

Journée LIG-ENS Cachan,  
3 novembre 2010, MSH (domaine universitaire)  
contacts : Gaelle Calvary et Denis Trystram

## programme

9h00 - Accueil (commun) par Hervé Martin, directeur du LIG

	<b>Groupe n°1</b>	<b>Groupe n°2</b>
9h30-10h30	Roland Groz	Rachid Echahed
10h30-11h30	Frédéric Prost	Denis Trystram
11h30-12h30	Hubert Garavel	Simon Perdrix
12h30-14h	Déjeuner	
14h-15h	Simon Perdrix	Alexandre Termier
15h-16h	Jean-Bernard Stéfani	Eric Gaussier
16h-17h	Jérôme David	Yves Ledru
17h00	Synthèse et cloture	

Jérôme David

« Flexibilité et inter-opérabilité dans le Web sémantique »

Le web sémantique est vision du web dans lequel les ressources (texte, images, vidéos, etc.) sont annotées grâce à des vocabulaires formels (théories logiques) appelés ontologies. Même il est recommandé de réutiliser au maximum les ontologies, il n'y a aucune raison qu'il n'y ait, à l'échelle du web, qu'une seule ontologie par domaine. Ainsi, il n'est a priori pas possible de lier deux jeux de données qui sont annotés par 2 ontologies différentes. Afin de résoudre ce problème, il est nécessaire d'aligner les ontologies c.-à-d. trouver un ensemble de correspondances entre les concepts issus de deux ontologies.

Dans ce contexte, le projet EXMO s'intéresse à diverses questions tant théoriques qu'appliquées. Par exemple, EXMO a proposé une sémantiques de alignements (et réseaux d'alignements), diverses mesures de distance entre ontologies basées sur leur contenu ou sur les alignements existants. Depuis peu le projet s'intéresse à l'application de ses recherches notamment pour la publication et le liage de données (Linked Data), ou encore sur l'utilisation des technologies du web sémantique sur les terminaux mobiles afin de faciliter leur interaction avec le milieu ambiant.

Rachid Echahed

« Reasoning on Graph Transformation »

Graph rewrite systems generalize string rewrite systems (also known as Chomsky's grammars) as well as term (tree) rewrite systems. They allow one to easily encode complex data structures (as graphs) and their transformation (by means of graph rewrite rules). Graph transformation is omnipresent in computer modeling and constitutes a very high level framework well suited to model natural processes like in biology or chemistry. Contrary to term and string rewrite systems, graph transformation do not benefit yet from well established standard definitions, proof techniques and implementations. Very recently, we developed new logics dedicated to specify static and dynamic properties of graph transformation. The aim of this talk is to present and discuss new formal methods dedicated to reason on graph transformation.

Hubert Garavel

« Modélisation et vérification de systèmes parallèles complexes »

Eric Gaussier

« Accès à l'information dans les grandes collections de documents »

L'apprentissage statistique s'est imposé en quelques années comme une technologie clé pour le traitement et l'analyse des grandes masses de données, qu'elles proviennent de bases d'entreprises ou de données diffusées sur le web. Il est aujourd'hui employé dans une multitude de domaines sous la forme d'applications clé en main. En contrepartie, la croissance des données, leur complexification, la multiplication des besoins exprimés par les nouvelles applications liées aux nouveaux moyens d'accès aux données (web, mobilité, sites collaboratifs, etc.) génèrent quantité de nouveaux problèmes de traitement de données pour lesquels l'apprentissage n'a pas encore de réponse satisfaisante. Nous aborderons dans cet exposé quelques-uns de ces problèmes fondamentaux, en mettant en évidence les verrous scientifiques et les principales directions de recherche associées. Ces problèmes concernent la classification avec des milliers ou des dizaines de milliers de classes, la modélisation probabiliste de données textuelles ainsi que la modélisation jointe images et textes.

Roland Groz

« Rétro-Conception de modèles d'automates étendus de composants logiciels en boîte-noire »

L'ingénierie logicielle à base de modèles formels, et en particulier les techniques de validation basées sur des spécifications formelles, permettent d'outiller les méthodes de preuve ou de test. Malheureusement, on est souvent confronté en pratique à des composants logiciels développés sans modèles. Une piste pour réintégrer du formalisme dans les processus de développement consiste à reconstruire des modèles à partir de l'observation des systèmes.

Depuis quelques années, l'inférence d'automates apparaît comme une piste prometteuse dans la validation de systèmes dont certains composants restent en boîte noire (travaux de Peled, Steffen,

Pasareanu etc).

L'exposé présentera des recherches menées dans ce cadre, pour valider l'intégration de composants avec des modèles d'automates étendus, ou pour enrichir les techniques de vérification de la sécurité (frelatage intelligent - smart fuzzing, vérification de protocoles cryptographiques...).

Yves Ledru

« Applications de la modélisation formelle des réglementations »

De nombreuses activités humaines ne sont possibles que parce qu'il existe des réglementations qui les régissent. C'est notamment le cas dans le transport aérien où des réglementations internationales permettent d'assurer la sûreté et la sécurité des vols commerciaux.

La qualité des réglementations et la conformité de leur application sont essentielles pour garantir sûreté et sécurité. Dans cet exposé, nous présenterons deux exemples tirés du projet EDEMOI, qui s'est intéressé à la formalisation des réglementations qui régissent la sécurité des aéroports. Ces exemples montrent comment la modélisation formelle permet de mettre en évidence des hypothèses implicites d'une réglementation, de garantir qu'une réglementation européenne respecte la réglementation internationale, et de générer des tests de conformité pour l'évaluation d'un aéroport.

Simon Perdrix

« Informatique Quantique »

Des phénomènes mis en évidence par la physique quantique dans le comportement des particules élémentaires sont désormais considérés sous l'angle de leur exploitation pour représenter, traiter et communiquer l'information.

Des résultats algorithmiques (algorithme de factorisation de Shor, algorithme de recherche de Grover), cryptographiques (distribution quantique de clés BB84, téléportation) et des avancées expérimentales soulignent l'intérêt d'un fondement quantique des sciences de l'information. Nous présenterons ce domaine de recherche notamment à travers les activités sur ce sujet de l'équipe CAPP du LIG.

Frédéric Prost

« Calculs et dépendances »

Savoir qui fait quoi est une question importante dans la vie courante ... en informatique également. Nous nous attacherons dans cet exposé à présenter l'étude des "interférences" en programmation. Quelles sont les influences des données les unes par rapport aux autres ? Nous parlerons des bases théoriques, des méthodes pour mener à bien des analyses de non-interférences et aussi des applications que de telles analyses peuvent amener. Notamment en considérant des problèmes d'optimisations (repérer le code mort, ie le code qui n'a d'interférence avec personne), de confidentialité (on ne veut pas d'interférences entre les données secrètes et publiques), d'informatique quantique (l'intrication de données est intrinsèque au modèle de calcul).

Jean-Bernard Stéfani

Programming dynamic distributed systems - process calculi, types, languages and proof techniques

The talk will present challenges involved in designing languages (and their formal semantics) for programming distributed systems that can evolve and reconfigure themselves at run-time. The talk will present current work in the Sardes team around process calculi with localities, types for component assemblages, behavioral theory for higher-order concurrent languages, coinductive proof techniques and their formal support using the Coq proof assistant, programming languages for multicore and distributed computing.

Alexandre Termier

« Fouille de données en parallèle »

La fouille de données consiste à extraire automatiquement des connaissances nouvelles dans de grands volumes de données. La découverte de motifs fréquents en particulier, est le domaine qui s'intéresse à l'analyse de répétitions de certains motifs dans les données, avec de nombreuses

applications aussi bien commerciales que scientifiques.

Nos recherches s'efforcent d'une part de rendre ces méthodes d'analyse accessibles à un public plus large par le biais d'algorithmes génériques. D'autre part, nous travaillons à diminuer le temps d'exécution de nos algorithmes, en particulier par la mise au point d'algorithmes parallèles.

Denis Trystram

« Comment optimiser les ressources avec un grand nombre d'utilisateurs ? »

La présentation proposée ici se situe dans le contexte du développement récent des systèmes de calcul parallèles à grande échelle.

Les applications destinées à tourner sur ces systèmes sont décomposées en tâches qui sont exécutées sur les différentes ressources disponibles. Typiquement, ce genre de systèmes sont composés "d'organisations" possédant leurs propres grappes d'ordinateurs ou machines multi-cœurs. Ces ordinateurs sont partagés par tous les utilisateurs. Un utilisateur soumet ses tâches localement à son organisation, et un système d'ordonnancement global décide quelle machine va exécuter une tâche (éventuellement une machine d'une autre organisation).

Même si toutes les machines sont disponibles pour tous les utilisateurs, les utilisateurs d'une organisation veulent aussi tirer le meilleur parti de leurs propres ressources. Autrement dit, ils veulent que leurs performances ne se dégradent pas à cause du partage de ces ressources. Afin d'améliorer la coopération entre les organisations (incitation à exécuter des tâches venant de l'extérieur) les tâches locales doivent être exécutées en priorité sur le système de chaque organisation. La distribution de tâches dans le système entre les machines disponibles est un problème fondamental. Bien que chaque utilisateur soumette des tâches localement (dans sa propre organisation), il est nécessaire d'optimiser l'affectation des tâches sur toute la plate-forme (c'est-à-dire, globalement) afin d'obtenir la meilleure performance. De ce fait, il faut nécessairement déterminer un compromis entre la qualité de l'utilisation des ressources et le "contentement" des utilisateurs. Nous discuterons d'une part les limites actuelles des approches classiques pour répondre efficacement à ces questions, puis, nous proposerons rapidement une approche nouvelle utilisant l'optimisation multi-objectifs. Enfin, nous ouvrirons sur une approche plus prospective basée sur la théorie des jeux.