LANFOR 01

Compte-rendu par les organisateurs:
Sandrine Julia (Nice) et Géraud Sénizergues (Bordeaux 1)

La réunion du groupe **LANFOR** (**LAN**gages **FOR**mels) de l' AFIT s'est tenue à Buis-les-Baronnies (Drôme, France) du 16/01/01 (soir) au 19/01/01 (midi).
V. Bruyère a accepté de donner trois cours sur le théorème de Kleene et ses extensions.

****

Liste des participants

AUGROS Xavier	I3S, Nice	augros@essi.fr
BRUYERE Véronique	UMH, Mons	vero@sun1.umh.ac.be
CALBRIX Hugues	LAPSC, Lyon	calbrix@jonas.univ-lyon1.fr
CHAMPARNAUD Jean-Marc	LIFAR , Rouen	jmc@dir.univ-rouen.fr
DE-LA-HIGUERA Colin	EURISE, St Etienne	Colin.Delahiguera@univ-st-etienne.fr
DUCHAMP Gérard	LIFAR ,Rouen	ged@dir.univ-rouen.fr
GAREL Emmanuelle	IRISA, Rennes 1	garel@insa-rennes.fr
JANODET Jean-Christophe	EURISE, St Etienne	janodet@univ-st-etienne.fr
JULIA Sandrine	I3S, Nice	julia@unice.fr
LEPELLETIER Régis	LIFAR, Rouen	Regis.Lepelletier@dir.univ-rouen.fr
LOMBARDY Sylvain	ENST, Paris	lombardy@inf.enst.fr
LY Olivier	LaBRI, Bdx 1	ly@labri.u-bordeaux.fr
MORVAN Christophe	IRISA, Rennes 1	Christophe.Morvan@irisa.fr
PATROU Bruno	LIFAR, Rouen	Bruno.Patrou@dir.univ-rouen.fr
PRIEUR Christophe	LIAFA, Paris	prieur@liafa.jussieu.fr
REGNIER Mireille	INRIA, Paris	Mireille.Regnier@inria.fr
SEEBOLD Patrice	LaRia, Amiens	seebold@laria.u-picardie.fr
SENIZERGUES Géraud	LaBRI, Bdx 1	ges@labri.u-bordeaux.fr
TERLUTTE Alain	LIFL,Lille 3	terlutte@univ-lille3.fr
ZIADI Djelloul	LIFAR ,Rouen	Djelloul.Ziadi@dir.univ-rouen.fr

****

Résumés

Le théorème de Kleene et ses extensions
V. Bruyère
(En collaboration avec O. Carton)
(Université de Mons-Hainaut, Belgique)
pages personnelles: `http://sun1.umh.ac.be/~vero/index.html`

Résumé Dans ces exposés, on considère des mots indicés par des ordres linéaires, généralisant ainsi les mots finis, les mots infinis, les mots sur les ordinaux. On introduit des automates et des expressions rationnelles adaptés aux mots indicés par des ordres linéaires. On présente un théorème de Kleene affirmant que ces deux concepts sont équivalents pour les ordres dénombrables dispersés.

Code et factorisations selon des trajectoires
S. Augros
(En collaboration avec M. Ahmad)
(I3S, Université de Nice)
pages personnelles: `http://deptinfo.unice.fr/~augros`

Résumé En préparation.

Un langage caractéristique pour les ω-puissances rationnelles
S. Julia
(I3S, Université de Nice)
pages personnelles: `http://deptinfo.unice.fr/~julia`

Résumé Notre étude porte sur les langages de mots infinis, aussi appelés ω-langages. Considérons une ω-puissance L^ω sur un alphabet Σ et appelons $\chi(L^\omega)$ le langage suivant : $\chi(L^\omega) = \{u \in \Sigma^+, uL^\omega \subseteq L^\omega \text{ et } u^\omega \in L^\omega\}$. On sait déjà que ce langage est rationnel et que c'est un majorant de l'ensemble des ω-générateurs de L^ω, lui-même ω-générateur de L^ω dès qu'il est semi-groupe.
On démontre que deux ω-puissances rationnelles sont égales si elles ont le même langage χ. On en déduit que deux ω-langages, à la fois ω-puissances et adhérences sont égaux si et seulement si ils ont le même ensemble Per_χ où Per_χ est l'ensemble des ω-mots périodiques avec une période dans χ.

Topologie et transducteurs de mots infinis

Christophe Prieur
(LiAFA, Université Paris 7)
pages personnelles: `http://www.liafa.jussieu.fr/~prieur/`

Résumé La “manipulation” des mots infinis nécessite de pouvoir définir la notion d'“approximation”, qui mène à des considérations topologiques. En pratique, il s'agit de pouvoir travailler sur des préfixes de ces mots.

En premier lieu, on montre comment calculer l'adhérence d'une relation rationnelle (*i.e.* définie par transducteur — ou automate avec sortie), qui se trouve être elle-même rationnelle. La taille d'un transducteur qui la réalise peut être quadratique par rapport au transducteur de départ (ce qui n'est pas le cas pour l'adhérence des langages rationnels).

Enfin la propriétété de continuité uniforme d'une fonction rationnelle, qui peut être vue comme permettant la mise en parallèle de plusieurs fonctions, peut se décider en temps quadratique.

Expressions rationnelles, dérivées et automates : cas Booléen

J.-M. Champarnaud
(En collaboration avec D. Ziadi)
(LIFAR, Université de Rouen)

Résumé On s'intéresse aux relations qui existent entre les différents automates produits par conversion d'une expression rationnelle : automate "pas à pas", ε-automate, automate des positions, automate des continuations, automate des dérivées partielles, automate des prébases. On rappelle les résultats classiques (isomorphisme entre automate "pas à pas", des positions et des continuations; isomorphisme entre automate des dérivées partielles, et des prébases, ...). On introduit le calcul des c-dérivées et l'automate des c-continuations. On montre que l'automate des positions est isomorphe à cet automate, et que l'automate des dérivées partielles s'en déduit par quotient. Il en découle une construction en $O(n^2)$ de l'automate des dérivées partielles, où n est le nombre d'occurrences de symboles dans l'expression.

Expressions rationnelles, dérivées et automates : du Booléen aux multiplicités

D. Ziadi
(En collaboration avec J.-M. Champarnaud)
(LIFAR, Université de Rouen)

Résumé Le passage du Booléen aux multiplicités constitue un programme de recherche défini en commun par les équipes Automates et Calcul Symbolique du LIFAR. Nous considérons ici l'extension des résultats obtenus en matière de conversion d'expression rationelle en automate. Ce travail est en cours de réalisation; nous en donnerons les principaux axes. Un premier point à fixer est la "bonne" définition de la notion de K-expression rationnelle. L'automate des positions d'une K-expression construit par Caron et Flouret (CIAA'2000) suppose par exemple que l'étoile est définie sur les seules expressions à terme constant nul. Nous proposons d'autres définitions possibles. Nous montrons ensuite comment transformer la structure ZPC (représentation implicite de l'automate des positions) pour réaliser le calcul rapide du K-automate des positions. Nous abordons enfin l'extension des constructions basées sur les dérivées.

Mots en variables non-commutatives :
interactions entre l'Informatique Théoriques et les autres sciences

G. Duchamp
(LIFAR, Rouen)

Résumé
L'Informatique Théorique a su développer des modèles pour comprendre le comportement des machines. Ce faisant, elle a du inventer des structures nouvelles qui permettent de résoudre des problèmes dans les autres sciences.

Décomposition des $\mathsf{Z}/2\mathsf{Z}$ modules de fonctions booléennes
R. Lepelletier
(LIFAR, Université de Rouen)
pages personnelles: `http://regis.lepelletier.free.fr`

Résumé Soit f une fonction booléenne à n variables. On cherche à étudier les décompositions du module de générateur f sous l'action à gauche de la $\mathsf{Z}/2\mathsf{Z}$-algèbre du groupe symétrique, l'action du groupe étant donnée par permutation des variables. On utilise pour ce faire la structure de $\mathsf{Z}/2\mathsf{Z}$-automate naturelle dont les lettres sont les transpositions simples.
L'algorithme que nous proposons permet d'effectuer la décomposition du module engendré par cette fonction en sous-modules indécomposables, en résolvant le système linéaire donné par les règles des états retournants. Cet algorithme, à peine modifié permet de déterminer l'orbite de la fonction par le groupe symétrique.

Les graphes rationnels
C. Morvan
(IRISA, Université Rennes 1)
pages personnelles: `http://www.morvan.fr.st/`

Résumé Dans cet exposé, nous avons introduit la famille des graphes rationnels (qui sont reconnus par des transducteurs avec sorties étiquetés). Cette famille est une extension de familles de graphes infinis déjà connus (les graphes *context-free* de Muller et Schupp, les graphes *équationnels* de Bruno Courcelle ou encore les graphes *préfixe-reconnaissables* de Didier Caucal).
Au cours de l'exposé nous avons exhibé certaines de leurs propriétés telles que l'indécidabilité de leur théorie du premier ordre, des caractérisations interne et externe.
Enfin, nous nous sommes intéressés à leurs traces (à savoir l'ensemble des étiquettes des chemins partant d'un sommet pour arriver à un sommet), nous avons montré que ces traces formaient une sous famille des langages contextuels contenant les langages contextuels croissants.

Graphes infinis automatiques et DOL de graphes

Olivier LY
(LaBRI, Université Bordeaux 1)
pages personnelles: `http://dept-info.labri.u-bordeaux.fr/~ly/`

Résumé Le concept central de notre travail est celui de *bout* d'un graphe. Les bouts d'un graphe représentent le comportement du graphe "à l'infini". Nous montrons que le problème de savoir si un graphe automatique infini possède un seul ou bien plusieurs bouts , est récursivement *indécidable*. Le coeur de la preuve est une analyse des propriétés topologiques globales du graphe des configurations d'une machine de Turing auto-stabilisante. Remarquons que ce même problème devient décidable pour la sous-classe des graphes automatiques qui sont les graphes de Cayley de groupes hyperboliques au sens de Gromov (cela a été prouvé récemment). Le résultat d'indécidabilité permet aussi de montrer l'indécidabilité de la connexité de tous les graphes engendrés par un $D0L$-system de graphes. Ce second résultat repose sur un lien très étroit entre suites de graphes engendrées par un $D0L$-system et suites des sphères d'un graphe automatique. Enfin, nous prouvons que les $D0L$-system de graphes peuvent simuler les grammaires "à remplacements d'hyperarcs" de graphes; ce résultat contribue donc à clarifier la frontière entre le décidable et l'indécidable pour le problème de la connexité.

Puissance des langages rationnels

H. Calbrix
(Université Lyon 1)

Résumé Pour tout langage L sur un alphabet fini, on définit $puis(L)$ la puissance du langage L comme le langage formé des puissances non nulles des mots de L, c'est à dire l'ensemble $\{u^k \mid u \in L, k > 0\}$. La puissance d'un langage rationnel n'est généralement pas rationnelle, même avec un alphabet à une lettre. Le but de cet exposé est de donner une caractérisation des langages rationnels sur un alphabet à une lettre dont la puissance reste rationnelle. Ce théorème de caractérisation utilise le théorème de Dirichlet qui affirme que toute progression arithmétique d'entiers, dont le premier terme et la raison sont premiers entre eux, contient une infinité de nombres premiers. La première caractérisation de ces langages a été donnée par Thierry Cachat. On ne connait encore aucune caractérisation de tels langages dans le cas où l'alphabet comporte plusieurs lettres.

Langages et Statistiques des Mots dans le Génome
M. Régnier
(InRIA, Roquencourt)
Résumé En préparation.

De Prouhet à Arshon, en passant par Thue et Morse
P. Seebold
(LaRia, Université Amiens)

Résumé En 1851, Prouhet a décrit une règle pour diviser les n^p premiers entiers positifs en sous-ensembles de n^{p-1} éléments, règle utilisée pour résoudre un problème de théorie des nombres (un cas particulier du problème de Tarri-Escott).

En fait, l'algorithme de Prouhet construit un mot infini sans chevauchement sur n lettres. Ce mot est engendré par un morphisme π_n qui généralise celui sur 2 lettres introduit indépendamment par Thue en 1912 et Morse en 1921.

En 1937, Arshon a donné un algorithme pour construire un mot infini sans carré sur n lettres, a_n. Nous prouvons que la construction d'Arshon peut être réalisée en utilisant le morphisme de Prouhet. Berstel a montré que quand n est impair, a_n est engendré par un tag-système mais pas par un morphisme. Nous prouvons qu'au contraire quand n est pair, a_n est engendré par un morphisme. Sur l'alphabet à n lettres $\{0, \ldots, n-1\}$, ce morphisme vaut π_n pour les lettres de valeur paire et $\widetilde{\pi_n}$ pour celles de valeur impaire.

Inférence Grammaticale
C. De La Higuera
(EURISE, Université de St Etienne)
pages personnelles: `http://www.univ-st-etienne.fr/eurise/cdlh.html`
transparents: `http://www.univ-st-etienne.fr/eurise/cdlh/lanforcdlh.ps`

Résumé L'Inférence Grammaticale relève du domaine général de l'Intelligence Artificielle, et plus particulièrement de l'Apprentissage Automatique. Son objet est de fournir des théories et des algorithmes pour apprendre, à partir d'exemples, quelles sont les règles qui gouvernent la façon dont des éléments primitifs s'enchaînent pour former des séquences cohérentes. L'exemple le plus naturel est celui où les éléments primitifs sont des mots d'une langue : l'inférence grammaticale a alors pour but de trouver les règles de la grammaire de la langue à partir d'exemples de phrases dans cette langue. Si ce cas particulièrement complexe n'est résolu que partiellement à l'heure actuelle, on trouve dans des domaines moins "naturels" de bons sujets d'application. En Reconnaissance des Formes, l'Inférence Grammaticale donne des outils de description de séquences de formes primitives ; en Biochimie, elle permet de modéliser les séquences génétiques ; en Traitement Automatique de la Parole, elle sert à modéliser les grammaires qui régissent les enchaînements de sons élémentaires ou de mots formant une phrase à reconnaître. L'apprentissage de grammaires formelles à partir d'exemples nécessite des algorithmes rapides (taille importante des ensembles de données), robustes (ces données peuvent être bruitées) et fiables (la convergence de l'algorithme est une donnée importante). Dans cet exposé nous présenterons le domaine, ses techniques, modèles et applications, et verrons les principaux algorithmes utilisés.

Automates finis à états résiduels

A. Terlutte
(LIFL, Université Lille 3)
pages personnelles: `http://www.grappa.univ-lille3.fr/~terlutte/`

Résumé Les langages réguliers sont usuellement representés à l'aide d'automates finis déterministes ou non. Les automates finis déterministes (AFD) possèdent de bonnes propriétés, notamment l'existence d'une forme minimale unique dont la construction est basée sur des composants naturels du langage (les langages résiduels). Cependant, la taille de l'AFD minimal peut être beaucoup plus importante qu'une représentation minimale du langage. Les automates finis non déterministes (AFN), quant à eux, ont moins de propriétés intéressantes, mais sont souvent beaucoup plus petits que les AFD, et ce dans des proportions qui peuvent être exponentielles.

Nous présenterons une nouvelle classe d'automates, la classe des automates finis à états résiduels (AFER), qui partage certaines propriétés essentielles avec les AFD et qui possède notamment une forme canonique dont la taille peut être beaucoup plus petite que l'AFD minimal. En un mot, un AFER est un automate fini non déterministe dont chaque état définit un langage résiduel du langage. Nous décrirons également quelques algorithmes et opérateurs utiles au sein de cette classe.