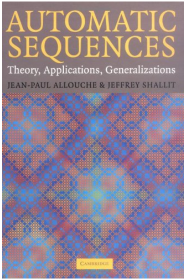
Combinatorics on words: Properties of automatic words

Anna FRID

Aix-Marseille Université, September 2020

The main source

Jean-Paul Allouche, Jeffrey Shallit, *Automatic Sequences — Theory, Applications, Generalizations*. Cambridge Univ. Press, 2003.



Formal definition

$$\Sigma_k = \{0, \ldots, k-1\}.$$

Definition

An infinite word $\mathbf{u} = u[0]u[1] \cdots u[n] \cdots$ over an alphabet Δ is *k-automatic* for a $k > 1$, if there exists a DFAO $A = (Q, \Sigma_k, \delta, q_0, \Delta, \lambda)$ such that $u[n] = \lambda(\delta(q_0, a))$ for all $n \geq 0$ and for all k-ary representations $a \in \Sigma_k^*$ of n.

Formal definition

$$\Sigma_k = \{0, \ldots, k-1\}.$$

Definition

An infinite word $\mathbf{u} = u[0]u[1] \cdots u[n] \cdots$ over an alphabet Δ is *k-automatic* for a $k > 1$, if there exists a DFAO $A = (Q, \Sigma_k, \delta, q_0, \Delta, \lambda)$ such that $u[n] = \lambda(\delta(q_0, a))$ for all $n \geq 0$ and for all k-ary representations $a \in \Sigma_k^*$ of n.

So, the input of A is a k-ary representation of n and the output is $u[n]$.

Cobham theorem

Theorem (Cobham,1972)

An infinite word $\mathbf{w}$ *is* k*-automatic if and only if* $\mathbf{w} = \psi(\varphi^\omega(a))$ *for some* a*, where* φ *is a* k*-uniform morphism and* ψ *is a coding.*

Versions of the automatic definitions

- There is no problem with leading zeros.

Versions of the automatic definitions

- There is no problem with leading zeros.
- The definition does not depend on if we read the k-ary representation from left to right or from right to left.

Versions of the automatic definitions

- There is no problem with leading zeros.
- The definition does not depend on if we read the k-ary representation from left to right or from right to left.
- A word **u** over an alphabel Δ is k-automatic iff for all $a \in \Delta$, the language $F_a = \{(n)_k | u[n] = a\}$ is regular.

Versions of the automatic definitions

- There is no problem with leading zeros.
- The definition does not depend on if we read the k-ary representation from left to right or from right to left.
- A word $\mathbf{u}$ over an alphabel Δ is k-automatic iff for all $a \in \Delta$, the language $F_a = \{(n)_k | u[n] = a\}$ is regular.
- An infinite word obtained from a k-automatic word by changing a finite number of symbols is k-automatic.

Periodicity

An infinite word **w** is *ultimately t-periodic* if $\mathbf{w} = uv^{\omega}$ with $|v| = t$.

Periodicity

An infinite word $\mathbf{w}$ is *ultimately t-periodic* if $\mathbf{w} = uv^\omega$ with $|v| = t$.

Example: $abbccc\ abac\ abac\ abac \cdots = abbccc(abac)^\omega$

Periodicity

An infinite word $\mathbf{w}$ is *ultimately t-periodic* if $\mathbf{w} = uv^\omega$ with $|v| = t$.

Example: $abbccc\ abac\ abac\ abac \cdots = abbccc(abac)^\omega$

Lemma

An ultimately t-periodic word is k-automatic for every k.

Periodicity

An infinite word $\mathbf{w}$ is *ultimately t-periodic* if $\mathbf{w} = uv^\omega$ with $|v| = t$.

Example: $abbccc\ abac\ abac\ abac \cdots = abbccc(abac)^\omega$

Lemma

An ultimately t-periodic word is k-automatic for every k.

PROOF. It is sufficient to consider a pure period v (that is, $u = \varepsilon$).

$$Q = \{0, 1, \ldots, t-1\};$$

$$\delta(q, b) = (kq + b) \pmod{t};$$

$$\tau(q) = w[q].$$

Another Cobham theorem

Integers k and l are called *multiplicatively dependent* if $k^r = l^s$ for some integer r, s.

Theorem (Cobham, 1969)

Let k and l be multiplicatively independent integers. Then an infinite word which is k- and l-automatic is ultimately periodic.

k-kernel

The *k-kernel* of an infinite word $\mathbf{w} = w[0] \cdots w[n] \cdots$ is the set of infinite words of the form

$$\mathbf{w}_m^i = w[i]w[i + k^m] \cdots w[i + nk^m] \cdots ,$$

where $i = 0, \ldots k^n - 1$.

k-kernel

The *k-kernel* of an infinite word $\mathbf{w} = w[0] \cdots w[n] \cdots$ is the set of infinite words of the form

$$\mathbf{w}_m^i = w[i]w[i+k^m] \cdots w[i+nk^m] \cdots ,$$

where $i = 0, \ldots k^n - 1$.

Theorem

A word is k-automatic if and only if its k-kernel is finite.

Frequencies of letters

Frequency of every factor of a k-automatic word, if it exists, is rational.

Frequencies of letters

Frequency of every factor of a k-automatic word, if it exists, is rational.

This is one of tools to prove that an infinite word is *not* automatic.

Is a word automatic?

How can we prove that a word is k-automatic? Build an automaton (or morphisms).

Is a word automatic?

How can we prove that a word is k-automatic? Build an automaton (or morphisms).

How can we prove that a word is **NOT** k-automatic?

Is a word automatic?

How can we prove that a word is k-automatic? Build an automaton (or morphisms).

How can we prove that a word is **NOT** k-automatic?

- Reduce to (non)-regular languages (pumping lemma etc.);

Is a word automatic?

How can we prove that a word is k-automatic? Build an automaton (or morphisms).

How can we prove that a word is **NOT** k-automatic?

- Reduce to (non)-regular languages (pumping lemma etc.);
- Find a non-periodic sequence $\{w_{[uv^i]_k}\}_{i=1}^{\infty}$;

Is a word automatic?

How can we prove that a word is k-automatic? Build an automaton (or morphisms).

How can we prove that a word is **NOT** k-automatic?

- Reduce to (non)-regular languages (pumping lemma etc.);
- Find a non-periodic sequence $\{w_{[uv^i]_k}\}_{i=1}^{\infty}$;
- Find a symbol of irrational frequency;

Is a word automatic?

How can we prove that a word is k-automatic? Build an automaton (or morphisms).

How can we prove that a word is **NOT** k-automatic?

- Reduce to (non)-regular languages (pumping lemma etc.);
- Find a non-periodic sequence $\{w_{[uv^i]_k}\}_{i=1}^{\infty}$;
- Find a symbol of irrational frequency;
- Sometimes it is not clear!

Decidability and enumeration

Theorem (Charlier, Rampersad, Shallit, 2011)

If we can express a property of a k-automatic sequence $\mathbf{w}$ using quantifiers, logical operations, integer variables, the operations of addition, subtraction, indexing into $\mathbf{w}$, and comparison of integers or elements of $\mathbf{w}$, then this property is decidable.

Decidability and enumeration

Theorem (Charlier, Rampersad, Shallit, 2011)

If we can express a property of a k-automatic sequence $\mathbf{w}$ using quantifiers, logical operations, integer variables, the operations of addition, subtraction, indexing into $\mathbf{w}$, and comparison of integers or elements of $\mathbf{w}$, then this property is decidable.

Another meta-theorem from the same paper

Many sequences related to k-automatic words are *k-regular*.

Decidability and enumeration

Theorem (Charlier, Rampersad, Shallit, 2011)

If we can express a property of a k-automatic sequence $\mathbf{w}$ using quantifiers, logical operations, integer variables, the operations of addition, subtraction, indexing into $\mathbf{w}$, and comparison of integers or elements of $\mathbf{w}$, then this property is decidable.

Another meta-theorem from the same paper

Many sequences related to k-automatic words are *k-regular*.

Examples: number of factors of length n; number of palindromes of a given length; indices where these palindromes start, etc.

Walnut software

The `Walnut` software, written by H. Mousavi and updated by A. Baranwal (students of Jeffrey O. Shallit), solves problems and answers questions, posed in first-order logic, about automatic and related sequences.

Walnut software

The `Walnut` software, written by H. Mousavi and updated by A. Baranwal (students of Jeffrey O. Shallit), solves problems and answers questions, posed in first-order logic, about automatic and related sequences.

Available for free, used in many papers.

Walnut software

The Walnut software, written by H. Mousavi and updated by A. Baranwal (students of Jeffrey O. Shallit), solves problems and answers questions, posed in first-order logic, about automatic and related sequences.

Available for free, used in many papers.

So, the most interesting questions about k-automatic words are now those which cannot be solved by it!