

## Séance anniversaire Graphes, arbres et automates en informatique musicale

Vendredi 20 mai 2011

de 10h à 18h

Ircam, Salle I. Stravinsky  
1, place I. Stravinsky 75004 Paris  
(Entrée libre dans la mesure des places disponibles)

Cette séance conclusive du séminaire MaMuX est consacrée aux structures de graphes, d'arbres et d'automates en informatique et en musique. La séance, clôturant dix ans d'existence du séminaire, accueillera également un concert d'improvisation assistée par ordinateur à l'aide du logiciel OMAX (avec la participation de Vincent Lê Quang, saxophoniste et Professeur au CNSMDP, classe d'Improvisation Générative).

### Programme :

- 10h00 - 10h10 **Moreno Andreatta** et **Carlos Agon** - Présentation de la journée
- 10h10 - 10h50 **Thierry Lecroq** - An Efficient Motif Search Algorithm based on a Minimal Forbidden Patterns Approach
- 10h50 - 11h30 **Arnaud Lefevre** - Computing Abelian periods in words
- 11h40 - 12h20 **Fivos Maniatakos** - Multi Factor Graph (MFG) et applications à l'improvisation d'ordinateur
- 12h20 - 13h00 **Cyril Gavaille** - Oracles pour les arbres et les graphes
- 
- 14h30 - 15h10 **Alessandra Carbone** - Encoding of three-dimensional information in DNA sequences
- 15h10 - 15h50 **David Auber** - Treemap Géographique
- 15h50 - 16h30 **Pierre Roy** - Chaînes de Markov contraintes pour la génération contrôlée de séquences temporelles
- 16h50 - 17h30 **Gérard Assayag** et **Benjamin Lévy** - OMax : théorie et pratique
- 17h30 - 18h00 **Concert OMAX**, avec la participation de Vincent Lê Quang (saxophoniste), Georges Bloch (CNSMDP/Ircam), Gérard Assayag (Ircam/CNRS) et Benjamin Lévy (Ircam-CNRS / UPMC).
- Cocktail de clôture

### Résumés

**Thierry Lecroq** (université de Rouen, LITIS) - An Efficient Motif Search Algorithm based on a Minimal Forbidden Patterns Approach

Finding common motifs from a set of strings coding biological sequences is an important problem in Molecular Biology. Several versions of the motif finding problem have been proposed in the literature and for each version, numerous algorithms have been developed. However, many of these algorithms fall under the category of heuristics. In this talk, we concentrate on the Simple Motif Problem (SMP) and we propose two exact algorithms, called SMS-Forbid and SMS-H-Forbid for this version of motif finding problem. These algorithms are based on a Minimal Forbidden Patterns Approach.

**Arnaud Lefevre** (université de Rouen) - Computing Abelian periods in words

In the last couple of years many works have been related to Abelian complexity and efficient algorithms for Abelian pattern matching have been designed. However, apart of the greedy off-line algorithm, neither efficient nor on-line algorithm is known for computing all the Abelian periods of a given word. In this talk we present several and efficient off-line and on-line algorithms for computing all the Abelian periods of a given word. (Joint work with G. Fici, T. Lecroq and E. Prieur).

**Fivos Maniatakos** (doctorant Ircam / UPMC) - Multi Factor Graph (MFG) et applications à l'improvisation d'ordinateur

Nous présenterons une nouvelle structure informatique pour l'alignement des séquences sous le nom de Multi Factor Graph (MFG). L'avantage de MFG par rapport aux structures existantes conçues pour "exact matching" (automates et arbres des suffixes) est le coût réduit en espace pour sa construction ( $<2N$ ) via un algorithme linéaire en temps et efficace concernant la gestion de mémoire. Dans cet exposé nous présenterons le travail fondamental mené autour de MFG (formalismes, algorithmes des constructions, complexité et requêtes primitives). Puis nous ferons le lien avec les applications musicales basées sur des modèles séquentiels, notamment l'improvisation d'ordinateur et plus généralement la synthèse fondée sur les données. Dans ce cadre, nous évoquerons des stratégies permettant de contrôler la génération du contenu musical à l'aide des requêtes primitives liées à MFG.

**Cyril Gavaille** (LaBRI) - Oracles pour les arbres et les graphes

Je présenterai quelques résultats récents autour de la représentation efficace (codage) des distances dans les arbres et plus généralement dans les graphes. En deuxième partie j'évoquerai la construction d'oracle répondant à des requêtes plus complexes, prenant en compte, par exemple, la suppression de sommets ou d'arêtes.

**Alessandra Carbone** (Université Pierre et Marie Curie) - Encoding of three-dimensional information in DNA sequences

Given a family of protein sequences and the associated distance tree, we shall explain how important information related to the functional activity (protein interactions), the folding dynamics and the conformational changes of a protein is encoded in DNA sequences. A fine reading of the conservation and co-evolution signals between residues in the sequences lead to predictions that are confirmed experimentally. The role of the topology of the distance tree in the encoding will be highlighted.

**David Auber** (Equipe MABioVis, LaBRI) - Treemap Géographique

Au cours de cette présentation je présenterai les méthodes de représentation de données hiérarchiques. Je terminerai par une nouvelle technique basée sur l'utilisation de courbe fractale. Cette méthode permet la représentation des arborescences en utilisant la métaphore des cartes géographiques.

**Pierre Roy** (Sony Lab) - Chaînes de Markov contraintes pour la génération contrôlée de séquences temporelles

Les chaînes de Markov ont de nombreuses applications pour la modélisation de processus temporels tels que la structure de textes ou de mélodies. L'hypothèse de mémoire limitée permet d'utiliser les chaînes de Markov pour engendrer des séquences à l'aide d'algorithmes gloutons (ou *greedy*). Dans de nombreuses applications, telles que la composition ou l'improvisation musicales, il est essentiel de contrôler la structure des séquences engendrées. Les contraintes de contrôle induisent en général des dépendances entre éléments distants, ce qui viole l'hypothèse de Markov. Il est de ce fait impossible de contrôler la génération de séquences markoviennes à l'aide d'algorithmes gloutons.

Nous montrons qu'en revisitant les chaînes de Markov dans le contexte de la satisfaction de contraintes il est possible de contrôler la génération de séquences. Nous introduisons pour ce faire une classe de contraintes, appelée Markov constraints, et démontrons deux résultats :

- 1) La génération de séquences Markoviennes avec des contraintes de contrôle arbitraires peut être vue comme un problème d'optimisation de contraintes (*Constraint Optimization Problem*) et montrons qu'une implémentation efficace des "Markov Constraints" permet de le résoudre;
- 2) Si l'on se restreint à un type particulier de contraintes de contrôle, il est possible de compiler les dépendances entre éléments distants de la séquence dans une chaîne de Markov, non homogène. Il est dans ce cas possible de contrôler la génération de séquence à l'aide d'algorithmes gloutons;

Nous illustrons le premier résultat avec un problème de composition musicale: la génération d'une grille d'accords de Blues complexe et montrons que le second résultat permet de construire des applications d'improvisation en jazz.

**Gérard Assayag** (Ircam/CNRS) et **Benjamin Lévy** (doctorant Ircam / UPMC) - OMax : théorie et pratique

Basé sur l'algorithme de l'Oracle des Facteurs (Allauzen et al. 1999 & Lefebvre et al. 2000) le logiciel d'improvisation OMax est maintenant dans sa quatrième version. De la cartographie motivique d'une séquence musicale encodée dans le graphe de suffixe, l'on passe à la variation musicale en navigant et recombinant des éléments du discours. Après quelques rappels sur le fonctionnement interne d'OMax, nous présenterons les aspects visuels et interactifs d'OMax en interaction avec un musicien. Puis nous laisserons place à la musique avec une petite performance avec le saxophoniste Vincent Lê Quang.

#### Références bibliographiques :

- [1] Cyril Allauzen, Maxime Crochemore, Mathieu Raffinot (1999), « Factor oracle: a new structure for pattern matching », SOFSEM'99, vol. 1725, pp. 291-306, Springer-Verlag.
- [2] Arnaud Lefebvre, Thierry Lecroq (2000), « Computing repeated factors with a factor oracle », IWOC'A2000, p. 145-158.

#### Planning du séminaire :

- Vendredi 8 octobre 2010 : modèles formels de la pensée contrapuntique. Avec la participation de Julien Junod, Karim Haddad et Moreno Andreatta
- Vendredi 10 décembre 2010 : systèmes d'intervalles généralisés et théorie de l'homométrie. Avec la participation de Daniele Ghisi, John Mandereau et Guillaume Lachaussée.
- Samedi 11 décembre 2010 : école mathématique pour musiciens et autres non-mathématiciens animée par Pierre Cartier
- Vendredi 14 janvier 2011 : music and mathematics as seen by composers. Tom Johnson with five young composers (Christopher Adler, Steve Gisby, Brian Parks, Samuel Vriezen, Michael Winter)
- Vendredi 4 février 2011 : mathesis et subjectivité. Autour de la logique musicale. Avec la participation de Hugues Dufourt.
- Samedi 5 février 2011 : école mathématique pour musiciens et autres non-mathématiciens animée par Pierre Cartier
- Vendredi 11 mars 2011 : musicologie computationnelle et composition algorithmique. Hommage à Marcel Mesnage. Séance organisée en collaboration et avec le soutien de la SFAM.
- Vendredi 1er avril 2011 : block designs en composition. Avec la participation de Tom Johnson, Gilbert Delor et Franck Jedrzejewski
- Samedi 30 avril 2011 : Séance conclusive de l'école mathématique pour musiciens et autres non-mathématiciens animée par Pierre Cartier (avec la participation d'Annick Lesne, CNRS/UPMC)
- Vendredi 20 mai 2011 : graphes et arbres en informatique musicale. Avec la participation de Gérard Assayag, David Aubert, Georges Bloch, Alessandra Carbone, Cyril Gavoille, Thierry Lecroq, Arnaud Lefebvre, Benjamin Lévy, Vincent Lê Quang, Fivos Maniatakos, Pierre Roy.

#### Planning de la prochaine saison du séminaire (2011-2012) :

- Vendredi 7 octobre 2011
- Vendredi 4 novembre 2011
- Vendredi 2 décembre 2011
- Vendredi 6 janvier 2012
- Vendredi 3 février 2012
- Vendredi 9 mars 2012
- Vendredi 30 mars 2012
- Vendredi 4 mai 2012

#### Nouveaux contacts :

Pour tout renseignement, contacts et propositions pour la saison 2011-2012 :

John Mandereau (John.Mandereau[at]ircam.fr)

Jean Bresson (Jean.Bresson[at]ircam.fr)

Le Séminaire est organisé par L'Equipe Représentations Musicales de l'IRCAM, en collaboration avec Guerino Mazzola (MultiMediaLab de Université de Zürich / School of Music, University of Minnesota), Franck Jedrzejewski (CEA Saclay - INSTN/UESMS), Thomas Noll (Escola Superior de Musica de Catalunya) et avec le soutiens du CNRS (UMR STMS - Sciences et technologies de la musique et du son).

