LANFOR 01

report by the organisers:
Sandrine Julia (Nice) and Géraud Sénizergues (Bordeaux 1)

The meeting of the group **LANFOR** (**FOR**mal **LAN**guages) of AFIT was held in Buis-les-Baronnies (Drôme, France) from 16/01/01 (evening) up to 19/01/01 (noon).

V. Bruyère accepted to give three invited lectures on "Kleene's theorem and its generalisations".

****

List of 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

****

Abstracts

Automata on linear orderings
Véronique Bruyère
(Joint work with O. Carton)
(Mons-Hainaut university, Belgique)
personal pages: `http://sun1.umh.ac.be/~vero/index.html`

Abstract In this talk, we consider words indexed by linear orderings. These extend finite, (bi-)infinite words and words on ordinals. We introduce automata and rational expressions for words on linear orderings. We prove that for countable scattered linear orderings they are equivalent. This result extends Kleene's theorem.

Code and factorisations along trajectories
S. Augros
(Joint work with M. Ahmad)
(I3S, Nice university)
personal pages: `http://deptinfo.unice.fr/~augros`

Abstract This talk is about decipherability of a message constructed by the shuffle of two sources of informations. For exemple, a movie has an audio part and a video part which are interlaced to be sent on a network. To play it, the receiver has to reconstruct those two parts. Here, we propose a model, based on regular languages, to get the different parts of the message .

A characteristic language for rational ω-powers
S. Julia
(I3S, Nice university)
personal pages: `http://deptinfo.unice.fr/~julia`

Abstract Consider a rational ω-power L^ω over a finite alphabet Σ and call $\chi(L^\omega)$ the following language : $\chi(L^\omega) = \{u \in \Sigma^+, uL^\omega \subseteq L^\omega \text{ and } u^\omega \in L^\omega\}$. It is already known that this language is rational and an upper bound of the set of ω-generators of L^ω, itself ω-generator of L^ω as soon as it is a semigroup.
We prove that two rational ω-powers are equal if and only if they have the same language χ. We deduce that two ω-languages, both ω-powers and adherence are equal if and only if they have the same set Per_χ where Per_χ denotes the set of periodic ω-words with a period in χ.

Topology and transducers of infinite words
Christophe Prieur
(LiAFA, Paris 7 university)

Abstract Handling infinite words implies approximating them, which leads to topological issues. Practically, this means to be able to do computations on prefixes of words.

First of all, one shows how to compute the adherence of a rational relation (*i.e.* a relation defined by a transducer, which is an automaton with an output function). This adherence happens to be also rational but the size of a transducer realizing it may be quadratic in the size of the one realizing the initial relation (which is not the case for rational languages).

Finally the property of continuity of a rational function can be decided in quadratic time. In away, this property is the ability for functions to be run in parallel.

Expressions rationnelles, dérivées et automates : cas Booléen
J.-M. Champarnaud
(Joint work with D. Ziadi)
(LIFAR, Rouen university)

Abstract 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
(Joint work with J.-M. Champarnaud)
(LIFAR, Rouen university)

Abstract 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; noue 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.

Words in non commutative variables:
some interactions between Theoretical Computer Science et other scientific areas

G. Duchamp
(LIFAR, Rouen university)

Abstract
Theoretical Computer Science needed to develop models in order to understand the behaviour of computers. Doing this, one had to invent new structures allowing to solve some problems of other sciences.

Decomposition of $\mathsf{Z}/2\mathsf{Z}$ modules of boolean functions

R. Lepelletier
(LIFAR, Rouen university)

Abstract Let f be a boolean function of n variables. Our concern here is to study the decompositions of the module generated by f under the action of the $\mathsf{Z}/2\mathsf{Z}$-algebra of the symmetric group, the action of the group being given by permutations of variables. To this end we use the natural structure of $\mathsf{Z}/2\mathsf{Z}$-automaton whose transition letters are the simple transpositions.
The proposed algorithm allows to decompose the module generated by this function into indecomposable factors by means of the resolution of the linear system given by the rules of the returning states. This algorithm, hardly modified, allows also to compute the orbit of the function by the symmetric group.

Rational graphs

C. Morvan
(IRISA, Rennes 1 university)

Abstract In this talk we defined a family of infinite graphs: the *Rational graphs*. They are the rational subsets of the *partial semigroup* $X^* \times A \times X^*$ (where X is a finite alphabet and A a finite set of labels). This construction is analogous to the construction of rational languages in a free monoid. This family of graphs is a proper extension of the push-down graphs of Muller and Schupp, the equational graphs of Courcelle and the prefix-recognizable graphs of Caucal.

Along the talk, we stated basic properties, as well as an internal and an external (up to isomorphism) characterizations of these graphs. We also considered a subfamily of these graphs, with X an alphabet containing only one letter. This subfamily has a decidable first order theory.

Finally we introduced the traces of these graphs (*i.e.*, the set of labels of path in a rational graph leading from a vertex to another vertex) and we showed that they form an *Abstract Family of Languages* contained into the context-sensitives languages strictly containing linear bounded context-sensitives languages.

Infinite automatic graphs and graph DOL-systems

Olivier LY
(LaBRI, Bordeaux 1 university)

Abstract We deal with the concept of end, which is a classical mean to understand the behavior of a graph at infinity. In this respect, we show that the problem of deciding whether an infinite automatic graph has more than one end is recursively undecidable. The proof involves the analysis of some global topological properties of the configuration graph of a self-stabilizing Turing machine. Note that this question has been recently proved to be decidable for the sub-class consisting of the automatic graphs which are the Cayley graphs of word-hyperbolic groups. This result is applied to show the undecidability of connectivity of all the finite graphs produced by iterating a graph $D0L$-system. This comes from the fact that any $D0L$-sequence of finite graphs is encoded by an automatic structure in the sense that it is the sequence of the spheres of an automatic graph; and converse is true for a large class of automatic graphs. We also prove that the graph $D0L$-systems with which we deal can emulate hyperedge replacement systems for which the above connectivity problem is decidable; this result thus contributes in clearing up the decidability boundary of the connectivity question considered here.

Puissance des langages rationnels

H. Calbrix
(Lyon 1 university)

Abstract Pour tout langage L sur un alphabet fini, on defini $puis(L)$ 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 ou l'alphabet comporte plusieurs lettres.

Langages et Statistiques des Mots dans le Génome

M. Régnier
(InRIA, Roquencourt)

Abstract

From Prouhet to Arshon, through Thue and Morse

P. Seebold
(LaRia, Amiens university)

Abstract In 1851, in order to solve a problem of number theory (a particular case of the Tarri-Escott problem), Prouhet gave a rule to divide the first n^p positive integers in subsets of n^{p-1} elements.

The Prouhet algorithm constructs an n-letter overlap-free infinite word. This word is generated by a morphism, say π_n, which generalizes the binary morphism introduced, independently, by Thue in 1912 and by Morse in 1921.

In 1937, Arshon described an algorithm of constructing an n-letter square-free infinite word, say a_n. We prove that the construction of Arshon can be realized using the Prouhet morphism. When n is odd, Berstel has proved that a_n is generated by a tag-system, but not by a morphism. When n is even ($n \geq 4$) we show that a_n is generated by a morphism. On the n-letter alphabet $\{0, \ldots, n-1\}$, this morphism is π_n for even letters and $\widetilde{\pi_n}$ for odd letters.

Grammatical Inference

C. De La Higuera
(EURISE, St Etienne university)
personal pages: `http://www.univ-st-etienne.fr/eurise/cdlh.html`
slides: `http://www.univ-st-etienne.fr/eurise/cdlh/lanforcdlh.ps`

Abstract 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.

Finite automata with residual states

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

Abstract We introduce a subclass of non deterministic finite automata (NFA) that we call Residual Finite State Automata (RFSA): a RFSA is a NFA all the states of which define residual languages of the language it recognizes. We prove that for every regular language L, there exists a unique RFSA that recognizes L and which has both a minimal number of states and a maximal number of transitions. Moreover, this canonical RFSA may be exponentially smaller than the equivalent minimal DFA but it also may have the same number of states as the equivalent minimal DFA, even if minimal equivalent NFA are exponentially smaller. We provide an algorithm that computes the canonical RFSA equivalent to a given NFA. We study the complexity of several decision and construction problems linked to the class of RFSA: most of them are PSPACE-complete.