



---

## Thèses et HDR 2007-2008

---

Laboratoire d'InfoRmatique en Image et Systèmes d'information

LIRIS UMR 5205 CNRS / INSA de Lyon / Université Claude Bernard Lyon 1 / Université Lumière Lyon 2 / Ecole Centrale de Lyon





---

# Sommaire

---

2007

---

## Département Données, Connaissances, Services

### *Habilitation à diriger des recherches*

De l'adaptation à la prise en compte du contexte – Une contribution aux systèmes d'information pervasifs

**Lafortest Frédérique** .....3

### *Thèses de doctorat*

Contributions aux techniques de Prise de Décision et de Valorisation Financière

**Albert-Lorincz Hunor**.....5

Réingénierie des applications Web vers le Web sémantique: Approche dirigée par l'analyse de formulaires HTML

**Benslimane Sidi Mohamed** .....7

Intégration de l'information géographique dans les entrepôts de données et l'analyse en ligne : de la modélisation à la visualisation

**Bimonte Sandro** .....9

Exploration de données SAGE par des techniques de fouille de données en vue d'extraire des groupes de synexpression impliqués dans l'oncogénèse

**Blachon Sylvain**.....11

Grid Caching

**Cardenas Yonny** .....13

Adaptation d'applications pervasives dans les environnements multi-contextes

**Chari Tarak**.....15

Services pervasifs contextualisés : modélisation et mise en œuvre

**Dejene Ejigu Dedefa**.....17

Découvertes de motifs pertinents par l'implémentation d'un réseau bayésien : application à l'industrie aéronautique

**Fauré Clément** .....21

Découverte de correspondances sémantiques entre ressources hétérogènes dans un environnement coopératif	
<b>Ferreira da Silva Catarina</b> .....	<b>23</b>
Modèles et outils génériques pour la résolution des problèmes liés à la répartition de ressources sur grille	
<b>Gossa Julien</b> .....	<b>25</b>
Approximation Sémantique des Requêtes dans les Bases de Données Multimédia	
<b>Hammiche Samira</b> .....	<b>29</b>
Conception Coopérative d'Ontologies Pré-Consensuelles : Application au domaine de l'Urbanisme	
<b>Keita Abdel Kader</b> .....	<b>31</b>
Evolution de second ordre et algorithmes évolutionnaires : l'algorithme RBF-Gene	
<b>Lefort Virginie</b> .....	<b>35</b>
Une modélisation évolutionniste du liage temporel : A evolutionary model of temporal binding	
<b>Meunier David</b> .....	<b>37</b>
Médiation Sémantique Orientée Contexte pour la composition de Services Web	
<b>Mrissa Michaël Nicolas</b> .....	<b>39</b>
Reorganizing Dynamic Information Resources in Complex Web according to Complex Adaptive Systems Perspective	
<b>Rattrout Amjad</b> .....	<b>41</b>
Co-construction de sens par négociation pour la réutilisation en situation de l'expérience tracée	
<b>Stuber Arnaud</b> .....	<b>45</b>
Vers un accès sémantique aux données. Approche fondée sur RDF	
<b>Tanasescu Adrian</b> .....	<b>47</b>

## Département Image

### *Habilitations à diriger des recherches*

Modélisation de l'interaction des ions de haute énergie avec la matière inerte, la matière vivante et la matière en mouvement	
<b>Beuve Michaël</b> .....	<b>51</b>
Algorithmique pour l'analyse et la modélisation en géométrie discrète	
<b>Coeurjolly David</b> .....	<b>53</b>

### *Thèses de doctorat*

Tatouage Informé et Compression Multi-Sources	
<b>Dikici Cagatay</b> .....	<b>55</b>
Contribution à la restauration des images de documents	
<b>Drira Fadoua</b> .....	<b>57</b>

Real-Time Video Analysis from a Mobile Platform: Moving Object and Obstacle Avoidance <b>Limongiello Alessandro</b> .....	59
Calcul efficace et direct des représentations de maillages 3D utilisant les harmonies sphériques <b>Mousa Mohamed</b> .....	61
Solutions visuo-haptiques pour le prototypage virtuel, applications au secteur automobile <b>Ortega Michaël</b> .....	63
Variable Resolution Transform-based Music Feature Extraction and their Applications for Music Information Retrieval <b>Paradzinets Aliaksandr</b> .....	65
Prise en compte du mouvement respiratoire pour la reconstruction d'images tomodensitométriques <b>Rit Simon</b> .....	67
Interactive Light Transport with Virtual Point Lights <b>Segovia Benjamin</b> .....	69
Représentations hiérarchiques et discriminantes pour la reconnaissance des formes, identification des personnes et analyse des mouvements dans les séquences d'images <b>Thome Nicolas</b> .....	71

2008

---

## Département Données, Connaissances, Services

### *Habilitations à diriger des recherches*

Semantic and Privacy Aware Methods for Data Access <b>Benbernou Salima</b> .....	75
Apprivoiser la vie, modélisation individu-centrée de systèmes biologiques complexes <b>Beslon Guillaume</b> .....	77

### *Thèses de doctorat*

Découverte et composition dynamique de services web par approche multi-agents <b>Bova Rosanna</b> .....	79
Interactive and Opportunistic Knowledge Acquisition in Case-Based Reasoning <b>Cordier Amélie</b> .....	81
Modèles, langage et outils pour la réutilisation de profils d'apprenants <b>Eyssautier Carole</b> .....	83
Analyse de traces d'activité pour la modélisation cognitive : Application à la conduite automobile <b>Georgeon Olivier</b> .....	85

Confrontation de points de vue-opinions sur des corpus documentaires <b>Gesche Samuel</b> .....	87
Similarité sémantique inter-ontologies basée sur le contexte <b>Hoffmann Patrick</b> .....	89
Towards the conception of extended information systems organized around a web service platform <b>Le Blevec Yann</b> .....	91
Système d'accès personnalisé à l'information : application au domaine médical <b>Mirvat El Makssoud épouse Abbas Karine</b> .....	93
Accès personnalisé à l'Information - Adaptation au contexte de l'Utilisateur <b>Naderi Hassan</b> .....	95
Données et Métadonnées de Qualité pour les Champs Continus : Modèles de Terrain et Photogrammétrie <b>Pozzoli Alice</b> .....	97
Structuration relationnelle des politiques de contrôle d'accès représentation, raisonnement et vérification logiques <b>Thion Romuald</b> .....	99

## Département Image

### *Thèses de doctorat*

Apprentissage a contrario et architecture efficace pour la détection d'évènements visuels significatifs <b>Burrus Nicolas</b> .....	103
Compression adaptative de surfaces par ondelettes géométriques <b>Roudet Céline</b> .....	105
Géométrie discrète sur grilles irrégulières isothétiques <b>Vacavant Antoine</b> .....	107
Recognition of Emotions in Audio Signals <b>Xiao Zhongzhe</b> .....	109

2007

Département  
Données, Connaissances, Services





# De l'adaptation à la prise en compte du contexte – Une contribution aux systèmes d'information pervasifs

**Laforest Frédérique**

**HDR**

Soutenue le 10/12/2007

à l'INSA de Lyon - Université Claude Bernard, Lyon 1

## **Jury**

Pr. Corine Cauvet, LSIS, Aix-Marseille,	examineur
Pr. Bruno Defude, INT, Evry,	rapporteur
Pr. André Flory, LIRIS, Lyon,	président du jury
Pr. Maurice Laville, HCL, Lyon,	examineur
Pr. Hervé Martin, LSR-LIG, Grenoble,	rapporteur
Pr. Aris Ouksel, UIC, Chicago,	examineur
Pr. Florence Sèdes, IRIT, Toulouse,	rapporteur

Contact : frederique.laforest@liris.cnrs.fr

## **Résumé**

Le terminal mobile est aujourd'hui considéré comme un nouveau composant de l'architecture des systèmes d'information, et les nouvelles conditions d'usage induites forment le socle des systèmes d'information pervasifs. Les maîtres mots des systèmes pervasifs sont : sensibilité au contexte, smartness, scalabilité, invisibilité et pro-action. Tous ces mots-clés sous-tendent des procédures d'adaptation. La sensibilité au contexte permet de fournir les informations impulsant l'adaptation. Smartness et invisibilité sont des contraintes à appliquer entre autres à l'adaptation. La pro-action a pour but de prévoir les services et données qui seront nécessaires ; c'est une forme d'adaptation prévisionnelle.

Dans nos travaux, nous nous sommes d'abord intéressés aux documents comme paradigme d'interaction avec l'utilisateur. Cette approche nous a permis de ne plus restreindre le document à un seul outil de consultation de l'information, mais de l'étendre selon trois directions : saisie souple d'informations sous forme semi-libre, recherche de documents, et partage de documents. Dans un second temps, nous avons travaillé plus spécifiquement sur l'accès adaptatif aux données dans des systèmes pervasifs. Nous avons abordé l'adaptation d'interfaces graphiques à des terminaux multiples et la personnalisation des interfaces. Nous avons ensuite étudié l'adaptation dans le cadre plus général de la sensibilité au contexte, en proposant une méthode et une plate-forme pour l'adaptation d'applications legacy sur leurs dimensions interfaces graphiques, données et services.

Ce mémoire résume mon activité de recherche sur l'adaptation et les systèmes pervasifs durant une dizaine d'années. Il couvre le travail de plusieurs doctorants et de nombreux étudiants de master.

## **Abstract**

Mobile terminals are today considered as new components of information systems architectures, the induced new situations of use have conducted to pervasive information systems. The key words of pervasive information systems are context sensibility, smartness, scalability, invisibility and pro-action. All these keywords imply adaptation processes. Sensibility to the context provides information at the root of adaptation. Smartness and invisibility are constraints that must be applied to adaptation. Pro-action aims at foreseeing services and data that will be required ; it is a kind of predictive adaptation.

In our works, we have first studied documents as a paradigm for user interaction. We do not restrict documents use for information reading, but extend their interest in three directions : flexible information capture, documents search and documents sharing. Secondly, we have more specifically worked on the adaptative access to information in pervasive information systems. We have first worked on the adaptation of user interfaces to multiple terminals and the personnalization of user interfaces. We have then studied

the adaptation in a more global way, by providing a method and a platform for the adaptation of legacy applications on three dimensions : user interfaces, data and services.

My HDR dissertation summarizes my research activity on adaptation and pervasive systems done in the last ten years. It covers the works of some PhD students and many master students.

## Publications

- Frédérique Laforest and Youakim Badr. Modèle de couplage de documents structurés et de bases de données : le projet DRUID. Ingénierie des Systèmes d'Information, n° spécial Bases de données semi-structurées, 8(5-6):109--126, 2003.
- Tarak Chaari and Frédérique Laforest. Génération et adaptation automatiques des interfaces utilisateurs pour des environnements multi-terminaux : le projet SEFAGI. Ingénierie des Systèmes d'Information, n° spécial Systèmes d'information pervasifs, 9(2):11--38, 2004.
- Tarak Chaari, Dejane Ejigu, Frédérique Laforest, and Vasile-Marian Scuturici. A comprehensive approach to model and use context for adapting applications in pervasive environments. Int. Journal on Systems and Software}, 3, 2007.
- Yann Gripay, Frédérique Laforest, and Jean-Marc Petit. Towards Action-Oriented Continuous Queries in Pervasive Systems. In Bases de données Avancées 2007 (BDA'07), October 2007.

# Contributions aux techniques de Prise de Décision et de Valorisation Financière

THESE

Albert-Lorincz Hunor

Soutenue le 06/07/2007

à l'INSA de Lyon

## Jury

Dr. Guillaume Beslon (INSA Lyon)	examineur
Pr. Jean-François Boulicaut (INSA Lyon)	directeur de thèse
Pr. Arnaud Giacometti (Université de Tours)	rapporteur
Pr. Bart Goethals (Université d'Anvers, Belgique)	examineur
Pr. Alain Mille (Université Lyon 1)	examineur
Pr. Pascal Poncelet (Ecole des Mines d'Alès)	rapporteur
Denis Walez (Nomura Int. Plc)	examineur

Contact : alhunor@gmail.com

## Résumé

Nous travaillons dans le contexte général de la prise de décision financière. Nous considérons d'abord le problème algorithmique de l'extraction de motifs séquentiels, par exemple depuis des séries temporelles discrètes décrivant des évolutions de cours. Nous introduisons la famille des contraintes bien partitionnées qui permettent une décomposition hiérarchique des espaces de recherche. Le cas particulier de la conjonction d'une contrainte de fréquence et d'une contrainte d'expression régulière est détaillé. On peut alors réaliser des stratégies d'élagages adaptatives qui trouvent des compromis entre l'exploitation active de contraintes non anti-monotones (e.g., les expressions) sans pour autant se priver des possibilités d'élagage des contraintes anti-monotones (e.g., la fréquence). Nous présentons ensuite deux développements dédiés à la prise de décision financière. Nous proposons d'abord d'améliorer le pouvoir prédictif des indicateurs techniques en introduisant des signatures pour les configurations de marché et ainsi améliorer les performances des automates de « trading ». Ensuite, nous étudions la valorisation d'une classe particulière de produits dérivés où l'un des contreparties a le droit de rompre l'accord à une série de dates prédéterminées. Il est alors nécessaire de calculer des espérances conditionnelles à un futur état de monde, ce qui se fait traditionnellement par une double simulation Monte Carlo très gourmande en temps de calcul. Nous proposons une nouvelle technique baptisée « neighborhood Monte Carlo » qui est plus que 20 fois plus rapide que les méthodes précédentes.

## Abstract

This thesis investigates and develops tools for financial decision making. Our first contribution is aimed at the extraction of frequent sequential patterns from, for example, discretized financial time series. We introduce well partitioned constraints that allow a hierarchical structuration of the search space for increased efficiency. In particular, we look at the conjunction of a minimal frequency constraint and a regular expression constraint. It becomes possible to build adaptive strategies that find a good balance between the pruning based on the anti-monotonic frequency and the pruning based on the regular expression constraint which is generally neither monotonic nor anti-monotonic. Then, we develop two financial applications. At first, we use frequent patterns to characterise market configurations by means of signatures in order to improve some technical indicators functions for automated trading strategies. Then, we look at the pricing of Bermudan options, i.e., a financial derivative product which allows terminating an agreement between two parties at a set of pre-defined dates. This requires computing double conditional expectations at a high computational cost. Our new method, called "neighbourhood Monte Carlo" can be up to 20 times faster than the traditional methods.

## Publications

- H. Albert-Lorincz, J-F. Boulicaut. Mining frequent sequential patterns under regular expressions: a highly adaptive strategy for pushing constraints. Proceedings 3rd SIAM International Conference on Data Mining SDM'03, San Francisco, USA, May 2003. pp. 316-320.
- H. Albert-Lorincz, J-F. Boulicaut. A framework for frequent sequence mining under generalized regular expression constraints. Proceedings 2nd International Workshop on Knowledge Discovery in Inductive Databases KDID'03 co-located with ECML-PKDD 2003, Catvat-Dubrovnik, Croatia, September 2003. pp. 2-16.
- H. Albert-Lorincz, J-F. Boulicaut. Amélioration des indicateurs techniques pour l'analyse du marché financier. Actes 6e Journées Francophones d'Extraction et Gestion des Connaissances EGC'06, Lille, France, Janvier 2006. Cepadues RNTI-E-6, pp. 693-704.

# Réingénierie des applications Web vers le Web sémantique: Approche dirigée par l'analyse de formulaires HTML

**Benslimane Sidi Mohamed**

**THESE**

Soutenue le 17/06/2007

à l'Université de Sidi Bel Abbes (Algérie)

## Jury

Mustapha Kamel Rahmouni (Université d'Oran)	président
Ahmed Lehirech (Université de Sidi Bel Abbes)	examineur
Azzedine Chikh (Université de Tlemcen)	examineur
Fethallah Tebboune (Université de Sidi Bel Abbes)	examineur
Djamal Benslimane, (Université Lyon I)	encadrant
Mimoun Malki, (Université de Sidi Bel Abbes)	encadrant

Contact : [smbensli@bat710.univ-lyon1.fr](mailto:smbensli@bat710.univ-lyon1.fr) [Sidimohamed.benslimane@gmail.com](mailto:Sidimohamed.benslimane@gmail.com)

## Résumé

L'accroissement des technologies du Web et le développement rapide de ses applications, ont rendu l'information disponible n'importe où et n'importe quand. Cette émergence du Web dans tous les domaines, a permis à de nombreuses entreprises de rendre accessibles sur le Web une variété de leurs services, suscitant ainsi un besoin de partage et d'interopérabilité. Cela nécessite une infrastructure permettant à des agents logiciels d'exploiter, de composer et de raisonner sur les contenus constituant les ressources Web. La prolifération et la disponibilité des ontologies, qui est souvent au cœur de cette infrastructure, sont cruciales pour le succès de cette démarche qui s'inscrit dans le cadre du Web sémantique. Néanmoins leur construction demeure si coûteuse qu'elle entrave le progrès des activités du Web sémantique.

Pour contribuer à résoudre ce problème, nous nous sommes intéressés, dans cette thèse, à l'automatisation de la génération d'ontologie à partir d'application Web, en étudiant notamment l'impact que peut avoir les techniques de réingénierie dans une telle tâche. Nous présentons une approche semi-automatique de construction d'une ontologie OWL à partir d'un schéma de base de données relationnelle enrichi par la sémantique extraite des formulaires HTML d'une application Web. L'objectif de cette démarche est de réduire le coût de construction des ontologies, et de rendre exploitables par les machines, les bases de données relationnelles disponibles sur le Web.

## Abstract

The rapid growth of the Internet makes information available anywhere and anytime. As a result, most businesses run Web-based front-end databases upon which online services are offered to end-users. The next generation of the Web, the semantic Web, seeks to offer data in a usable form for automatic reasoning. To this purpose, it is necessary to make existing database content ready-to-use for semantic Web applications, which use ontologies to formally define the semantics of their data. The availability and proliferation of ontologies are crucial to the success of the semantic Web. Lately, ontologies have become the focus of research in several areas including knowledge engineering and management, information retrieval and integration, agent systems, the semantic Web, etc. Therefore, a large number of initiatives focus on building ontologies through automatic or semi-automatic processes. Nevertheless building ontologies is so costly that the progress of the semantic Web enhancement can get refrained. The major difficulties in building ontology reside in the manual work that could be prone to error. Therefore, the use of a semi-automatic ontology extraction technique is attractive.

In this thesis we present a semi-automatic reverse engineering approach that uses a relational database's HTML forms and a set of transformation rules to produce to an OWL ontology. Our approach proceeds by extracting semantics from the HTML-forms that relational databases expose to the external world. Prior to building the ontology in a semi-automatic way, the extracted semantics is combined with the proper semantics that underpins these relational databases. The main reason for this construction is to make the relational database information that is available on the Web machine-processable and reduce the time consuming task of ontology creation.

## Publications

### *Publications dans des Journaux/Revue internationale*

- S.M. Benslimane, M Malki, D. Bouchiha, D. Benslimane. "OntoWer: An Ontology based Web Application Reverse-Engineering approach". International Review on Computers and Software (IRECOS) 1(1):52-58, Praise Worthy Prize ed, ISSN 1828-6003. 2006.  
URL : [http://www.praiseworthyprize.com/IRECOS\\_vol\\_1\\_n\\_1.html](http://www.praiseworthyprize.com/IRECOS_vol_1_n_1.html)
- S.M. Benslimane, M Malki, M. Rahmouni, D. Benslimane. "Extracting Personalised Ontology from Data-Intensive Web Application: An HTML Forms-Based Reverse Engineering Approach". International Journal of INFORMATICA 18(4):511-534, 2007. IOS Press, ISSN 0868-495.  
URL: <http://www.mii.lt/Informatica/htm/INFO688.htm>
- S.M. Benslimane, D. Benslimane, M Malki, Z. Maamar, P. Thiran, Y. Amghar, M. Hacid. "Ontology development for the Semantic Web: An HTML form-based reverse engineering approach". International Journal of Web Engineering (JWE), 6(2):143-164, Rinton Press, ISSN 1540-9589. 2007. URL: <http://www.rintonpress.com/journals/jweonline.html#v6n2>
- S.M. Benslimane, M Malki, M. Rahmouni, A. Rahmoun. "Towards Ontology Extraction from Data-Intensive Web Sites: An HTML Forms-Based Reverse Engineering Approach". The International Arab Journal of Information Technology (IAJIT), 5(1):34-44, ISSN 1683-3198. 2008. URL : <http://www.iajit.org/>

### *Publications dans des conférences internationales*

- S.M. Benslimane, M Malki, M. RAHMOUNI, D. Benslimane. "Building domain-specific ontology from data-intensive Web site: An HTML forms-based reverse engineering approach" IEEE SITIS 2005, The International Conference On Signal-Image Technology & Internet-Based Systems, November 27th - December 1st, Cameroon. 2005. ISBN: 2-9525435-0, pp. 246-251.
- S.M. Benslimane, M. Malki, Amar Bensaber Djaml, Automated Migration of Data-Intensive Web Pages into Ontology-Based Semantic Web: A Reverse Engineering Approach. In, Meersman R., Tari Z. et al.,(eds.), On the Move to Meaningful Internet Systems 2005: DOA, ODBASE and CoopIS, vol. 2, LNCS 3761, pp. 1640 - 1649, 2005. Springer Verlag.
- S.M. Benslimane, D. Benslimane, M Malki, Y. Amghar, H. Saliah. "Acquiring OWL ontologies from data-intensive web sites". The Sixth International Conference on Web Engineering (ICWE'06), ACM ed. Palo Alto, California, USA. pp. 361-368. 2006.
- S.M. Benslimane, D. Benslimane, M Malki, Y. Amghar, F. Gargouri. « Construction d'une ontologie à partir d'une base de données relationnelle : approche dirigée par l'analyse des formulaires HTML ». INFORSID'06, Hammamet (Tunisie). pp. 991-1010. 2006.
- S.M. Benslimane, D. Benslimane, S Suwanmanee, Z. Maamar, H. Saliah Towards owl ontologies from relational databases: An HTML-form driven approach. . 22nd International Conference on Computers and Their Applications (CATA-2007), Honolulu, Hawaii U.S.A., pp. 219-224. 2007.

# Intégration de l'information géographique dans les entrepôts de données et l'analyse en ligne : de la modélisation à la visualisation

THESE

**Bimonte Sandro**

Soutenue le 18/12/2007

à l'INSA de Lyon

## Jury

M. Michel Schneider, Université Blaise Pascal, Clermont-Ferrand	rapporteur
M. Esteban Zimányi, Université Libre de Bruxelles, Belgique	rapporteur
M. Hervé Martin, Université Joseph Fourier, Grenoble	
Mme Anne Tchounikine, INSA-Lyon,	co-directrice de thèse
M. Robert Laurini, INSA-Lyon	directeur de thèse

Contact : maryvonne.miquel@liris.cnrs.fr

## Résumé

Les systèmes d'entrepôts de données et l'OLAP (On Line Analytical Processing) sont des solutions pour l'analyse décisionnelle. L'information géographique est très fréquemment présente dans les données, mais généralement sous-employée dans le processus décisionnel.

Le couplage de systèmes OLAP et de Systèmes d'Informations Géographiques au sein de systèmes OLAP Spatial (SOLAP) est une voie prometteuse. La majorité des solutions SOLAP existantes réduisent l'information géographique à la seule composante spatiale, limitant ainsi les capacités d'analyse du paradigme spatio-multidimensionnel. Nous proposons un modèle formel (GeoCube) et une algèbre associée, qui reformule les concepts du SOLAP afin d'introduire les aspects sémantiques et spatiaux de l'information géographique dans l'analyse multidimensionnelle. Cela se traduit par une modélisation des mesures sous forme d'objets géographiques organisés en hiérarchies, dans une vision complètement symétrique entre mesures et dimensions.

Nous proposons une algèbre qui fournit classiquement les opérateurs de forage et de coupe, mais aussi un opérateur qui permet de permuter mesure et dimension et des opérateurs de navigation au sein de la hiérarchie de mesures. L'introduction de nouveaux opérateurs de modification dynamique de l'hypercube par l'application d'opérateurs spatiaux, permet de concilier analyse OLAP et analyse spatiale. Nous avons réalisé un prototype web conforme à GeoCube. Pour décrire nos solutions, nous utilisons des données environnementales de la lagune de Venise. Enfin, nous proposons un nouveau paradigme de visualisation et d'interaction pour l'analyse des mesures géographiques.

## Abstract

Data warehouse and OLAP systems are decision-making solutions. Integration of spatial data into OLAP systems is an important challenge. Indeed, geographic information is always present implicitly or explicitly into data, but generally it is not exploited into the decisional process. Spatial OLAP (SOLAP) systems, which are the integration of OLAP and Geographic Information Systems (GIS), are a promising way. Most of SOLAP solution reduces geographic information to its spatial component, limiting the analysis capabilities of the spatio-multidimensional paradigm. We propose a formal model (GeoCube) and its associated algebra. GeoCube reformulates main SOLAP concepts in order to introduce semantic and spatial aspects of geographic information into the multidimensional analysis. So, we model measures and dimension members as geographic and or complex objects. A measure can belong to one or more hierarchies. We propose an algebra which provides the drill and slice operators, an operator to invert measures and dimension, and two operators to navigate into the hierarchy of the measure. The algebra permits to introduce the spatial analysis methods into multidimensional analysis through some new operators which

change dynamically the structure of the hypercube. We have realized a web prototype (GeWOlap) based on GeoCube. We describe our works using environmental data of Venice lagoon pollution. Finally, we propose a new visualization and interaction paradigm to analyze geographic measures.

## Publications

- Introduction de l'analyse spatiale et de l'information géographique dans l'analyse multidimensionnelle. S Bimonte, A. Tchounikine, M. Miquel, R. Laurini. Revue des Nouvelles Technologies de l'Information (RNTI), Editions Cépaduès (à paraître)
- Dall'OLAP Spaziale all'OLAP Geografico : Un nuovo paradigma per l'analisi multidimensionale di dati geografici.. A. Tchounikine, M. Miquel, S Bimonte. L'informazione geografica a supporto della pianificazione territoriale. Murgante B. (a cura di) FrancoAngeli Milano, 133-159 p. 2008.
- Vers l'intégration de l'analyse spatiale et multidimensionnelle. S Bimonte, A. Tchounikine, M. Miquel, R. Laurini. Dans Colloque International de GEOmatique et d'Analyse Spatiale (SAGEO), Clermont-Ferrand, France. pp. 1-12. 2007
- Supporting Geographical Measures Through A New Visualization Metaphor In Spatial OLAP. S Bimonte, A. Tchounikine, S. Di Martino, F. Ferrucci. Dans 9th International Conference on Enterprise Information Systems (ICEIS07), Funchal, Madeira - Portugal. 2007.
- Spatial OLAP: Open Issues and a Web Based Prototype. S Bimonte, A. Tchounikine, M. Miquel. Dans 10th AGILE International Conference on Geographic Information Science (short paper), Aalborg. pp. 1-11. 2007.
- GR-OLAP: On Line Analytical Processing of GRid monitoring information. J. Gossa, S Bimonte. Encyclopedia of Database Technologies and Applications, Second Edition, Eds. Laura C. Rivero, Jorge Horacio Doorn, Viviana E. Ferraggine. IDEA Group Inc. 2007
- GeWOlap: A Web Based Spatial OLAP Proposal. S Bimonte, P. Wehrle, A. Tchounikine, M. Miquel. Dans Second International Workshop on Semantic-based Geographical Information Systems, Robert Meersman, Zahir Tari, Pilar Herrero et al ed, Montpellier, France. pp. 1596-1605. Springer-Verlag . 2006
- GeoCube, a Multidimensional Model and Navigation Operators Handling Complex Measures: Application in Spatial OLAP. S Bimonte, A. Tchounikine, M. Miquel. Dans Fourth Biennial International Conference on Advances in Information Systems, Erich Neuhold, Tatyana Yakhno ed. Izmir, Turquie. pp. 100-109. Springer-Verlag . 2006
- GeOlaPivot Table: a Visualization Paradigm for SOLAP Solutions. S Bimonte, D.M. Sergio, F. Filomena, A. Tchounikine. Dans Visual Languages and Computing Workshop of the 2006 International Conference on Distributed Multimedia Systems (DMS'2006), Grand Canyon, USA. 2006
- Panorama de travaux autour de l'intégration de données spatio-temporelles dans les hypercubes. A. Tchounikine, M. Miquel, R. Laurini, T Ahmed, S Bimonte, V. Baillot. 1ère journée francophone sur les Entrepôts de Données et l'Analyse en ligne (EDA 2005), Lyon, Revue des Nouvelles Technologies de l'Information : 21-23. 2005
- Towards a Spatial Multidimensional Model. S Bimonte, A. Tchounikine, M. Miquel. Dans DOLAP 05, ACM Eighth International Workshop on Data Warehousing and OLAP, Song and J. Trujillo ed. Bremen, Allemagne. pp. 39-46. ACM . 2005.



# Exploration de données SAGE par des techniques de fouille de données en vue d'extraire des groupes de synexpression impliqués dans l'oncogénèse

**Blachon Sylvain**

**THESE**

Soutenue le 06/07/2007

à l'INSA de Lyon

## Jury

Pr. Jean-François Boulicaut (INSA Lyon)	co-directeur
Pr. Bruno Crémilleux (Université de Caen)	examineur
Dr. Olivier Gandrillon (Université Lyon1/CNRS CGMC)	co-directeur
Dr. Jean-Jacques Kupiec (INSERM Paris)	rapporteur
Pr. Jean-Daniel Zucker (Université Paris 13)	rapporteur

Contact : sylvain.blachon@gmail.com

## Résumé

Avec le développement de techniques à haut débit, l'accumulation de grandes quantités de données permet au biologiste de se poser de nouvelles questions. L'essentiel de notre contribution réside dans l'étude et l'interrogation des données SAGE humaines issues du « Cancer Genome Anatomy Project ». Nous avons étudié en profondeur ces données et les questions biologiques que nous pouvions nous poser à partir de ces données. Répondre à ces questions a nécessité différentes méthodes d'extraction de connaissances à partir des données et la conception de scénarios originaux de découverte de connaissances. Leur mise en œuvre a reposé sur l'utilisation de différents algorithmes d'extraction de motifs développés par les partenaires de l'ACI Bingo « Bases de Données Inductives pour la Génomique ». Nous avons étudié le problème du codage booléen de propriétés d'expression à partir de données SAGE et nous avons particulièrement travaillé au post traitement des motifs extraits et à leurs interprétations (classification de motifs similaires pour contrôler le nombre de « quasi synexpression groups » - QSGs - à interpréter, exploitation de sources externes comme Gene Ontology pour identifier des motifs pertinents). L'impact de ces éléments de méthodes a été validé sur une interprétation de deux QSGs extraits des données SAGE humaines.

## Abstract

The development of high throughput methods that enable to collect molecular biology data gives rise to new opportunities for the biologists. Our main contribution concerns an in-depth study of the human SAGE data collected by the Cancer Genome Anatomy Project. We have studied in depth these data and the kind of questions that could be addressed thanks to them. This has been based on various data-driven knowledge discovery processes also called KDD scenarios. The practical application of these scenarios has been possible thanks to the many pattern discovery algorithms designed by the partners of ACI Bingo « Bases de Données Inductives pour la Génomique ». Among others, we studied Boolean encoding of gene expression properties when starting from SAGE gene expression raw data. We made a specific contribution to the post-processing of large pattern collections, and we interpreted the biological meaning of some of the extracted patterns (clustering similar patterns to control the number of quasi-synexpression groups - QSGs – that have to be interpreted, using data sources like Gene Ontology to support the discovery of relevant patterns). The impact of these gene expression data analysis methods has been validated on the interpretation of two QSGs that hold in the studied human SAGE data.

## Publications

- Blachon S., Pensa R.G., Besson J., Robardet C., Boulicaut J.F. and Gandrillon O., (2007) Clustering formal concepts to discover biologically relevant knowledge from gene expression data. Révision demandée et

soumise à In Silico Biology, 25 pages.

- Becquet, C., Blachon, S., Jeudy, B., Boulicaut, J.F. and Gandrillon, O. (2002). Strong association rule mining for large gene expression data analysis: a case study on human SAGE data. *Genome Biology* 3(12):research0067.1-0067.16 .
- Blachon, S., Robardet, C., Boulicaut J.F. et Gandrillon, O. (2004). Extraction de connaissances dans les données d'expression SAGE humaines. *Informatique pour l'Analyse du Transcriptome*. Chapitre 7, pages 207-230. *Traité IC2*. Hermes Lavoisier.
- Klema J., Soulet A., Crémilleux B., Blachon S., Gandrillon O. Mining Plausible Patterns from Genomic Data. *Proc. 19th IEEE Int. Symp. on Computer-Based Medical Systems CBMS'06*, IEEE Computer Society Press, Salt Lake City, USA, June 2004. pp. 183-188.
- Hébert, C., Blachon, S., Crémilleux, B. (2005). Mining delta-strong characterization rules in large SAGE data. *Proc. Discovery Challenge co-located with ECML-PKDD 2005, Porto, PT, October 2005*. pp. 90-101.
- Rioult, F., Robardet, C., Blachon, S., Crémilleux, B., Gandrillon, O., Boulicaut J.F. Mining concepts from large SAGE gene expression matrices. *Proc. 2nd Int. Workshop on Knowledge Discovery in Inductive Databases KDID'03 co-located with ECML-PKDD 2003, Catvat-Dubrovnik, HR, September 2003*. pp. 107-118.
- C'06, Lille, F, janvier 2006, Cepadues RNTI-E-6, pp. 569-580.

Soutenue le 10/12/2007

à l'INSA de Lyon

### Jury

Norman PATON, University of Manchester	rapporteur
Jean-Marc NICOD, Université de Franche-Comté	rapporteur
Lionel BRUNIE, INSA de Lyon	directeur de thèse
Jean-Marc PIERSON, Université. Paul Sabatier	co-directeur de thèse
Abdelkader HAMEURLAIN, IRIT Toulouse	examineur
Claudia RONCANCIO, ENSIMAG Grenoble	examinatrice

Contact : lionel.brunie@liris.cnrs.fr

### Abstract

Grids support multiple models of distributed computation and need to operate large data entities in a distributed way. A significant quantity of these data are used only for a limited period of time. Consequently, temporary data management in grid is highly critical and global coordination of the storage network resources is necessary. In this context, caching is recognized as one of the most effective techniques to manage temporary data and collaborative cache is traditionally proposed to scale cache capabilities in distributed environments.

This thesis proposes an approach for the design and implementation of collaborative cache systems in grids that supports capabilities for monitoring and controlling cache interactions. Our approach permits to compose and evaluate high-level collaborative cache functions in a flexible way. Our proposal is based on a multilayer model that defines the main functions of a collaborative grid cache system. These capabilities are implemented by a set of common and standard operations that support the data access, the monitoring and the configuration of a cache. This model and the provided specification are used to build a flexible and generic software infrastructure for the operation and control of collaborative caches.

This infrastructure is composed of a group of autonomous cache elements called Grid Cache Services (GCS). The GCS is a local administrator of temporary storage and data which is implemented as a grid service that provides the cache capabilities defined by the model. It implements interfaces and associated cache operations. We implemented a prototype of the GCS with Globus Toolkit 4 middleware and we have made wide area performance measurements for our GCS prototype. We study a possible configuration for a group of GCS that constitutes a basic management system of temporary data called Temporal Storage Service (TSS). TSS uses a group of caches that collectively provide storage capabilities to store temporary data. The results of a simulation experiment demonstrate that group caches (GCS) operating according to their activity can provide supplementary storage space with minimal cost.

### Publications

#### Journals

- Management of a Cooperative Cache in Grids with Grid Cache Services. Y. Cardenas, J. Pierson, L. Brunie. *Concurrency and Computation: Practice and Experience*, special issue, Volume 19, Issue 16 (November 2007), pp. 2141 – 2155, John Wiley & Sons, Ltd. ISSN 1532-0634. 2007
- GGM Efficient Navigation and Mining in Distributed Geno-Medical Data. J. Pierson, J. Gossa, P. Wehrle, Y. Cardenas, E.S. Mahmoud, C. Sebastien, L. Brunie, C. Dhaenens, H. Abdel Kader, N. Melab, M. Miquel, F. Morvan, T. El Gazali, A. Tchounikine. *IEEE Transactions on NanoBioscience* 6(2):110-116, IEEE, ISSN 1536-1241. 2007.

### *International Workshops*

- Management of Cooperative Cache in Grids. Y. Cardenas, J. Pierson, L. Brunie. 2nd International Workshop on Data Management in Grids DMG – In conjunction with the 32nd International Conference on Very Large Data Bases ,VLDB 2006, Seoul, Korea. pp. 38-49. 2006.

### *International conference papers (In proceedings)*

- Temporal Storage Space for Grids. Y. Cardenas, J. Pierson, L. Brunie. Second International Conference on High Performance Computing and Communications (HPCC 2006), Munich, Germany. pp. 803-812. Lecture Notes in Computer Science . Springer Springer Berlin / Heidelberg. ISBN 978-3-540-39368-. ISSN 0302-9743. 2006.
- Uniform Distributed Cache Service for Grid Computing. Y. Cardenas, J. Pierson, L. Brunie. In 16th DEXA: In 2th International Workshop on Grid and Peer-to-Peer Computing Impacts on Large Scale Heterogeneous Distributed Database Systems., IEEE ed. Copenhagen, Denmark. pp. 351-355. IEEE Computer Society Los Alamitos, California. ISBN 0-7695-2424-9. ISSN 1529-4188. 2005.

### *National conference papers (In proceedings)*

- Service de Cache pour les Grilles de Calcul. Y. Cardenas, L. Brunie. Dans actes du Premier atelier sur Extraction et Gestion Parallèles Distribuées de Connaissances (EGPDC'2005), 5èmes Journées d'Extraction et Gestion des Connaissances(EGC'2005), Association EGC, Paris, 2005.

### *Contributions to books*

- Service de Cache pour les Grilles de Calcul. Y. Cardenas, J. Pierson, L. Brunie. Dans Chapitre 6, RNTI E-5 "Extraction des connaissances: Etat et perspectives", Editions Cépaduès, ISBN 2-85428-707.X 2005.

### *Posters*

- Ragtime Rhône-Alpes Grille pour le Traitement d'Informations Médicales.Y. Cardenas, K Hassan, L Seitz. Dans 16èmes Rencontres Régionales de la Recherche, Région Rhône-Alpes, Grenoble, France, Janvier 2006.

# Adaptation d'applications pervasives dans les environnements multi-contextes

THESE

Chaari Tarak

Soutenue le 28/09/2007

à l'INSA de Lyon

## Jury

Philippe ANIORTÉ,	rapporteur
Florence SÈDES,	rapporteur
Augusto CELENTANO,	
Paul RUBEL,	
André FLORY	co-directeur
Frédérique LAFOREST	co-directrice

Contact : frederique.laforest@liris.cnrs.fr

## Résumé

Les systèmes pervasifs ont pour objectif de rendre l'information disponible partout et à tout moment. Ces systèmes doivent pouvoir être utilisés dans différents contextes selon l'environnement de l'utilisateur, son profil et le terminal qu'il utilise. L'un des problèmes majeurs de ce type de systèmes concerne donc l'adaptation au contexte d'utilisation. Dans ce travail de thèse, nous proposons une stratégie complète, générique et évolutive d'adaptation d'applications au contexte d'utilisation sur trois volets: (i) les services offerts à l'utilisateur, (ii) les données renvoyées par ces services et (iii) leur présentation à l'utilisateur. L'adaptation des services consiste à intercepter les appels vers les services originaux de l'application et à modifier leur comportement à l'aide d'un ensemble d'opérateurs d'adaptation fonctionnelle. L'adaptation des données consiste à transformer ou à remplacer chaque objet multimédia renvoyé par les services de l'application qui n'est pas utilisable dans la situation contextuelle en question. L'adaptation de la présentation se base sur un processus de génération automatique du code de l'interface utilisateur qui garantit l'interaction avec les données et les services adaptées. La stratégie que nous avons élaborée atteint deux objectifs : (i) intégrer la sensibilité au contexte dans l'application de façon incrémentale et (ii) garantir l'adaptation à partir d'une description simple des services offerts à l'utilisateur. Nous avons mis en œuvre cette stratégie en développant une plateforme d'adaptation d'applications au contexte d'utilisation. Nous avons utilisé les technologies Java, OSGi et les services Web pour réaliser cette plateforme. Nous avons également validé nos modèles et notre approche d'adaptation sur une application médicale de suivi de patients dialysés à domicile.

## Abstract

Pervasive systems aim to make information available anywhere and at anytime. These systems should be used in different contexts depending on the environment of the user, her/his profile and her/his device. Consequently, one of the main problems of this type of information systems is the adaptation to context. In this PHD thesis, we propose a complete, generic, and evolutionary strategy that guarantees the adaptation of applications to context on three facets: (i) the services offered to the user, (ii) the data returned by these services and (iii) their presentation to the user. Service adaptation consists of modules that intercept the application's service calls and modifying their behaviour using a list of functional adaptation operators. Data adaptation consists in transforming or replacing the non-usable multimedia service outputs in the considered context situation. Presentation adaptation consists in automatically generating the complete code of the user interface that guarantees the interaction with the adapted data and services. Our adaptation strategy has achieved two goals: (i) integrate context incremental awareness in the application and (ii) guarantee the adaptation starting from a simple description of the services offered to the user. We have validated this strategy by developing a platform that guarantees the adaptation of applications to

context. We have used Java, OSGi and Web service technologies to implement this platform. We have also successfully tested our adaptation approach on a home health care application of dialysed persons. In essence, we believe that this work is a step ahead towards building adaptive and evolutionary pervasive information systems.

# Services pervasifs contextualisés : modélisation et mise en œuvre

Dejene Ejigu Dedefa

THESE

Soutenue le 12/12/2007

à l'INSA de Lyon

## Jury

Prof. Bruno DEFUDE INT (Paris)	rapporteur
Prof. Lionel BRUNIE (INSA de Lyon)	directeur
Dr. Marian SCUTURICI (INSA de Lyon)	co-directeur
Prof. Aris OUKSEL (Université d'Illinois, Chicago)	examineur
Prof. Jean-Marc PETIT (INSA de Lyon)	examineur
Dr. Thierry DELOT (Université de Valenciennes)	examineur
Dr. Richard CHBEIR (Université de Bourgogne)	examineur

Contact : Lionel.Brunie@insa-lyon.fr

## Résumé

Les systèmes pervasifs visent à intégrer des services fournis par des dispositifs répartis communicants. De tels environnements ont comme objectif d'optimiser l'interaction de l'utilisateur avec les dispositifs intégrés, par exemple en permettant à l'utilisateur d'accéder à l'ensemble des informations disponibles et en adaptant celles-ci aux conditions matérielles effectives (qualité de service réseau, caractéristiques du matériel de connexion). Cela impose aux applications d'adapter dynamiquement leur fonctionnement aux caractéristiques de l'environnement (notion de "contexte d'exécution").

Pour réaliser cette adaptation il est important de disposer d'un mécanisme efficace de capture et gestion du contexte et d'un mécanisme de raisonnement approprié. La gestion du contexte comprend la représentation, l'agrégation, l'interprétation, le stockage et le traitement des données contextuelles. Le raisonnement est le processus de déductions des nouveaux faits à partir des données contextuelles observées.

Dans cette thèse, nous proposons un modèle sémantiquement riche pour la collaboration, la représentation et la gestion du contexte. Nous utilisons un modèle de représentation du contexte fondé sur une approche hybride utilisant des ontologies et des bases de données relationnelles (nommé modèle HCoM : Hybrid Context Management model). Le modèle HCoM utilise l'ontologie pour la modélisation et la gestion des métadonnées riches en sémantique du contexte, et le schéma de la base de données relationnelles pour la modélisation et la gestion des données brutes du contexte. Les deux sont liés à travers des relations sémantiques construits dans l'ontologie. La séparation de ces deux éléments de modélisation nous permet d'extraire, charger, partager et utiliser seulement les données du contexte relevant afin de limiter la quantité de données dans l'espace de raisonnement.

Les éléments constitutifs du modèle HCoM sont les données contextuelles brutes, l'ontologie et les règles d'inférence. Ces éléments sont organisés dans un modèle que nous appelons EHRAM: Entité, Hiérarchie, Relation, Axiome et Métadonnée. EHRAM est mappé à un schéma de base de données relationnelle pour la représentation des données contextuelles et permet une représentation compatible avec les langages à bas de balises pour son ontologie et ses règles d'inférence. Cette richesse de modélisation nous permet de sélectionner de manière efficace les informations contextuelles pertinentes et ainsi d'améliorer les performances du processus de raisonnement mis en œuvre dans l'analyse du contexte d'exécution.

Nous présentons également la plateforme logicielle d'intégration de services pervasifs que nous avons développée (nommé CoCA : Collaborative Context-Aware service Plateform). Cette plateforme s'appuie sur la méthodologie et les modèles de représentation et de gestion du contexte proposés dans la thèse. Elle permet une interaction "contextualisée" des services fournis par les dispositifs participants, offrant en

particulier des mécanismes d'adaptation au contexte et de déclenchement proactif ou réactif de services en réponse à une évolution du contexte. Cette plateforme implémente le protocole JXTA dans ses composants de collaboration et utilise la librairie JENA pour le raisonnement (déclaration et interprétation des règles d'analyse du contexte).

Des démonstrateurs ont été développés et testés illustrant l'utilisation de la plate-forme dans trois cas d'utilisation liés à des domaines applicatifs variés : les réseaux sociaux, l'hôpital intelligent, l'adaptation d'IHM au contexte. Les résultats obtenus illustrent la performance, la robustesse et l'extensibilité de l'approche proposée.

*Mots clés :*

Informatique Pervasif, Réactivité au Contexte, Contexte Modélisation, Raisonnement Sémantique, Ontologie du Contexte, Informatique Collaborative.

## **Abstract**

Pervasive or ubiquitous computing aims to integrate computing and computing appliances into the environment rather than having computers as distinct objects. This can be realized through applications that adapt their behavior to every changing environment. Such systems need to ensure that the adaptive behavior experienced is useful, relevant, nondistracting and consistent with individual and organizational goals. Such adaptation needs proper capturing, management and reasoning of constantly changing context. Context capturing involves extracting relevant context data about selected entities in the environment. Context management deals with representation, aggregation, interpretation, storage and processing of context data. Context reasoning is the process of drawing inferences or conclusions (unknowns) from known facts using information from the various sources of context.

The computationally intensive characteristics of context reasoning process, the presence of handheld or wearable, tiny and resource hungry computing devices, and the lack of a semantically rich context model have been a bottleneck for the development of such applications. Moreover, most of the current context-aware systems are based on ad-hoc models of context, which causes lack of the desired formality and expressiveness. They do not separate processing of context semantics from processing and representation of context data and structure.

In this thesis, we propose a semantically rich and a collaborative context representation and management model that uses a hybrid of ontology and database management approaches (called HCoM model: Hybrid Context Management model). HCoM model uses ontology for modeling and management of context semantics and relational database schema for modeling and management of context data. These two modeling elements are linked to each other through the semantic relations built in the ontology. Separation of the two context modeling elements allows us to extract, load, share and use only relevant context data into the reasoner in order to limit the amount of context data in the reasoning space. By doing this, we considerably improve the performance of the reasoning process. The building blocks of the HCoM model are context data, context ontology, and deduction rules. These data elements are organized into a context representation structure (called EHRAM: Entity, Hierarchy, Relation, Axiom and metadata). EHRAM is a graphical context representation structure that serves as a context conceptualization model. EHRAM is mapped to a standard relational database schema for representation of its context component and is serialized to markup languages for representation of its ontology and rule component.

We also present a domain independent context-aware middleware platform (called CoCA: Collaborative Context-Aware service platform) under which our proposed context management model is implemented and used. CoCA uses data organized into the HCoM model as its data source and provides reasoning and decision services based on changing contexts. It triggers proactive and/or reactive actions and provides a collaboration interface between the pervasive peers. CoCA collaboration is based on JXTA protocols and its reasoning is based on Jena framework.

To evaluate the scalability and extensibility of the proposed model, reusability of the platform and performance of the collaboration process, we have developed a test case of the use of our context model in the platform using data from multiple scenarios: Community based network in a campus, smart hospital and adaptation of HCI to context. Results obtained from our experiment show that compared to other



related works in the domain, our approach gives a robust, extensible and scalable model and platform for the development of context-aware applications in pervasive environment.

*Keywords:*

Pervasive Computing, Context-Aware Computing, Context Modeling, Semantic Reasoning, Context Ontology, Collaborative Computing.

## **Publications**

- Hybrid Approach to Collaborative Context-Aware Service Platform for Pervasive Computing. E Dejene, V. Scuturici, L. Brunie. *Journal of Computers (JCP)* 3(1):40-50, Academy Publisher, ISSN 1796-203X. 2008.
- A Comprehensive Approach to Model and Use Context for Adapting Applications in Pervasive Environments. T Chaari, E Dejene, F. Laforest, V. Scuturici. *Int. Journal of Systems and software* 80(12):1973-1992, Elsevier. 2007.
- Semantic Approach to Context Management and Reasoning in Ubiquitous Context-Aware Systems. E Dejene, V. Scuturici, L. Brunie. Dans *The Second IEEE International Conference on Digital Information Management(ICDIM 2007)*, INSA, Lyon. pp. 500-5005. *Proceedings of ICDIM'07* . ISBN 1-4244-1476-8. 2007.
- A Collaborative Context-Aware Service Platform for Pervasive Computing. E Dejene, V. Scuturici, L. Brunie. Dans *Fourth Annual IEEE International Conference on Information Technology: New Generations, ITNG 07*, pp. 297-302, Las Vegas, USA, , . 2007.
- An Ontology-Based Approach to Context Modeling and Reasoning in Pervasive Computing. E Dejene, V. Scuturici, L. Brunie. *Fifth Annual IEEE International Conference on Pervasive Computing and Communications Workshops, PerComW 07*, pp.14-19, New York,USA, 2007.
- Modeling and Using Context in Adapting Applications to Pervasive Environments . T Chaari, E Dejene, F. Laforest, V. Scuturici. Dans *ICPS'06 : IEEE International Conference on Pervasive Services 2006 26-29 June 2006*, Lyon, . 2006.
- Positioning Support in Pervasive Environments. V. Scuturici, E Dejene. Dans *ICPS'06 : IEEE International Conference on Pervasive Services 2006 26-29 June 2006*, Lyon, . 2006.



# Découvertes de motifs pertinents par l'implémentation d'un réseau bayésien : application à l'industrie aéronautique

Fauré Clément

THESE

Soutenue le 20/11/2007

à l'INSA de Lyon

## Jury

Pr. Jean-François Boulicaut (INSA Lyon)	co-directeur
Dr. Jean Charlet (INSERM Paris)	rapporteur
Sylvie Delprat (Centre de Recherche EADS)	examineur
Pr. Bart Goethals (Université d'Anvers, Belgique)	examineur
Pr. François Jacquenet (Université de Saint-Etienne)	rapporteur
Pr. Alain Mille (Université Lyon 1)	co-directeur

Contact : clement.faure@gmail.com

## Résumé

L'étude d'un processus opérationnel se heurte souvent à l'analyse de données hétérogènes et volumineuses. L'environnement associé à ce processus évoluant constamment dans le temps, on constate inévitablement l'apparition de différences entre ce qui était attendu et ce qui est réellement observé. À partir des données recueillies et des connaissances du domaine, il faut alors être capable de déceler ces différences, et ainsi faire évoluer le modèle utilisé. Dans cette optique, nous proposons un processus de découverte de connaissances intégrant la définition et l'exploitation d'un réseau bayésien pour faciliter l'analyse d'un ensemble concis de règles d'association. L'évolution de ce modèle est pilotée par la découverte de règles pertinentes, elles-mêmes rendues plus accessibles par l'exploitation des propriétés de ce modèle. Enfin, nous montrons une application pratique de nos propositions au domaine des interruptions opérationnelles dans l'industrie aéronautique

## Abstract

The study of an operational process often runs up against the analysis of heterogeneous and large data. While the environment associated with this process evolves constantly, one inevitably notices the appearance of differences between what was expected and what is really observed. By using the collected data and available expertise, it is then necessary to detect these differences, and thus to update the model being used. Accordingly, we propose a knowledge discovery process that integrates the definition and the exploitation of a bayésian network to facilitate the analysis of a concise set of association rules. The evolution of this model is controlled by the discovery of relevant rules, themselves made more accessible by the exploitation from the properties of this model. Finally, we show a practical application of our proposals to the field of operational interruptions in the aircraft industry.

## Publications

- C. Faure, S. Delprat, J-F. Boulicaut, A. Mille. Iterative bayésian network implementation by using annotated association rules. Proc. 15th Int. Conf. on Knowledge Engineering and Knowledge Management EKAW'06, Podebrady, CZ, October 2006, Springer LNAI 4248, pp. 326-333.
- C. Fauré, S. Delprat, A. Mille, J-F. Boulicaut. Construction itérative d'un modèle de connaissance par l'exploitation de règles d'association. Actes 17ème Journées Francophones Ingénierie des Connaissances IC'06, Nantes (F), juin 2006, pp. 1-10.
- C. Fauré, S. Delprat, A. Mille, J-F. Boulicaut. Utilisation des réseaux bayésiens dans le cadre de l'extraction de règles d'association. Actes 6ème Journées Francophones Extraction et Gestion de Connaissances EGC'06, Lille, F, janvier 2006, Cepadues RNTI-E-6, pp. 569-580



# Découverte de correspondances sémantiques entre ressources hétérogènes dans un environnement coopératif

Ferreira da Silva Catarina

THESE

Soutenue le 07/12/2007

à l'Université Claude Bernard, Lyon 1

## Jury

Mme Chantal REYNAUD (Université de Paris XI)	présidente
Mme Corine CAUVET (Université Paul Cézanne Aix-Marseille 3)	rapporteur
M. Aris M. OUKSEL (Université d'Illinois, Chicago, USA)	rapporteur
M. Paulo RUPINO DA CUNHA, (Université de Coimbra, Portugal)	examineur
M. Djamel BENSLIMANE (Université Lyon 1)	examineur
Mme Parisa GHODOUS (Université Lyon 1)	directeur
M. Celson LIMA (Centre Scientifique et Technique du Bâtiment)	directeur
M. Lionel MÉDINI (Université Lyon 1)	directeur

Contact : alvescaterina@yahoo.com

## Résumé

Dans le travail coopératif les experts sont amenés à utiliser des ressources sémantiques. Ces ressources sémantiques, sont en général représentées à l'aide de concepts et de relations et peuvent être modélisées sous forme d'ontologies. Chaque expert dans un domaine d'application utilise fréquemment une ou plusieurs ressources sémantiques et peut également avoir accès à d'autres ressources sémantiques mises à disposition par ses partenaires. Cependant, ces ressources sémantiques sont hétérogènes entre elles à différents niveaux ce qui rend difficile l'interopérabilité syntaxique, structurelle et sémantique. Puisque les deux premiers types d'interopérabilité ont déjà été traités par ailleurs, ce travail de thèse met l'accent sur l'interopérabilité sémantique des ressources sémantiques.

Nous proposons une méthodologie semi-automatique pour aider à la découverte de correspondances sémantiques entre concepts de ressources sémantiques différentes. Pour réaliser ce travail, nous transformons ces ressources sémantiques en OWL-DL, afin d'utiliser les services de moteurs d'inférences fondés sur les logiques de description. Nous proposons aussi un procédé pour créer des équivalences entre relations de ressources sémantiques différentes, appelé unification de relations, et qui vise à découvrir d'autres correspondances sémantiques. Les résultats de ce travail ont été testés à l'aide d'un prototype. L'application a été réalisée avec des ressources sémantiques issues du secteur européen de la construction. Pour aider à l'évaluation des correspondances sémantiques, nous proposons aussi une modélisation des connaissances du domaine à l'aide de la logique floue.

## Abstract

In collaborative work experts are encouraged to use semantic resources. These resources are generally represented with the aid of concepts and relationships and can be modelled in the form of ontologies. Each expert in a field of application often uses one or more semantic resources, and may also have access to other semantic resources made available by its partners. However, these resources are heterogeneous at different levels making difficult syntactic, structural and semantic interoperability. In this work, we emphasize semantic interoperability.

We propose a methodology to help semiautomatic discovery of semantic mappings between concepts of different semantic resources. To accomplish this work, we transform these semantic resources in OWL-DL to use the services of inference engine based on description logics. We also propose a process for creating equivalence for relations between different semantic resources, called unification of relations; it aims to

discover other semantic mapping. The results of this work were tested by a prototype. The implementation was carried out with resources from within the industry European construction sector. To assist in the evaluation of semantic mappings, we also propose a domain knowledge modelling based on fuzzy logic

## Publications

- Using an Inference Engine to Detect Conflicts in Collaborative Design. M. Lima Dutra, C. Ferreira da Silva, P. Ghodous, R. Jardim-Gonçalves. Dans 14th International Conference on Concurrent Enterprising (ICE 2008), Lisbonne, Portugal, Juin 2008.
- Semantic Interoperability of Heterogeneous Semantic Resources. C. Ferreira da Silva, L. Médini, S. Abdul Ghafour, P. Hoffmann, P. Ghodous. *Electronic Notes in Theoretical Computer Science* 150(2):71-85, Elsevier. 2006.
- Improving Mappings Discovery for Semantic Interoperability. C. Ferreira da Silva, L. Médini, C. Lima, P. Ghodous. Dans 13th ISPE INTERNATIONAL CONFERENCE ON CONCURRENT ENGINEERING: RESEARCH AND APPLICATIONS, Leading the Web in Concurrent Engineering, Antibes, France. pp. 227-234. *Frontiers in Artificial Intelligence and Applications* 143. IOS Press Nieuwe Hemweg 6B 1013 BG Amsterdam Nethe. ISBN 1-58603-651-3. ISSN 0922-6389. 2006.
- Assessing the quality of mappings between Semantic Resources in Construction. C. Lima, C. Ferreira da Silva, J.P. Pimentão. Dans *Intelligent Computing in Engineering and Architecture - 13th EG-ICE Workshop 2006*, Ian F.C. Smith ed. Ascona, Switzerland. pp. 416-427. *Lecture Notes in Computer Science, Lecture Notes in Artificial Intelligence* 4200. Springer Berlin / Heidelberg <http://www.springerlink.com/content/978-3-540-46246-0/>. ISBN 978-3-540-46246-. ISSN 3-540-46246-5. 2006.
- ENHANCING SEMANTIC INTEROPERABILITY AMONG SEMANTIC RESOURCES FOR CONSTRUCTION. C. Ferreira da Silva, C. Lima, L. Médini, P. Ghodous, A. Zarli. Dans *Joint International Conference on Computing and Decision Making in Civil and Building Engineering*, Hugues Rivard (Canada), Hani Melhem (USA), Edmond Miresco (Canada) ed. Montreal, Canada. pp. 2487-2496. ISBN 2-921145-58-8. 2006.
- Interoperability among Semantic Resources in Construction: Is it Feasible?. C. Lima, C. Ferreira da Silva, P. Sousa, J.P. Pimentão. Dans 22nd CIB-W78 Conference Information Technology in Construction, Raimar J. Scherer, Peter Katranuschkov, Sven-Eric Schapke ed. Dresde, Allemagne. pp. 285-292. CIB-w78 -304. CIB publications Institute for Construction Informatics, Technische Universität Dresden, Germany. ISBN 3-86005-478-3. 2005.
- Towards a framework for managing standards-base semantic e-Resources in the European Construction Industry. C. Lima, G. Storer, A. Zarli, C. Ferreira da Silva. Dans *Construction Research Congress 2005 BROADENING PERSPECTIVES (CRC 2005)*, Iris D. Tommelein ed. San Diego, California, EUA. pp. 1-10. American Society of Civil Engineers ASCE International Headquarters 1801 Alexander Bell Drive Reston, VA 20191-4400 USA. ISBN 0-7844-0754-1. 2005.
- Interoperability of heterogeneous semantic resources. C. Ferreira da Silva, L. Médini, S. Abdul Ghafour, P. Hoffmann, P. Ghodous. Dans *International Workshop on Database Interoperability (InterDB 2005) in connection with the 7th International Conference on Coordination Models and Languages*, Philippe Thiran, Tore Risch, Djamal Benslimane ed. Namur, Belgique. pp. 74-89. 2005.
- A Framework to Support Interoperability Among Semantic Resources. C. Lima, C. Ferreira da Silva, C. Le Duc. Dans *First International Conference on Interoperability of Enterprise Software and Applications (INTEROP-ESA'2005)*, Dimitri Konstantas, Jean-Paul Bourrières, Michel Léonard, Nacer Boudjlida ed. Genève, Suisse. pp. 87-98. Springer-Verlag London . ISBN 1-84628-151-2. 2005.
- Setting up the Open Semantic Infrastructure for the European Construction Sector - the FUNSIEC Project. C. Lima, B. Fiès, C. Ferreira da Silva, S. Barresi. Dans *European Conferences on Product and Process Modeling in the Building Industry (ECPM)*, Istanbul, Turquie. 2004.

# Modèles et outils génériques pour la résolution des problèmes liés à la répartition de ressources sur grille

THESE

Gossa Julien

Soutenue le 03/12/2007

à l'INSA de Lyon

## Jury

Michelle Sibilla, IRIT Toulouse	rapporteur
Franck Cappello, INRIA LRI Orsay	rapporteur
Nouredine Melab, LIFL Lille	
Thomas Ludwig, Ruprecht-Karls-Universität Heidelberg	
Jean-Marc Pierson, IRIT Toulouse	co-directeur
Lionel Brunie, INSA Lyon	co-directeur

Contact : lionel.brunie@liris.cnrs.fr

## Résumé

Depuis plusieurs années, les intergiciels de grille n'ont eu de cesse de gagner en complexité. Grâce à l'adoption des SOA (architectures orientées services), ils permettent aujourd'hui de concevoir des architectures logicielles complexes, de très haut niveau et naturellement distribuées. Les acteurs (administrateurs, développeurs et utilisateurs) sont ainsi quotidiennement confrontés à de nombreux problèmes liés à la nature répartie des ressources : Comment sélectionner une ou plusieurs ressources pour s'acquitter d'une tâche donnée ? Comment déployer une ou plusieurs ressources, voire toute une architecture logicielle ? Comment composer plusieurs ressources ? etc.

La résolution de ces problèmes est difficile car les acteurs travaillent à un niveau très élevé, depuis lequel l'infrastructure matérielle est complètement abstraite. Elle s'avère également particulièrement complexe du fait de la diversité des ressources (du simple fichier à un service web impliquant communications et calculs), de la diversité des objectifs (performances, équilibrage de charge, qualité des réponses, aspects financiers...), de la prise en compte des performances de l'infrastructure et également de la diversité des problèmes (sélection, composition, déploiement...). Or cette résolution est cruciale car elle conditionne les performances des plateformes aussi bien que la quantité de travail des acteurs.

Notre approche propose un service web globus appelé le Network Distance Service. NDS permet de déclarer intuitivement une vaste étendue des problèmes sous forme de graphe grâce à une notion de distance adaptable aux ressources et objectifs et adaptative aux performances de l'infrastructure matérielle. Des algorithmes classiques de la théorie des graphes sont ensuite utilisés pour calculer les solutions : les plus courts chemins sont utilisés pour résoudre les problèmes de sélection et de composition ; un algorithme de clustering est utilisé pour résoudre les problèmes de déploiement.

Deux algorithmes originaux ont également été développés grâce à cette approche : FReDi, capables de piloter dynamiquement des réplicas, et MRKM, capable de prendre les décisions inhérentes au déploiement d'une architecture logicielle complète.

Enfin, un entrepôt de données a été conçu dans le but d'améliorer les capacités d'analyse des outils de surveillance.

Notre approche est illustrée par l'architecture logicielle GGM, ses différentes infrastructures matérielles cibles et les différents problèmes de distribution rencontrés. Des expérimentations grande échelle ont été menées sur la plateforme Grid5000 et montrent l'obtention des solutions optimales pour des temps de calcul très courts et un travail utilisateur limité.

## Mots clés :

Grille, Architectures Orientées Services (SOA), surveillance (monitoring), problèmes liés à la répartition

(sélection, déploiement, composition)

## Abstract

Since many years, the complexity of grid middlewares never ceased to increase. Thanks to the adoption of SOA (Services Oriented Architectures), they allow to design highly complex, high level, and naturally distributed software architectures. The actors (administrators, developers and users) are thus confronted to many problems related to the distributed nature of resources : How to select one replicated resource to fulfill one given task? How to deploy one or several resources, or even a complete software architecture? How to compose several resources ? etc.

Solving these problems is difficult as actors are working at the highest level, from which the infrastructure is fully abstracted. Another difficulty is due to the diversity of resources (from simple files to complex web services implying communications as well as computations), to the diversity of goals (performances, load balancing, response quality, financial aspects... ), to the infrastructure performances and to the diversity of problem (selection, deployment, composition...). However, such solving is crucial as it conditions the platform performances as well as the quantity of actors work.

Our proposal is a Globus web services called the Network Distance Service (NDS). NDS allows to model intuitively a large variety of problems as graphs thanks to a notion of distance adaptable to resources and goals and adaptive to infrastructure performances. Classical graph theory algorithms are then used to solve the problems : a shortest-path algorithm is used to solve selection and composition problems ; a clustering algorithm is used to solve deployment problem.

Two genuine algorithms have also been developed thanks to this approach : FReDi, in order to drive dynamically some replicas, and MRKM, in order to deploy a whole software architecture.

Finally, one data warehouse has been designed in order to improve the analysis capabilities of monitoring systems.

Our approach is illustrated by the GGM software architecture, its different target infrastructures and the different distribution problem encountered. Large scale experiments have been conducted on the Grid5000 platform and shows that optimal solutions have been computed in very short solving times and with limited user work.

### *Keywords:*

grid, Services Oriented Architectures (SOA), monitoring, distribution problems (selection, deployment, composition)

## Publications

### *Revue internationale*

- Jean-Marc Pierson, Julien Gossa, Pascal Wehrle, Yonny Cardenas, El Samad Mahmoud, Cahon Sebastien, Lionel Brunie, Clarisse Dhaenens, Hameurlain Abdel kader, Nouredine Melab, Maryvonne Miquel, Franck Morvan, Talbi El gazali, and Anne Tchounikine. Ggm efficient navigation and mining in distributed genomedical data. IEEE Transactions on NanoBioscience, May 2007.
- Mahmoud El Samad, Julien Gossa, Franck Morvan, Abdelkader Hameurlain, Jean-Marc Pierson, and Lionel Brunie. Monitoring service for large scale dynamic query optimization in grid environment. International Journal of Web and Grid Services (IJWGS), 2007.

### *Chapitres d'ouvrages*

- Julien Gossa and Sandro Bimonte. Gr-olap : On line analytical processing of grid monitoring information, August 2007. Encyclopedia of Database Technologies and Applications, Second Edition, Eds. Laura C. Rivero, Jorge Horacio Doorn, Viviana E. Ferraggine. IDEA Group Inc.
- Julien Gossa and Jean-Marc Pierson. Nds : Network distance service. In Extraction et Gestion Parallèles Distribuées de Connaissances (EGPDC 06), January 2006.

### *Conférences internationales*

- Julien Gossa, Jean-Marc Pierson, and Lionel Brunie. Supporting decision-making in pervasive grids with



advanced network distances. In LNCS, editor, The 15th International Euro-Par Conference European Conference on Parallel and Distributed Computing, 2007.

- Julien Gossa, Jean-Marc Pierson, and Lionel Brunie. Fredi: Flexible replicas displacer. In IEEE CS Press, editor, 5th International Conference on Networking, April 2006.
- Julien Gossa, Jean-Marc Pierson, and Lionel Brunie. Adapted distance-based decision-making support, a mandatory tool to make the grid pervasive. In IEEE, editor, IEEE International Conference on Pervasive Services (ICPS'07), July 2007.
- Julien Gossa. End-to-end distance computation in grid environment by nds, the network distance service. In IEEE, editor, 4th European Conference on Universal Multiservice Networks (ECUMN'07), February 2007.

#### *Atelier international*

- Julien Gossa. Evaluation of network distances properties by nds, the network distance service. In Create-Net IEEE, editor, GridNets '06, October 2006.

#### *Conférences nationales*

- Julien Gossa, Jean-Marc Pierson, and Lionel Brunie. (dé)placement dynamique de réplicas dans un environnement pervasif. In Iere Journées Francophones Mobilité et Ubiquité, UbiMob '04, 2004.
- Julien Gossa and Romuald Thion. Gestion de la sécurité des utilisateurs mobiles: une adaptation du filtrage de paquet. In ACM, editor, Se Journées Francophones Mobilité et Ubiquité, UbiMob '06, pages 119-122, September 2006.



# Approximation Sémantique des Requêtes dans les Bases de Données Multimédia

**Hammiche Samira**

**THESE**

Soutenue le 15/06/2007

à l'Université Claude Bernard Lyon1

## Jury

Thèrèse Libourel, Université Montpellier II	rapporteur
Mohand Boughanem Université ,Paul Sabatier, Toulouse	rapporteur
Omar Boucelma, Université Aix-Marseille 3 .	examineur
Jean-Marie Pinon, INSA de Lyon .	examineur
Mohand-Said Hacid, Université Lyon1 .	directeur de recherche
Salima Bennernou, Université Lyon 1	co-directrice de thèse

Contact : mohand-said.hacid@liris.cnrs.fr

## Résumé

La disponibilité croissante des données multimédia sur le Web a fait émerger plusieurs préoccupations relatives à leur représentation et à leur exploitation. Les données multimédia sont caractérisées par leur complexité structurelle (en termes de médias et d'unités logiques) et par la subjectivité de l'interprétation assignée. Par conséquent, il est difficile d'avoir des réponses précises et complètes aux requêtes des utilisateurs. Pour résoudre ces problèmes, il est nécessaire de décrire le contenu des données multimédia et de définir des stratégies de recherche adéquates basées sur des langages de requêtes sur le contenu sémantique.

Dans notre travail de thèse, nous nous intéressons à la description et à l'interrogation efficace de données multimédia. Pour permettre une description riche de ces données, nous utilisons le standard MPEG-7, qui permet une description compréhensible et riche du contenu multimédia. Cependant, comme MPEG-7 utilise XML Schema pour définir ses structures de description, cela lui vaut certaines limites dont la manque d'une sémantique formelle dans MPEG-7, le manque de moyens d'intégrer des connaissances du domaine et le manque d'un mécanisme d'inférence pour déduire les informations implicites. Afin de résoudre ces problèmes, nous proposons d'étendre MPEG-7 d'un niveau conceptuel où des connaissances du domaine sont définies pour associer un cadre sémantique au vocabulaire MPEG-7, et permettre l'intégration des connaissances du domaine dans les descriptions MPEG-7. Les connaissances sont formalisées en utilisant un langage logique hybride, qui combine le pouvoir expressif des logiques de description et des règles Datalog.

Concernant l'interrogation de données multimédia décrites moyennant le standard MPEG-7, nous utilisons le standard d'interrogation des documents semi-structurés, XQuery. Cependant, comme ce dernier est complexe à utiliser pour un utilisateur novice et qu'il n'est doté d'aucun mécanisme d'inférence, nous exploitons le formalisme des règles, utilisé pour formaliser les connaissances du niveau conceptuel, comme support flexible pour la formulation des requêtes. Ainsi, les requêtes sont formulées en utilisant une conjonction de concepts, de relations et d'attributs, puis elles sont traitées et réécrites en plusieurs requêtes sémantiquement équivalentes. La réécriture se base sur les règles d'inférence et le raisonnement de subsomption fournis par le formalisme logique. Les requêtes réécrites sont par la suite transformées en XQuery, puis évaluées sur la base des descriptions MPEG-7. La transformation se fait moyennant un ensemble de règles de correspondance ainsi que des algorithmes de transformation, que nous définissons.

## Publications

### *Conférences Internationales*

- S. Hammiche, B. Lopez, S. Benbernou, M-S Hacid, A. Vakali : "Domain Knowledge Based Queries for

Multimedia Data Retrieval". 4th special Workshop on Multimedia Semantics (WMS 06). Chania, Crete, Greece, June 2006.

- S. Hammiche, S. Benbernou, A. Vakali : "A Logic Based Approach for the Multimedia Data Representation and Retrieval" . 7th IEEE International Symposium on Multimedia (ISM'05). Irvine, California, USA. December 12- 14, 2005.
- S. Hammiche, S. Benbernou, M-S Hacid and A. Vakali, "Semantic Retrieval of Multimedia Data". In proceeding of the Second ACM International Workshop on Multimedia Databases, pp 36-44. Washington DC. USA. November 13, 2004.

#### *Conférences Nationales*

- B. Lopez, S. Hammiche, S. Sebahi, M-S Hacid : "Annotation et Navigation de Données Archéologiques". 7ème Journées d'Extraction et Gestion de Connaissances (EGC 2007). Namur, Belgique. 23 – 26 Janvier 2007.
- S. Hammiche : "Interrogation de données multimédias fondée sur les connaissances du domaine". 24ème congrès INFORSID'06 (Informatique des Organisations et Systèmes d'Information et de Décision). Hammamet. Tunisie. 1- 3 Juin 2006. Papier accepté en tant que Papier Jeune Chercheur.
- S. Hammiche, S. Benbernou et M-S Hacid, "Réécriture de requêtes multimédias : approche basée sur l'usage d'une ontologie". 5ème Journées d'Extraction et Gestion de Connaissances (EGC 2005), pp 389-394. Paris. France. 19-21 Janvier 2005.
- S. Hammiche, "Appariement d'arbres pour l'interrogation de données multimédias". 22ème congrès INFORSID'04 (Informatique des Organisations et Systèmes d'Information et de Décision). Forum jeunes chercheurs, pp 547-548. Biarritz. France. 25-28 Mai 2004.

# Conception Coopérative d'Ontologies Pré-Consensuelles : Application au domaine de l'Urbanisme

**Keita Abdel Kader**

**THESE**

Soutenue le 06/06/2007

à l'INSA de Lyon

## Jury

ZIMMERMANN Monique ( INSA de Lyon)	présidente
BOUCELMA Omar (Université Aix-Marseille 3)	rapporteur
LAS CASAS Giuseppe (Université de la Basilicate)	rapporteur
LAURINI Robert (INSA de Lyon)	directeur
LIBOUREL Thérèse (Université Montpellier II)	rapporteur
ROUSSEY Catherine (Université Lyon I)	co-directrice

Contact : akakeita@yahoo.fr

## Résumé

Une ontologie possède deux définitions différentes suivant le domaine auquel on s'intéresse, en l'occurrence la philosophie ou l'informatique :

-Du point de vue philosophique, une ontologie se définit comme la science de ce qui existe.

- Du point de vue informatique et plus particulièrement ingénierie des connaissances, la définition la plus communément admise est celle de Studer [STUD98] : "Une ontologie est une spécification formelle, explicite d'une conceptualisation partagée".

La majorité des travaux et des outils existants concernent la création des ontologies, une fois le consensus réalisé (que nous intitulons ontologies post-consensuelles).

Ainsi, une des étapes clés (importantes) dans le processus de développement d'une ontologie est le consensus entre les différents acteurs. Mais le problème principal est comment atteindre le consensus ?

L'une des caractéristiques essentielles des ontologies est qu'elles fournissent des données consensuelles sur un domaine donné. De nos jours, les ontologies englobent un ensemble si riche de la connaissance que leur compréhension (complète) dépasse celle de n'importe quel développeur ou concepteur seul ou même d'une petite équipe de concepteurs. Le développement d'une ontologie à grande échelle doit être le fruit d'un effort commun de plusieurs experts de domaine et concepteurs de logiciels. Ainsi l'approche collaborative pour le développement d'ontologie est la mieux adaptée pour cette tâche par rapport aux autres approches non collaboratives, vues les caractéristiques mêmes d'une ontologie.

Cependant la plupart des méthodes et méthodologies ne proposent pas de directives pour atteindre le consensus quand les ontologies sont construites de manière collaborative.

La construction des ontologies dans un environnement collaboratif a été et demeure un sujet de recherche continu et ouvert. Il existe un certain nombre d'outils pour le développement collaboratif d'ontologies. Mais aucun n'offre un support complet et adaptable pour toutes les exigences du développement collaboratif.

L'analyse des outils et environnements existants (pour le développement collaboratif des ontologies) ne vérifie pas seulement le fait que le développement collaboratif est la nature inhérente de la conception d'ontologie, mais indique aussi que le développement collaboratif d'ontologie n'est bien soutenu par aucun des outils ou environnements existants. Ceci présente un nouveau challenge pour trouver des outils mieux adaptés pour le développement collaboratif d'ontologie.

L'inexistence (en notre connaissance) d'ontologie pour l'urbanisme et l'aménagement du territoire, ainsi

que la complexité et la dimension du domaine de l'urbanisme ajoute à ce challenge un intérêt particulier.

#### Contributions

Au contraire de la majorité des travaux et des outils existants qui concernent la création des ontologies, une fois le consensus réalisé (ontologies post-consensuelles), nos travaux, réalisés dans le cadre du projet européen *Towntology*, se situent dans les premières phases de développement d'une ontologie, lorsque le consensus n'est pas encore atteint. Ce travail essaye de répondre à la question d'atteindre le consensus en proposant une méthodologie pour créer une ontologie pré-consensuelle en partant des opinions de divers acteurs. Par exemple, nous insistons sur la nécessité de garder différentes définitions pour le même concept.

Plus exactement, nous proposons :

- Une méthodologie de développement d'ontologie pré-consensuelle et une démarche permettant lors de la phase d'acquisition des connaissances à des experts du domaine (dans notre cas des urbanistes) de modéliser les différences de points de vue sur le sens des termes. Notre méthodologie fait intervenir les experts du domaine le plus longtemps possible en leur permettant d'aller vers le consensus. Nous proposons une extension de XML pour stocker notre ontologie.
- Un outil visuel permettant d'éditer, de naviguer dans l'ontologie et de l'interroger.
- Un support pour le développement collaboratif d'ontologie proposant l'ontologie créée comme ontologie d'ancrage pour le travail collaboratif.

#### *Mots clés :*

Ontologies, conception coopérative des ontologies, *Towntology*, Groupware, pré-consensus, urbanisme.

#### **Abstract**

An ontology has two different definitions according to the field of interest, in fact philosophy or computer science:

- From the philosophical point of view, an ontology is defined as being the science of what exists.
- From the point of view of computer science and more particularly, knowledge engineering, the definition most commonly accepted is that of (Studer, 1998) : "an ontology is an explicit formal specification of a shared conceptualization of a field ".

The majority of works and the existing tools relate to the creation of ontologies, once the consensus is carried out (which we entitle as post-consensual ontologies).

Thus, one of the key stages in the process of developing an ontology is the consensus between the various actors. But the principal problem is how to reach this consensus.

One of the essential characteristics of ontologies is that they provide consensual data on a given field. Nowadays, ontologies include a so rich set of knowledge that their full comprehension exceeds that of any single developer or creator or even that of a small team of creators. The development of an ontology on a large scale must be the result of a joint effort from several experts in various fields and software developers. So the collaborative approach for the development of ontology is best adapted for this task compared to the other non collaborative approaches, because of the characteristics, themselves, of an ontology.

However the majority of the methods and methodologies do not propose recommendations to reach the consensus when ontologies are built with a collaborative way.

The construction of ontologies with a collaborative environment was and remains a continuous and open subject of research. There exists a certain number of tools for the collaborative development of ontologies. But none offers a complete and adaptable support for all the requirements of the collaborative development.

The analysis of the existing tools and groupware products (for the collaborative development of ontologies) does not check only the fact that the collaborative development is the inherent nature of the design of ontology, but also states that the collaborative development of ontology is well supported by none of the existing tools or environments. This presents a new challenge to find tools better adapted for the

collaborative development of ontology.

The inexistence (as we know) of ontology for urban planning and regional planning, as well as the complexity and the dimension of the domain of urban planning add a particular interest to this challenge.

#### Contributions

In contrast of the majority of research works and existing tools which relate to the creation of ontologies once the consensus is carried out (post-consensual ontologies), our research works, produced in the context of the European project TOWNTOLLOGY, focus on the first phases of development of an ontology, when the consensus is not yet reached. This work tries to answer the question of reaching the consensus by proposing a methodology to create a pre-consensual ontology on the basis of the opinions of various actors. For example, we insist on the need for keeping various definitions for the same concept.

We propose more exactly:

- A methodology of development of pre-consensual ontology and a step allowing the experts of the field (in our case, the urban planners) to model the differences in the various points of view about the meaning of the terms during the phase of knowledge acquisition. Our methodology makes the experts of the field intervene as long as possible by enabling them to tend to some consensus. We propose an extension of XML to store our ontology.
- A visual tool allowing to edit, to navigate in ontology and to query it.
- A support for the collaborative development of ontology proposing the designed ontology as an anchoring ontology for collaborative work.

#### *Keywords:*

ontologies, coopérative development of ontologies, Towntology, Groupware, pré-consensus, urban planning

#### **Publications**

- KEITA A.K., LAURINI R., ROUSSEY C., ZIMMERMAN M. "Towards an Ontology for Urban Planning: The Towntology Project". In Proceedings of the 24th UDMS Symposium, Chioggia, October 27-29, 2004, pp 12.-24.
- ROUSSEY C., KEITA A.K., LAURINI R., BEAULIEU C., TARDY Y. ZIMMERMANN M. (2004) "Le projet Towntology : Un retour d'expérience pour la construction d'une ontologie urbaine". "Revue Internationale de Géomatique", vol. 14, 2, 2004 pp. 217-237
- TELLER J., KEITA A.K., ROUSSEY C., LAURINI R. "Urban Ontologies for an Improved Communication in Urban Civil Engineering Projects". At International Conference on Spatial Analysis and GEOmatics, Research & Developments (SAGEO 2005), Avignon, France, June, 20th-23rd, 2005.
- KEITA A. K., C. ROUSSEY, R LAURINI. "Un outil d'aide à la construction d'ontologies pré-consensuelles : le projet Towntology". In actes du 24ème congrès de Informatique des Organisations et Systèmes d'Information et de Décision (INFORSID), Tunis 31 Mai-4 Juin 2006, pp 911-926.





# Evolution de second ordre et algorithmes évolutionnaires : l'algorithme RBF-Gene

Lefort Virginie

THESE

Soutenue le 05/07/2007

à l'INSA de Lyon

## Jury

Dr. Guillaume Beslon (INSA Lyon)	co-directeur
Dr. Steffen Christensen (Université de Carleton, Canada)	examineur
Pr. Philippe Collard (Université de Sophia-Antipolis)	rapporteur
Pr. Joël Favrel (INSA Lyon)	co-directeur
Dr. Evelyne Lutton (INRIA Rocquencourt)	rapporteur
Pr. José Maria Peña (Universidad politécnica de Madrid, Espagne)	examineur

Contact : guillaume.beslon@iris.cnrs.fr

## Résumé

On parle d'évolution de second ordre (ou de sélection indirecte) lorsque les individus sont sélectionnés non pour leur seule adaptation à l'environnement mais aussi pour leur capacité à évoluer « mieux ». Bien qu'un tel mécanisme soit a priori très intéressant en évolution artificielle, la structure des algorithmes évolutionnaires interdit généralement celui-ci car les processus évolutifs sont figés. Nous avons ainsi proposé un nouvel algorithme évolutionnaire, RBFGene. Il possède un niveau intermédiaire, le protéome (composé de « protéines »), entre le phénotype d'un individu et son génotype lui permettant de faire varier la structure du génome sans modifier son phénotype sachant que ces variations auront une influence sur les reproductions futures. Nous montrons qu'une sélection de second ordre est bien à l'œuvre dans l'algorithme et qu'elle permet de façonner les génomes, en modifiant les tailles des zones non codantes et l'ordre des gènes.

## Abstract

Second order evolution (or indirect selection) corresponds to a situation where the individuals are not only selected on their fitness to an environment, but also on their ability to evolve « better ». Even if such a mechanism seems a priori very interesting in artificial evolution, it is not permitted by the structure of evolutionary algorithms because the evolutionary processes are fixed. Therefore, we propose a new evolutionary algorithm, RBFGene. It includes an intermediate level, the proteom (made of « proteins »), between the phenotype of an individual and its genotype, that allows for changes in the structure of the genome without changing the phenotype. These modifications can thereafter have an influence on later reproductions. We show the existence of an indirect selection in our algorithm, acting on genomes by changing the size of the non coding sequences or the order of the genes.

## Publications

- W. Banzhaf, G. Beslon, S. Christensen, J. Foster, F. Képès, V. Lefort, J. Miller, M. Radman, J. Ramsden. >From artificial evolution to computational evolution: a research agenda. *Nature Review Genetics* 7(9):729-735, September 2006.
- V. Lefort, C. Knibbe, G. Beslon, J. Favrel. A bio-inspired genetic algorithm with a self-organizing genome: The RBF-Gene model. *Proceedings Genetic and Evolutionary Computation GECCO'04 Part II*, Springer LNCS 3103, pp. 406-407
- C. Knibbe, V. Lefort, F. Chaudier, G. Beslon, J-M Fayard. Self adaptation of genome size in artificial organisms. *Proceedings 8th European Conference in Artificial Life ECAL'05*, Springer LNAI 3630, pp. 423-432.

- V. Lefort, C. Knibbe, G. Beslon, J. Favrel. Simultaneous optimization of weights and structure of an RBF Neural Network. Proc. 7th Int. Conf. Artificial Evolution EA'05 Revised and Selected Papers, Springer LNCS 3871, pp. xxx-xxx.
- V. Lefort, C. Knibbe, G. Beslon, J. Favrel. Introducing proteins into genetic algorithms. Proceedings Complex Systems, Intelligence and Modern Technology Applications CSIMTA'04, pp. 181-186.
- V. Lefort, G. Beslon. Optimisation de réseaux de neurones RBF pour la régression via un algorithme évolutionnaire. Actes 8ème Conférence Extraction et Gestion des Connaissances EGC'08, RNTI, Cépaduès. pp. 331-336.
- C'06, Lille, F, janvier 2006, Cepadues RNTI-E-6, pp. 569-580.

# Une modélisation évolutionniste du liage temporel : A evolutionary model of temporal binding

THESE

Meunier David

Soutenue le 19/10/2007

à l'Université Lumière, Lyon 2

## Jury

Docteur Hugues BERRY	examineur
Docteur Olivier BERTRAND	examineur
Docteur Agnès GUILLOT	examineur
Professeur Hélène PAUGAM-MOISY	directeur
Professeur Marc SCHOENAUER	rapporteur
Docteur Catherine TALLON-BAUDRY	rapporteur

Contact : helene.paugam-moisy@liris.cnrs.fr

## Résumé

L'hypothèse du liage temporel par synchronie suscite un intérêt important en neurobiologie, car elle permet d'expliquer comment différentes structures du cerveau peuvent établir entre elles un lien fonctionnel, en rapport avec une fonction cognitive. Cependant, il n'existe pas de modèle permettant de faire communiquer différents groupes de neurones par le biais de leurs émissions.

Nous avons développé un modèle de réseau de neurones impulsionnels, dont la topologie est modifiée par un algorithme évolutionniste. Le critère de performance utilisé pour l'algorithme évolutionniste est évalué par l'intermédiaire du comportement d'un individu contrôlé par un réseau de neurones impulsionnels, et placé dans un environnement virtuel. L'utilisation du neurone impulsionnel, ayant la propriété de détection de synchronie, oblige l'évolution à construire un système utilisant cette propriété au niveau global, d'où l'émergence de la synchronisation neuronale à large-échelle. Les propriétés topologiques et dynamiques du réseau de neurones ne sont pas prises en compte dans le calcul de la performance, mais sont étudiées a posteriori, en comparant les individus avant et après évolution.

D'une part, grâce aux outils de la théorie des réseaux complexes, nous montrons l'émergence d'un certain nombre de propriétés topologiques, notamment la propriété de réseau "petitmonde". Ces propriétés topologiques sont similaires à celles observées au niveau de l'anatomie des systèmes nerveux en biologie. D'autre part, au niveau de la dynamique, nous établissons que la propriété de synchronisation neuronale à large-échelle, résultant de la présentation d'un stimulus, est présente chez les individus évolués. Pour ce faire, nous nous appuyons sur les outils classiquement utilisés en électrophysiologie, et nous les étendons pour pouvoir interpréter la grande quantité de données obtenue à partir du modèle.

Le modèle montre que l'on peut construire des réseaux de neurones basés sur l'hypothèse du liage temporel en ayant recours à l'évolution artificielle, en se basant sur un critère de performance écologique, c.à.d. le comportement de l'individu dans son environnement. D'autre part, les outils développés pour l'analyse des propriétés du modèle peuvent être utilisés dans d'autres domaines, en premier lieu en électrophysiologie. En effet, à cause des progrès techniques sur les enregistrements électrophysiologiques, la quantité de données se rapproche singulièrement de celle issue du modèle.

## Abstract

The hypothesis of temporal binding by synchrony raises a major challenge in neurobiology, because it explains how different cerebral areas can establish a functional relation between each others, linked to a cognitive process. However, there is no model allowing different groups of neurons to communicate by the mean of their emissions.

We have developed a model of spiking neuron network, where the topology is modified by an

evolutionary algorithm. The fitness criterion used for the evolutionary algorithm is evaluated by the mean of the behaviour of an individual, controlled by a spiking neuron network, and placed in a virtual environment. The use of a spiking neuron model, with the property of synchrony detection, compels evolution to build a system using this property at a global level, and results in the emergence of large-scale neural synchronisation. Topological and dynamical properties are not taken into account in the fitness computation, but are studied a posteriori by comparing individuals before and after evolution.

On the one hand, with the help of complex network tools, we show the emergence of several topological properties, including "small-world" network properties. These topological properties are similar to those obtained from the anatomy of biological nervous systems. On the other hand, at the dynamical level, we show that the property of large-scale neural synchronisation, induced by the presentation of a stimulus, is found for evolved individuals. To show this result, we use some tools classically used in electrophysiology, and extend these tools in order to analyze the big quantity of data obtained from the model.

The model shows that it is possible to build artificial neural networks based on the hypothesis of temporal binding, by using an ecological fitness criterion (the behaviour of an individual in a virtual environment) for artificial evolution. The tool extended for analyzing the model properties can be used in other domains, including electrophysiology. The quantity of data in this domain and those obtained with the model are indeed similar, because of recent technical innovations in electrophysiological recordings.

## Publications

### *Revue nationale avec comité de lecture :*

- Meunier, D., Paugam-Moisy, H. (2005) "Simulation d'un protocole d'amorçage intermodal sur un réseau de neurones impulsifs". *Revue d'Intelligence Artificielle*, 19(1-2) :375-388, HERMES - Lavoisier.

### *Conférences internationales :*

- Meunier, D., Paugam-Moisy, H. (2008) "Neural Networks and Computational Neuroscience : A tutorial". ESANN'2008, Europ. Symp. on Artificial Neural Networks, 367-378, M. Verleysen Ed., Bruges, BELGIQUE.
- Meunier, D., Paugam-Moisy, H. (2006) "Cluster detection algorithm in neural networks". ESANN'2006, Europ. Symp. on Artificial Neural Networks, 19-24, M. Verleysen Ed., Bruges, BELGIQUE.
- Meunier, D., Paugam-Moisy, H. (2005) "Evolutionary supervision of a dynamical neural network allows learning with on-going weights". IJCNN'2005, Int. Joint Conf. on Neural Networks. 1493-1498, IEEE, INNS. Montréal, CANADA.
- Meunier, D., Paugam-Moisy, H. (2004) "A "spiking" Bidirectional Associative Memory for modeling intermodal priming". NCI'2004, Int. Conf. on Neural Networks and Computational Intelligence. 25-30, ACTA Press. Grindelwald, SUISSE.

### *Conférences nationales :*

- Meunier, D., Martinez, R., Paugam-Moisy, H. (2006) "Assemblées Temporelles dans les Réseaux de Neurones Impulsifs". NeuroComp'06, 1ère de conférence française de neurosciences computationnelles, 185-189, Pont-à-Mousson, FRANCE.
- Meunier, D., Martinez, R., Paugam-Moisy, H. (2006) "Emergence de bandes de fréquences dans un réseau de neurones impulsifs". NeuroComp'06, 1ère de conférence française de neurosciences computationnelles, 187-190, Pont-a-Mousson, FRANCE.

# Médiation Sémantique Orientée Contexte pour la composition de Services Web

Mrissa Michaël Nicolas

THESE

Soutenue le 15/11/2007

à l'Université Claude Bernard, Lyon 1

## Jury

Nadine Cullot

Khalil Drira

Marie-Christine Fauvet

Chirine Ghedira

Djamal Benslimane

Zakaria Maamar

directrice

directeur

Contact : michael.mrissa@iut.univ-lyon1.fr

## Résumé

L'adoption des services Web constitue une avancée majeure dans le développement des systèmes d'information interopérables. En particulier, la composition de services permet de répondre aux besoins de plus en plus complexes des utilisateurs, par la combinaison de plusieurs services Web au sein d'un même processus métier.

Cependant, malgré cette large adoption des services Web, de nombreux obstacles empêchent leur réconciliation sémantique lors de la composition.

L'interprétation consistante des données échangées entre services Web composés est gênée par des différences de représentation et d'interprétation sémantiques. Dans ce mémoire, nous nous intéressons aux hétérogénéités sémantiques des données échangées entre les services Web engagés dans une composition.

Notre contribution évolue autour de la notion de contexte pour décrire la sémantique des données. Nous étudions dans quelle mesure le contexte peut enrichir l'échange des données entre services Web. Nous proposons un modèle orienté contexte pour représenter la sémantique des données, ainsi qu'une approche de médiation qui tire avantage de notre modèle pour résoudre les hétérogénéités sémantiques entre services Web composés.

## Abstract

The adoption of Web services is a keystone in the development of interoperable systems. Particularly, Web services composition allows answering to the increasingly complex users' needs, by combining several Web services into a common business process.

However, despite this widespread adoption of Web services, several obstacles still hinder their semantic reconciliation when being composed. Consistent understanding of data exchanged between composed Web services is hampered by different semantic interpretations and representations.

In this report, we focus on the semantic heterogeneities of data exchanged between Web services engaged in a composition.

Our contribution revolves around the notion of context in order to describe data semantics. We evaluate to which extent context enriches data exchange between Web services. We propose a context-based model to describe the semantics of data, and a mediation approach that takes advantage of our model to solve semantic heterogeneities between composed Web services.

## Publications

- A Context-based Mediation Approach to Compose Semantic Web Services Michael Mrissa, Chirine Ghedira, Djamel Benslimane, Zakaria Maamar, Florian Rosenberg, Schahram Dustdar In ACM Transactions On Internet Technology (ACM TOIT), ACM, Vol. 8, No. 1, 2007.
- A Context Model for Semantic Mediation in Web Services Composition Michael Mrissa, Chirine Ghedira, Djamel Benslimane, Zakaria Maamar In the 25th International Conference on Conceptual Modeling (ER2006) November 6-9 2006, Tucson, Arizona, USA.
- Context and Semantic Composition of Web Services Michael Mrissa, Chirine Ghedira, Zakaria Maamar, Djamel Benslimane In the 17th International Conference on Database and Expert Systems (DEXA) September 4-8, 2006, Krakow, Poland.
- Towards Context-based Mediation for Semantic Web Services Composition Michael Mrissa, Chirine Ghedira, Djamel Benslimane, Zakaria Maamar In International Workshop on Evaluation and Evolution of Component Composition (EECC) held in conjunction with the Eighteenth International Conference on Software Engineering and Knowledge Engineering (SEKE'06) July 5-7, 2006, San Francisco, California, USA.
- Towards a Semantic- and Context-based Approach for Composing Web Services Michael Mrissa, Djamel Benslimane, Zakaria Maamar, Chirine Ghedira International Journal of Web and Grid Services (IJWGS), InderScience, 2005, Vol. 1, No. 3/4, pp. 268-286.
- A Mediation Framework for Web Services in a Peer-to-Peer Environment Michael Mrissa, Chirine Ghedira, Djamel Benslimane, Zakaria Maamar, Jacques Fayolle In Proceedings of the 3rd ACS/IEEE International Conference on Computer Systems and Applications (AICCSA'05) January 3-6, 2005, Cairo, Egypt.
- A Mediation Framework for Web Services in a Distributed Healthcare Information System Michael Mrissa, Chirine Ghedira, Djamel Benslimane, Zakaria Maamar, Jacques Fayolle In Proceedings of the IDEAS'04-DH Conference (With IEEE IDEAS 2004): Workshop on Medical Information Systems September 1-3, 2004, Beijing, China, p. 15-22, ISBN 0-7695-2289-0.

# Reorganizing Dynamic Information Resources in Complex Web according to Complex Adaptive Systems Perspective

**Ratrout Amjad**

**THESE**

Soutenu le 14/12/2007

à l'Université Claude Bernard, Lyon 1

## Jury

BOISSIER Olivier, Ecole des Mines de Saint Etienne,	rapporteur
Giovanna Di Marzo Serugendo, Brikbeck( university of London)	rapporteur
JAYYOUSI Rashid, Al Quds University	
MILLE Alain, Université Lyon 1 – LIRIS	
HASSAS Salima, Université Lyon 1 – LIESP	directrice

Contact : alain.mille@liris.cnrs.fr

## Résumé

De nos jours, le monde évolue de différentes manières, et cette évolution amène une grande diversité. De l'évolution vers la diversité, et vice versa, on peut observer une sorte d'émergence qui se produit dans le monde créant des environnements riches et dynamiques. En l'occurrence, l'espace d'information représente un environnement particulier riche en connaissance. Le Web permet un accès à cet espace d'information. Par conséquent, l'évolution, la diversité et l'émergence peuvent être observé dans le cadre de l'environnement dynamique du Web.

Dans ce travail, nous nous intéressons aux ressources distribuées d'information, et plus précisément les ressources dynamiques auto organisables dans l'environnement du Web complexe. Le system d'information du Web contient trois composants principaux (espaces) : contenus, usages et liens (structure), qui représentent, individuellement un moyen d'accès spécifique. Notre but est de combiner ces différents espaces de manière inter- relationnelle pour aboutir à un system auto organisable au sein du Web. Pour cela, nous proposons un modèle qui combine le système multi agents (MAS) et le system complexe adaptatif (CAS) pour faire en sorte de remonter le défi de l'auto organisation.

Dans cette perspective, nous analysons le Web en tant que CAS, comme il a été défini par J. Holland dans son modèle (ECOH) en utilisant des propriétés et des mécanismes définis pour comprendre le comportement du système complexe du Web. Nous ne limitons pas notre travail à l'étude de la matière à travers laquelle interagissent les composants du Web, mais également la manière dont ils agissent pour former le comportement du system Web en entier. Cela étant dit, on ne peut pas décrire le tout sans décrire chaque composant, chaque composant du Web doit être décrit en relation avec d'autres composants. De plus, nous utilisons MAS afin de simplifier la complexité qui existe dans le Web. MAS représente une collection de techniques calculatoires dans laquelle les agents pris individuellement et leurs interactions sont simulés explicitement et des propriétés d'émergence sont observées. Le but d'utiliser MAS est de créer des modèles qui sont suffisamment simple et dont les mécanismes d'émergences puissent être compris et suffisamment élaborés pour monter des comportements intéressants. Le fait d'adapter un modèle de la vie naturelle (colonie de fourmis) nous aide à améliorer la propriété auto organisatrice des ressources d'information dynamiques dans le Web, où les fourmis simulent des agents dans les composants du Web. L'évaluation des corrélations entre les espaces (contenu, lien et usage) est mesurée en utilisant les règles de la similarité définis par Menczer dans Infospider. Nous mesurons la complexité pour chaque composant de manière séparée, puis nous créons des combinaisons entre les composants. Par la suite, nous organisons l'espace du Web en utilisant ces corrélations.

En tant qu'application, le but voulu par cette méthode est d'être particulièrement utile pour la recherche et la découverte de tendances considérées comme émergentes dans la system Web. Nous montrons

comment cette combinaison entre MAS et CAS, nous donne un accès permanent à des informations consistantes nouvelles, émergentes, et potentiellement de grande valeur qui pouvaient être cachées auparavant pour l'utilisateur. Les valeurs des similarités et les caractéristiques communes sont utilisés afin de classer les pages et les noter selon leur intérêt en temps réel en respectant l'ordre de parcours de l'utilisateur.

Pour résumer, nous proposons une approche pour étudier le Web complexe en tant que source distribuées d'information. Le modèle proposé, "Webcomp", dédié au Web, permet au précédent d'agir en tant que system auto organisable en utilisant des espaces réels et des systèmes virtuels entrecoupés. Ce modèle représente une nouvelle génération qui facilite l'exploration et l'exploitation du Web et ouvre de nouveaux horizons pour bâtir plus de modèles Web auto organisables.

## Abstract

Nowadays, the world is evolving in many ways and this evolution comes together with extreme diversity. From evolution to diversity and vice versa, we can observe a kind of emergence that occurs in the world creating rich and dynamic environments. For instance, information space represents a particular environment of rich knowledge. The World Wide Web (Web) provides accessibility to this information space. Consequently, evolution, diversity and emergence, can be observed in the web dynamic environment.

In this work, we focus on the dynamic distributed resources of information, more precisely the self organizing dynamic resources in the complex Web environment. Considering the Web information system, it contains three main components (spaces): contents, usage and link (structure), which individually represents a specific access mean. Our goal is to combine these different spaces in an interrelated way to achieve a self organising system in the Web. Therefore, we propose a model that combines multi agents systems (MAS) and complex adaptive systems (CAS) in order to tackle the self organisation challenge considering these interrelated spaces.

In this respect, we analyze the web as a CAS as defined by J. Holland in his model (ECOH) using defined properties and mechanisms to understand the behaviour of the complex web system. We do not limit our work to study how the web components interact but also how they act together to form the behaviour of the whole web system. Considering that we cannot describe the whole without describing each component, each web component must be described in relation to other components. In addition, we use MAS to simplify the complexity existing in the web. MAS represent a collection of computational technique in which individual agents and their interactions are explicitly simulated and emergent properties observed. The aim of using MAS is to design models that are sufficiently simple and that the mechanisms of emergence can be understood and yet elaborate enough to show interesting behaviour. Adapting a model from the nature life (Ants system) helps us to improve the self organization of the dynamic information resources in the web, where the "ants" are simulated as agents in the web components. The evaluation of the correlations between the spaces (content, link, and usage) is measured by using the similarity rules defined by Menczer in Infospider. We measure the complexity for each component separately, and then we make combinations between the components. Later on, we reorganize the web space using these correlations.

As application, this method intends to be particularly useful for searching and discovering tendencies considered emergent in the web systems. We show how this combination, between MAS and CAS, give us a permanent access to new relevant, emerging, and possibly valuable information that was previously hidden from the user. Similarities values and common characteristics are used in order to classify pages and rank them in a real time manner with respect to user sequential hits.

To resume, we propose an approach to study complex web as a distributed source of information. The proposed model "Webcomp", dedicated to the web, allows the later to act as a self organizing system using real space and virtual interweaved spaces. This model represents a new generation that facilitate the web exploitation and exploration and opens new horizons to build more self organizing web models.



## Publications

- The web from a complex adaptive systems perspective. M Rupert, A Rattrout, S Hassas (Journal of Computer and System Sciences) Journal ELSEVIER 2007
- The Web and Complex Adaptive Systems. M Rupert, S Hassas, A Rattrout. Dans 20th International Conference on Advanced Information Networking and Applications - Volume 2 (AINA'06), 2006. IEEE Computer Society Press ed. Vienna University of Technology, Vienna, Austria. pp. 200-204. 2006.
- Organisation du contenu du Web selon la perspective des Systèmes Complexes Adaptatifs. M Rupert, A Rattrout, S Hassas. Dans CORIA'2006(Conférence en Recherche d'Informations et Applications), ARIA ed. Lyon France. pp. 283-294. ISBN 2-9520326-6-1. 2006.
- Reorganizing dynamic information resources in the complex Web. A Rattrout, M Rupert, S Hassas. Dans ACIT2005, JORDANIE- AMMAN. pp. 103-110. ISSN ISSN: 1812-0857. 2005.
- Poster Juillet 2003 UCBL1 (Organisation dynamique de sources d'informations distribuées selon la perspective des systèmes complexes)



# Co-construction de sens par négociation pour la réutilisation en situation de l'expérience tracée

Stuber Arnaud

THESE

Soutenue le 04/12/2007

à l'Université Claude Bernard Lyon1

## Jury

Amal El Fallah Seghrouchni, LIP6, Université Pierre et Marie Curie Paris 6	examinateur
Catherine Garbay, CNRS, CLIPS-IMAG, Grenoble	examinateur
Salima Hassas, LIESP, Université Claude Bernard Lyon 1	directeur
Ramon Lopez de Mantaras, CSIC, IIIA, Barcelone	rapporteur
Alain Mille, LIRIS, Université Claude Bernard Lyon 1	directeur
Jean-Paul Sansonnet, CNRS, LIMSI-CNRS, Orsay	rapporteur

Contact : alain.mille@liris.cnrs.fr

## Résumé

Le travail présenté se place dans la problématique de la capitalisation des connaissances au sein d'une communauté d'acteurs humains interagissant avec un outil informatique commun, support de leur activité. Notre approche est d'associer un agent assistant personnalisé à chaque acteur humain. Un tel agent assistant est chargé d'effectuer des suggestions de traces d'expérience qui sont proches de la situation courante rencontrée par l'acteur humain.

Pour traiter des situations qui ne peuvent être définies a priori, nous nous sommes concentré sur la problématique de co-construction de sens par négociation entre l'acteur humain et son agent assistant, afin d'élaborer des connaissances issues de l'expérience tracée et permettant de la remobiliser.

Notre principale contribution se place dans le mécanisme de négociation de sens, celui-ci s'effectue en transposant les principes d'émergence de langage de forme lexicale, tels qu'ils sont présentés par Luc Steels. Cette approche nous a conduit à l'étude d'une reformulation symbolique des traces d'expérience, afin que l'agent assistant puisse établir, en négociant avec l'acteur humain, des expressions symboliques pour désigner les situations d'assistance.

En perspective, nous voyons l'échange d'expérience entre acteurs humains au travers du système se faisant via l'émergence de langages entre agents assistants résultants de la mise en commun de leurs expériences et langages propres. Par ailleurs, notre transposition des mécanismes d'émergence de langage de forme lexical met en perspective la problématique d'émergence de grammaire, en premier lieu pour enrichir l'expressivité du « discours » des agents assistants.

## Publications

### *Conférences internationales :*

- « Combining MAS and Experience Reuse for assisting collective task achievement » Arnaud Stuber, Salima Hassas, Alain Mille dans « Workshop 'From structured cases to unstructured problem solving episodes for Experience-based Assistance' - 5th International Conference on Case-based Reasoning (ICCB'03) » pages 314-323, Lorraine McGinty (ed.) - Trondheim, Norway, June 2003
- « Langage games for meaning negotiation between human and computer agents » Arnaud Stuber, Salima Hassas, Alain Mille Engineering Societies in the Agents' World (ESAW 2005) - 26-28 october, 2005 - Kusadasi, Aydin, TURKEY – LNAI 3963, Springer Verlag
- « Tracing experience as a potential support for negotiation of meaning between human and computer agents » Arnaud Stuber, Salima Hassas, Alain Mille pour poster dans ECCS'05 : European Conference on

*Conférences nationales :*

- « Assistance contextuelle à la réutilisation collective de l'expérience pour une activité coopérative de conception » Stuber A., Hassas S., Mille A. In C. Le Bozec, M.-C. Jaulent & E. Zapletal (eds), Actes de l'Atelier Raisonement à partir de Cas, RaPC'2002, Paris, Juin 2002, pp 1-10.
- « Combiner le paradigme multiagents et la réutilisation de l'expérience pour assister la réalisation collective de tâches. » Stuber A., Hassas S., Mille A., in. Jean Lieber (eds), Actes de l'Atelier Raisonement à partir de Cas, RaPC'2003, plateforme AFIA 2003, Laval, (Juillet 2003). (8 pages )
- « Combiner le paradigme Multi-agents et le Raisonement partir d'expérience pour assister la réalisation collective de tâche. » Stuber A., Hassas S., Mille A. Journée thématique « Traces, interactions, co-constructions collectives et relation à la cognition » - Action Spécifique CoMETE, RTP 38, CNRS- STIC, Paris, 19/12/2003
- « Combiner le paradigme multiagents et la réutilisation de l'expérience pour assister la réalisation collective de tâches. » Stuber A., Hassas S., Mille A. Actes de l'Atelier Raisonement à partir de Cas, RaPC'2004, Villetaneuse, 25-26 mars 2004
- « L'expérience tracée comme support potentiel à la négociation de sens entre agents informatiques et humains » Stuber A., Hassas S., Mille A. Actes de l'Atelier Raisonement à partir de Cas, RaPC'2005, Plateforme AFIA 2005, Nice, 31 mai 2003

# Vers un accès sémantique aux données. Approche fondée sur RDF

Tanasescu Adrian

THESE

Soutenue le 19/06/2007

à l'Université Claude Bernard Lyon1

## Jury

Zohra Bellahsene, Université Montpellier II,	rapporteur
Chantal Soulé-Dupuy, Université Toulouse 1,	rapporteur
Djamel Abdelkader Zighed, Université Lumière Lyon 2, .	examineur
Raphaël Troncy, CWI (Centrum voor Wiskunde en Informatica), Pays-Bas,	examineur
Mohand-Saïd Hacid, Université Lyon .	directeur de recherche

Contact : Mohand-Said Hacid

## Résumé

La plupart du contenu du Web aujourd'hui est orienté vers une utilisation par l'homme. Une approche alternative consiste à représenter le Web dans une forme plus facilement compréhensible par les machines et d'utiliser des techniques intelligentes afin de profiter de ses représentations. Les efforts qui s'inscrivent dans cette approche visant à révolutionner le Web sont réunis dans ce qu'on appelle aujourd'hui l'initiative du Web Sémantique. Dans ce contexte nous nous orientons vers l'utilisation de RDF (Resource Description Framework) pour la description sémantique de documents en général et multimédia en particulier. Nos recherches se développent autour du web sémantique, plus exactement autour de la recherche sémantique d'informations exprimées à l'aide de RDF. Ce langage fournit le formalisme nécessaire à la représentation des données pour le Web sémantique. Néanmoins, seul un format de représentation des connaissances ne suffit pas et nous avons besoin de langages d'interrogation et de supports pour le raisonnement sur ces connaissances afin de rendre les applications capables d'exploiter les documents RDF. Les efforts pour fournir et surtout normaliser ces éléments (i.e. SPARQL – langage d'interrogation pour RDF) sont en plein déroulement au sein du W3C, l'organisme international de normalisation du web. Dans ce contexte, nous étudions les différentes approches d'intégration de données de ce type mais aussi le processus d'interrogation de données RDF.

Une des problématiques clés de nos travaux concerne la qualité de service dans la recherche d'informations. Si l'on considère qu'un utilisateur cherche des informations détaillées sur un sujet qui doivent satisfaire certaines conditions, la plupart du temps, il aimerait obtenir des réponses correctes et les plus complètes possible et cela indépendamment de l'endroit où se trouve l'information. En pratique, d'une part différentes sources de données peuvent contenir des informations, se recouvrant partiellement mais complémentaires, sur le même objet recherché. D'autre part, différentes sources de données peuvent décrire un même objet dans deux visions différentes suivant le contexte dans lequel les descriptions sont considérées. Par exemple, un site de vente de peintures en ligne ne décrit pas forcément une toile avec les mêmes propriétés qu'utiliserait un historien de l'art. Un autre facteur qui doit être pris en considération est la possibilité que des documents, décrivant une même ressource, peuvent être en contradiction sous certains aspects de leurs descriptions. Dans ce cas se pose la question de la validité et de l'exactitude des sources de données, et surtout des précautions à prendre afin d'éviter de combiner des informations qui se révèlent contradictoires.

Considérant ces hypothèses, nous proposons de mettre les bases d'une approche permettant de fournir, pour une requête de l'utilisateur, des réponses plus complètes et en même temps cohérentes. Cette approche consiste essentiellement en une combinaison sémantique des informations en provenance de plusieurs graphes. Pour la mise en place de l'approche nous adressons, principalement, les points suivants:

- Déterminer si les descriptions RDF sont en contradiction. La base de documents étant constitué par l'utilisateur, cette étape est effectuée à l'insertion de chaque nouveau document. A l'aide d'une mesure de "similarité" que nous avons défini, nous déterminons avec quels documents déjà existants le nouveau document est en contradiction. Ainsi, nous réduisons, au moment d'une interrogation, le nombre de combinaisons possibles entre des documents qui, à priori, constitueraient des réponses partielles, candidates à la construction d'une réponse complète.
- Utiliser les ontologies pour déterminer les relations entre les concepts. Tous les documents RDF contenus dans la base de données font référence à des ontologies à l'aide des espaces de noms (namespaces). Ces ontologies décrivent les classes et les propriétés utilisées dans la description des documents. En exploitant les hiérarchies décrites dans les ontologies nous pouvons déterminer si deux concepts sont équivalents ou si un concept est un subsumé par un autre. Nous utilisons ces informations, d'une part, dans le calcul de la similarité entre documents et, d'autre part, dans la combinaison de graphes RDF visant la construction d'une meilleure réponse à une interrogation.
- Construire des réponses cohérentes et plus complètes. Il s'agit d'un algorithme de recherche dont les étapes principales sont: la décomposition de la requête en plusieurs sous-requêtes minimales, la recherche des graphes qui répondent à chaque sous-requête et, implicitement, pour chaque graphe du nombre de sous-requêtes auxquelles il répond, l'enrichissement les éléments des graphes répondant à une partie des sous-requêtes par des éléments d'autres graphes répondant aux autres sous-requêtes.
- L'aboutissement de ce travail de recherche nous a permis de réaliser une plateforme de stockage et interrogation de documents RDF dont le moteur de recherche est intégralement construit sur la base de notre approche. Cette plateforme, nommée SyRQuS, est entièrement basée sur le Web et permet l'interrogation de la base de données à l'aide du langage de requête SPARQL. Les résultats qu'elle fournit sont présentés, d'une part, sous la forme d'un affichage HTML détaillé, intégré à la page Web et, d'autre part, sous forme de documents XML. De plus, cette plateforme permet aussi une navigation visuelle simplifiée dans les descriptions contenues dans la base de données, partant des classes d'objets les plus générales vers les objets de ces classes, puis d'objet en objet en fonctions des relations définies entre eux par les descriptions RDF.

2007

Département  
Image





# Modélisation de l'interaction des ions de haute énergie avec la matière inerte, la matière vivante et la matière en mouvement

HDR

Beuve Michaël

Soutenue le 05/12/2007

à l'Université Claude Bernard Lyon1

## Jury

M Amaldi Ugo, Université de Milan	rapporteur
M Balanzat Emmanuel, CIRIL-CNRS	rapporteur
M Delingette Hervé, INRIA	rapporteur
M Ille Bernard, IPNL, université Lyon1	
M Moretto Philippe, Université Bordeaux 1, CENBG	
M Mueller Alex, CNRS-IN2P3, IPNO Lyon	
M Rémillieux Joseph, Université Lyon 1	invité
M Shariat Behzad Université Lyon 1, LIRIS	

Contact : mbeuve@liris.cnrs.fr

## Résumé

Les travaux de modélisation et de simulation qui sont synthétisés dans ce manuscrit abordent trois aspects très pluridisciplinaires de l'interaction des ions de hautes énergies avec la matière: physique, radiobiologie et informatique. Le premier volet se concentre sur des aspects de physique fondamentale. Il vise à décrire et à quantifier les mécanismes d'interaction d'un ion avec un matériau et les modifications de structure qui résulte de cette interaction.

Les deux autres volets de ce manuscrit sont orientés vers l'amélioration des traitements du cancer par faisceau d'ions légers (Hadronthérapie). L'objectif est de mieux simuler la dose physique et la dose biologique déposées dans les organes par les faisceaux d'ions. Alors que la dose physique traduit les transferts d'énergie entre l'ion et les tissus, la dose biologique représente les effets biologiques induits. Cette dernière ne peut se déduire simplement de la dose physique et doit être modélisée. Le volet radiobiologie du manuscrit rassemble mes travaux de modélisation descriptive et prédictive des effets radiobiologiques des ions. Une évaluation expérimentale des modèles prédictifs y est également présentée.

Quelque soit le raffinement apporté à la modélisation de la dose physique et de la dose biologique, si le mouvement des organes n'est pas intégré, les calculs dosimétriques seront inappropriés. Les systèmes de contention permettent de réduire certains mouvements externes (tête, cou, membres...), mais pas les mouvements internes. Le dernier volet de ce manuscrit, le volet informatique, présente mes travaux de modélisation des mouvements internes des organes, notamment les mouvements respiratoires. Ces modèles ont été construits de telle sorte qu'ils puissent intégrer les spécificités du transport des ions dans la matière

## Abstract

The work of modeling and simulation that is summarized in this manuscript deals with three multidisciplinary aspects of the interaction of high-energy ions with matter: physics, radiobiology and computers. The first part focuses on aspects of fundamental physics. It aims at describing and quantifying the mechanisms at work during the interaction of an ion with a material and the structural changes that are induced.

The other two parts are directed towards the improvement of cancer treatment by light-ion beam (Hadrontherapy). The aim is to better simulate the physical dose and the biological dose deposited into the organs by the ion beams. While the physical dose reflects the transfer of energy between the ion and the

tissues, the biological dose refers to biological effects. The latter cannot simply be deduced from the physical dose and should be modeled. The radiobiology part of the manuscript gathers my contribution to the development of descriptive and predictive models. An experimental evaluation of the predictive models is also presented.

Whatever the level of refinement to the modeling of physical and biological dose, if tissue motion is not integrated, dose calculations will be inappropriate. Contention systems can reduce some external movements (head, neck, members ...), but hardly internal motions. The last part of this manuscript is thus dedicated to computer science and focus on the modeling of internal-organ motions, including respiratory motions. These models were constructed in such way that they can integrate the specificities of ion transport in matter.

## Publications

- "Influence of the spatial and temporal structure of the deposited-energy distribution in swift-ion-induced sputtering" M. Beuve, N. Stolterfoht, M. Toulemonde, C. Trautmann, H. M. Urbassek PHYSICAL REVIEW B 68, 125423 (2003)
- "Radiobiological parameters and local effect model predictions for head and neck squamous carcinoma cells exposed to high linear energy transfer ions " M. Beuve, G. Alphonse, M. Maalouf, A. Colliaux, P. Battiston Montagne, P. Jalade, E. Balanzat, A. Demeyer, M Bajard, C. Rodriguez International Journal of Radiation Oncology Biology Physics 71 (2008) 635-642
- "An Approach to Convert 4D Geometry into a 4D CT Scan" Villard Pierre-Frederic, Beuve Michael, Shariat Behzad In Proceedings of WSCG 2006, Plzen (Czech Republic), Union Agency, February 2006.

Soutenue le 05/12/2007

à l'Université Claude Bernard Lyon 1

## Jury

Mme. Borgefors Gunilla, CBA, Uppsala Universitet, Suède	rapporteur
Mr. Braquelaire Achille, LaBRI, Université Bordeaux 1	rapporteur
M. Maître Henri, TSI, ENST, Paris	rapporteur
Mme Montanvert Annick, GIPSA-Lab, UPMF, Grenoble	examineur
M. Devillers Olivier, INRIA, Sophia-Antipolis ,	examineur
M. Péroche Bernard, Professeur, LIRIS, Université Lyon 1	examineur

Contact : david.coeurjolly@liris.cnrs.fr

## Résumé

Le contexte général de mes activités de recherche est la géométrie discrète. Cette thématique s'intègre, au moins d'un point de vue historique, dans l'analyse de formes dans des images numériques. En effet, de nombreux systèmes d'acquisition de données images fournissent des données organisées sur une grille régulière, appelées données discrètes. Que ce soit pour une visualisation ou pour l'extraction de mesures sur ces objets discrets (paramètres de formes), les axiomes et théorèmes de la géométrie euclidienne ne sont pas directement applicables. Une approche classique consiste à une transposition de ces théorèmes et mesures dans l'espace discret. Ces différentes re-définitions donnent lieu au paradigme mathématique et informatique qu'est la géométrie discrète. Dans ce contexte, nos contributions portent sur l'analyse des modèles et objets fondamentaux (grille, droite, plan, cercle, ...) permettant la définition d'algorithmes de reconstruction géométrique. Dans un second temps, nous nous sommes intéressés à des algorithmes performants pour l'analyse volumique d'objets discrets (transformation en distance, axe médian,...), ainsi qu'à leurs généralisations.

## Abstract

The context of the work presented in this thesis is the digital geometry. This research area is devoted to the automatic analysis of objects in digital images. Indeed, many acquisition systems provide regular data in dimension 2 and 3. In order to design efficient analysis algorithms, Euclidean geometry theorems and algorithms should be redesigned to consider specificities of the digital space. In this framework, we have first investigated fundamental objects (grids, straight lines, circles,...), allowing us to define efficient geometrical reconstruction algorithms. Then, we have considered high performance algorithms to perform volumetric analysis of digital objects (distance transformation, medial axis,...).

## Publications

### *Edited Books and Proceedings*

- Géométrie discrète et images numériques. D. Coeurjolly, A. Montanvert, J.m. Chassery. Ouvrage collectif, Traité IC2, Hermès, pp-416 2007.
- Discrete Geometry for Computer Imagery, 14th IAPR International Conference. D. Coeurjolly, I. Sivignon, L Tougne, F. Dupont. (DGCI 2008), , Lyon, France. 552p. LNCS 4992. Springer Verlag . ISBN 978-3-540-79125-. 2008.

### *Selected publications in Journals*

- A framework for dynamic implicit curve approximation by an irregular discrete approach. A. Vacavant, D. Coeurjolly, L Tougne. Graphical Models 71(3):113-124. 2009.

- Gift-Wrapping based Preimage Computation Algorithm. Y. Gerard, D. Coeurjolly, F Feschet. Pattern Recognition (), Elsevier. 2009.
- Minimum Decomposition of a Digital Surface into Digital Plane Segments is NP-Hard. I. Sivignon, D. Coeurjolly. Discrete Applied Mathematics 157(3):558-570, Elsevier. 2008.
- Finding a Minimum Medial Axis of a Discrete Shape is NP-hard. D. Coeurjolly, J. Hulin, I. Sivignon. Theoretical Computer Science 206(1-2):72-79, Elsevier. 2008.
- Discrete bisector function and Euclidean skeleton in 2D and 3D. M. Couprie, D. Coeurjolly, R. Zrour. Image and Vision Computing 25(10):1519-1698, Elsevier. 2007.
- Optimal Separable Algorithms to Compute the Reverse Euclidean Distance Transformation and Discrete Medial Axis in Arbitrary Dimension. D. Coeurjolly, A. Montanvert. IEEE Transactions on Pattern Analysis and Machine Intelligence 29(3):437-448. 2007.
- Digital Planarity - A Review. V. Brimkov, D. Coeurjolly, R. Klette. Discrete Applied Mathematics 155(4):468-495, Elsevier. 2007.
- Supercover model, digital straight line recognition and curve reconstruction on the irregular isothetic grids. D. Coeurjolly, L. Zerarga. Computer and Graphics 30(1):46-53. 2006.

**Dikici Cagatay**

Soutenue le 03/12/2007

à l'INSA de Lyon

### Jury

Jean Marc CHASSERY, LIS/ENSIEG de Grenoble  
Bulent SANKUR, Université de Bogazici, Turkey  
Christine GUILLEMOT, l'INRIA/IRISA Rennes  
William PUECH, Université de Montpellier  
Fabrice MERIAUDEAU, LE2I, Le Creusot  
Florent DUPONT, UCB Lyon-1  
Atilla BASKURT, LIRIS, INSA de Lyon  
Khalid IDRISSEI, LIRIS, INSA de Lyon

Contact : cagatay.dikici@lss.supelec.fr

### Résumé

Les avancées technologiques qu'ont connu les télécommunications, le multimédia et les systèmes mobiles ont ouvert la porte à l'émergence, puis au développement de nouveaux services tels que la vidéoconférence, la protection des contenus, etc. tout en utilisant des systèmes à faible puissance. D'où la nécessité de disposer de nouvelles techniques de codage à complexité réduite. Les techniques de codage exploitant la présence d'une information parallèle peuvent constituer une solution potentielle permettant de déporter la complexité de codage vers le décodeur. Celles-ci s'appliquent notamment à deux principes de codage :

- 1) Le codage de source distribuée (Distributed Source Coding DSC) pour compresser un signal donné, sachant qu'un autre signal corrélé à celui d'origine est disponible au niveau du décodeur.
- 2) La dissimulation de données informée (Informed Data Hiding IDH) permettant d'insérer un message dans un signal hôte, ce dernier n'étant connu qu'au codeur.

Pour chacune de ces deux applications, nous proposons des solutions qui approchent les limites théoriques. Nous combinons pour cela des techniques performantes tant de codage canal, de type LDPC, que de quantification de type Treillis (TCQ). Par ailleurs, nous étudions les limites théoriques pouvant être atteintes par IDH, dans le cas où une version bruitée du signal hôte est disponible au décodeur.

Enfin, exploitant la forte dualité qui existe entre DSC et IDH, nous proposons un schéma pratique hybride complet mettant en œuvre les deux techniques, ainsi qu'une étude théorique de la fonction débit / distorsion et de la capacité d'un tel système.

### Abstract

Technological advances in the fields of telecommunications, multimedia and the diverse choice of portable handhelds during the last decade, derive to create novel services such as video-conferencing, content protection etc. running on low-power devices. Hence alternative low complexity coding techniques need to be developed for replacing conventional ones. Coding with state information, a potential solution to shifting the encoder complexity to the decoder, has two main applications:

- 1) Distributed Source Coding(DSC) for compressing a source given a correlated version of it is available only to the decoder.
- 2) Informed Data Hiding(IDH) for embedding a watermark to a host signal where the host signal is available only to the encoder.

For each problem stated above, practical code designs that operate close to the theoretical limits are proposed. The combination of good error correcting codes such as Low Density Parity-Check (LDPC) Codes and good quantization codes such as Trellis Coded Quantization (TCQ) are used at the design of the proposed capacity approaching codes.

Moreover, the theoretical achievable rate limits for a relaxed IDH setup, such that a noisy observation of the host signal is available to the decoder is derived. Finally, motivated by the strong duality between DSC and IDH, a hybrid scheme that uses both data hiding and compression is proposed. In addition to the derivation of theoretical channel capacity and rate distortion function, a complete framework is proposed.

## Publications

### *Conférences Internationales avec comité de lecture*

- Dikici C., K. Idrissi and A. Baskurt, “Joint Data-Hiding and Source Coding of Still Images”, Proc. of European Signal Processing Conf. EUSIPCO 2006, Florence, Sept 2006
- Dikici C. , K. Idrissi and A. Baskurt, “Joint Data-Hiding and Source Coding with Partially Available Side Information”, Proc. of SPIE 2006 E. Imaging, San José, Jan 2006
- Dikici C, R. Guermazi, K. Idrissi, A. Baskurt, “Distributed Source Coding of Still Images”, Proc. of European Signal Processing Conf. EUSIPCO 2005, Antalya , Sept 2005
- Dikici C., K. Idrissi and A. Baskurt, “ Dirty-paper writing based on LDPC codes for Data Hiding”, International Workshop on Multimedia Content Representation, Classification and Security (MRCS 2006), Istanbul, Oct 2006
- Dikici C., C. Guillemot, C. Fontaine , K. Idrissi, and A. Baskurt. “Dirty Paper Coding with Partial State Information.” Dans IEEE Proc. of International Symposium on Image/Video Communications over fixed and mobile networks (ISIVC), Bilbao, Jul 2008.

### *Conférences Nationales avec comité de lecture*

- Dikici, C., K. Idrissi et A. Baskurt, “Tatouage Informé pour le Codage Distribué”, Proc. of CORESA’06, Caen, Nov 2006.

# Contribution à la restauration des images de documents

**Drira Fadoua**

**THESE**

Soutenue le 12/12/2007

à l'INSA de Lyon

## Jury

M. Christian JUTTEN, U.J.F de Grenoble	examineur
M. Adel ALIMI, ENI-Sfax-Tunisie	examineur
M. Abdel BELAID, Université de Nancy2	rapporteur
M. Christian OLIVIER, Université de Poitiers	rapporteur
M. Hubert EMPTOZ, I.N.S.A de Lyon	directeur
M. Frank LEBOURGEOIS, I.N.S.A de Lyon	directeur

Contact : hubert.emptoz@liris.cnrs.fr

## Résumé

La numérisation massive de documents patrimoniaux ouvre de nouvelles perspectives de Recherche comme la restauration des documents dégradés. Ces dégradations sont dues au mauvais état de conservation et à la numérisation. De telles altérations sur des images aussi complexes ne peuvent pas être traitées par une simple transposition des outils traditionnels de restauration. Cette thèse traite donc de la restauration des dégradations des images des documents anciens, et plus particulièrement des dégradations du support papier et des formes des lettres car elles nuisent à la lisibilité des documents numérisés et empêchent tout traitement automatisé.

Pour les dégradations du support papier, assimilées à un problème de superposition de couches, nous proposons deux méthodes. La première est une méthode de segmentation récursive issue de l'algorithme des nuées dynamiques appliquée sur les données décorréélées par l'analyse en composantes principales. La deuxième méthode est une amélioration de l'algorithme du Mean-Shift qui réduit sa complexité de calcul.

Nous proposons, ensuite, de traiter les dégradations des formes des lettres par des méthodes de diffusion. Tout d'abord, nous présentons les modèles existants les plus adaptés à la restauration des images de document. Ensuite, nous proposons un modèle de diffusion capable de filtrer les contours et de renforcer la continuité des traits tout en préservant les singularités.

Pour concevoir une approche capable de restaurer à la fois les dégradations du support papier et des formes des lettres, nous proposons une combinaison entre la segmentation et la diffusion. Cette combinaison permet de réaliser une segmentation en couches des images de documents tout en préservant la forme des contours des lettres.

### Mots clés :

Images de documents dégradés - restauration d'image - traitement par EDP - séparation recto/verso - segmentation colorimétrique - paradigme local/global.

## Abstract

The massive digitization of heritage documents raised new prospects for Research like the restoration of the degraded documents. These degradations are due to the bad conditions of conservation and even to the digitization process. Images of old and degraded documents cannot be the retored directly by classical approaches. Hence, we propose in this thesis to develop and analyze document image restoration algorithms. We are mainly interested in foreground/ background degradations, since they harm the legibility of the digitized documents and limit the processing of these images.

For background degradations, considered as a problem of the superposition of layers, we propose two-based segmentation methods. The first is a recursive approach that relies on the k-means clustering

algorithm and the principal component analysis. The second method is an improvement of the initial algorithm of MeanShift in an attempt to reduce its complexity.

For foreground degradations, we propose to tackle the problem with PDE-based diffusion approaches. This solution has many useful features that are relevant for use in character restoration. Our comparative study of existing methods allows us to select the best approaches well adapted to our problem. We propose also a new diffusion method preserving singularities and edges while smoothing.

Our previously proposed solutions, the diffusion and the Mean-Shift algorithms, are used with success in a joint iterative framework to solve foreground and background degradations. This framework generates segmented images with more reduced artefacts on the edges and on the background than those obtained in separate application of each method.

*Keywords:*

Degraded Document images - image restoration - PDE based approaches background/ foreground separation - colorimetric segmentation - local/global paradigm.

## **Publications**

### *Contributions à des ouvrages*

- F. Drira, F. Lebourgeois, H. Emptoz, A Modified Mean Shift Algorithm For Efficient Document Image Restoration, Signal processing for image enhancement and multimedia processing Ed. Springer Verlag, ISBN 978-0-387-72499-7, on-line copy <http://www.springerlink.com/content>, 2008.

### *Conférences Internationales avec comité de lectures et actes*

- F. Drira, F. Lebourgeois, H. Emptoz, OCR Accuracy Improvement Through a PDE-based Approach, 9th International Conference on Document Analysis and Recognition (ICDAR), Septembre 2007, Curitiba, Brazil, pp. 1068-1072
- F. Lebourgeois, H. Emptoz, A coupled Mean Shift-Anisotropic Diffusion Approach for Document Image Segmentation and restoration, 9th International Conference on Document Analysis and Recognition (ICDAR), Septembre 2007, Curitiba, Brazil pp. 814-818
- V. Eglin, F. Lebourgeois, S. Bres, H. Emptoz, Y. Leydier, I. Moalla, F. Drira, Computer assistance for Digital Libraries : Contributions to Middle-ages and Authors' Manuscripts exploitation and enrichment, Second IEEE International Conference on Document Image Analysis for Libraries (DIAL2006), Lyon, France. pp. 265-280, 2006.
- F. Drira, Towards restoring historic documents degraded over time, Second IEEE International Conference on Document Image Analysis for Libraries (DIAL2006), mars 2006, Lyon, France, pp. 350-357
- F. Drira, F. Lebourgeois, H. Emptoz, Restoring Ink Bleed-Through Degraded Document Images Using a Recursive Unsupervised Classification Technique, 7th IAPR Workshop on Document Analysis Systems (DAS2006), Nelson, New Zealand. Lecture Notes in Computer Science, Vol. 3872, Springer, pp.38-49, 2005
- F. Drira, H. Emptoz, A Recursive Approach For Bleed-Through removal, First International Workshop on Camera-Based Document Analysis and Recognition (CBDAR2005), Koichi Kise, David S. Doermann ed. Seoul, Korea, pp. 119-126, 2005

### *Conférences nationales avec comité de lecture et actes*

- F. Drira, F. Lebourgeois, H. Emptoz, A Modified Mean Shift Algorithm For Efficient Document Image Restoration, IEEE/ACM international conference on signal processing and Internet based systems , Hammamet, Tunisia. pp. 686-695, 2006
- F. Drira, F. Lebourgeois, H. Emptoz, Séparation recto/verso des images de documents anciens par une approche «aveugle», Colloque International Francophone sur l'Ecrit et le Document, CIFED'06, octobre 2006, Fribourg, Suisse, pp. 199-204



# Real-Time Video Analysis from a Mobile Platform: Moving Object and Obstacle Avoidance

Limongiello Alessandro

THESE

Soutenue le 13/06/2007

à l'Universitat di Alicante (Espagne)

## Jury

P. Foggia (IT)

F. Escolano (ESP)

C. De la Higuera (FR)

M. Vento (IT)

JM Jolion (FR)

directeur

directeur

Contact : alimongiello@unisa.it

## Résumé

Nous présentons un système de vision pour la navigation autonome d'une plateforme mobile. Ce système est en mesure d'interagir avec l'espace immédiatement environnant, en reconnaissant les obstacles et les objets en mouvement et en construisant une vision stable du monde extérieur. Le système est partagé en 3 composants : la représentation de l'espace environnant ; la détection et l'évitement des obstacles et l'analyse comportementale. La contribution majeure de ce travail concerne la représentation « perceptive » de l'espace, c'est-à-dire une représentation qui est comparée à l'objectif final de la navigation autonome. Cette représentation est basée sur le paradigme de la vision stéréo et elle permet de déterminer dans la scène les obstacles et les objets en mouvement. Notre méthode fournit la profondeur moyenne par région. L'estimation de la position des régions est suffisamment précise pour la navigation et le système est assez rapide pour les applications en temps réel.

## Abstract

We introduce a vision system for autonomous navigation of a mobile platform. This system is able to interact with its immediate environment by recognizing obstacles and moving objects and by means of a stable representation of external world. This system is made of 3 components: external world representation, obstacles detection and avoidance, behavioral analysis. The main contribution of this work lies in the perceptive representation of the external world, e.g. a representation compared to the final goal of autonomous navigation. This representation is based on the stereovision paradigm and is able to determine in the scene obstacles and moving objects. Our approach returns the depth of any region. The location estimation of regions is precise with respect to navigation requirements and the system is fast enough for real time applications.

## Publications

### *International Journals*

- P. Foggia, A. Limongiello, F. Tufano, M. Vento, "Learning Graphs from examples: an application to the prediction of the toxicity of chemical compounds". *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 20, No. 6, pp. 883-896, 2006.

### *Chapters of International Journals*

- D. Conte, P. Foggia, C. Guidobaldi, A. Limongiello, M. Vento, "An Object Tracking Algorithm Combining Different Cost Functions", *Lecture Notes in Computer Science*, vol. 3212, Springer, Berlin, pp. 614-622, 2004.
- L.P. Cordella, A. Limongiello, C. Sansone, "Network Intrusion Detection by a Multi-stage Classification System", *Lecture Notes in Computer Science*, vol. 3077, Springer, Berlin, pp. 324-333, 2004.

### *International Conferences*

- P. Foggia, J.-M. Jolion, A. Limongiello, M. Vento, "Stereo Vision for Obstacle Detection: a Graph-Based Approach", to appear in Proceedings of the 6th IAPR-TC-15 Workshop on Graph-based Representations in Pattern Recognition (GbR-07), Alicante (Spain), June 11-13, 2007.
- P. Foggia, A. Limongiello, M. Vento, "Stereo Vision for Obstacle Detection: a Region-Based Approach", Proceedings of the 1st International Workshop on Robot Vision in conjunction with the 2nd International Conference on Computer Vision Theory and Applications (VISAPP-07), Barcelona (Spain), pp. 36-45, March 8-11, 2007.
- P. Foggia, J.-M. Jolion, A. Limongiello, M. Vento, "A new approach for Stereo Matching in Autonomous Mobile Robot applications", Proceedings of International Joint Conference on Artificial Intelligence (IJCAI-07), Hyderabad (India), pp. 2103-2108, January 6-12, 2007.
- P. Foggia, A. Limongiello, M. Vento, "A moving object and obstacle detection system in real-time AVG and AMR applications", Proceedings of ICGST International Conference on Automation, Robotics and Autonomous Systems (ARAS-05), ISBN 21969/2005, ICGST 2005, Cairo (Egypt), pp. 19-24, December 19-21, 2005.
- M. Vento, P. Foggia, A. Limongiello, G.M. Rispoli, G. Greco, M. Amodio, C. Colella, M. de' Gennaro, O. Marino, P.G. Cappelletti, A. Colella, M. Colella, "Risultati dell'indagine conoscitiva e restituzione virtuale degli affreschi ianuariari delle Catacombe di Capodimonte", Proceedings of Convegno di studi internazionale – San Gennaro XVII Centenario del martirio (305-2005), Napoli, (in press), September 21-23, 2005.
- G.M. Rispoli, G. Greco, M. Amodio, C. Colella, M. de' Gennaro, O. Marino, P.G. Cappelletti, A. Colella, M. Colella, M. Vento, P. Foggia, A. Limongiello, "Analisi integrata e restituzione virtuale di monumenti della pittura antica: l'Ipogeo di Caivano", Proceedings of XXI Convegno internazionale SULLE PITTURE MURALI: Riflessioni, Conoscenze, Interventi, Bressanone, pp. 657-666, ed. Arcadia Ricerche - Marghera (VE), July 12-15, 2005.
- P. Foggia, A. Limongiello, M. Vento, "A real-time stereo-vision system for moving object and obstacle detection in AGV and AMR applications", Proceedings of IEEE CAMP'05 International Workshop on Computer Architecture for Machine Perception, ISBN 0-7695-2255-6, IEEE Computer Society, Palermo (Italy), pp. 58-63, July 4-6, 2005.

### *Chapters of National Journals*

- M. Vento, A. Limongiello, P. Foggia, "Acquisizione e fruizione mediante musei virtuali e internet di opere d'arte", in Collana Fonti e Documenti per la Storia Atellana - IPOGEO DI CAIVANO, Gioia Rispoli, Giovanna Greco, Carmine Colella e Mario Vento (editors), ed. Istituto di Studi Atellani, vol. 10, pp. 89-129, Novembre, 2005.

### *National Conference*

- P. Foggia, C. Guidobaldi, A. Limongiello, M. Vento, "An Image Server for the realization of Web Museums", Atti del IX Convegno della Associazione Italiana Intelligenza Artificiale, Perugia, September 15-17, 2004.
- P. Foggia, A. Limongiello, M. Vento, "Virtual museums: methodologies, tools and open issues", Atti del IX Convegno della Associazione Italiana Intelligenza Artificiale, Perugia, September 15-17, 2004.
- D. Conte, P. Foggia, A. Limongiello, M. Petretta, F. Tufano, M. Vento, "Object detection and tracking for video surveillance", Atti del 2° Convegno del Gruppo Italiano Ricercatori in Pattern Recognition, Perugia, September 15-17, 2004.
- D. Conte, L.P. Cordella, P. Foggia, A. Limongiello, C. Sansone, M. Vento, "Acquisizione e Fruizione su Internet di Opere d'arte", Atti del Convegno Contesti virtuali e fruizione dei beni culturali, Napoli, May 22-23, 2003.

# Calcul efficace et direct des représentations de maillages 3D utilisant les harmonies sphériques

Mousa Mohamed

THESE

Soutenue le 10/09/2007

à l'Université Claude Bernard Lyon1

## Jury

Stefanie Hahmann, ENSIMAG-INPG,  
Marc Neveu, université de Bourgogne,  
Francis Schmitt, ENST,  
Bruno Lévy, LORIA-INRIA Lorraine,  
Samir Akkouche, université Lyon1,  
Raphaëlle Chaine, université Lyon 1,

rapporteur  
rapporteur  
examineur  
examineur  
directeur de thèse  
co-encadrante

Contact : samir.akkouche@liris.cnrs.fr

## Résumé

Dans le domaine de l'informatique graphique, les harmoniques sphériques jouent un rôle important dans des applications aussi variées que l'illumination globale, le rendu, le filtrage de surface et la reconnaissance de forme. La transformée en harmoniques sphériques permet de réaliser l'analyse fréquentielle de fonctions définies sur une sphère. Nous nous intéressons ici au cas particulier des fonctions sphériques définissant un maillage (maillage d'une surface étoilée, BRDF linéaire par morceaux, fonction indicatrice de l'intersection entre un objet et une sphère 3D) et nous offrons un algorithme permettant de calculer les coefficients de la transformée en harmoniques sphériques directement à partir des triangles ou des arêtes composant le maillage. Cet algorithme s'accompagne d'une analyse permettant de contrôler la précision et l'erreur géométrique commise suite à des opérations de filtrage. On évite ainsi les problèmes d'erreur systématiques liés à une voxelisation régulière de l'espace dans lequel est plongée la forme 3D. Nous avons appliqué ces résultats dans une variété d'applications telles que le filtrage, le transfert de texture géométrique, la reconstruction de surface.

## Publications

- M. Mousa, R. Chaine, and S. Akkouche. Frequency-based representation of 3D point-based surfaces using the spherical harmonics. In ICCVG'06, International Conference on Computer Vision and Graphics, Warsaw, Poland, 2006.
- M. Mousa, R. Chaine, and S. Akkouche. Direct spherical harmonic transform of a triangulated mesh. Journal of Graphics Tools, 11 :17–26, 2006.
- M. Mousa, R. Chaine, and S. Akkouche. Frequency-based representation of 3D models using spherical harmonics. In WSCG'06 : the 14th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, volume 14, pages 193–200, Plzen, Czech Republic, January 30 - February 3 2006.
- M. Mousa, R. Chaine, S. Akkouche, and E. Galin. Efficient spherical harmonics representation of 3D objects. submitted for publication.
- M. Mousa, R. Chaine, and S. Akkouche. Une représentation fréquentielle des modèles 3D utilisant les harmoniques sphériques. In 18ièmes journées de l'Association Française d'Informatique Graphique, 3ièmes journées de l'Association Chapitre Français d'Eurographics, volume 18, pages 79–89, 2005.
- M. Mousa, R. Chaine, and S. Akkouche. Direct spherical harmonic transform of a triangulated mesh. Technical Report RR-LIRIS-2005-015, LIRIS : Lyon Research Center for Images and Intelligent Information Systems, 2005.



# Solutions visuo-haptiques pour le prototypage virtuel, applications au secteur automobile

THESE

Ortega Michaël

Soutenue le 26/02/2007

à l'INRIA Rhone Alpes

## Jury

Bruno ARNALDI, INSA Rennes,	rapporteur
Edouard GENTAZ, CNRS,	rapporteur
Roland BLACH, Fraunhofer IAO (Allemagne),	examinateur
Joëlle COUTAZ, UJF,	présidente du jury
Jean Lorisson, PSA	invité
Bernard Péroche, Université Lyon 1,	directeur de thèse
Sabine Coquillart, INRIA,	co-directrice de thèse

Contact : [bernard.peroche@liris.cnrs.fr](mailto:bernard.peroche@liris.cnrs.fr)

## Résumé

Certaines entreprises, comme PSA Peugeot Citroën, font de plus en plus appel à une nouvelle approche du prototypage: le prototypage virtuel. L'efficacité des solutions de réalité virtuelle dédiées passe par la qualité de simulation et d'intégration de différents retours sensoriels dans un même système.

Nous présentons ici une nouvelle configuration visuo-haptique préservant au mieux les facteurs de performance des retours sensoriels. En complément, nous proposons une nouvelle méthode de rendu haptique pour la manipulation d'objets rigides avec propriétés de surfaces (lissage, textures, ...) et suppression des effets indésirables des méthodes traditionnelles (frottements incontrôlés, effets de collage, ...). La configuration est testée et évaluée grâce à un premier scénario industriel concret.

Nous étudions également une deuxième configuration visuo-haptique à base de visiocasques semi-transparents pour la manipulation en aveugle. Un deuxième scénario industriel est mis en place.

## Abstract

Now a day, more and more enterprises like PSA Peugeot Citroën need a new prototyping approach called virtual prototyping. The efficiency of virtual reality dedicated solutions is closely linked with the simulation and integration quality of different sensory feedbacks in a same system. Here, we present a new visuo-haptic configuration which preserves the performance factors of integrated sensory feedbacks. In addition, we propose a new haptic rendering method for rigid objects manipulation with surface properties (smooth, textures ...) and without the well known undesirable effects from traditional methods (uncontrolled friction, sticking effects ...). The configuration is evaluated with a first industrial concrete scenario. We also study a second visuo-haptic configuration using a video see-through head mounted display for blind manipulation applications. A second industrial scenario is presented.

## Publications

- M. ORTEGA and S. COQUILLART, Prop-based haptic interaction with co-location and immersion: an automotive application, in Proc. of Haptic Audiovisual Environments and their Applications, 2005
- M. ORTEGA, S. REDON and S. COQUILLART, A six degree-of-freedom god-object method for haptic display of rigid bodies, in IEEE VR'06
- M. ORTEGA, S. REDON and S. COQUILLART, A six degree-of-freedom god-object method for haptic display of rigid bodies with surface properties, in IEEE TVCG Journal, 2007



# Variable Resolution Transform-based Music Feature Extraction and their Applications for Music Information Retrieval

**Paradzinets Aliaksandr**

**THESE**

Soutenue le 19/12/2007

à l'Ecole Centrale de Lyon

## Jury

Myriam Desainte-Catherine

Gaël Richard

François Pachet

Liming Chen

encadrant

Contact : aliaksandr.paradzinets@ec-lyon.fr

## Résumé

Dans le secteur de loisirs il y a un nombre considérable d'enregistrements numériques musicaux produits, diffusés et échangés qui favorise la demande croissante de services intelligents de recherche de musique. La navigation par contenu devient cruciale pour permettre aux professionnels et également aux amateurs d'accéder facilement aux quantités de données musicales disponibles. Ce travail présente les nouveaux descripteurs de contenu musical et mesures de similarité qui permettent l'organisation automatique de données musicales (recherche par similarité, génération automatique des playistes) ainsi que l'étiquetage (classification automatique en genres). Ce travail s'intéresse au problème de la construction des descripteurs du point de vue musical en complément des caractéristiques spectrales de bas-niveau. Plusieurs aspects d'analyse musicale, telles que l'analyse du signal où une nouvelle technique de transformation fréquentielle à résolution variable est proposée et décrite. Le traitement de niveau plus haut touche aux aspects de l'extraction des connaissances musicales. Cette thèse présente les algorithmes de détection de coups (beats) et d'extraction de fréquences fondamentales multiples. Les deux algorithmes sont basés sur la transformation à résolution variable proposée. Les informations issues de ces algorithmes sont utilisées dans la construction des descripteurs musicaux, représentés sous forme d'histogrammes (nouvel histogramme rythmique 2D qui permet d'estimer directement le tempo, et les histogrammes de succession et profil de notes). Deux applications majeures qui utilisent les caractéristiques mentionnées sont décrites et évaluées dans cette thèse.

## Abstract

As a major product for entertainment, there is a huge amount of digital musical content produced, broadcasted, distributed and exchanged. There is a rising demand for content-based music search services. Similarity-based music navigation is becoming crucial for enabling easy access to the ever-growing amount of digital music available to professionals and amateurs alike. This work presents new musical content descriptors and similarity measures which allow automatic musical content organizing (search by similarity, automatic playlist generating) and labeling (automatic genre classification). The work considers the problem of content descriptor building from the musical point of view in complement of low-level spectral similarity measures. Several aspects of music analysis are considered such as music signal analysis where a novel variable resolution transform is presented and described. Higher level processing touches upon the musical knowledge extraction. The thesis presents algorithms of beat detection and multiple fundamental frequency estimation which are based on the variable resolution transform. The information issued from these algorithms is then used for building musical descriptors, represented in form of histograms (novel 2D beat histogram which enables a direct tempo estimation, note succession and note profile histograms etc.). Two major music information retrieval applications, namely music genre classification and music retrieval by similarity, which use aforementioned musical features are described and evaluated in this thesis.

## Publications

- A Paradinets, O. Kotov, H Harb, L. Chen., Continuous Wavelet-like Transform Based Music Similarity Features for Intelligent Music Navigation. CBMI07, Bordeaux, France. 2007.
- Kotov O., Paradinets A., Bovbel E. Musical Genre Classification using Modified Wavelet-like Features and Support Vector Machines. EuroIMSA 2007, Chamonix, France.
- Paradinets A., Harb H., Chen L., "Use of Continuous Wavelet-like Transform in Automated Music Transcription", EUSIPCO2006
- Parshin V., Paradinets A., Chen L. Multimodal Data Fusion for Video Scene Segmentation, VIS2005



# Prise en compte du mouvement respiratoire pour la reconstruction d'images tomодensitométriques

THESE

Rit Simon

Soutenue le 21/09/2007

à l'Université Lumière, Lyon 2

## Jury

Mme. Isabelle Bloch	rapporteur
M. Pierre Grangeat	rapporteur
Mme. Isabelle Magnin	examineur
M. Grégoire Malandain	examineur
M. Marcel Van Herk	examineur
M. Serge Miguet	directeur
M. David Sarrut	co-directeur
Mme. Chantal Ginestet	invitée
M. Francois Spriet	invité

Contact : simon.rit@gmail.com

## Résumé

Une image tomодensitométrique (TDM) représente la carte 3D des coefficients d'atténuation linéaire d'un faisceau de rayons X. Elle est obtenue par reconstruction à partir de projections acquises sous différentes incidences autour du patient supposé statique. Si le patient respire et le mouvement n'est pas pris en compte, les images TDM du thorax sont perturbées par d'importants artefacts, tels que du flou, des traits et des bandes. L'objectif de cette thèse est de proposer des méthodes de correction et de les appliquer dans le cadre de la radiothérapie du cancer du poumon, à partir de séquences de projections acquises en salle de traitement par un faisceau à géométrie conique embarqué sur un accélérateur linéaire. La première méthode envisagée s'appuie sur un signal respiratoire représentatif de la position du patient dans le cycle respiratoire et permettant de sélectionner pour la reconstruction les projections correspondant à une même phase. Pour l'appliquer, nous avons mis au point une méthode d'extraction automatique du signal respiratoire depuis les projections, sans appareil externe. Par ailleurs, une étude quantitative comparative a été menée sur données simulées pour évaluer l'impact de l'algorithme de reconstruction et des différents paramètres de sélection des projections. Nous obtenons ainsi des images TDM sans flou mais d'une qualité limitée par l'utilisation d'un faible nombre de projections. D'autres approches modifient l'algorithme de reconstruction pour compenser le mouvement respiratoire à partir d'un modèle réaliste quelconque supposé connu, ce qui permet d'utiliser toutes les projections acquises. Nous avons proposé deux méthodes de reconstruction avec compensation du mouvement. La première est analytique mais basée sur une heuristique. Elle permet l'élimination du flou mais laisse des artefacts de traits et bandes. La seconde résout le problème algébriquement par une méthode itérative. La composition des transformations en jeu, le mouvement respiratoire et la projection conique, est modélisée discrètement via deux nouvelles approches, l'une avant et l'autre arrière. L'intégralité des artefacts est alors visuellement éliminée. Les contributions ont été évaluées et validées grâce à une plateforme comprenant des séquences de projections réelles et simulées. Les résultats permettent d'envisager l'obtention d'images TDM 4D avec peu d'artefacts sur lesquelles le praticien peut se baser à chaque séance pour obtenir une information anatomique précise. Il faudra cependant résoudre les problématiques laissées ouvertes liées au temps de calcul et au modèle de mouvement.

## Abstract

A computed tomography (CT) represents the 3D map of the linear attenuation coefficients of a X-ray beam. If the patient breath and the motion is not taken into account, the CT images of the thorax are disturbed by

strong artifacts such as blur, streaks and bands. The objective of this thesis is to propose methods to correct these artifacts and to apply them in the context of the radiotherapy of the lung cancer to sequences of cone-beam projections acquired with a CT scanner mounted on the gantry of a linear accelerator. The first method uses a respiratory signal to select for the reconstruction the projections corresponding to a same phase. To apply it, we proposed a method to extract automatically the respiratory signal from the cone-beam projections. A quantitative analysis was then performed on simulated data to evaluate the impact of the reconstruction algorithm and of the different selection parameters of the cone-beam projections. We obtain thus CT images without blur but with a quality limited due to the small number of projections used. Other approaches modify the reconstruction algorithm to compensate for the respiratory motion using a realistic model, which allows to use all the acquired projections. We proposed two motion compensated reconstruction methods. The first one is analytic and based on a heuristic. The second one solves the problem algebraically from a discrete formulation of the transformations at stake via two new approaches, one forward and the other backward.

## Publications

- Chaieb, E.; Rit, S.; Giroud, E.; Ayadi, M.; Ginestet, C. & Sarrut, D. (2006), 'Evaluation des différents protocoles pour l'acquisition d'images 4D sur un scanner non prévu a cet effet' Journées Scientifiques de la Société Française de Physique Médicale (SFPM)', Lyon, France.
- Desbat, L.; Rit, S.; Clackdoyle, R.; Mennessier, C.; Promayon, E. & Ntalampeki, S. (2007), 'Algebraic and analytic reconstruction methods for dynamic tomography' IEEE Engineering in Medicine and Biology Conference (EMBC)', Lyon, France, 726--730.
- Rit, S. (2007), 'Prise en compte du mouvement respiratoire pour la reconstruction d'images tomodensitométriques', PhD thesis, Université Lumière Lyon 2.
- Rit, S. & Sarrut, D. (2007), 'Cone-beam projection of a deformable volume for motion compensated algebraic reconstruction' Conf Proc IEEE Eng Med Biol Soc', Technical report, LIRIS, Université Lumière Lyon 2, 5 Av. P. Mendès-France, 69676 Bron, France. simon.rit@gmail.com, Lyon, France, 6544--6547.
- Rit, S.; Sarrut, D.; Boldea, V. & Ginestet, C. (2006), 'Extraction du signal respiratoire a partir de projections cone-beam', Traitement du signal 23(3--4), 307--319.
- Rit, S.; Sarrut, D.; Boldea, V. & Ginestet, C. (2005), 'Extraction du signal respiratoire a partir de projections cone-beam du thorax pour l'imagerie TDM 4D en radiothérapie' Journées Scientifiques de la Société Française de Physique Médicale (SFPM)', Avignon, France.
- Rit, S.; Sarrut, D.; Boldea, V. & Ginestet, C. (2005), 'Extraction du signal respiratoire a partir de projections cone-beam du thorax pour l'imagerie TDM 4D en radiothérapie' Congrès de la Société Française de Radiothérapie Oncologique (SFRO)', Paris, France.
- Rit, S.; Sarrut, D. & Desbat, L. (In press), 'Comparison of analytic and algebraic methods for motion-compensated cone-beam CT reconstruction of the thorax', IEEE Transactions on Medical Imaging.
- Rit, S.; Sarrut, D. & Ginestet, C. (2007), 'Comparison of gated and dynamic cone-beam CT reconstruction methods' International Conference on the Use of Computers in Radiation Therapy (ICCR)', Toronto, Canada.
- Rit, S.; Sarrut, D. & Ginestet, C. (2005), 'Respiratory signal extraction for 4D CT imaging of the thorax from cone-beam CT projections' Medical Image Computing and Computer-Assisted Intervention (MICCAI)', Palm Springs, USA, 556--563.
- Rit, S.; Sarrut, D. & Ginestet, C. (2005), 'Respiratory signal extraction for 4D CT imaging of the thorax from cone-beam CT projections', Int J Radial Oncol Biol Phys 'Proceedings of the 47th Annual ASTRO Meeting', S533-S534.
- Rit, S.; Sarrut, D. & Miguet, S. (2007), 'Gated cone-beam CT imaging of the thorax: a reconstruction study' SPIE Medical Imaging', San Diego, California, USA.

**Segovia Benjamin**

Soutenu le 26/10/2007

à l'Université Claude Bernard Lyon1

### Jury

G. DRETTAKIS, INRIA,	rapporteur
N. HOLZSCHUCH, INRIA,	rapporteur
J.M. JOLION, INSA Lyon,	président du jury
M. PAULIN, Université Paul Sabatier,	examineur
R. LABAYRADE , ENTPE	invité
J.C. IEHL, université Lyon 1,	co-directeur de thèse
B. PEROCHE, université Lyon 1,	directeur de thèse

Contact : [bernard.peroche@liris.cnrs.fr](mailto:bernard.peroche@liris.cnrs.fr)

### Abstract

Since about twenty years, the field of computer graphics has been striving towards increased realism. The goal is to achieve photorealistic outputs, which actually involves to precisely describe lighting phenomena, their intrinsic models and the light transport equations and to finally propose numerical schemes to solve them. In the field of off-line rendering, efficient and aggressive numerical methods have already been explored and set up. Therefore, even if there is still room for improvement, the most challenging and appealing problems are today related to real-time rendering which consists in ensuring to display at least 20 frames per second with some interactivity offered to the user.

In this Ph.D. thesis, I therefore focus on the real-time or, if not possible, interactive simulation of light transport phenomena. During the three years while I tried to propose new ideas, I thus mainly aimed at efficiently using current existing hardware systems, GPUs or CPUs and designing new numerical schemes which may be easily implemented in such systems. It is finally the most common role of numericians: making a subtle mixture of mathematical and algorithmic methods and efficiently combining them.

### Publications

- B. Segovia, J.C. Iehl and B. Péroche, Metropolis Instant Radiosity, Computer Graphics Forum (Proceedings of Eurographics'07), 26(3): 425--434, 2007.
- B. Segovia, J.C. Iehl, R. Mitanchey and B. Péroche, Bidirectional Instant Radiosity, Rendering Techniques 2006: 389--398, 2006.
- B. Segovia, J.C. Iehl, R. Mitanchey and B. Péroche, Non-interleaved Deferred Shading of Interleaved Sample Pattern}, SIGGRAPH/Eurographics Workshop on Graphics Hardware'06: 53--60, 2006.



# Représentations hiérarchiques et discriminantes pour la reconnaissance des formes, identification des personnes et analyse des mouvements dans les séquences d'images

**Thome Nicolas**

**THESE**

Soutenue le 11/07/2007

à l'Université Lumière Lyon 2

## Jury

Jenny Benois-Pineau, Professeur, Université Bordeaux 1

Jean-Baptiste Ducatez, Industriel, Foxstream

Jean-Michel Jolion, Professeur, INSA de Lyon

Serge Miguet, Professeur, Université Lyon 2

Patrick Pérez, INRIA, Rennes

directeur de recherche

Monique Thonnat, INRIA, Sophia-Antipolis

directeur de recherche

Contact : serge.miguet@liris.cnrs.fr

## Résumé

Trois contributions principales ont été proposées pour développer un système de lecture de plaques minéralogiques capable de fonctionner de manière robuste dans des situations réelles d'analyse. Tout d'abord, une méthode combinant les avantages des approches statistiques et structurelles a été mise en place pour la reconnaissance optique de caractères. D'autre part, l'aspect de temporel de la vidéo est explicitement modélisé, en utilisant une distance de Levenstien probabiliste pour moyenniser les résultats de reconnaissance. Enfin, des boucles de rétro-action sont introduites aux points sensibles de l'algorithme, pour accroître la précision du système tout en permettant une reconnaissance indépendante de la nationalité.

Une méthode de suivi visuel d'humain a été proposée, consistant à apprendre l'apparence des personnes en ligne dans la vidéo, afin de contraindre la mise en correspondance et rendre le suivi efficace dans des situations complexes (occlusions, personnes sortant du champ de vue, etc). Ainsi, un modèle d'apparence articulé est mis à jour pour chaque personne suivie, en vue d'une identification ultérieure. La formation d'un descripteur pour chaque membre rend l'étape de reconnaissance plus efficace qu'avec les modèles d'apparence globaux existants. Elle nécessite par ailleurs l'étiquetage des membres individuels à partir de la silhouette, qui est assuré par une technique de mise en correspondance de graphes adaptée des Shock Graphs. La mise en correspondance entre un modèle topologique de la structure articulée humaine et les segments extraits du squelette de l'image assure une identification des membres indépendante du point de vue ou de la pose de la personne, ce qui permet d'apprendre l'apparence dans des configurations très génériques.

L'interprétation du mouvement, appliqué à la détection de chutes, est assurée par approche temps réel et multi-vues. Partant d'un unique descripteur de verticalité image, il est montré que 2 caméras suffisent dans les conditions standard d'analyse pour assurer une reconnaissance correcte des poses, laquelle est effectuée par une fusion d'informations utilisant des opérateurs de logique floue. La détection de chutes est alors formulée comme un problème d'inférence sur de la séquence temporelle des poses, utilisant un modèle de Markov caché hiérarchique (LHMM), qui offre la possibilité de détecter conjointement des mouvements brusques comme les chutes, tout étant robuste aux erreurs de bas niveau.

## Abstract

The field of Looking-at-People has known a huge increase over the last decade. However, developing robust algorithms to analyse image sequences remains an open problem. In this Phd, the problem has been

tackled in the following aspects : Licence Plate Recognition and Human Motion Analysis, in a monocular and real time context. The proposed approach for Licence Plate Recognition decomposes into the following steps : plate detection, binarization, character segmentation, shape classification and temporal averaging.

Regarding visual analysis of human motion, we can classify the approaches in two main groups : Bottom Up and Top Down approaches. Our approach can be considered intermediate with respect to the former archetypes. Firstly, we extract moving regions from each image and track them over time using a simple dynamical matching approach. This step is applicable to any type of object and is very fast. Additionally, it makes it possible to detect split and merge situations. When the tracking can be successfully achieved by this strategy, we analyse the shape silhouette at a finer scale, aiming at detecting and labelling body parts. We propose to perform this task by matching an image graph generated from the skeleton silhouette and a model graph derived from the articulated human body structure. Only morphological and topological information is encoded in the graph, making the approach independent of the pose of the person, the viewpoint, the geometry or the appearance of the limbs. The body parts labelling makes it possible to update an appearance model for each limb, capturing shape, color and texture properties. The appearance model is used in difficult situations to identify people.

Regarding human motion analysis, we propose a real-time and multi-view approach dedicated to fall detection. We prove that two cameras are sufficient to discriminate standing from lying poses. Then, we make use of a Layered Hidden Markov Model (LHMM) to classify the pose sequences. The analysis is performed at two different temporal scales, making it possible to detect sudden changes as falls, as well as being robust to noise in fast motions by incorporating high level verification.

## Publications

- A Real-Time, Multi-View Fall Detection System : a LHMM-Based Approach; Nicolas Thome , Serge Miguet et Sébastien Ambellouis, IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Event Analysis in Videos, Volume 18, Issue 11, p.1522–1532, ISSN : 1051-8215, November 2008.
- Learning Articulated Appearance Models for Tracking Humans : a Spectral GraphMatching Approach, Nicolas Thome , Djamel Mérad et Serge Miguet, Processing : Image Communication, vol. 23, issue 10, .769–787, December 2008.
- A Combined Statistical-Structural Strategy for Alphanumeric Recognition, Nicolas Thome et Antoine Vacavant, 3rd International Symposium on Visual Computing- Lake Tahoe, Nevada, California, November 26–28 2007. LNCS, Springer, volume 4842, pp. 529–538.
- A HHMM-Based Approach for Robust Fall Detection, Nicolas Thome et Serge Miguet, International Conference on Control, Automation, Robotics and Vision (IEEEICARCV) - Grand Hyatt, Singapore, 5–8 December 2006, pp. 1–8. ISBN : 1-4244-0342-1.
- Human Body Part Labeling and Tracking Using GraphMatching Theory, Nicolas Thome , Djamel Merad et Serge Miguet, International Conference on Advanced Video and Signal based - Sydney, 21–24 November 2006, pp. 32–38. ISBN : 0-7695-2688-8.
- A Robust AppearanceModel for Tracking HumanMotions, Nicolas Thome et Serge Miguet, Conference on Advanced Video and Signal based Surveillance (AVSS) - Como, 15–16 September 2005, pp. 528–533. ISBN : 0-7803-9385-6.

2008

Département  
Données, Connaissances, Services





Soutenue le 26/11/2008

à l'Université Claude Bernard, Lyon1

#### Jury

Pr.Elisa Bertino (Purdue University, USA)  
Pr.Nicole Bidoit-Tollu (Université Paris XI)  
Pr.Claude Godard (Université de Nancy 1))  
Pr.Mohand Saïd Hacid (Université Claude Bernard Lyon1)  
Pr.Gaetan Hains (Université Paris 12)  
Pr.Marco Pistore (IST Trento, Italy)

Contact : sbenbern@liris.univ-lyon1.fr

#### Résumé

Comment trouver des informations pertinentes (réponses aux requêtes) dans un document, une bases de données, ou sur le Web? c'est un problème qui est largement encours d'investigation. Plusieurs méthodes et algorithmes ont été proposés. Cependant, des efforts supplémentaires sont requis afin de prendre en compte les aspects sémantiques des données et requêtes. Mes recherches concernent principalement la conception de frameworks où les aspects sémantiques sont considérés dans la modélisation, la requête et le contrôle de l'accès aux données. A cet effet, je me suis intéressée aux problèmes d'accès à différents types de données et proposant des raisonnements tels que: classement de documents lors de la comparaison sémantique de leur contenu, recherche sémantique de contenu multimédia, découverte dynamique de mapping dans un environnement pair à pair, découverte sémantique de services web pour la composition automatique de services web et le partage de business processes moyennant les techniques du Web 2.0.

En outre, lors de la prise en compte de données dans le domaine du calcul orienté service (SOC), il y a une variété de données personnelles collectées à partir des clients tels que les propriétaires de voitures, patients, et utilisées par des fournisseurs de services. Comme le mauvais usage de données personnelles est en croissance, la privacy est devenue un problème crucial pour les fournisseurs de services et les demandeurs de services et des législateurs. A cet effet, je me suis intéressée aux problèmes de la confidentialité dans l'interaction des services et dans le contrôle de l'usage des données privées.

#### Abstract

How to find relevant information (query answers) in a document, a database, or on the Web ? This is an issue which is extensively under investigation. Several methods and algorithms have been proposed. However, additional efforts are required in order to accommodate semantic aspects of data and queries. Our research mainly deals with the design of frameworks where semantics can be considered in the process of modeling, querying and controlling access to data. We investigated the data semantic issues to access different types of data and provide some reasoning services : ranking documents when comparing semantically their contents, semantic retrieval of multimedia content, mapping discovery in a peer to peer environment, semantic discovery of web services to automate composition of distributed process, semantic and sharing of personal business processes.

Furthermore, when handling data in the area of service oriented computing, there is a large amount of personal data that are collected from clients/individuals, such as taxpayers, car owners, patients, benefit recipients, and processed by government and other public service providers. As the number of inappropriate usage and leakage of personal data is increasing, privacy concerns is becoming one of the most important concerns of service requesters, service providers and legislators. In this setting, we

addressed the privacy issue in services interactions and a way to handle and control the private data usage.

## Publications

- V. Andrikopoulos, S Benbernou, M.P. Papazoglou. "Managing the Evolution of Service Specifications", The 20th International Conference on Advanced Information Systems Engineering (CAiSE'08)), Montpellier. LNCS 5074 Springer 2008.
- S Hammiche, L.B Lopez, S Benbernou, M. Hacid. " Query Rewriting for Semantic MultimediaData Retrieval.", Studies in Computational Intelligence 116/2008():351-372, Springer. 2008.
- S Benbernou, H. Meziane, M. Hacid. Run Time Monitoring for Privacy Compliance. International Conference on Service Computing, ICSOC' 07, Asit Dan, Schahram Dustdar, Bernd Krämer ed. Vienna, Austria. pp. 351-364, LNCS 4749. Springer-Verlag 2007.
- k. Mokhtari, S Benbernou, M. Said, E Coquery, M. Hacid, F. Leyman. "Verification of Privacy Timed Properties in Web Service Protocols.", 2008 IEEE international Conference on Service Computing, SCC'08 IEEE ed. Honolulu.
- S Hammiche, L.B Lopez, S Benbernou, M. Hacid, A. Vakali. "Domain Knowledge Based Queries for Multimedia Data Retrieval. " Journal of Digital Information Management 5(2):75-81 2007.
- N. Guermouche, S Benbernou, E Coquery, M. Hacid. Privacy-Aware Web Service Protocol Replaceability. IEEE International Conference on Web Services ICWS'07, Liang Jie Zhang Kenneth Birmann, Jia Zhang ed. Salt Lake City, USA. pp. 1048-1056.
- S Benbernou, M. Hacid Dynamic Web Service Call for Data Integration. JOURNAL OF SOFTWARE (JSW) ISSN : 1796-217X Volume : 1 Issue : 1 pp:1-10 Date : July 2006,
- S.Benbernou, M.S.Hacid. "Resolution and constraint propagation for semantic web services discovery". Distributed and Parallel Databases International Journal 18(1):65-81, Kluwer Academic. 2005.

# Apprivoiser la vie, modélisation individu-centrée de systèmes biologiques complexes

HDR

Beslon Guillaume

Soutenue le 14/11/2008

à l'INSA de Lyon

## Jury

Gilles Bernot (Université de Nice)  
Paul Bourguin (Ecole Polytechnique)  
Joël Favrel (INSA Lyon)  
Jean-Michel Fayard (INSA Lyon)  
François Képès (Génopole Evry)  
Alain Mille (UCBL)  
Michel Morange (ENS, Paris)  
Michel Morvan (ENS-Lyon, EHES)

Contact : guillaume.beslon@iris.cnrs.fr

## Résumé

Les systèmes biologiques sont des systèmes complexes au sens où ils sont composés d'un très grand nombre d'éléments hétérogènes en interaction et que la dynamique de ces interactions donne naissance à une unité structurale. L'objectif de la biologie des systèmes est d'identifier les mécanismes structurant par lesquels cette unité structurale émerge de la multitude des interactions. Pour cela, nous proposons des approches de modélisation individu-centrée destinées à produire des hypothèses explicatives.

Nous présentons ici les travaux réalisés dans trois domaines de la biologie : les neurosciences computationnelles, la biologie évolutive et la modélisation cellulaire. Dans les trois cas, nous présentons les questionnements abordés, les modèles développés et les hypothèses que nous avons pu produire.

## Abstract

Biological systems are complex systems because they are made up of a large number of interacting heterogeneous components. While interacting, these components create a structural unity that may be, for instance, a cell, an organism or a species. The central objective of systems biology is to identify the mechanisms by which the interactions between a biological system's components lead to the emergence and maintenance of such a global structural unity. We propose here to use individual-based modeling in order to explore the behavior of complex biological systems and to produce new hypothesis.

We present the models we have developed in the field of computational neurosciences, evolutionary biology and cell models. In these three fields, we present the scientific questions we address, the models we developed and the main results we obtained.

## Publications

- Modelling Evolution of Regulatory Networks in Artificial Bacteria. Y. Sanchez-Dehesa, D.P. Parsons, J.M. Pena, G. Beslon. *Mathematical Modelling of Natural Phenomena (MMNP)* 3(2):27-66, Vitaly Volpert, ISSN 0973-5348. 2008.
- The Topology of the Protein Network Influences the Dynamics of Gene Order : From Systems Biology to a Systemic Understanding of Evolution. C. Knibbe, J.M. Fayard, G. Beslon. *Artificial Life* 14(1):149-156. 2008.
- Large multi-protein structures modeling and simulation: the need for mesoscopic models. A Coulon, G. Beslon, O Gandrillon. *Methods in Molecular Biology series*, Humana press 2008.
- A Long-Term Evolutionary Pressure on the Amount of Noncoding DNA. C. Knibbe, A Coulon, O. Mazet,

J.M. Fayard, G. Beslon. *Molecular Biology and Evolution* 24(10):2344-2353. 2007.

- Evolutionary coupling between the deleteriousness of gene mutations and the amount of non-coding sequences. C. Knibbe, O. Mazet, F. Chaudier, J.M. Fayard, G. Beslon. *Journal of Theoretical Biology* 244(4):621-630, Elsevier, ISSN 0022-5193. 2007.
- Modélisation cellulaire pour l'émergence de structures multiprotéiques auto-organisées. A Coulon, H. Soula, O. Mazet, O Gandrillon, G. Beslon. *Technique et Science Informatiques* 26(1-2):123-148. 2007.
- From artificial evolution to computational evolution: a research agenda. W. Banzhaf, G. Beslon, S. Christensen, J. Foster, F. Képès, V. Lefort, J.F. Miller, M. Radman, J. Ramsden. *Nature Review Genetics (NRG)* 7(9):729-735, ISSN 1471-0056. 2006.
- Spontaneous Dynamics of Asymmetric Random Recurrent Spiking Neural Networks. H. Soula, G. Beslon, O. Mazet. *Neural Computation* 18(1):60-79, MIT Press. 2006.
- Modeling the emergence of multi-protein dynamic structures by principles of self-organization through the use of 3DSpi, a Multi-Agent-based Software. H. Soula, C. Robardet, F. Perrin, S. Gripon, G. Beslon, O Gandrillon. *BMC Bioinformatics* 6(228). 2005.

# Découverte et composition dynamique de services web par approche multi-agents

**Bova Rosanna**

**THESE**

Soutenue le 06/06/2008

à l'Université Claude Bernrad Lyon 1

## Jury

Vincenzo D'Andrea, Université de Trento, Italie	rapporteur
- Farouk Toumani, Université Blaise Pascal de Clermont Ferrand	rapporteur
- Boualem Benatallah, UNSW de Sydney, Australia	examineur
- Omer Rana, Université de Cardiff, Grande-Bretagne	examineur
- Salima Hassas, Université Lyon 1 de recherche	directrice
- Salima Benbernou, Université Lyon 1	co-directrice

Contact : salima.benbernou@liris.cnrs.fr

## Résumé

Les activités typiques d'une structure qui réalise de la composition de services Web sont : la définition de la structure d'un service Web composite, la sélection de services Web courants à lier au service Web composite, l'orchestration de l'exécution d'un service Web composite. Beaucoup de travaux de recherche existent dans le domaine général de la découverte, sélection et composition de services Web. En général, ces approches se basent sur la description de techniques de matching. Dans cette thèse, on présente un approche qui est complémentaire à tous les approches classiques sur la sélection de services. En particulier, on propose d'ajouter la connaissance sur comment l'activité de sélection de services Web a été conduite dans le passé. Cette connaissance est utilisée pour améliorer l'efficacité de cette activité même. Pour cet objectif, on définit le concept de Mémoire de la Tâche. La mémoire de la tâche représente la connaissance sur les configurations de services et le contexte dans lequel ces configurations ont été considérées les plus appropriées par les utilisateurs. En s'inspirant de la nouvelle vague apportée par le Web 2.0 qui favorise un nouveau paradigme dans lequel les fournisseurs et les utilisateurs finals (y compris les utilisateurs pas experts) facilement et librement partagent de l'information et de services sur le Web, nous introduisons le concept de Forum de Tâches. Essentiellement, un forum de tâches offre un moyen de collecter et partager des définitions de tâches et des mémoires de tâches, d'un domaine spécifique, entre les utilisateurs. Cela permet aux utilisateurs de réutiliser et personnaliser des définitions partagées au lieu de formuler des définitions de tâches à partir de zéro. Un nouveau prototype, appelé YouMaché, a été implémenté pour supporter le partage des tâches et la sélection de services en utilisant les forums de tâches et les mémoires de tâches. L'implémentation de ce prototype s'appuie sur les données des services et des agents logiciels.

## Abstract

The typical activities of a Web service composition framework include : defining the structure of a composite Web service; selecting actual Web services to bind to a composite Web service; orchestrating a composite Web service execution; and handling runtime exceptions. A large body of research exists in the general area of Web services discovery, selection and composition. Generally, these approaches rely on the description of matching techniques (e.g., whether descriptions of services and requests are compatible). Descriptions refer to meta-data such as service capabilities and non-functional properties (e.g., quality of service properties). In this dissertation, we present an approach complementary to all classical Web service selection approaches. We first propose to add on top of Web service composition frameworks the knowledge on how the activity of Web service selection is carried out in the past. Then, to use this knowledge to improve the effectiveness of this activity. For this purpose, we introduce the concept of task memory. A task memory represents knowledge about the selected service configurations and the context

in which these configurations are considered most appropriate by users. By leveraging the emerging wave of innovations in Web 2.0 that promotes a new paradigm in which both providers and end-users (including non-expert users) can easily and freely share information and services over the Web, we introduce the concept of task forum. Essentially, a task forum provides a means to collect and share domain specific task definitions and task memories among users. This allows users to reuse and customize shared definitions instead of developing definitions of tasks from scratch. In addition, a task forum uses a publish and subscribe interaction model to support massive service selection recommendations. The techniques presented in this dissertation are implemented by reusing existing service composition techniques to support the definition and execution of composite services. A new prototype called YouMach e is implemented to support tasks sharing and service selection using tasks forums and task memories. The implementation of this prototype relies on data services and software agents.

## Publications

- WS-Advisor: A Task Memory for Service Composition Frameworks.. R Bova, P.H-Y Paik, S Hassas, S Benbernou, B. Benatallah. Dans BDA (23èmes journées de bases de données avancés), Marseille. 2007.
- Task Memories & Task Forums: A Foundation for Sharing Service-based Personal Processes. R Bova, P.H-Y Paik, B. Benatallah, L. Zeng, S Benbernou. Dans International Conference on Service Oriented Computing (ICSOC 2007) , Asit Dan, Schahram Dustdar, Bernd Krämer ed. Vienna, Austria. pp. 365-376. LNCS 4749. Springer Verlag . 2007
- WS-Advisor: A Task Memory for Service Composition Frameworks. R Bova, P.H-Y Paik, S Hassas, S Benbernou, B. Benatallah. Dans International Conference on Computer Communications and Networks (ICCCN'07), IEEE ed. Honolulu, Hawaii USA. 2007.
- On Embedding Task Memory in Services Composition Frameworks. R Bova, P.H-Y Paik, S Hassas, S Benbernou, B. Benatallah. Dans ICWE'07 (International Conference on Web Engineering), Piero Fraternali, Luciano Baresi, Geert-Jan Houben, ed. Como (Italy). LNCS 4607. Springer Verlag . 2007.
- An Immune System-Inspired Approach for Composite Web Services Reuse. R Bova, S Hassas, S Benbernou. (Workshop "AI for Service Composition" (ECAI 06)), , Trento (Italy). 8p. 2006.
- Réutilisation de services web composites par la métaphore du système immunitaire. R Bova, S Hassas, S Benbernou. (Atelier "Conception des systèmes d'information et services Web" (INFORSID 06)), , Hammamet (Tunisie). 12p. 2006.

# Interactive and Opportunistic Knowledge Acquisition in Case-Based Reasoning

Cordier Amélie

THESE

Soutenue le 13/11/2008

à l'Université Claude Bernard, Lyon 1

## Jury

Eyke Hüllermeier

examineur

Agnar Aamodt

rapporteur

David Leake

rapporteur

Sylvie Després

examineur

Jean Lieber

examineur

Alain Mille

directeur

Béatrice Fuchs.

directrice

Contact : amelie.cordier@liris.cnrs.fr

## Résumé

Jeune discipline à la croisée de l'informatique, de l'intelligence artificielle et des sciences cognitives, l'ingénierie des connaissances vise à modéliser les connaissances d'un domaine pour les opérationnaliser dans un outil informatique. Pour cela, elle propose des outils théoriques, des modèles et des méthodologies empiriques pour instrumenter les systèmes et permettre l'échange de connaissances entre l'utilisateur et l'outil informatique. Le travail développé ici traite de l'ingénierie des connaissances d'une catégorie de systèmes en particulier : les systèmes de raisonnement à partir de cas (RàPC). Un système de RàPC assiste un utilisateur dans sa tâche de résolution de problème en lui proposant une adaptation à la situation courante d'une précédente expérience. C'est en particulier au système en interaction "utilisateur - outil de RàPC" que nous nous intéressons ici.

La problématique étudiée peut donc être exprimée ainsi : quelles méthodes et outils développer pour instrumenter efficacement le système apprenant "utilisateur - outil de RàPC" ? Cette problématique soulève un questionnement sur les connaissances du raisonnement et conduit à une analyse au niveau connaissance de tels systèmes. Un autre volet d'analyse porte sur les interactions entre l'utilisateur et l'artefact informatique pendant les phases de résolution de problème. Ces aspects sont étudiés à plusieurs niveaux dans les différentes contributions présentées dans cette thèse. Nos différentes expériences et expérimentations nous ont conduits à proposer, comme première contribution, une formalisation à un niveau général de l'apprentissage interactif de connaissances en RàPC (FIKA). Cette formalisation repose sur les échecs de raisonnement qui, puisqu'ils permettent de mettre en évidence les lacunes dans les connaissances disponibles, sont utilisés pour guider le processus d'apprentissage. Deux extensions de ce modèle général ont été proposées : IakA et FrakaS. IakA raffine les principes proposés par FIKA pour permettre leur mise en œuvre immédiate dans une certaine catégorie de systèmes où les connaissances peuvent être représentées selon un modèle donné (cas et connaissances d'adaptation représentées par des opérateurs d'adaptation). Ces principes ont été implantés et expérimentés dans une application développée à des seules fins expérimentales. FrakaS propose des méthodes et outils similaires pour une autre catégorie de systèmes où les connaissances du domaine sont utilisées pour guider l'adaptation. Ces principes ont, quant à eux, été implantés dans un prototype inspiré d'une application réelle.

IakA et FrakaS, les deux extensions de l'approche FIKA, présentent des forces et des limites, une réflexion quant à leur intégration possible a donc été menée. Une première tentative pratique d'intégration a été mise en œuvre dans une application de RàPC permettant l'adaptation de recettes de cuisine : le logiciel Taaable.

## Abstract

As a young discipline at the junction of computer science, artificial intelligence and cognitive sciences, knowledge engineering aims at modelling knowledge of a specific domain to operationalise them in a computer system. To this end, it offers theoretical tools, models and empirical methodologies to support knowledge sharing between the user and the system. The work developed here is related to knowledge engineering of a particular type of system: case-based reasoning systems (CBR). A CBR system assists a user in his problem solving task by retrieving a previous successful problem solving experience and by adapting it to the current situation. In this work, we are mainly interested in the interacting system “user - CBR tool”.

The main research question we address here can be formulated as: what methods and tools have to be developed to support knowledge acquisition in the learning system “user - CBR tool”. This issue raises the question of the knowledge of the reasoning process and leads to an analysis at the knowledge level of CBR systems. Another part of the analysis aims at studying the interactions between the user and the CBR tool during the problem solving phases. These aspects are studied at several levels in the different contributions presented in this thesis. Our different experiences and experiments lead us to propose, as a first contribution, a formalisation at general level of interactive knowledge learning in CBR (FIKA). This formalisation relies on the reasoning failures which, as they allow to highlight the gaps in the available knowledge, are used to guide the learning process. Two extensions of this general model have been proposed: IakA and FrakaS. IakA refines the principles proposed in FIKA to permit their immediate implementation in a particular type of system where knowledge can be represented according to a given model (cases and adaptation knowledge in the form of adaptation operators). These principles have been implemented and experimented with in an application developed exclusively for this purpose. FrakaS proposes similar methods and tools for another type of system where domain knowledge is used to guide adaptation. As for these principles, they have been implemented in a prototype inspired by a real world application.

We have conducted a study of strengths and limits of FrakaS and IakA and we have investigated possible ways to combine them. A first practical implementation has been made in a CBR application allowing the adaptation of cooking recipes, the project Taaable.

## Publications

- Opportunistic Acquisition of Adaptation Knowledge and Cases - The IakA approach. Amélie Cordier, Béatrice Fuchs, Leonardo Lana de Carvalho, Jean Lieber, Alain Mille. European Conference on Case-Based Reasoning (ECCBR), Trier, Germany. pp. 150-164. 2008.
- Taaable: Text Mining, Ontology Engineering, and Hierarchical Classification for Textual Case-Based Cooking. Fadi Badra, Rokia Bendaoud, Rim Bentebitel, Pierre-Antoine Champin, Julien Cojan, Amélie Cordier, Sylvie Després, Stéphanie Jean-Daubias, Jean Lieber, Thomas Meilender, Alain Mille, Emmanuel Nauer, Amedeo Napoli, Yannick Toussaint. Computer Cooking Contest Workshop at European Conference on Case-Based Reasoning (ECCBR), Trier, Germany. pp. 219-228. 2008.
- Interactive Knowledge Acquisition in Case Based Reasoning. Amélie Cordier, Béatrice Fuchs, Jean Lieber, Alain Mille. Knowledge Discovery and Similarity in Case-Based Reasoning workshop. International Conference on Case-Based Reasoning, David Wilson, Deepak Khemani ed. Belfast. pp. 85-94. 2007.
- On-Line Domain Knowledge Management for Case-Based Medical Recommendation. Amélie Cordier, Béatrice Fuchs, Jean Lieber, Alain Mille. Case Based Reasoning in the Health Sciences workshop. International Conference on Case-Based Reasoning, David Wilson, Deepak Khemani ed. Belfast. pp. 285-294. Belfast. 2007.
- Failure Analysis for Domain Knowledge Acquisition in a Knowledge-Intensive CBR System. Amélie Cordier, Béatrice Fuchs, Jean Lieber, Alain Mille. International Conference on Case-Based Reasoning - ICCBR'07. LNAI 4626, Springer-Verlag. pp. 463-477. Rosina Weber and Michael Richter ed. Belfast. 2007.
- Engineering and Learning of Adaptation Knowledge in Case-Based Reasoning. Amélie Cordier, Béatrice Fuchs, Alain Mille. 15th International Conference on Knowledge Engineering and Knowledge Management EKAW'06, S. Staab and V. Svatek ed. Podebrady, Czech Republic. pp. 303-317. LNAI 4248. Springer-Verlag Berlin Heidelberg. 2006.



Soutenue le 26/05/2008

à l'Université Joseph Fourier – Grenoble 1

## Jury

Mme Monique Grandbastien (Université Nancy 1)	rapporteur
Mr Pascal Leroux (Université du Maine)	rapporteur
Mr Jean-Marc Labat (Université Paris 6)	président
Mr Alain Mille (Université Lyon 1)	examinateur
Mr Jean-Pierre Peyrin (Université Grenoble 1)	directeur
Mme Stéphanie Jean-Daubias (Université Lyon 1)	co-directeur
Mr Jean-Philippe Pernin (Institut National de Recherche Pédagogique)	co-directeur

Contact : carole.eyssautier@liris.cnrs.fr

## Résumé

Cette thèse en informatique se place dans le domaine des Environnements Informatiques pour l'Apprentissage Humain. Durant une activité d'apprentissage, de nombreuses informations peuvent être recueillies et interprétées pour constituer un profil, en termes de connaissances, compétences, conceptions ou comportements d'un apprenant ou d'un groupe d'apprenants. Ces profils peuvent différer de part leur intention pédagogique, leur contenu, leur structuration, leur présentation. Par ailleurs, les acteurs de la situation d'apprentissage concernés par ces profils sont divers.

Il n'existe pas à l'heure actuelle de solution technique permettant de réutiliser ces profils hétérogènes. Cette thèse cherche donc à proposer des modèles et des outils permettant la réutilisation pour les différents acteurs de profils d'apprenants créés par d'autres.

Dans nos travaux, nous proposons le modèle de processus de gestion de profils REPro (Reuse of External Profiles). Pour permettre la réutilisation de profils hétérogènes, nous proposons de les réécrire selon un formalisme commun qui prend la forme d'un langage de modélisation de profils, le langage PMDL (Profiles MoDeling Language). Nous définissons ensuite un ensemble d'opérateurs permettant la transformation des profils ainsi harmonisés, ou de leur structure, tels que l'ajout d'éléments dans le profil, ou la création d'un profil de groupe à partir de profils individuels. Ces propositions ont été mises en œuvre au sein de l'environnement EPROFILEA du projet PERLEA (Profils d'Élèves Réutilisés pour L'Enseignant et l'Apprenant), avant d'être mises à l'essai auprès d'enseignants en laboratoire

## Abstract

This PhD thesis in computer science belongs to the computer supported learning field. Various information on the learner, or on a group of learner, related to the learning – knowledge, skills, conceptions or behavior – are collected to compose a profile. Existing profiles differ in their intent, content or structure. Moreover, various learning actors are involved and concerned by these profiles.

As of today, there is no technical solution in place for reusing these heterogeneous profiles. This PhD thesis aims at providing models and tools allowing the various actors to reuse learners' profiles created by others.

In this work, we propose the REPro model (Reuse of External Profiles) to represent the profiles management process. To allow reuse of external profiles, we rewrite them according to a common formalism through a profile modeling language, the PMDL language (Profiles MoDeling Language). We also introduce some transformation operators on the harmonised profiles or their structure, such as the addition of elements to a profile or the creation of a group profile from individual profiles. This approach

has been implemented in the EPROFILEA environment (PERLEA project) and experimented in laboratory with teachers.

## Publications

- Carole EYSSAUTIER-BAVAY, Stéphanie JEAN-DAUBIAS, Jean-Philippe PERNIN, "A model of learners profiles management process", 14th International Conference on Artificial Intelligence in Education AIED'09, Brighton, UK, 6-10 July 2009. (à paraître)
- Carole EYSSAUTIER-BAVAY, Stéphanie JEAN-DAUBIAS, Jean-Philippe PERNIN, "Un modèle de processus de gestion de profils d'apprenants", EIAH'09, Le Mans, France, 23-26 juin 2009. (à paraître)
- Stéphanie JEAN-DAUBIAS, Carole EYSSAUTIER-BAVAY, Marie LEFEVRE, "Uniformisation de la structure de profils d'apprenants issus de sources hétérogènes", EIAH'09, Le Mans, France, 23-26 juin 2009. (à paraître)
- Carole EYSSAUTIER-BAVAY, "Réutilisation informatique de données hétérogènes sur l'apprentissage", Le Mensuel de l'Université, Mars 2009. <http://www.lemensuel.net/2009/03/21/reutilisation-informatique-de-donnees-heterogenes-sur-lapprentissage/>
- Carole EYSSAUTIER-BAVAY, Magali OLLAGNIER-BELDAME, "Médiatiser la prise de conscience métacognitive en contextes d'apprentissage", 1ères Rencontres Jeunes Chercheurs en EIAH, Evry, France, 11 et 12 mai 2006.
- Stéphanie JEAN-DAUBIAS, Carole EYSSAUTIER-BAVAY, "An environment helping teachers to track students' competencies", Workshop LeMoRe of AIED'05 conference, Amsterdam, Netherlands, July 2005.
- Stéphanie JEAN-DAUBIAS, Carole EYSSAUTIER-BAVAY, "Aider l'enseignant pour le suivi des compétences des apprenants", Poster, Annexes aux actes du colloque EIAH, Montpellier, France, mai 2005.
- Carole EYSSAUTIER-BAVAY, "Le portfolio en éducation, concept et usages", Actes du colloque Tice Méditerranée, Revue ISDM n°18 (Information Sciences for Decision Making), Nice, France, 13 p., 26-27 novembre 2004.
- Carole EYSSAUTIER, Stéphanie JEAN-DAUBIAS, "A Device Helping Learners to Self-Assess Themselves", Proceedings of CALIE'04, Grenoble, France, pp. 185-190, 2004.
- Kevin Keenoy, Mark Levene, Sara de Freitas, Birkbeck, Márta Turcsányi-Szabó, Eszter Bodnár, Bruno Emans, Judith Schoonenboom, Jean-Philippe Pernin, Carole Eyssautier, Ann Jones, Andrew Brasher, Jenny Waycott, Lydia Montandon, "Personalised and Collaborative Trails of Digital and Non-Digital Learning Objects", Deliverable D22.4.2, Kaleidoscope Network of Excellence, Dec. 2004.

# Analyse de traces d'activité pour la modélisation cognitive : Application à la conduite automobile

Georgon Olivier

THESE

Soutenue le 19/02/2008

à l'Université Lumière Lyon 2

## Jury

Guy BOY, EURISCO Int.

rapporteur

Bernard CADET, Université de Caen Basse Normandie

rapporteur

Alain MILLE, Université Claude Bernard Lyon I.

Robert MARTIN, Université Lumière Lyon II.

Bruno BACHIMONT, Université de Technologie de Compiègne.

Andry RAKOTONIRAINY, Queensland University of Technology.

Thierry BELLET, INRETS.

Claude BASTIEN, Université de Provence.

Contact : olivier.georgon@gmail.com

## Résumé

Cette thèse s'inscrit dans le champ de l'analyse de l'activité en ergonomie cognitive, appliquée à l'étude du conducteur automobile. Elle combine deux approches : l'ergonomie et l'ingénierie des connaissances. En ergonomie, nous partons d'une prise en considération d'une cognition « située et écologique » et nous nous référons aux notions de « représentation mentale », de « conscience de la situation », de « schémas cognitifs », et de différents « niveaux de contrôle » de l'activité, plus ou moins implicites ou explicites. En Ingénierie des connaissances nous nous basons sur les travaux du « raisonnement à partir d'expérience tracée », pour proposer un « système à base de traces » fondé sur des approches d'ingénierie des connaissances et de « knowledge discovery ». Nous avons effectué un travail de terrain avec une expérimentation de recueil de données de 22 participants sur route ouverte avec un véhicule instrumenté. L'activité de conduite est observée, instrumentée, et verbalisée, et ces trois sources fournissent des « traces », qui constituent des inscriptions informatiques de l'activité. Nous définissons ces traces comme un ensemble de données interprétables par un ergonome pour comprendre le flux temporel des aspects pertinents pour un opérateur, de son interaction avec son environnement. Nous avons modélisé ces traces afin d'obtenir une représentation conceptuelles de la conduite automobile, qui nous conduit à produire des modèles cognitifs du conducteur. Parallèlement nous avons développé un atelier informatique permettant cette modélisation. Nous présentons cet atelier nommé Abstract (Analysis of Behavior and Situation for menTal Representation Assessment and Cognitive acTivity modeling). Il est composé (1) d'un système de collecte, (2) d'un système de manipulation des traces permettant de les transformer et de les visualiser, ainsi que d'un éditeur d'ontologie permettant de définir les modèles de trace, (3) d'un système de documentation et d'indexation. Nos résultats sont constitués (1) de l'atelier informatique développé, qui capitalise le travail de modélisation, (2) de notre méthodologie de construction de connaissances à partir des traces, (3) des modèles de schémas cognitifs de changements de voie sur autoroute que nous avons produits.

## Abstract

How can we describe the mental processes and states carried out by a human when he or she performs an activity? What are the mental processes or state? How can we infer them from the behavior and statements? How to describe them? Our goal is to provide answers to these questions, in the context of car driving. To this end, we have collected "traces of activity" during an experiment made with an instrumented vehicle. They are sequences of events that describe the interaction of the driver with his environment. They contain data from sensors or from assessments made by the ergonomist or by the driver. We have

developed a progressive process for modeling these traces. It is based on a knowledge engineering tool, which we have specifically designed and implemented. This process allows us to produce models of the activity at different levels of abstraction. From these models, we show how an ergonomist can explain the activity, in reference to their own theoretical explicative framework. The framework that we have chosen comes from cognitive psychology. It proposes to describe the mental processes and states of the subject as "cognitive schemas" carried out at different "levels of control". Our approach allows us to search and find the implementation of cognitive schemas in the activity and to propose models of these schemas. From an epistemological viewpoint, this approach is based on a "constructivist" positioning, i.e. evolutionist, pragmatic and driven by the ergonomist. Our results consist of the methodology that we have developed, the software tool that we have built, the models of the activity that we have produced, and the cognitive schemas that we have modeled.

## Publications

### *Conférences*

- Wynn, T., Georgeon, O., Henning, M. J., Richardson, J. H., Bellet, T. & Krems, J. (2008). Cognitive activity modelling: a case study of lane change schemas and sensation seeking. European Conference on Human Centered Design for Intelligent Transport Systems. Lyon, France. April 4th 2008.
- Georgeon, O., Henning, M. J., Bellet, T. & Mille, A. (2007). Creating Cognitive Models from Activity Analysis: A Knowledge Engineering Approach to Car Driver Modeling. International Conference on Cognitive Modeling (pp. 43–48). Ann Arbor, MI.
- Henning, M. J., Georgeon, O. & Krems, J. (2007). The quality of behavioral and environmental indicators used to infer the intention to change lanes. 4th International Driving Symposium on Human Factors in Driver Assessment. Stevenson, Washington USA.
- Georgeon, O., Bellet, T. & Mille, A. (2007). Construction de représentations mentales à partir de traces d'activité. Colloque des jeunes chercheurs en sciences cognitives. Lyon, France.
- Georgeon, O., Mille, A. & Bellet, T. (2006). Analyzing behavioral data for refining cognitive models of operator. Philosophies and Methodologies for Knowledge Discovery, Seventeenth international Workshop on Database and Expert Systems Applications (pp. 588-592). Krakow, Poland: IEEE Computer Society Press.
- Georgeon, O., Mille, A. & Bellet, T. (2006). Abstract: un outil et une méthodologie pour analyser une activité humaine médiée par un artefact technique complexe. Ingénierie des connaissances. Semaine de la connaissance. Nantes, France.
- Georgeon, O., Bellet, T., Mille, A., Letisserand, D. & Martin, R. (2005). Driver behaviour modelling and cognitive tools development in order to assess driver situation awareness. International Workshop on Modelling Driver Behaviour in Automotive Environments (pp. 236-241). Ispra, Italy: Office for Official Publication of the European Communities.
- Bellet, T., Bailly, B., Mayenobe, P., Georgeon, O., Tattegrain-Veste, H., Martin, R., Mille, A. & Trassoudaine, L. (2005). Cognitive modelling and simulation of drivers mental model. Humanist workshop on Modelling Driver Behaviour in Automotive Environments. Ispra, Italy.
- Georgeon, O., Bellet, T., Mille, A. & Martin, R. (2005). Traçage de l'activité de conduite automobile pour la modélisation cognitive du conducteur. Colloque des jeunes chercheurs en sciences cognitives. Bordeaux, France.

### *Chapitre de livre*

- Bellet, T., Bailly, B., Mayenobe, P. & Georgeon, O. (2007). Cognitive modelling and computational simulation of drivers mental activities. Critical Issues in Advanced Automotive Systems and Human-Centred Design. P. C. Cacciabue and C. Re. London: Springer Verlag.

# Confrontation de points de vue-opinions sur des corpus documentaires

Gesche Samuel

THESE

Soutenue le 01/12/2008

à l'INSA de Lyon

## Jury

Mlle Sylvie Calabretto, INSA de Lyon	directrice
M. Guy Caplat, INSA de Lyon	examineur
M. Gilles Falquet, Université de Genève	examineur
Mme Catherine Garbay, CNRS	présidente
Mlle Elisabeth Murisasco, 'Université du Sud Toulon-Var	rapporteur
M. Manuel Zacklad, Université de Technologie de Troyes	rapporteur

Contact : samuel.gesche@insa-lyon.fr / sylvie.calabretto@insa-lyon.fr

## Résumé

Cette thèse se concentre sur l'aspect enrichissant de la confrontation entre des positions d'experts. La diversité de ces positions génère une hétérogénéité qui est fondamentalement bonne et génératrice de créativité et d'avancées. Afin de mener une confrontation de points de vue qui soit enrichissante, nous avons choisi d'outiller plus particulièrement la recherche de cette diversité. Mais parmi toutes les formes d'hétérogénéité que l'on peut trouver, la divergence de point de vue est la plus difficile à localiser. Elle est à la fois similarité, parce que l'on parle de la même chose, et différence, parce que l'on n'a pas le même avis dessus. Nous présentons donc une méthode qui s'appuie beaucoup sur l'expert lui-même, qui est plus apte à mener ce « jeu des différences » que n'importe quel algorithme. Nous attachons une grande importance à ce que l'expert retire de la confrontation, plus qu'au résultat lui-même : nous n'essayons pas au premier chef d'intégrer des connaissances.

Nous appréhendons la confrontation de points de vue dans un contexte indépendant du corpus, et autant que faire se peut, du formalisme dans lequel le point de vue est exprimé. Notre approche tire principalement parti de l'alignement d'ontologies, dont les lacunes sont complétées par le paradigme d'assistance par ordinateur, l'exploitation de graphes peu formalisés et la communication homme-machine.

### Mots clés :

Confrontation, Alignement, Point de Vue, Opinion, Modèle, Ontologie, Sérendipité.

## Abstract

This thesis deals with the issue of enrichment while confronting expert viewpoints. These viewpoints bring much diversity, which is basically good for generating creativity and advances in Science. In order for the confrontation of viewpoints to enrich the expert, we chose to focus our research on the search for this diversity. Indeed, among all forms of heterogeneity that are part of viewpoint matching, locating difference of opinion is one of the toughest tasks. It comes from the fact that difference of opinion is a difference on the one hand (since the expert have different theories), and a similarity on the other hand (since they talk about the same subject). We present here a method for taking advantage of the knowledge of the expert, which is the best "tool" we could find for dealing with this "find the differences game", way better than any algorithm, especially since we focus more on the benefit the expert gains in confrontation than in the result itself: our aim is not to integrate knowledge.

Our approach of confrontation does not depend on the documents used, or to a certain extent on the formalism used to express it. We base it on ontology matching methods, using other paradigms to overcome the intrinsic limits of ontology matching. Such paradigms include the computer-aided paradigm, the use of poorly formalized graph languages and human/computer communication.

*Keywords:*

Confrontation, Matching, Viewpoints, Opinion, Model, Ontology, Serendipity.

**Publications**

- C. BERDIER, S.CALABRETTO, G.CAPLAT, S. GESCHE. "Managing heterogeneity in urban ontologies". In. 3rd Workshop of COST Action C21: Construction of multilingual ontologies for Urban Civil Engineering projects.
- S. GESCHE. "Exploitation de l'hétérogénéité entre points de vue-opinion". In Actes du XXVème congrès INFORSID, INFORSID'2008, Fontainebleau, France, 27-30 mai 2008, pp. 101-116.
- S. GESCHE, G. CAPLAT., S. CALABRETTO. "Annotation Multi-Documents de Points de Vue Exprimés"
- In Hypermedia Hypertext: Products, Tools and Methods, H2PTM'07, Hammamet, Tunisie, 29-31 Octobre 2007, pp. 69-73.
- S. GESCHE, G. CAPLAT., S. CALABRETTO. "Managing an Expert System among Others". In Proceedings of the 16th International Conference on Systems Science, ICSS'2007, Wroclaw, Poland, september 04-06, 2007, pp.193-199.
- S. GESCHE, G. CAPLAT., S. CALABRETTO. "Managing Difference of Opinion in Semantic Structures". In Proceedings of the International Workshop On Semantically Aware Document Processing and Indexing held in cooperation with ACM SIGWEB, SADPI'07, Montpellier, France, may 21-22, 2007, pp. 79-86.
- S. GESCHE, S. CALABRETTO, G. CAPLAT. "Un modèle pour la Confrontation d'opinions numérisées sous Porphyry". In Actes du Colloque International sur le Document Electronique, CIDE'2006, Fribourg, Suisse, 18-22 septembre 2006, pp. 253-267.
- S. GESCHE, S. CALABRETTO, G. CAPLAT. "Confrontation de Points de Vue dans le Système Porphyry". In Actes des sixièmes journées Extraction et Gestion des Connaissances, EGC'2006, Lille, France, 17-20 janvier 2006, Cépaduès-Éditions 2006, ISBN 2-85428-718-5, Volume 2, pp. 725-726.

# Similarité sémantique inter-ontologies basée sur le contexte

THESE

Hoffmann Patrick

Soutenue le 16/12/2008

à l'Université Claude Bernard, Lyon 1

## Jury

M. Kokou YÉTONGNON  
M. Aris M. OUKSEL  
M. Djamel BENSLIMANE  
M. Shaw C. FENG  
Parisa GHODOUS  
Ram D. SRIRAM

rapporteur  
rapporteur  
examineur  
examineur  
directrice  
directeur

Contact : hoffmannp@free.fr

## Résumé

Cette thèse étudie l'intérêt du contexte pour améliorer l'interopérabilité entre ontologies hétérogènes, d'une manière qui permette leur évolution indépendante. Lors de collaborations, les organisations échangent leurs données, qui sont décrites par des concepts définis dans des ontologies. L'objectif est d'obtenir un service d'évaluation de tels concepts, basé sur le contexte.

Nous proposons une méthodologie pour déterminer, modéliser et utiliser le contexte. En l'appliquant, nous découvrons trois usages du contexte qui contribuent à améliorer la réconciliation d'ontologies : Nous proposons de désambiguïser les sens pragmatiques possibles des concepts en comparant les "perspectives" avec lesquelles les concepts ont été développés ; de personnaliser en considérant le contexte des agents, constitué d'une sélection pertinente parmi les domaines et tâches de l'organisation ; d'évaluer la pertinence des données associées au concept pour la tâche qui a suscité le besoin en interopérabilité.

## Abstract

This thesis studies how a context-based approach could improve interoperability between heterogeneous ontologies developed and evolving autonomously. The hypothesis is that during collaborations, organizations exchange their data whose meaning is described by concepts defined by ontologies. The objective is to propose a context-based evaluation of such concepts.

We propose a methodology to determine, model and use the context. Applying this methodology, we uncover three usages of context that may be used together to improve ontology reconciliation : We propose to disambiguate among the possible pragmatic meanings of the concepts by comparing the "perspectives" with or for which the concepts have been developed; to personalize this comparison by considering the agent's context, made of a relevant selection of the agent's company domains and tasks; to evaluate the pertinence of the data associated with the concept for the task that triggered the interoperability need.

## Publications

### *Articles de revue avec comité de lecture*

- C Ferreira da Silva, L Médini, S Abdul Ghafour, P. Hoffmann, P Ghodous: Semantic Interoperability of Heterogeneous Semantic Resources. *Electronic Notes in Theoretical Computer Science* 150(2):71-85, Elsevier. 2006.

### *Article dans actes de conférence*

- P Hoffmann, S. C. Feng, G. Ameta, P. Ghodous, Lihong Qiao, R. Sriram: Towards an Approach for Multiple-View Semantic Model in Product Development. In 15th ISPE INTERNATIONAL CONFERENCE ON

CONCURRENT ENGINEERING: RESEARCH AND APPLICATIONS (CE 2008), August 2008.

- L Médini, N. Lumineau, C. Ferreira Da Silva, P. Hoffmann, P. Ghodous : Découverte de correspondances sémantiques par inférences en environnement P2P. (Atelier "Passage à l'échelle des techniques de DEcouverte de CORrespondances", conférence EGC'07), Namur, Belgique. 10p. 2007.
- P Hoffmann, L Médini, P Ghodous : Using Context to improve Semantic Interoperability. In 13th ISPE INTERNATIONAL CONFERENCE ON CONCURRENT ENGINEERING: RESEARCH AND APPLICATIONS, Leading the Web in Concurrent Engineering, Parisa Ghodous, Rose Dieng-Kuntz, Geilson Loureiro ed. Antibes, France. Frontiers in Artificial Intelligence and Applications 143. IOS Press Nieuwe Hemweg 6B 1013 BG Amsterdam Netherland. ISBN 1-58603-651-3. ISSN 0922-6389. 2006.
- C Ferreira da Silva, L Médini, S Abdul Ghafour, P Hoffmann, P Ghodous: Interoperability of heterogeneous semantic resources. In International Workshop on Database Interoperability (InterDB 2005) in connection with the 7th International Conference on Coordination Models and Languages, Philippe Thiran, Tore Risch, Djamal Benslimane ed. Namur, Belgique. pp. 74-89. 2005.



# Towards the conception of extended information systems organized around a web service platform

Le Blevec Yann

THESE

Soutenue le 19/12/2008

à l'Université Claude Bernard Lyon1

## Jury

M. Kokou YETONGNON, Université de Bourgogne	examineur
M. Bruno DEFUDE, Institut TELECOM SudParis	examineur
M. Yamine AIT AMEUR, Université de Poitiers	examineur
M. Djamel BENSLIMANE, Université Claude Bernard Lyon 1	directeur
Mme Chirine GHEDIRA, Université Claude Bernard Lyon 1	co-directrice
M. Xavier DELATTE, DSI Alcatel Vacuum Technology France	invité

Contact : djamal.benslimane@liris.cnrs.fr / chirine.ghedira@liris.cnrs.fr

## Résumé

Avec les services web, les entreprises ont désormais accès à une technologie simple et standardisée leur permettant d'ouvrir à faible coût leurs systèmes d'information. De part nos travaux, nous avons essayé d'aider ces entreprises à exposer par le biais d'Internet leurs processus métier pour les rendre accessibles à des partenaires extérieurs (clients / fournisseurs). Voilà pourquoi, avec un intérêt tout particulier pour les questions de sécurité et de fiabilité, nous avons focalisé notre approche sur les deux problématiques suivantes: comment exposer de manière sécurisée des services au travers d'un réseau public? Comment maîtriser le nombre grandissant de services fournis tout en maintenant un niveau de service constant? En effet, plus les services web seront utilisés grâce notamment à de nouveaux modèles de consommation et plus les acteurs demanderont de la transparence au niveau de l'exécution de ces services, mettant en avant un fort besoin en terme de qualité de service (QoS).

Dans ce contexte, nous avons proposé trois contributions complémentaires. Notre premier travail consiste en la création d'une architecture sécurisée permettant l'exposition de service métier au travers d'Internet. La solution proposée offre un compromis réaliste entre le niveau de sécurité obtenu et le niveau global de complexité du système. Après quoi, une fois la question obligatoire de sécurité traitée, nous avons défini une plateforme de QoS permettant à une architecture orientée service d'évaluer dynamiquement des métriques de QoS personnalisables provenant de sources de mesure hétérogènes. Un prototype a été implémenté afin de valider les concepts fondateurs de la plateforme. Finalement, nous nous sommes également intéressés aux différents modes de consommation des services web. A partir de technologies actuelles, plusieurs modèles d'interaction ont été formalisés définissant comment des services d'entreprises peuvent être intégrés à des applications web composites.

## Publications

### *Journaux internationaux*

- Yann Le Blevec, Chirine Ghedira, Djamel Benslimane, and Xavier Delatte. Service-oriented computing: Bringing business systems to the web. *IT Professional*, 9(3):19-24, 2007.

### *Conférences internationales*

- Yann Le Blevec, Chirine Ghedira, and Djamel Benslimane. Providing web services with quality of service features to business partners. In *NOTERE 2007 : 7eme Conférence Internationale sur les NOuvelles TEchnologies de la RÉpartition*, Marrakech, Maroc, June 2007.
- Yann Le Blevec, Chirine Ghedira, Djamel Benslimane, Xavier Delatte, and Zahi Jarir. Exposing web services to business partners: Security and quality of service issue. In *1st International Conference on Digital Information Management*, pages 69-74, 2006.



# Système d'accès personnalisé à l'information : application au domaine médical

Mirvat El Makssoud épouse Abbas Karine

THESE

Soutenue le 09/12/2008

à l'Insa de Lyon

## Jury

Pr Paul RUBEL ( INSA de Lyon)

Pr Dominique RIEU (Université de Grenoble)

Pr Jean-Louis CAVARERO (université Nice-Sofia Antipolis )

Pr Michel GOURGAND (université de Clermont Ferrand)

Pr André FLORY (INSA de Lyon),

directeur

Pr Christine VERDIER (Université de Grenoble),

directrice

Contact : abbaskarine@hotmail.com

## Résumé

Le travail mené dans le cadre de cette thèse repose sur une problématique centrale : la personnalisation de l'accès à l'information. En effet, face à la croissance considérable des données, à l'hétérogénéité des rôles et des besoins et au rapide développement des systèmes mobiles, il devient important de proposer un système personnalisé capable de fournir à l'utilisateur une information pertinente et adaptée à ses besoins et son utilisation quotidienne. Ce système doit prendre en considération les différentes caractéristiques de l'utilisateur ainsi que l'ensemble des situations contextuelles qui influence son comportement lors de son interaction avec le système d'information.

Le système d'accès personnalisé à l'information que nous proposons est basé sur le principe de la gestion des profils. Un modèle générique du profil est proposé afin de couvrir l'ensemble des facettes de la personnalisation. Ce modèle est capable de collecter des informations sur l'utilisateur et son contexte d'utilisation et de représenter tout besoin de personnalisation une fois modélisé.

Le système de personnalisation est fondé principalement sur trois éléments : les profils, le contexte et les services. Les profils sont des conteneurs des connaissances sur l'utilisateur. Le contexte définit un ensemble de paramètres qui caractérisent l'environnement d'utilisation du système. Les services sont des programmes autonomes capables d'exécuter les tâches de la personnalisation. Le processus de la personnalisation est déclenché lorsque l'utilisateur envoie une requête et appliqué sur des données structurées extraites des différentes sources de données en réponse à cette requête.

## Mots clés :

accès personnalisé – système d'information – modèle générique des profils – contexte – services – interface adaptative – personnalisation des données – gestion de la personnalisation – schéma XML – utilisateur

## Abstract

The thesis work relays on a central problematic : the personalisation access to information. Indeed, with the considerable growth of data, the heterogeneity of roles and needs and the spread development of mobile systems, it becomes important to propose a personalised system to get relevent information. This system allows to user to provide relevent and adapted information. This system takes into account the different user characteristics as well as the different contextual situations which influence his behavior during the information access process.

The personalised access system we propose is based on the profiles management. A generic profile model is defined to cover all personalisation facets. The model is able to collect information on user and his context of use and to represent all personalisation needs.

The personalised system is mainly founded on three elements : the profiles, the context and the services. The profiles are containers of knowledge on users. The context defines a set of parameters characterising the user environment when the system is used. The services are autonomous programs able to execute the personalisation tasks. The personalisation process starts when the user sends a request that produces the extraction of data useful for it.

*Keywords:*

personalised access – information system– generic profile model – context – services – adaptive interface – data personalisation – management of personalisation needs– XML Schema – user

## **Publications**

*Revue nationale avec comité de lecture*

- Abbas K., Verdier C., «conception d'une structure globale du dossier médical pour les réseaux de soins », revue Santé et Systémique, Hermès Lavoisier, 2007

*Congrès internationaux avec comité de lecture*

- Abbas K., Verdier C., Flory A., « Personnalisation de l'information dans les systèmes de santé », 9ème Conférence Internationale sur la Science des Systèmes de Santé (ICSSHC 2008), 3-5 septembre 2008, Lyon, France.
- Abbas K. Verdier C., Flory A. « Adaptive Context for Medical Information Systems » MeDeHis workshop co-located with ICDIM'07, IEEE, ACM Siggap, ISBN: 1-4244-1476-8, Lyon, France, pp 591-96 October 2007
- Abbas K., Verdier C., Flory A., « Exploiting profile modeling for web-based information systems », in 8th International Conference on Web Information Systems Engineering (WISE 07), PAWI 07 Workshop, Springer LNCS 4832, ISBN 978-3-540-77009-1, 3-7 december, 2007, Nancy, France

*Congrès nationaux avec comité de lecture*

- Abbas K., Verdier C., Flory A., « Gestion des connaissances centrée sur le modèle d'utilisateur », dans Giseh 2006, 14-16 septembre 2006, centre Henri Tudor, Luxembourg

*Autres revues*

- Abbas K., Verdier C., Flory A., « Gestion des connaissances centrée sur le modèle d'utilisateur », revue Gestion Hospitalière, N°466, p 328-335, mai 2007

*Posters*

- Abbas K., Verdier C., Flory A., « Adapted and Contextual information in Medical Information Systems”, International Conference on Enterprise Information Systems - ICEIS 2008, 12-16, June 2008, Barcelona, Spain.
- Abbas K., Verdier C., « Design of a medical database transformation algorithm». International Conference on Enterprise Information Systems - ICEIS 2006, 23 - 27, May 2006, Paphos – Cyprus.

Soutenue le 23/01/2008

à l'INSA de Lyon

## Jury

A. Lbath, LIG - Université Joseph Fourier,	rapporteur
J. Lemaitre, LISIS - Université du Sud Toulon-Var	examineur
J-M. Pinon, LIRIS - INSA de Lyon,	directeur
I. Roxin, LASELDI - Université de Franche-Comté, Montbéliard	examineur
B. Rumpler, LIRIS - INSA de Lyon,	co-directrice
J. Savoy, Université de Neuchâtel, Suisse	rapporteur

Contact : [beatrice.rumpler@insa-lyon.fr](mailto:beatrice.rumpler@insa-lyon.fr)

## Résumé

Les informations disponibles sur l'Internet se développent à un rythme tel que bientôt les méthodes de recherche textuelle utilisant la fréquence des termes ne seront plus suffisantes. Un courant de pensée est consacré à la personnalisation de la recherche, à savoir prendre en compte des traits spécifiques et le contexte de l'utilisateur pour répondre à sa requête. Nous pensons que le profil, la communauté, et le contexte de l'utilisateur sont les trois concepts essentiels à envisager pour faire face au problème de la croissance du World Wide Web. La thèse étudie la combinaison de ces trois courants de pensée.

Dans la première partie de cette thèse, nous développons un système de recherche d'information personnalisé et collaboratif (appelé PERCIRS), qui utilise les deux premiers concepts (le profil et la communauté). PERCIRS crée une liste classée de documents pertinents, pour la requête  $q$  de l'utilisateur  $U$ . Cette classification se fait sur la base des documents sélectionnés pour les requêtes semblables à  $q$  par les utilisateurs au profil similaire à celui de  $U$ . Le choix de la méthode de recherche d'utilisateurs similaires joue un rôle important dans l'efficacité de PERCIRS. À cette fin, nous avons proposé trois catégories de formules pour calculer la similarité entre deux profils d'utilisation: formule fondée sur l'égalité, formule fondée sur la similarité et enfin une formule fondée sur les graphes. Afin de trouver la catégorie optimale, nous avons proposé deux mécanismes d'évaluation: fondés sur les concepts de la catégorisation et de la classification. Ces deux mécanismes s'appuient sur les formules de calcul de profil utilisateur fondées sur les graphes.

PERCIRS étant un système de recherche d'information (SRI) personnalisé (en raison de la prise en compte des profils d'utilisateur), il ne peut pas être évalué par des mécanismes d'évaluation tels que Cranfield (par exemple TREC). Par conséquent, dans cette thèse, nous proposons un nouveau mécanisme qui permet de l'évaluer en même temps que les autres SRI classiques tels que BM25 – Okapi.

Dans la deuxième partie de ce travail de thèse, le contexte de l'utilisateur est utilisé pour adapter aux préférences de l'utilisateur un document trouvé par PERCIRS. Nous proposons d'adapter physiquement et sémantiquement un document selon le profil de l'utilisateur et selon le profil de contexte. Un mécanisme est également proposé pour naviguer dans des documents adaptés en fonction des préférences de l'utilisateur.

## Abstract

The information found on the Internet is growing at such a rapid rate that soon methods of searching through text using terms frequencies will not be enough. One stream of thought is devoted to personalizing the search, which means taking into account the specific traits and background of a user when answering that user's query. We believe that user's profile, user's community, and user's context are the three main issues to tackle the problem of the ever growing World Wide Web. The thesis investigates the combination

of these three streams of thought.

In the first part of this thesis, we develop a PERSONalized and Collaborative Information Retrieval System (called PERCIRS), which makes use of the first two factors (profile and community). PERCIRS creates a ranked list for the query  $q$  of the user  $U$  based on the documents selected by similar users to  $U$  which have been considered as relevant to the similar queries to  $q$ . The effectiveness of method for finding the similar users plays an important role in PERCIRS's effectiveness. To this aim, we have proposed three categories of formulas to calculate the similarity between two use profiles: equality-based, similarity-based and graph-based. In order to find the optimal category we proposed two evaluation mechanisms: categorization-based and classification-based. Both of these two mechanisms agreed on the graph-based user profile calculation formulas as the optimal formulas.

As PERCIRS is a personalized IR system (due to considering user profiles), it can not be evaluated by Cranfield like evaluation mechanisms (e.g. TREC). Hence in this thesis we propose a new mechanism which enables us to evaluate it with the other classical IR systems such as BM25-Okapi.

In the second part, the user's context is used to adapt a found document by PERCIRS to user's preferences. We propose to physically and semantically adapt a given document based on the user's profile and context's profile. A mechanism is also proposed to navigate into the adapted documents according to user's preference.

## Publications

- Graph-based Profile Similarity Calculation Method and Evaluation . H Naderi, B. Rumpler. Dans ACM ECIR 08, Springer in LNCS. ed. Glasgow , Scotland. pp. 637-641. ISSN 0302-9743. 2008. (lang: en) (équipe BDD, id: 3326)
- A Graph-based profile similarity calculation method for collaborative information retrieval. H Naderi, B. Rumpler, J.M. Pinon. Dans ACM SAC'08, ACM New York, USA ed. Fortaleza, Ceara Brazil. pp. 1127-1131. ISBN 978-1-59593-753-. 2008. (lang: en) (équipe BDD, id: 3280)
- An efficient Information Retrieval System by Incorporating the User Profile. H Naderi, B. Rumpler, J.M. Pinon. Adaptative Multimedia Retrieval : User, Context and Feedback LNCS 4398():247-257, Springer, ISSN 0302-9743. 2007. (lang: en) (équipe BDD, id: 2889)
- Three User Profile Similarity Calculation (UPSC) Methods and Their Evaluation. H Naderi, B. Rumpler. Dans IEEE- - ACM SIGAPP SITIS'07, Shanghai - Chine. pp. 239-245. IEEE Computer Society - CPS . ISBN 978-0-7695-3122-9. 2007. (lang: en) (équipe BDD, id: 3279)
- Physical Document Adaptation to user s context and user s profile. H Naderi, B. Rumpler. Dans 9th International Conference on Enterprise Information Systems. ACM ICEIS, Funchal - Madeira. pp. 92-97. 2007. (lang: en) (équipe BDD, id: 2761)
- PERCIRS: a PERSONalized Collaborative Information Retrieval System. H Naderi, B. Rumpler. Dans INFORSID, Hammamet - Tunisie. pp. 113-127. 2006. (lang: fr) (équipe BDD, id: 2783)

# Données et Métadonnées de Qualité pour les Champs Continus : Modèles de Terrain et Photogrammétrie

THESE

Pozzoli Alice

Soutenue le 27/06/2008

à l'INSA de Lyon

## Jury

Norbert Bartelme (Technical University of Graz)	rapporteur
Riccardo Barzaghi (Politecnico di Milano)	examineur
Alessandro Carosio (ETH de Zurich)	rapporteur
Christophe Claramunt (Ecole Navale de Brest)	rapporteur
Gianfranco Forlani (Università di Parma)	examineur
Pierre Grussenmeyer (INSA de Strasbourg)	rapporteur
Robert Laurini (NSA de Lyon)	co-directeur
Luigi Mussio (Politecnico di Milano)	co-directeur

Contact : [alice.pozzoli@polimi.it](mailto:alice.pozzoli@polimi.it)

## Résumé

Le sujet principal de cette thèse est le traitement des données en géomatique allant de l'acquisition des données photogrammétriques à la représentation cartographique. L'objectif est ainsi l'utilisation des techniques statistiques pour le traitement des données géomatiques afin de créer des modèles numériques des terrains en partant des données photogrammétriques.

La fonction principale de la Photogrammétrie est la transformation des données en partant de l'espace-image à l'espace-objet. Nous avons proposé une solution pratique pour l'orientation automatique à partir de trois images. Cette méthodologie d'orientation présente de nombreux avantages pour les applications environnementales et de surveillance, et elle est un puissant instrument que l'on peut utiliser à côté de méthodologies plus traditionnelles. Parmi diverses applications possibles, on a choisi de construire le relief d'un modèle hydraulique 3D qui représente la confluence de deux torrents dans une région montagneuse.

D'un point de vue informatique, nous avons proposé une description de données photo-grammétriques basée sur le format XML pour les données géographiques (extension de GML, Geographic Markup Language). L'objectif est d'optimiser l'archivage et la gestion des données géomatiques.

Enfin, un logiciel original a été produit, qui permet de modéliser les terrains en utilisant la photogrammétrie à trois images.

## Abstract

This thesis deals with data processing in Geomatics. It ranges from data acquisition in Photogrammetry to data representation as well as in Cartography. The objective of this research was to use statistical techniques of data processing for the creation of digital surface models starting from photogrammetric images. The main function of photogrammetry is the transformation of data coming from the image space to the object space. An easy solution for three image orientation is proposed. The orientation procedure described has some relevant advantages for environmental and monitoring applications, and makes it a very powerful tool in addition to more traditional methodologies. Among many different applications, an interesting project for the survey of a hydraulic 3D model of a stream confluence in the mountain area has been performed. From a computing point of view, we propose a description of the photogrammetric data based on the XML format for geographic data (Geographic Markup Language). The aim is to optimize the archiving and management of geo-data.

As a conclusion, an original software product which allows to model terrains starting from three-image photogrammetry has been developed and tested.

## Publications

- Mussio L., Pozzoli A. (2003): Quickly solutions particularly in close range photogrammetry. Int. Archives of Photogrammetry and Remote Sensing, vol. XXXIV, part 5W12. Ancona, p.273-278.
- Mussio L., Pozzoli A. (2003): Non- linear problems of analytical photogrammetry. Int. Archives of Photogrammetry and Remote Sensing, vol. XXXIV, part 6W11, Zagreb.
- Pozzoli A. (2003): Ricerca esaustiva e soluzione diretta dall'Orientamento Relativo all'Orientamento Assoluto, ASITA 2003, Verona (Italy).
- Pozzoli A., Mussio L., Scaioni M. (2004a): A Solution for the General Case of Three-image Orientation. Int. Archives of Photogrammetry and Remote Sensing, vol. XXXV, part B3. Istanbul, p. 992-997.
- Pozzoli A., Mussio L., Scaioni M. (2004b): A Project for the Survey of a Hydraulic Stream Confluence Model Based on a General Procedure for Three-image Orientation. In: E. Baltsavias (Ed), Commemorative volume for the 60th birthday of Prof. Dr. Armin Gruen. Zurich, p. 201-208.
- Pozzoli A., Mussio L. (2004): Modelling Techniques for Temporal GIS. In E. Fendel, M. Rumor (Ed's), Proc. Of UDMS'04-24th Urban Data Management Symposium. Chioggia (Italy), 7.II.2.
- Pozzoli A., (2004): Orientamento Esterno di tre Immagini Applicato al Rilievo di un Modello Idraulico, ASITA 2004, Roma (Italy), p. 1687-1692.
- Pozzoli A. (2005): Rilievo di una Confluenza di Torrenti in zona Alpina grazie all'Orientamento Automatico di tre Immagini. ASITA 2005, Catania (Italy).
- Pozzoli A., Mussio L. (2006): New Tools for Spatial – Temporal Database. In: E. Fendel, M. Rumor (Ed's) Proc. of UDMS'06 – 25th Urban Data Management Symposium. Aalborg, 9.I.3.



# Structuration relationnelle des politiques de contrôle d'accès représentation, raisonnement et vérification logiques

**Thion Romuald**

**THESE**

Soutenue le 17/06/2008

à l'INSA de Lyon

## Jury

Pr. Georges Gardarin, (PRISM-UVSQ)	examineur
Pr. Philippe Pucheral, (PRISM-UVSQ)	rapporteur
Pr. Florence Sèdes, (IRIT-Paul Sabatier)	rapporteur
Stéphane Coulonde (LIRIS-INSA Lyon)	directeur
André Flory (LIRIS-INSA Lyon)	directeur

Contact : romuald.thion@gmail.com

## Résumé

Le contrôle d'accès est le mécanisme qui définit et impose ce qu'il est permis et interdit de faire dans un système. Les politiques de contrôle d'accès sont des ensembles de faits et règlements structurés selon des modèles. Depuis l'initiative des modèles à rôles, de nombreuses générations d'organisation des droits d'accès ont vu le jour. Les politiques et les modèles sont ainsi devenus plus grands, plus complexes et soulèvent de nombreux problèmes de formalisation, de vérification et d'administration.

La thèse montre que les modèles de contrôle d'accès existants partagent de nombreux traits. Après avoir analysé et synthétisé ces éléments communs, nous proposons une structuration relationnelle pour concevoir, organiser et formaliser les droits d'accès. L'approche est fondée sur les dépendances de données : des classes de formules logiques qui imposent des contraintes sur les données relationnelles.

L'approche met à profit des résultats qui dépassent le cadre strict du modèle relationnel et résolvent des problèmes actuels d'expression, de vérification et de raisonnement sur les politiques de contrôle d'accès. Un des problèmes majeurs que nous abordons est de s'assurer qu'une politique est intègre, qu'elle satisfait bien aux propriétés imposées par le modèle.

La thèse met à profit des ponts existants entre les dépendances et des domaines de recherche connexes. Nous proposons ainsi une représentation permettant d'exprimer graphiquement les modèles ainsi qu'une méthode semi-automatisée pour assister la réorganisation de politiques existantes. Les récentes applications des dépendances de données offrent de riches perspectives d'applications au contrôle d'accès.

## Abstract

Access control is a mechanism which defines and controls the privileges of users in a system. Nowadays, it is one of the most common and pervasive mechanisms used for security enforcement in information systems. Access control policies are sets of facts and rules organized by mean of access control models. Since the role-based access control initiative, several access control models have been proposed in the literature. The policies and models have become larger and more complex, and several issues on formalization, verification and administration have appeared.

The PhD thesis shows that access control models share common characteristics. Upon analysis and synthesis of these traits, we propose a relational structuration for the design, organization and formalization of privileges. The framework is built upon data dependencies: fragments of first-order logic dedicated to express constraints between relational data.

Some results from the databases community benefit the approach by helping address current issues on expression, verification, and reasoning on access control policies. We focus particularly on the integrity property of policies: guaranteeing that the policies enforce the properties defined in the model.

The thesis profits from bridges between data dependencies, conceptual graphs and formal concepts analysis. Thus, we propose a graphical representation of the models and a semi-automated method for engineering the policies. Finally, we present perspectives for access control models based upon recent applications of data dependencies from the databases community.

## Publications

### *Articles de revue avec comité de lecture*

- Découverte automatisée de hiérarchies de rôles pour les politiques de contrôle d'accès . R Thion, S. Coulondre. Ingénierie des Systèmes d'Information (ISI) : Modèles, Formalismes et Outils pour les Systèmes d'Inf 13(4):107-131, Lavoisier. 2008.
- Contrôle d'accès logique au dossier patient informatisé. R Thion, S. Coulondre, A. Flory. Santé et systémique 10(1/2):83-104, Hermes-Lavoisier, ISBN 978-2-7462-1912-, ISSN 1280-7834. 2007.

### *Colloques/congrès*

- Découverte automatisée de hiérarchies de rôles pour les politiques de contrôle d'accès. R Thion. Dans XXVème congrès INFORSID - Jeune chercheur, INFORSID'07, Hermes ed. Perros-Guirec. pp. 139-154. ISBN 978-2-9527630-1-. 2007.
- Un Modèle Homogène pour la Confidentialité et l'Intégrité des Données Relationnelles. R Thion, S. Coulondre. Dans XXIIème Journée Base de Données Avancées, BDA'06, Lille. 2006.
- Modelling and inferring on role-based access control policies using data dependencies. R Thion, S. Coulondre. Dans 17th International Conference on Database and Expert Systems Applications, DEXA'06, Krakow, Poland. pp. 914-923. LNCS 4080. Springer . ISBN 3-540-37871-5. 2006.
- Representation and Reasoning on Role-Based Access Control Policies with Conceptual Graphs. R Thion, S. Coulondre. Dans 14th International Conference on Conceptual Structures, ICCS'06, Aalborg University, Denmark. pp. 427-440. LNCS 4068. Springer . ISBN 3-540-35893-5. 2006.

### *Contributions à des ouvrages*

- Access Control Models. R Thion. Cyber Warfare and Cyber Terrorism (Colarik Andrew M. and Janczewski Lech, Eds.), p. 318-328, Information Science Reference, May 2007, IDEA Group Publishing, ISBN 978-1-59140-991-5 2007.

2008

Département  
Image



# Apprentissage a contrario et architecture efficace pour la détection d'évènements visuels significatifs

THESE

**Burrus Nicolas**

Soutenue le 01/12/2008

à l'université Pierre Et Marie Curie

## Jury

Dr. Thierry BERNARD	directeur de thèse
Dr. Jacques BLANC-TALON	examineur
Dr. Bertrand COLLIN	examineur
Pr. Matthieu CORD	examineur
Pr. Jean-Michel JOLION	directeur de thèse
Pr. Alain MERIGOT	rapporteur
Pr. Lionel MOISAN	rapporteur

Contact : jean-michel.jolion@universite-lyon.fr

## Résumé

Pour assurer la robustesse d'un algorithme de détection, il est nécessaire de maîtriser son point de fonctionnement, et en particulier son taux de fausses alarmes. Cette tâche est particulièrement difficile en vision artificielle à cause de la grande variabilité des images naturelles, qui amène généralement à introduire des paramètres choisis a priori qui limitent la portée et la validité des algorithmes. Récemment, l'approche statistique a contrario a montré sa capacité à détecter des structures visuelles sans autre paramètre libre que le nombre moyen de fausses alarmes tolérées, en recherchant des entités dont certaines propriétés sont statistiquement trop improbables pour être le fruit du hasard.

Les applications existantes reposent toutefois sur un cadre purement analytique qui requiert un travail important de modélisation, rend difficile l'utilisation de caractéristiques multiples et limite l'utilisation d'heuristiques de recherche dirigées par les données. Nous proposons dans cette thèse d'assouplir ces restrictions en ayant recours à de l'apprentissage pour les quantités non calculables analytiquement. Nous illustrons l'intérêt de la démarche à travers trois applications : la détection de segments, la segmentation en régions homogènes et la détection d'objets à partir d'une base de photos. Pour les deux premières applications, nous montrons que des seuils de détection robustes peuvent être appris à partir d'images de bruit blanc. Pour la dernière, nous montrons que quelques exemples d'images naturelles ne contenant pas d'objets de la base suffisent pour obtenir un algorithme de détection fiable.

Enfin, nous remarquons que la monotonie du raisonnement a contrario permet d'intégrer incrémentalement des informations partielles. Cette propriété nous conduit à proposer une architecture "anytime" pour la détection d'objets, c'est-à-dire capable de fournir des détections progressivement au cours de son exécution, en commençant par les objets les plus saillants. L'algorithme peut donc être stoppé à tout moment pour satisfaire à des contraintes temporelles.

### Mots clés :

Vision par ordinateur, méthodes a contrario, apprentissage statistique, segmentation, reconnaissance d'objets, anytime

## Abstract

To ensure the robustness of a detection algorithm, it is important to get a close control of the false alarms it may produce. Because of the great variability of natural images, this task is very difficult in computer vision, and most methods have to rely on a priori chosen parameters. This limits the validity and applicability of the resulting algorithms. Recently, by searching for structures for which some properties are very unlikely to be due to chance, the a contrario statistical approach has proved successful to provide

parameterless detection algorithms with a bounded expected number of false alarms.

However, existing applications rely on a purely analytical framework that requires a big modeling effort, makes it difficult to use heterogeneous features and limits the use of data-driven search heuristics. In this thesis, we propose to overcome these restrictions by using statistical learning for quantities that cannot be computed analytically. The interest of this approach is demonstrated through three applications : segment detection, segmentation into homogeneous regions, and object matching from a database of pictures. For the two first ones, we show that robust decision thresholds can be learned from white noise images. For the last one, we show that only a few examples of natural images that do not contain the database objects are sufficient to learn accurate decision thresholds.

Finally, we notice that the monotonicity of a contrario reasoning enables an incremental integration of partial data. This property leads us to propose an architecture for object detection which has an “anytime” behavior : it provides results all along its execution, the most salient first, and thus can be constrained to run in limited time.

#### *Mots clés :*

Computer vision, a contrario reasoning, statistical learning, segmentation, object detection, anytime

#### **Publications**

- Image segmentation by a contrario simulation Nicolas Burrus and Thierry M. Bernard and Jean-Michel Jolion *Journal of Pattern Recognition*, À paraître
- Bottom-up and top-down object matching using asynchronous agents and a contrario principles Nicolas Burrus and Thierry M. Bernard and Jean-Michel Jolion 6th International Conference on Computer Vision Systems (ICVS 2008), Mai 2008
- Segmentation d'image par simulations a contrario Nicolas Burrus, Thierry M. Bernard et Jean-Michel Jolion 16e congrès francophone AFRIF-AFIA Reconnaissance des Formes et Intelligence Artificielle, Janvier 2008
- Smart retina as a contour-based visual interface Paul Nadrag and Antoine Manzanera and Nicolas Burrus Distributed Smart Cameras Workshop (DSC'06), Octobre 2006
- Adaptive Vision Leveraging Digital Retinas : Extracting Meaningful Segments Nicolas Burrus and Thierry M. Bernard Advanced Concepts for Intelligent Vision Systems International Conference (ACIVS'06), Septembre 2006

# Compression adaptative de surfaces par ondelettes géométriques

Roudet Céline

THESE

Soutenue le 24/11/2008

à l'Université Claude Bernard Lyon 1

## Jury

Marc ANTONINI, I3S Sophia Antipolis	rapporteur
Frédéric TRUCHETET, Université de Bourgogne, le Creusot	rapporteur
Marc DANIEL, Université de la Méditerranée, Marseille	examinateur
Patrick GIOIA, Orange Labs, Cesson-Sévigné	examinateur
Atilla BASKURT, INSA de Lyon	co-directeur
Florent DUPONT, Université Claude Bernard Lyon 1	co-directeur

Contact : croudet@liris.cnrs.fr

## Résumé

L'évolution de l'infographie et des techniques de numérisation a récemment ouvert la voie à une modélisation tridimensionnelle du monde qui nous entoure. Afin de s'adapter à l'hétérogénéité des ressources et médias manipulant ces objets 3D, des techniques basées sur l'analyse multirésolution sont généralement utilisées car elles fournissent une représentation "scalable" de ces modèles géométriques. C'est dans ce cadre de compression et de transmission progressive d'objets 3D (modélisées sous forme de maillages surfaciques) que se situe ce travail de thèse, réalisé dans le cadre du projet "CoSurf" (collaboration entre le laboratoire LIRIS et France Télécom R&D Rennes). Nous proposons ainsi une nouvelle méthode de compression hiérarchique s'appuyant sur une décomposition en ondelettes, outil d'analyse performant et robuste qui a fait ses preuves en termes de compression d'images et de vidéos. Notre méthode se démarque des techniques existantes, puisqu'elle s'appuie sur une segmentation préalable de la surface en régions d'amplitudes fréquentielles variables. Les partitions résultantes peuvent ainsi être traitées indépendamment durant les phases d'analyse multirésolution, de quantification et d'allocation binaire, de façon à s'adapter aux caractéristiques surfaciques locales des maillages et ainsi réduire les informations à coder. La contribution visuelle de chacune des partitions à l'ensemble de la surface est également un point important à considérer dans la phase d'optimisation des bits alloués à celles-ci, notamment pour des applications comme la transmission et la visualisation sélectives. D'autres applications telles que le tatouage, le filtrage ou le débruitage adaptatifs, l'indexation ou enfin la correction d'erreurs après transmission sur un canal bruité, pourraient bénéficier de ce concept générique que nous avons proposé.

### Mots clés :

Maillages surfaciques, ondelettes géométriques, analyse multirésolution, schéma lifting, compression progressive, segmentation, transmission sélective.

## Abstract

The recent advances in computer graphics and digitization allow access to an ever finer three-dimensional modelling of the world. The critical challenges with 3D models lie in their transmission and rendering, which must fit the heterogeneity of the end resources (network bandwidth, display terminals ...). In this context, this thesis investigates the progressive compression and transmission of 3D models, based on multiresolution analysis, to provide a scalable representation of these geometric models. This work is part of "CoSurf", a collaborative research project involving LIRIS laboratory and France Télécom R&D in Rennes. The proposed hierarchical compression method is based on a wavelet decomposition, which is a robust and competitive analysis tool incorporated in the JPEG2000 and MPEG4 standards. The innovation, compared to existing techniques, lies in a prior surface segmentation in regions, each having a different frequency

amplitude. Hence, it is possible to treat independently each resulting partition during the wavelet decomposition, quantization and binary allocation. The advantage is to fit the local features of the mesh surface, which can reduce the overall encoding information. It is finally important to consider each patch visual contribution to the entire surface, to optimize the bit allocation, essential for view-dependent transmission and visualization. Other applications like adaptive watermarking, filtering or denoising, indexing or error-resilient coding may benefit from this generic concept we proposed.

*Keywords:*

Meshes, geometric wavelets, multiresolution analysis, lifting scheme, progressive compression, 3-D mesh partitioning, view-dependent transmission.

## Publications

*Conférences internationales avec comité de lecture :*

- C. Roudet, F. Dupont, A. Baskurt. Semi-Regular 3D Mesh Progressive Compression and Transmission based on an Adaptive Wavelet Decomposition. Dans Proceedings of SPIE Volume 7248 - Wavelet Applications in Industrial Processing VI, San Jose, CA United States. Janvier 2009.
- C. Roudet, F. Dupont, A. Baskurt. Multiresolution mesh segmentation based on surface roughness and wavelet analysis. Dans Proceedings of SPIE Volume 6508 (2) - Visual Communications and Image Processing (VCIP) 2007, pp. 65082E.1-65082E.12. Janvier 2007, San Jose, California - USA.

*Conférences nationales avec comité de lecture :*

- C. Roudet, F. Dupont, A. Baskurt. Mesure de la concavité et de la convexité dans les maillages surfaciques 3D par décomposition en ondelettes. Dans COMpression et REprésentation des Signaux Audiovisuels (CORESA'09). 19-20 Mars 2009, Toulouse, France.
- C. Roudet, F. Dupont, A. Baskurt : Comparaison de schémas de décomposition en ondelettes pour un traitement local des maillages surfaciques triangulaires. Dans COMpression et REprésentation des Signaux Audiovisuels (CORESA'06). 9-10 Novembre 2006, Caen, France.

*Journées nationales sans acte :*

- C. Roudet, F. Dupont, A. Baskurt. Compression adaptative de maillages 3D à partir d'une segmentation basée sur les ondelettes de subdivision. Journées "Informatique et Géométrie" (JIG'08). 4-5 Juin 2008, Dijon, France.
- C. Roudet, F. Dupont, A. Baskurt. Analyse multirésolution adaptative pour la compression de maillages surfaciques triangulaires. Journée Compression Géométrique du GdR ISIS. 19 octobre 2006, Paris, France.
- C. Roudet, F. Dupont, A. Baskurt. Comparaison de schémas de décomposition en ondelettes pour les maillages surfaciques triangulaires. Journées "Informatique et Géométrie" (JIG'06). 1 et 2 juin 2006, Lyon, France.

*Rapports de contrat et internes :*

- C. Roudet, F. Dupont, A. Baskurt. Adapted Semi-Regular 3-D Mesh Coding Based on a Wavelet Segmentation. Rapport de recherche RR-LIRIS-2008-004, Janv. 2008. URL : <https://liris.cnrs.fr/publis?id=3346>
- C. Roudet, F. Dupont, A. Baskurt. Multiresolution mesh segmentation based on surface roughness and wavelet analysis. Rapport de recherche RR-LIRIS-2006-018, Nov. 2006. URL : <https://liris.cnrs.fr/publis?id=2597>
- C. Roudet. Etat de l'art en compression de maillages surfaciques et bilan de l'acquis FT/LIRIS. Rapport intermédiaire dans le cadre du projet CoSurf (Compression de Surfaces). LIRIS Lyon - France Télécom R&D Rennes, 17 octobre 2006.
- C. Roudet, F. Dupont, A. Baskurt. Comparaison de schémas de décomposition en ondelettes pour un traitement local des maillages surfaciques triangulaires. Rapport de recherche RR-LIRIS-2006-010 ; Juin 2006. URL : <https://liris.cnrs.fr/publis/?id=2399>



**Vacavant Antoine**

Soutenue le 04/12/2008

à l'Université Lumière Lyon 2

### Jury

Eric Andrès, SIC, Poitiers	rapporteur
Edouard Thiel, LIF, Marseille	rapporteur
Walter G. Kropatsch, TU-Wien, Vienne, Autriche	rapporteur
Annick Montanvert, LIS, Grenoble	examineur
Jean-Michel Jolion, LIRIS, Lyon s	examineur
Laure Tougne, LIRIS, Lyon	directrice
David Coeurjolly, LIRIS, Lyon	co-directeur

Contact : antoine.vacavant@liris.cnrs.fr

### Résumé

Les systèmes d'acquisition de données image en deux ou trois dimensions (2-D ou 3-D) fournissent généralement des données organisées sur une grille régulière, appelées données discrètes. Que ce soit pour la visualisation ou l'extraction de mesures, la géométrie discrète définit les outils mathématiques et géométriques pour de nombreuses applications. Dans cette thèse, nous nous intéressons à l'adaptation des algorithmes de la géométrie discrète aux grilles irrégulières isothétiques. Ce modèle de grille permet de représenter de manière générique les structurations d'images en pixels ou voxels de taille et de position variable (ou cellules) : les grilles anisotropes, très répandues en imagerie médicale, les décompositions hiérarchiques telles que quadtree/octree, les techniques de compression comme le run length encoding, etc. Nous proposons d'étendre deux méthodologies largement étudiées pour analyser les formes discrètes à cette représentation : la reconstruction d'objets binaires complexes et la transformée en distance. Pour réaliser la reconstruction topologique et géométrique d'objets irréguliers, nous construisons le graphe de Reeb discret irrégulier, qui résume la forme de l'objet traité. Ensuite, nous calculons une structure polygonale exacte. Ces deux représentations peuvent être rapidement mises à jour, lors du raffinement d'une cellule de l'objet, ou quand plusieurs cellules sont groupées ensemble. Nous proposons deux applications de ce système : la distinction de caractères ambigus dans un outil de reconnaissance de plaques minéralogiques, et l'approximation de courbes implicites planaires. La transformée en distance d'une image régulière peut être calculée par de nombreux algorithmes, dont le point commun est généralement le calcul d'un diagramme de Voronoï discret. Dans cette thèse, nous étudions le calcul de cette décomposition sur grilles irrégulières isothétiques, et nous proposons un algorithme optimal linéaire (en le nombre de cellules de la grille) en 2-D. Nous avons également développé deux techniques extensibles à n dimensions, dont la complexité (temps et espace) dépend principalement de l'irrégularité de la grille traitée.

*Mots clés* : Géométrie discrète, grilles irrégulières isothétiques, reconstruction polygonale, arithmétique d'intervalles, transformée en distance, diagramme de Voronoï.

### Abstract

The systems that perform the acquisition of two or three (2-D or 3-D) dimensional images usually provide data structured on a regular grid, also called discrete data. Those discrete objects are now efficiently handled: for measure extraction or visualization, discrete geometry defines geometrical and mathematical tools for many applications. In this PhD thesis, our goal is to adapt algorithms developed in discrete geometry to irregular isothetic grids. This generic grid model allows to represent the image structures based on pixels or voxels with variable position and size (or cells): anisotropic grids, commonly used in

medical imagery, hierarchical decompositions like quadtree/octree, the compression methods like the run length encoding, etc. We propose to extend two methodologies, widely studied for discrete shape analysis, to this representation: the reconstruction of complex binary objects and distance transformation. To process the topological and geometrical reconstruction of irregular objects, we build the irregular discrete Reeb graph, that sums up the shape of the input object. Then, we compute an exact polygonal structure. Those two representations can be quickly updated, when a cell of the object is refined, or when several cells are grouped together. We propose two applications of our system: the distinction of ambiguous characters in a software for licence plate recognition, and planar implicit curve approximation. The distance transformation of a regular image may be computed with many algorithms, their common point is generally that they build the discrete Voronoi diagram. In this PhD thesis, we study the computation of this decomposition on irregular isothetic grids and we propose an optimal and linear algorithm (in respect to the number of cells in the grid) in 2-D. We have also developed two algorithms extensible to  $n$  dimensions, whose complexity (time et space) principally depends on the irregularity of the input grid.

*Keywords:* Discrete geometry, irregular isothetic grids, polygonal reconstruction, interval arithmetic, distance transformation, Voronoi diagram.

## Publications

### *Article dans une revue internationale*

- Antoine Vacavant, David Coeurjolly et Laure Tougne. A framework for dynamic implicit curve approximation by an irregular discrete approach. Soumis à Graphical Models (en révision).

### *Articles dans une conférence internationale avec comité de relecture*

- Antoine Vacavant, David Coeurjolly and Laure Tougne. Distance Transformation on Two-Dimensional Irregular Isothetic Grids. In 14th International Conference on Discrete Geometry for Computer Imagery (DGCI 2008), LNCS 4292, 238-249. 16-18 April 2008, Lyon, France.
- Nicolas Thome and Antoine Vacavant. A Combined Statistical-Structural Strategy for Alphanumeric Recognition. In 3rd International Symposium on Visual Computing (ISVC 2007), LNCS 4842, (2):529-538. 26-28 November 2007, Lake Tahoe, Nevada.
- Antoine Vacavant, David Coeurjolly and Laure Tougne. Dynamic reconstruction of complex planar objects on irregular isothetic grids. In 2nd International Symposium on Visual Computing (ISVC 2006), LNCS 4292, (2):205-214. 6-8 November 2006, Lake Tahoe, Nevada.
- Antoine Vacavant, David Coeurjolly and Laure Tougne. Topological and geometrical reconstruction of complex objects on irregular isothetic grids. In the 13th International Conference on Discrete Geometry for Computer Imagery (DGCI 2006), LNCS 4245, pages 470-481. 25-27 October 2006, Szeged, Hungary.

### *Article dans une conférence nationale avec comité de relecture*

- Antoine Vacavant, David Coeurjolly et Laure Tougne. Reconstruction topologique et géométrique d'objets complexes sur grilles isothétiques irrégulières. Dans COmpression et REpresentation des Signaux Audiovisuels (CORESA 2006). 9-10 Novembre 2006, Caen, France.

### *Présentations dans une conférence nationale*

- Antoine Vacavant. Approximation de courbes implicites grâce à des outils discrets irréguliers. Dans les Journées Informatique et Géométrie (JIG 2008). 4-5 Juin 2008, Dijon, France.
- Antoine Vacavant. Outils de la géométrie discrète pour le tracé de courbes implicites par arithmétique d'intervalles. Dans les Rencontres Arithmétiques de l'Informatique Mathématique (RAIM 2007). 22-25 Janvier 2007, Montpellier, France.
- Antoine Vacavant et David Sarrut. Structuration sur grilles isothétiques irrégulières dans les simulations de radiothérapie. Dans les Journées Informatique et Géométrie. 13-14 Octobre 2005, Paris, France.

### *Rapport de recherche*

- Antoine Vacavant, David Coeurjolly et Laure Tougne. A framework for dynamic implicit curve approximation by an irregular discrete approach. Rapport de recherche RR-LIRIS-2007-020, 2007.

**Xiao Zhongzhe**

Soutenue le 25/01/2008

à l'Ecole Centrale de Lyon

### Jury

Mme. ANDRE-OBRECHT Régine, Université Paul Sabatier	rapporteur
M. BESACIER Laurent, Université Joseph Fourier	rapporteur
M. CHEN Liming, Ecole Centrale de Lyon	directeur
M. DELLANDREA Emmanuel, Ecole Centrale de Lyon	co-directeur
Mme. DOU Weibei, Université de Tsinghua	co-directrice

Contact : xiaozhongzhe@gmail.com

### Résumé

Les travaux de recherche réalisés dans le cadre de cette thèse de doctorat portent sur la reconnaissance automatique de l'émotion et de l'humeur au sein de signaux sonores. En effet, l'émotion portée par les signaux audio constitue une information sémantique particulièrement importante dont l'analyse automatique offre de nombreuses possibilités en termes d'applications, telles que les interactions homme-machine intelligentes et l'indexation multimédia. L'objectif de cette thèse est ainsi d'étudier des solutions informatique d'analyse de l'émotion audio tant pour la parole que pour les signaux musicaux.

Nous utilisons dans notre travail un modèle émotionnel discret combiné à un modèle dimensionnel, en nous appuyant sur des études existantes sur les corrélations entre les propriétés acoustiques et l'émotion dans la parole ainsi que l'humeur dans les signaux de musique. Les principales contributions de nos travaux sont les suivantes. Tout d'abord, nous avons proposé, en complément des caractéristiques audio basées sur les propriétés fréquentielles et d'énergie, de nouvelles caractéristiques harmoniques et Zipf, afin d'améliorer la caractérisation des propriétés des signaux de parole en terme de timbre et de prosodie. Deuxièmement, dans la mesure où très peu de ressources pour l'étude de l'émotion dans la parole et dans la musique sont disponibles par rapport au nombre important de caractéristiques audio qu'il est envisageable d'extraire, une méthode de sélection de caractéristiques nommée ESFS, basée sur la théorie de l'évidence est proposée afin de simplifier le modèle de classification et d'en améliorer les performances. De plus, nous avons montré que l'utilisation d'un classifieur hiérarchique basé sur un modèle dimensionnel de l'émotion, permet d'obtenir de meilleurs résultats de classification qu'un unique classifieur global, souvent utilisé dans la littérature. Par ailleurs, puisqu'il n'existe pas d'accord universel sur la définition des émotions de base, et parce que les états émotionnels considérés sont très dépendant des applications, nous avons également proposé un algorithme basé sur ESFS et permettant de construire automatiquement un classifieur hiérarchique adapté à un ensemble spécifique d'états émotionnels dans le cadre d'une application particulière. Cette classification hiérarchique procède en divisant un problème de classification complexe en un ensemble de problèmes plus petits et plus simples grâce à la combinaison d'un ensemble de sous-classifieurs binaires organisés sous forme d'un arbre binaire. Enfin, les émotions étant par nature des notions subjectives, nous avons également proposé un classifieur ambigu, basé sur la théorie de l'évidence, permettant l'association d'un signal audio à de multiples émotions, comme le font souvent les êtres humains.

L'efficacité de ces techniques de reconnaissance a été évaluée sur les ensembles de données Berlin et DES pour la reconnaissance de l'émotion dans la parole, et sur un ensemble de données construit au sein de notre laboratoire pour l'humeur dans les signaux de musique, dans la mesure où il n'existe pour le moment aucun jeu de données public.

## Abstract

This Ph.D thesis work is dedicated to automatic emotion/mood recognition in audio signals. Indeed, audio emotion is high semantic information and its automatic analysis may have many applications such as smart human-computer interactions or multimedia indexing. The purpose of this thesis is thus to investigate machine-based audio emotion analysis solutions for both speech and music signals.

Our work makes use of a discrete emotional model combined with the dimensional one and relies upon existing studies on acoustics correlates of emotional speech and music mood. The key contributions are the following. First, we have proposed, in complement to popular frequency-based and energy-based features, some new audio features, namely harmonic and Zipf features, to better characterize timbre and prosodic properties of emotional speech. Second, as there exists very few emotional resources either for speech or music for machine learning as compared to audio features that one can extract, an evidence theory-based feature selection scheme named Embedded Sequential Forward Selection (ESFS) is proposed to deal with the classic "curse of dimensionality" problem and thus over-fitting. Third, using a manually built dimensional emotion model-based hierarchical classifier to deal with fuzzy borders of emotional states, we demonstrated that a hierarchical classification scheme performs better than single global classifier mostly used in the literature. Furthermore, as there does not exist any universal agreement on basic emotion definition and as emotional states are typically application dependent, we also proposed a ESFS-based algorithm for automatically building a hierarchical classification scheme (HCS) which is best adapted to a specific set of application dependent emotional states. The HCS divides a complex classification problem into simpler and smaller problems by combining several binary sub-classifiers in the structure of a binary tree in several stages, and gives the result as the type of emotional states of the audio samples. Finally, to deal with the subjective nature of emotions, we also proposed an evidence theory-based ambiguous classifier allowing multiple emotions labeling as human often does.

The effectiveness of all these recognition techniques was evaluated on Berlin and DES datasets for emotional speech recognition and on a music mood dataset that we collected in our laboratory as there exist no public dataset so far.

## Publications

- Zhongzhe Xiao, Emmanuel Dellandrea, Weibei Dou, Liming Chen, "Automatic Hierarchical Classification of Emotional Speech," *ismw*, pp. 291-296, Ninth IEEE International Symposium on Multimedia Workshops (ISMW 2007), 2007
- Zhongzhe Xiao, Emmanuel Dellandrea, Weibei Dou, Liming Chen, Two-stage Classification of Emotional Speech, *International Conference on Digital Telecommunications (ICDT'06)*, p. 32-37, August 29 - 31, 2006, Cap Esterel, Côte d'Azur, France.
- Zhongzhe Xiao, Emanuel Dellandréa, Weibei Dou, Liming Chen., Features extraction and selection in emotional speech, *International Conference on Advanced Video and Signal based Surveillance (AVSS 2005)*. p. 411-416. September 2005, Como, Italy.
- Zhongzhe Xiao and Zaiwang Dong, Improved GIB Synchronization Method for OFDM Systems, *10th International Conference on Telecommunications ICT'2003*, pp 1417-1421, Volume II, February 23-28, 2003, Tahiti, Papeete – French Polynesia.

### *Research report:*

- Zhongzhe Xiao, Emmanuel Dellandréa, Weibei Dou, Liming Chen, A Dimensional Emotion Model Driven Multi-stage Classification of Emotional Speech, research report RR-LIRIS-2007-033, LIRIS UMR 5205 CNRS, 2007, submitted to *IEEE Transactions on Audio, Speech, and Language Processing*
- Zhongzhe Xiao, Emmanuel Dellandréa, Weibei Dou, Liming Chen, Hierarchical Classification of Emotional Speech, research report RR-LIRIS-2007-006, LIRIS UMR 5205 CNRS, 2007