cette thèse s'intéressera dans un premier temps au décodage à décision souple de codes de source à longueurs variables, et au décodage itératif de descriptions multiples. Elle se poursuivra par une prise en compte de la modulation dans la méthodologie de conception des modèles de représentation et de codage des signaux vidéo. Cette dernière approche, visant une optimisation globale de la chaîne de communication, évite la création d'un train binaire. Cette voie se trouve aujourd'hui confortée par l'émergence des concepts de radio logicielle, qui en permettant un traitement du signal en bande de base en tout logiciel devrait offrir une plus grande souplesse d'adaptation dynamique des fonctions de communication.

Sujet de thèse proposé pour la rentrée 2000-2001

## Analyse vidéo par maillages actifs 2D en présence de discontinuités

**Localisation :** Irisa, Rennes

**Equipe :** Temics

**Responsable :** M. Pateux (tél. direct : 02 99 84 73 60, email : [stephane.pateux@irisa.fr](mailto:stephane.pateux@irisa.fr))

**Mot-clés :** maillages actifs, MPEG4, segmentation.

**Sujet :** Cette thèse porte sur l'analyse de vidéos par l'intermédiaire de maillages actifs 2D. Les maillages actifs 2D sont une solution élégante afin de modéliser les déformations d'objets présents dans une scène vidéo, comme pourrait le faire un modèle 3D par facettes d'un objet (l'utilisation de maillages actifs 2D est prévue dans MPEG4 afin de pouvoir compresser efficacement certains types de vidéo). Toutefois, les déformations modélisées par ce type de maillage restent des déformations continues, et ne peuvent donc prendre en compte correctement les phénomènes d'occultation entre différents objets, ou bien encore d'auto-occultation au sein d'un même objet.

L'objectif de cette étude est de développer des outils d'analyses par maillages actifs 2D de vidéos en présence de discontinuités, et de gérer la structure du maillage en conséquence. Ainsi, les mailles d'un objet occulté par un objet en avant plan ne seront pas amenés à se déformer, mais à disparaître sous les mailles de l'objet en avant-plan, et pourront ainsi réapparaître ultérieurement (similairement à ce qui se ferait avec un maillage 3D des différents objets). La détection et la gestion de ces discontinuités, permet d'envisager ainsi différents cadres applicatifs: codage vidéo hybride, segmentation en objets vidéos (définition et/ou suivi), reconstruction 3D, ...

Dans cette thèse seront à développer les outils d'analyse par maillages, ainsi que de détection et de gestion des discontinuités; une phase de validation des différents outils sera effectuée au travers de différents cadres applicatifs (segmentation d'objets, codage vidéo, ...).



Sujet de thèse proposé pour la rentrée 2000-2001

## Indexation multimédia

**Localisation** : Irisa, Rennes

**Equipe** : Vista, Temics

**Responsable** : Patrick GROS, tél. 02 99 84 74 28, Patrick.Gros@irisa.fr

Laurent AMSALEG, tél. 02 99 84 74 44, Laurent.Amsaleg@irisa.fr

**Mots-clés** : indexation, algorithmique, complexité, base de données, images, multimédia

**Sujet** : La disponibilité toujours accrue de données multimédias numériques (images, sons, vidéos, texte) pose le problème de leur stockage et de leur indexation. Rien ne sert de conserver des documents si l'on ne peut les retrouver facilement. Les systèmes de gestion de bases de données proposent des techniques performantes pour l'indexation des données exactes, mais ces méthodes ne sont pas adaptées aux descripteurs issus de données telles les images ou les vidéos ou aux modèles statistiques représentant ces descripteurs. Dans ce cas, on est souvent amené à indexer des vecteurs qui sont à la fois imprécis et de grande dimension, ce qui nécessite la mise au point d'algorithmes différents.

Nous avons commencé un travail d'évaluation de quelques méthodes d'indexation récentes présentées dans la littérature. Deux axes complémentaires prolongent ce travail et forment le corps du sujet que nous proposons. Il est d'une part nécessaire de mieux comprendre les propriétés (distributions, corrélations, ...) des vecteurs qui décrivent les images pour identifier les caractéristiques que doit avoir une structure d'index efficace. Par ailleurs, les techniques modernes de description des images demandent d'enchaîner plusieurs interrogations consécutives avant de pouvoir retourner une réponse. Utiliser de cette façon les index multimédias actuels n'est pas viable, essentiellement pour des raisons de performance. Il est donc obligatoire de réfléchir aux manières d'enchaîner les interrogations, d'en limiter le nombre, afin d'optimiser les temps de réponse.

Le travail que nous proposons consiste à explorer ces questions, et se situe donc à l'intersection du domaine des bases de données et de celui de l'image. Il sera surtout consacré à des études algorithmiques et de complexité : il ne nécessite donc pas d'expertise pointue dans le domaine des bases de données ou dans celui des images et du multimédia. Par contre, une très bonne maîtrise de la programmation en C, C++ et une habitude du développement (sur station Unix de préférence) sont requises.

Sujet de thèse proposé pour la rentrée 2000-2001

## Méthodes d'analyses d'images échographiques tri-dimensionnelles

**Localisation :** Irisa, Rennes

**Equipe :** Vista

**Responsable :** C. BARILLOT (tél. direct : 02 99 84 75 05, email : Christian.Barillot@irisa.fr)

**Mot-clés :** modèles déformables, recalage d'images, réalité virtuelle, imagerie médicale 3D, imagerie cérébrale, imagerie échographique.

**Sujet :** Les Ultrasons sont une modalité d'imagerie médicale, peu coûteuse, d'une innocuité quasi totale pour le patient et qui permet de suivre le mouvement de structures anatomiques déformables en temps réel. Les ultrasons 3D risquent de bouleverser dans les prochaines années le diagnostic médical de nombreuses pathologies (en particulier les maladies cardio-vasculaires, première cause de mortalité dans les pays développés), ainsi que le contrôle en temps réel de nombreuses interventions chirurgicales. Au travers de cette thèse nous voulons développer une compétence dans ce domaine très prometteur.

Le but de cette thèse est de nous donner les moyens de mieux exploiter les images échographiques tridimensionnelles afin d'améliorer le diagnostic et le suivi thérapeutique, et dans ce cadre de développer des approches d'analyse d'image adaptées aux problèmes spécifiques de l'imagerie échographique.

Cette proposition de thèse fait suite à une phase préparatoire qui s'est déroulée dans le cadre d'une action de recherche concertée INRIA dans laquelle nous avons pu acquérir une expérience à la fois dans le domaine des systèmes d'acquisition 3D échographiques que dans le domaine du suivi des déformations dans des séquences d'images tridimensionnelles. Dans ce dernier cas, notre objectif final est de pouvoir intégrer les ultrasons 3D dans le cadre d'une imagerie interventionnelle (par exemple en neurochirurgie).

Sujet de thèse proposé pour la rentrée 2000-2001

## Utilisation des moments en vision robotique et réalité augmentée

**Localisation :** Irisa/Inria, Rennes

**Equipe :** Vista (<http://www.irisa.fr/vista>)

**Responsable :** F. Chaumette (tél. direct : 02 99 84 72 55, email : [chaumett@irisa.fr](mailto:chaumett@irisa.fr))

**Mot-clés :** Moments, asservissement visuel, calcul de pose.

**Sujet :** Les moments d'un objet (surface, centre de gravité, moments d'inertie, etc.) ont fait l'objet de nombreux travaux en vision par ordinateur, mais principalement dans le domaine de la reconnaissance de formes. Il s'agira ici d'exploiter les bonnes propriétés d'invariance et de découplage des moments par rapport aux divers mouvements 3D afin de généraliser à l'emploi d'objets complexes les techniques actuelles d'asservissement visuel et de localisation 3D. L'originalité de l'approche envisagée consiste à appréhender un objet sous sa forme globale (après une phase de segmentation spatiale) plutôt que d'en considérer des éléments caractéristiques. Les applications associées à ce travail se situent principalement dans le domaine de la vision robotique, mais aussi dans le domaine de la réalité augmentée où les aspects de localisation sont primordiaux.

Après les phases de modélisation puis de simulation, les expérimentations permettant de valider les travaux théoriques effectués seront menées sur une des cellules de vision robotique de l'IRISA (programmation en C ou C++ sous Linux).

Sujet de thèse proposé pour la rentrée 2000-2001

## Reconstruction 3D d'objets complexes par vision active

**Localisation** : Irisa/Inria, Rennes

**Equipe** : Vista (<http://www.irisa.fr/vista>)

**Responsables** : F. Chaumette (tél. direct : 02 99 84 72 55, email : chaumett@irisa.fr)

E. Marchand (tél. direct : 02 99 84 74 27, email : marchand@irisa.fr)

**Mot-clés** : vision dynamique, vision active.

**Sujet** : La reconstruction 3D d'objets simples à partir des informations fournies par une caméra mobile est un problème ne présentant plus de difficultés majeures. La démarche employée au sein de l'équipe VISTA consiste à commander par asservissement visuel les mouvements de la caméra afin d'obtenir une reconstruction précise et robuste. Le but de la thèse consistera à généraliser cette approche pour des objets complexes ne pouvant être représentés par des primitives géométriques paramétrées et comportant des surfaces gauches et/ou des contours d'occultation. Il s'agira de concevoir des stratégies globales d'exploration afin de déterminer tout d'abord un modèle grossier de l'objet, puis d'élaborer des stratégies locales de reconstruction optimale pour atteindre une précision maximale sur les zones d'intérêt.

Les expérimentations permettant de valider les travaux théoriques effectués seront menées sur une des cellules de vision robotique de l'IRISA (programmation en C ou C++ sous Linux).

Ce travail sera réalisé en collaboration avec le Cemagref Rennes dans le cadre du projet Capra, inscrit dans le Contrat de Plan Etat-Région Bretagne, pour la caractérisation de produits agro-alimentaires.