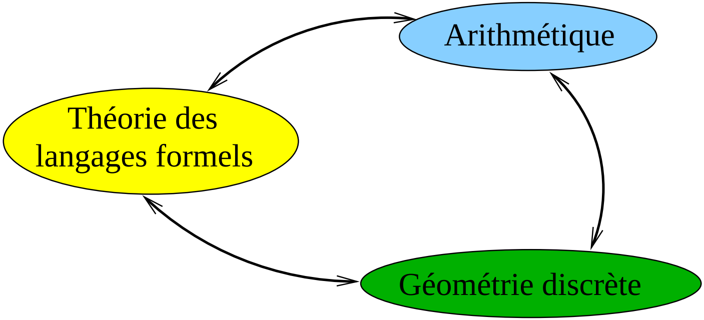
ARITHMÉTIQUE, AUTOMATES ET GÉOMÉTRIE DISCRÈTE

Jean-Paul Allouche, Valérie Berthé, Michel Rigo

Collège Belgique
mercredi 16 février 2011



Combinatoire des mots, analyse d'algorithmes, théorie ergodique, théorie des nombres, dynamique symbolique, vérification, logique, cryptographie, jeux combinatoires, structures automatiques, arithmétique des ordinateurs, ...

Un peu de théorie des langages formels

Définition

En informatique théorique, définition mathématique des notions de mot et de langage, étude de leurs propriétés syntaxiques et des opérations correspondantes.

Un alphabet : $A = \{a, b, c\}$, $B = \{0, 1\}$, ...

A^* ensemble des mots finis sur l'alphabet A,
$A^* = \{\varepsilon, a, b, c, aa, ab, ac, ba, bb, bc, ca, cb, cc, aaa, \ldots\}$

Hiérarchie de Chomsky (1956)

- langages récursivement énumérables,
  acceptés par machines de Turing
- langages "contextuels",
  acceptés machines de Turing bornées linéairement
- langages algérbiques,
  acceptés par automates à pile
- langages réguliers,
  acceptés par automates finis

$$\{m \in \{a,b\}^* \mid m \text{ contient un nombre premier de } a\}$$

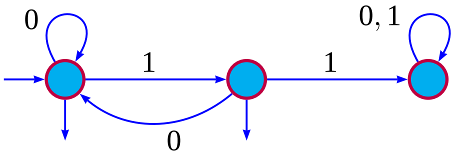
$$= \{aa, aba, bbaa, ababa, aabbba, \ldots\}$$

$$\{a^n b^n \mid n \geq 1\} = \{ab, aabb, aaabbb, \ldots\}$$

$$(aba)^* bb \cup a(bb)^* a = \{bb, aa, ababb, abba, abaababb, abbbba, \ldots\}$$

Les langages les plus “simples”

Les mots sur $\{0,1\}$ sans facteur 11



$$0^*\{10,0\}^*\{\varepsilon,1\}$$

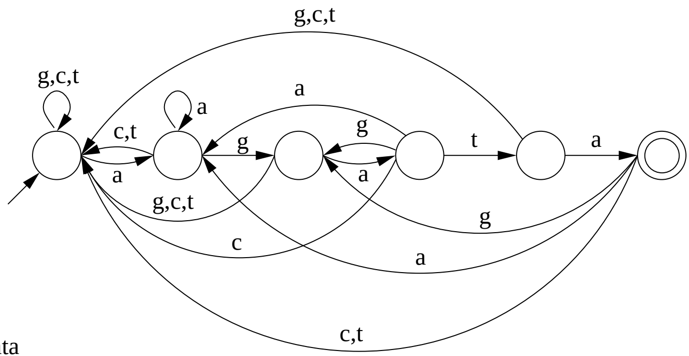
Théorème de Kleene (1956)

Un langage est *engendré* par une expression régulière si et seulement si il est *accepté* par un automate fini.

L’ensemble des langages réguliers sur A est la plus petite famille de langages contenant les langages finis et stable pour l’union, la concaténation et l’étoile de Kleene.

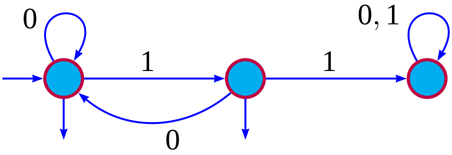
Une application

Recherche d'un mot dans un texte



agata

UN LIEN AVEC LA THÉORIE DES NOMBRES



16	8	4	2	1	
				1	1
			1	0	2
		1	0	0	4
		1	0	1	5
	1	0	0	0	8
	1	0	0	1	9
	1	0	1	0	10
1	0	0	0	0	16

Un lien avec la théorie des nombres

Soit $k \geq 2$ une base. Un ensemble $X \subseteq \mathbb{N}$ d'entiers est "*simple*", si l'ensemble des représentations en base k des éléments de X est un langage régulier.

Définition

Ensemble *k-reconnaissable* d'entiers.

- Dépendant de la base choisie ?
- Caractériser les ensembles k-reconnaissables ?
- Définir d'autres systèmes de numération ?

Un lien avec la théorie des nombres

Alan Cobham

http://recursed.blogspot.com/2010/04/alan-cobham.html

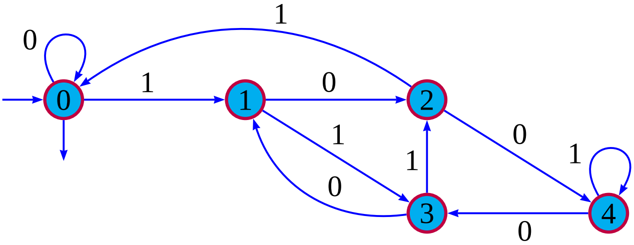
Un lien avec la théorie des nombres

Théorème de Cobham (1969)

Soient $k, \ell \geq 2$ multiplicativement indépendants.
Si $X \subseteq \mathbb{N}$ est k-reconnaissable et ℓ-reconnaissable, alors X est une union finie de progressions arithmétiques.

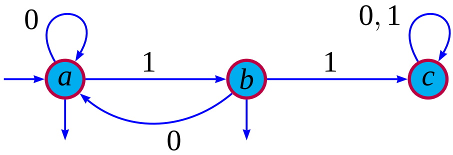
Exercice

Soit $k \geq 2$. Si $X \subseteq \mathbb{N}$ est une union finie de progressions arithmétiques, alors X est k-reconnaissable.



http://orbi.ulg.ac.be/handle/2268/39461

Suites automatiques



$$0, 1, 2, 4, 5, 8, 9, 10, 16, \ldots$$

$$f : a \mapsto ab, b \mapsto ac, c \mapsto cc, \qquad g : a, b \mapsto 1, c \mapsto 0$$

$$f(a) = ab,\ f^2(a) = f(a)f(b) = abac,\ f^3(a) = abacabcc, \ldots$$

$$g(f^\omega(a)) = 11101100111000001 \cdots$$

Proposition

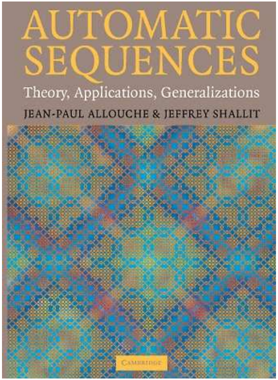
Un ensemble $X \subset \mathbb{N}$ est 2-reconnaissable si et seulement si sa suite caractéristique est l'image par codage d'un point fixe d'un morphisme 2-uniforme.

SUITES AUTOMATIQUES

THÉORÈME DE COBHAM (1972) *uniform tag systems*

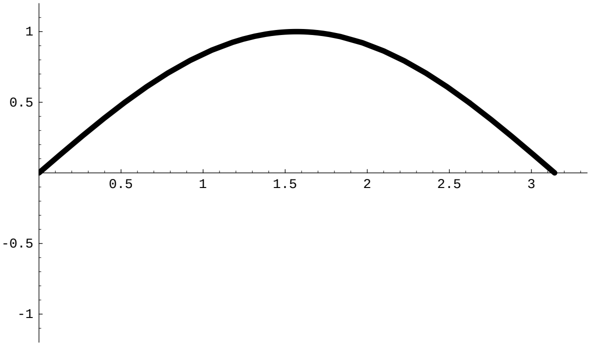
Une suite $(x_n)_{n\geq 0} \in A^{\mathbb{N}}$ est de la forme $g(f^{\omega}(a))$ où
g est un codage et f est un morphisme k-uniforme
si et seulement si elle est k-automatique.

Etude d'une large famille de mots infinis sur un alphabet fini...



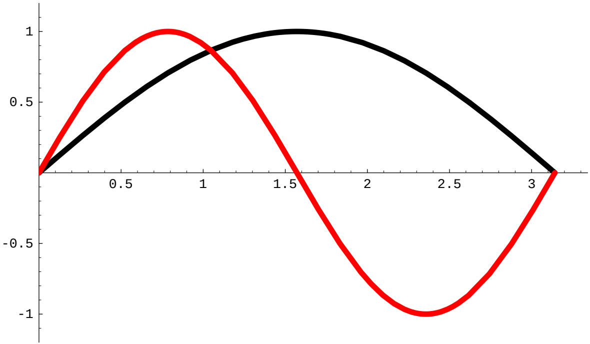
Une application

Etude du signe de $\sin(x)\sin(2x)\sin(4x)\cdots\sin(2^n x)$ sur $]0,\pi[$?



UNE APPLICATION

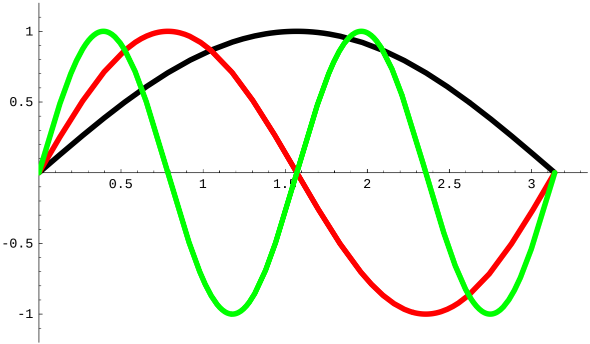
Etude du signe de $\sin(x)\sin(2x)\sin(4x)\cdots\sin(2^n x)$ sur $]0,\pi[$?



$+-$

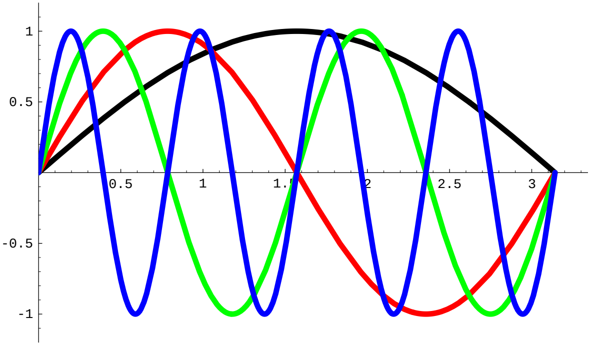
UNE APPLICATION

Etude du signe de $\sin(x)\sin(2x)\sin(4x)\cdots\sin(2^n x)$ sur $]0,\pi[$?



$+ - - +$

Etude du signe de $\sin(x)\sin(2x)\sin(4x)\cdots\sin(2^n x)$ sur $]0,\pi[$?



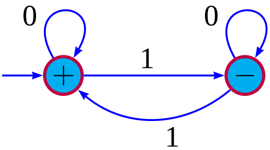
$+ - - + - + + -$

Etude du signe de $\sin(x)\sin(2x)\sin(4x)\cdots\sin(2^n x)$ sur $]0,\pi[$?

On obtient un préfixe du mot (de Prouhet–Thue–Morse) engendré par

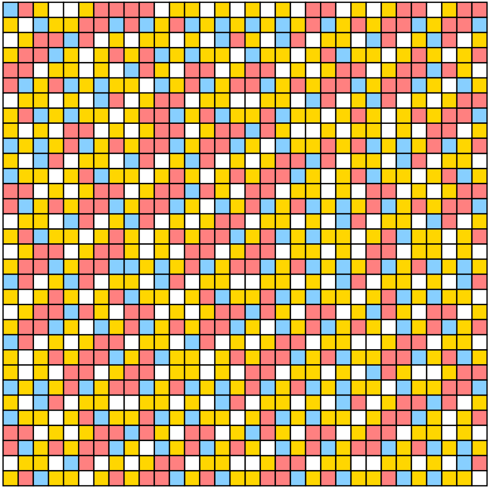
$$+ \mapsto +-$$

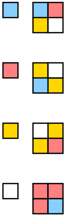
$$- \mapsto -+$$



J.-P. Allouche, J. Shallit, *The ubiquitous Prouhet-Thue-Morse sequence*. Sequences and their applications (Singapore, 1998), 1-–16, Springer Ser. Discrete Math. Theor. Comput. Sci., Springer, London, 1999.

En dimension supérieure





Une caractérisation des suites k-automatiques

Définition (S. Eilenberg – O. Salon)

Soient $k \geq 2$ et $x = (x_n)_{n\geq 0}$ une suite sur A. Le k-noyau de x est

$$N_k(x) = \{(x_{k^i n+r})_{n\geq 0} \mid i \geq 0,\ 0 \leq r < k^i\}$$

UNE CARACTÉRISATION DES SUITES k-AUTOMATIQUES

$0110100110010110100101100110100110010110\cdots$

$0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ \cdots$

$\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ 1\ 0\ 1\ 1\ 0\cdots$

$0\quad 1\quad 1\quad 0\quad 1\quad 0\quad 0\quad 1\quad 1\quad 0\quad \cdots$

$\ 1\quad 0\quad 0\quad 1\quad 0\quad 1\quad 1\quad 0\quad 0\quad 1\quad \cdots$

$\ \ 1\quad 0\quad 0\quad 1\quad 0\quad 1\quad 1\quad 0\quad 0\quad 1\ \cdots$

$\ \ \ 0\quad 1\quad 1\quad 0\quad 1\quad 0\quad 0\quad 1\quad 1\quad 0\cdots$

Le 2-noyau de la suite de Thue–Morse contient 2 éléments.

THÉORÈME

Soit $k \geq 2$. La suite $x = (x_n)_{n \geq 0}$ est k-automatique si et seulement si $N_k(x)$ est fini.