Critical constants for unique expansions in general alphabets

Marco Pedicini

Istituto per le Applicazioni del Calcolo "Mauro Picone"
Consiglio Nazionale delle Ricerche

Journées Numération, Prague 26 - 30 May 2008
Doppler Institute, Czech Technical University

Critical constants for unique expansions in general alphabets

Marco Pedicini

Istituto per le Applicazioni del Calcolo "Mauro Picone"
Consiglio Nazionale delle Ricerche

joint work with

Vilmos Komornik (Universitè de Strasbourg)

Anna Chiara Lai (Università di Roma "La Sapienza")

Journées Numération, Prague 26 - 30 May 2008
Doppler Institute, Czech Technical University

Expansions in alphabets with deleted digits

We consider generic alphabets

$$A = \{a_1, \ldots, a_J\}$$

of real numbers $a_1 < \cdots < a_J$.

Given a real number $q > 1$, by an expansion of a real number x we mean a sequence (c_i) of numbers $c_i \in A$ satisfying the equality

$$\pi_q(c) := \sum_{i=1}^{\infty} \frac{c_i}{q^i} = x.$$

In order to have an expansion, x must belong to the interval $[\frac{a_1}{q-1}, \frac{a_J}{q-1}]$

Interval Condition

This result was proved in [Pedicini2005]:

Theorem

Every $x \in [\frac{a_1}{q-1}, \frac{a_J}{q-1}]$ *has at least one expansion in base* q *if and only if*

$$1 < q \leq Q_A := 1 + \frac{a_J - a_1}{\max_{j>1}\{a_j - a_{j-1}\}} (\leq J). \quad (1)$$

Univoque sequences

Definition

A sequence $(c_i) \in A^\infty$ is called *univoque* in base q if

$$x := \sum_{i=1}^{\infty} \frac{c_i}{q^i}$$

has no other expansion in this base.

Example

The constant sequences $(a_1)^\infty$ and $(a_J)^\infty$ are univoque in every base q: they are called the *trivial unique expansions*.

Characterization theorem

Theorem

Assume the interval condition (1)*. An expansion* (c_i) *is unique in base* q *if and only if the following conditions are satisfied:*

$$\sum_{i=1}^{\infty} \frac{c_{n+i} - a_1}{q^i} < a_{j+1} - a_j \quad \textit{whenever} \quad c_n = a_j < a_J;$$

$$\sum_{i=1}^{\infty} \frac{a_J - c_{n+i}}{q^i} < a_j - a_{j-1} \quad \textit{whenever} \quad c_n = a_j > a_1.$$

Consequences of the characterization

Corollary

For every given set $C \subset A^\infty$ *there exists a number*

$$1 \le q_C \le Q_A$$

such that

$$q > q_C \implies \text{ every sequence } c \in C \text{ is univoque in base } q;$$
$$1 < q < q_C \implies \text{ not every sequence } c \in C \text{ is univoque in base } q.$$

Critical constants

Definition

The number q_C is called the *critical base* of C. If $C = \{c\}$ is a one-point set, then $q_c := q_C$ is also called the critical base of the sequence c.

If C is a nonempty finite set of eventually periodic sequences, then the supremum $\sup q_\alpha$ in the above proof is actually a maximum.

In the last case, it is possible that not all sequences $c \in C$ are univoque in base $q = q_C$.

Generalized Golden number

It is well-known that for the alphabet $A = \{0, 1\}$ there exist nontrivial univoque sequences in base q if and only if $q > \frac{1+\sqrt{5}}{2}$.

Generalized Golden number

It is well-known that for the alphabet $A = \{0, 1\}$ there exist nontrivial univoque sequences in base q if and only if $q > \frac{1+\sqrt{5}}{2}$. There exists a "generalized Golden number" for every alphabet:

Corollary

There exists a number $1 < G_A \leq Q_A$ such that

$$q > G_A \implies \text{there exist nontrivial univoque sequences;}$$
$$1 < q < G_A \implies \text{there are no nontrivial univoque sequences.}$$

Generalized Golden number

It is well-known that for the alphabet $A = \{0, 1\}$ there exist nontrivial univoque sequences in base q if and only if $q > \frac{1+\sqrt{5}}{2}$. There exists a "generalized Golden number" for every alphabet:

Corollary

There exists a number $1 < G_A \leq Q_A$ *such that*

$$q > G_A \implies \text{there exist nontrivial univoque sequences;}$$
$$1 < q < G_A \implies \text{there are no nontrivial univoque sequences.}$$

Definition

The number G_A is called the *critical base* of the alphabet A.

Invariance properties

Proposition

The critical base does not change if we replace the alphabet A

- *by* $A + b = \{a_j + b \mid j = 1, \dots, m\}$ *for some real number* b*;*
- *by* $bA = \{ba_j \mid j = 1, \dots, m\}$ *for some nonzero real number* b*;*
- *by the dual alphabet defined by* $D(A) = \{a_m + a_1 - a_j \mid j = 1, \dots, m\}$.

Invariance properties

Proposition

The critical base does not change if we replace the alphabet A

- *by* $A + b = \{a_j + b \mid j = 1, \ldots, m\}$ *for some real number* b*;*
- *by* $bA = \{ba_j \mid j = 1, \ldots, m\}$ *for some nonzero real number* b*;*
- *by the dual alphabet defined by* $D(A) = \{a_m + a_1 - a_j \mid j = 1, \ldots, m\}$*.*

We wish to establish the critical bases for ternary alphabets

$$A = \{a_1, a_2, a_3\}.$$

Restriction to ternary alphabets

By the Proposition above we may restrict ourselves without loss of generality to the case of alphabets

$$A_m = \{0, 1, m\}$$

with $m \geq 2$.
Interval condition (1) in the ternary case takes the form

$$1 < q \leq \frac{2m-1}{m-1}.$$

Lemma

An expansion (c_i) *is unique in base* q *for the alphabet* A_m *if and only if the following conditions are satisfied:*

$$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} < 1 \quad \text{whenever } c_n = 0; \tag{2}$$

$$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} < m-1 \quad \text{whenever } c_n = 1; \tag{3}$$

$$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} > \frac{m}{q-1} - 1 \quad \text{whenever } c_n = 1; \tag{4}$$

$$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} > \frac{m}{q-1} - (m-1) \quad \text{whenever } c_n = m. \tag{5}$$

Alphabet $A_3 = \{0, 1, 3\}$

Consider the periodic sequence $(c_i) = (31)^\infty$. By the periodicity of (c_i) we have for each n either $c_n = 3$ and $(c_{n+i}) = (13)^\infty$ or $c_n = 1$ and $(c_{n+i}) = (31)^\infty$.
In this case Theorem 4 contains only three conditions on q.
For $c_n = 3$ we have the condition

$$\sum_{i=1}^{\infty} \frac{3 - c_{n+i}}{q^i} < 2 \Longleftrightarrow \frac{2q}{q^2 - 1} < 2,$$

while for $c_n = 1$ we have the following two conditions:

$$\sum_{i=1}^{\infty} \frac{3 - c_{n+i}}{q^i} < 1 \iff \frac{2}{q^2 - 1} < 1$$

and

$$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q^i} < 2 \iff \frac{3}{q - 1} - \frac{2}{q^2 - 1} < 2.$$

They are equivalent approximatively to the inequalities $q > 1.61803$, $q > 1.73205$ and $q > 2.18614$ respectively, so that $q_c \approx 2.18614$.

Main result

In order to formulate our main result we introduce the quantities $P_m = P(m)$, $p_1 = p_1(k, h, m)$, $p_2 = p_2(k, h, m)$, $p = p(k, h, m)$ for every real number $m > 0$ and nonnegative integers k and h by the following formulae:

$$P(m) = 1 + \sqrt{\frac{m}{m-1}};$$

$$\pi_{p_1}\left(\left(m^k 1 (m^{k-1} 1)^h\right)^\infty\right) = m - 1;$$

$$\pi_{p_2}\left(\left((m^{k-1} 1)^h m^k 1\right)^\infty\right) = \frac{m}{p_2 - 1} - 1;$$

$$p = \max\{p_1, p_2\}.$$

Theorem

If $p(k, h, m) \le P(m)$ for some k and h, then $G_{A_m} = p(k, h, m)$.

Blocks

The blocks $S_{j,m} := m^{j-1}1$ $(j = 1, 2, \ldots)$ have the following property:

Lemma

If $2^k \le m \le 2^{k+1}$ and $1 < q \le P_m$ for some $m \ge 2$, then any nontrivial univoque sequence in base q has either the form

$$0^t 1 S_{\varepsilon_1,m} S_{\varepsilon_2,m} \ldots$$

or the form

$$m^t 1 S_{\varepsilon_1,m} S_{\varepsilon_2,m} \ldots$$

with some nonnegative integer t and some sequence (ε_i) of elements $\varepsilon_i \in \{k, k+1\}$.
Moreover, if $m \ge M \simeq 2.80194$, the largest root of the polynomial $m^3 - 4m^2 + 3m + 1$, then it cannot begin with a 0 digit.

Sketch of proof

Let (c_i) be a univoque sequence in some base $1 < q \leq P_m$. The lemma will follow from the following six properties:

- (c_i) does not contain any block of the form $m0$;
- (c_i) does not contain any block of the form 10;
- (c_i) does not contain any block of the form $0m$;
- if $m > M$, then (c_i) does not contain any block of the form 01;
- each 1 digit is followed by at least $k - 1$ consecutive m digits;
- each 1 digit is followed by at most k consecutive m digits.

Quasi-greedy and quasi-lazy expansions

The quasi-greedy expansion of a real number x in some base q is its lexicographically largest infinite expansion, while the quasi-lazy expansion of x is the conjugate of the quasi-greedy expansion of $\frac{m}{q-1} - x$ with respect to the conjugate alphabet $\{0, m-1, m\}$.

The following characterization of special quasi-greedy and quasi-lazy expansions follows from more general results in [Pedicini2005].

Lemma

Consider a sequence (c_i) with digits in $\{1, m\} \subsetneq A_m$.

Let $q > 1$ and set $x := \sum_{i=1}^{\infty} \frac{c_i}{q^i}$.

If $(c_{i+n}) \leq (c_i)$ whenever $c_n = 1$, then (c_i) is the quasi-greedy expansion of x in base q.

If $(c_{i+n}) \geq (c_i)$ whenever $c_n = 1$, then (c_i) is the quasi-lazy expansion of x in base q.

Our following result explains the introduction of $p = p(k, h, m)$ in Theorem 11.

Lemma

The critical base of the sequence $(S^h_{k,m} S_{k+1,m})^\infty$ *is equal to* $p(k, h, m)$.

Lemma

Assume that $p(k, h, m) \le P_m$. *If* $p_1 \ge p_2$, *then*

$$S_k(S^h_k S_{k+1})^\infty \le (\gamma_i) \quad \textit{and} \quad (\delta_i) = S_{k+1}(S^h_k S_{k+1})^\infty; \tag{6}$$

If $p_2 \ge p_1$, *then*

$$(S^h_k S_{k+1})^\infty = (\gamma_i) \quad \textit{and} \quad (\delta_i) \le m(S^h_k S_{k+1})^\infty. \tag{7}$$

In view of the preceding lemma this implies that

$$S_k(S_k^h S_{k+1})^\infty < (c_{n+i}) < S_{k+1}(S_k^h S_{k+1})^\infty \quad \text{whenever} \quad c_n = 1 \tag{8}$$

if $\pi_p((S_{k+1} S_k^h)^\infty) = m - 1$, and

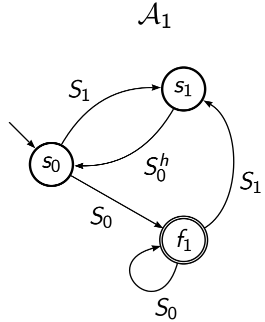
$$(S_k^h S_{k+1})^\infty < (c_{n+i}) < m(S_k^h S_{k+1})^\infty \quad \text{whenever} \quad c_n = 1 \tag{9}$$

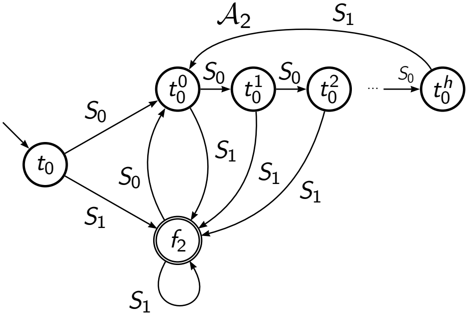
if $\pi_p((S_k^h S_{k+1})^\infty) = \frac{m}{p-1} - 1$.

For any fixed h we can define the following Büchi automata:

- Concerning the right gap $m - 1$:
  - $\mathcal{A}_1$ accepts (c_i) if it satisfies the left-hand side of condition (8);
  - $\mathcal{A}_2$ accepts (c_i) if it satisfies the right-hand side of condition (8);
- Concerning the left gap 1:
  - $\mathcal{A}'_1$ accepts (c_i) if it satisfies the left-hand side of condition (9);
  - $\mathcal{A}'_2$ accepts (c_i) if it satisfies the right-hand side of condition (9).

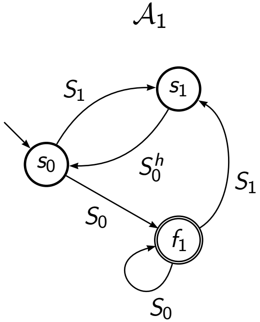
In these pictures we replaced S_{k+i} by S_i, so that S_0 denotes S_k and S_1 denotes S_{k+1}.



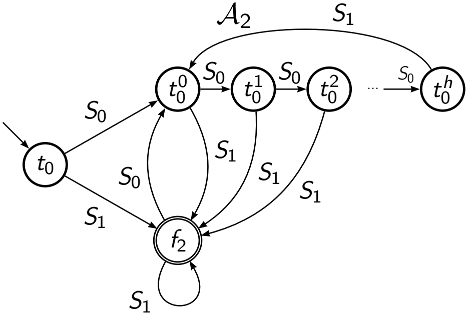


Automata and lexicographic condition (right gap)

In these pictures we replaced S_{k+i} by S_i, so that S_0 denotes S_k and S_1 denotes S_{k+1}.

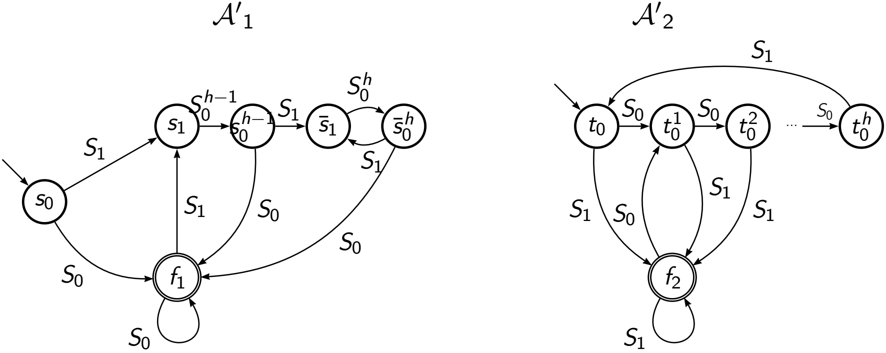


$S_0(S_0^h S_1)^\infty < (c_{n+i})$

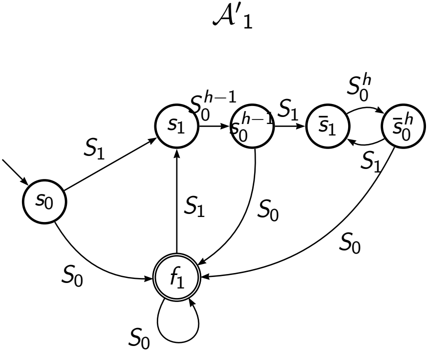


$(c_{n+i}) < S_1(S_0^h S_1)^\infty$

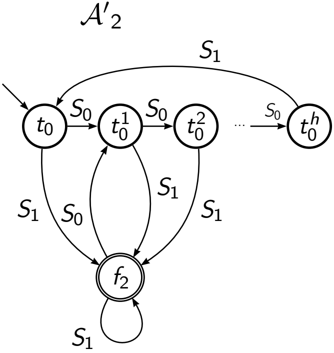
Automata and lexicographic condition (left gap)



Automata and lexicographic condition (left gap)



$(S_0^h S_1)^\infty < (c_{n+i})$



$(c_{n+i}) < m(S_0^h S_1)^\infty$

The proof of the theorem

Note that by Lemma 12 the choice of the alphabet $\{S_k, S_{k+1}\}$ for these automata does not imply a loss of generality.
We can reformulate the conditions (8) and (9) as follows. An expansion (c_i) is unique in base p only if, starting from the smallest n such that in case $\pi_p((S_{k+1}S_k^h)^\infty) = m - 1$

$$(c_{n+i}) \in L(\mathcal{A}_1) \cap L(\mathcal{A}_2) \quad \text{whenever} \quad c_n = 1,$$

while in case $\pi_p((S_k^h S_{k+1})^\infty) = \frac{m}{p-1} - 1$

$$(c_{n+i}) \in L(\mathcal{A}'_1) \cap L(\mathcal{A}'_2) \quad \text{whenever} \quad c_n = 1.$$

Intersection Automaton

The construction of the intersection automaton in the case of infinite words can be obtained A standard construction for the intersection automata in the case of infinite words can be applied:

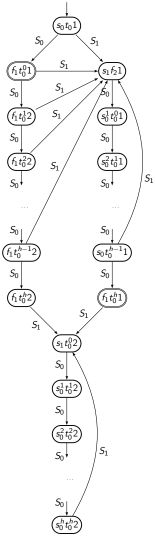
$$\mathcal{A} = (\bar{A}, S \times T \times \{1,2\}, (s_0, t_0, 1), \tau, f_1 \times T \times \{1\}).$$

Denoting by a a generic element of $\bar{A}$, the transition function τ is defined for every $s \in S$ and for every $t \in T$ by the following rules:

$$\begin{aligned}
&\tau((s,t,1),a) = (\tau_1(s,a), \tau(t,a), 1) \quad \text{if } s \neq f_1;\\
&\tau((s,t,2),a) = (\tau_1(s,a), \tau(t,a), 2) \quad \text{if } t \neq f_2;\\
&\tau((f_1,t,1),a) = (\tau_1(f_1,a), \tau(t,a), 2);\\
&\tau((s,f_2,2),a) = (\tau_1(s,a), \tau(f_2,a), 2).
\end{aligned} \tag{10}$$

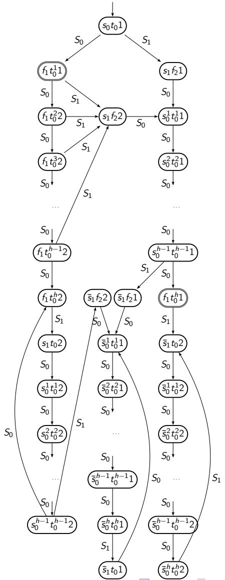
Intersection Automata (right gap)

We build the intersection automaton and we prove that no sequence is accepted, so that there are no other unique sequences in $L(\mathcal{A}_1) \cap L(\mathcal{A}_2)$.



Intersection Automaton (left gap)

The same construction can be done in order to build the intersection automaton for $\mathcal{A}'_1$ and $\mathcal{A}'_2$. Also in this case no sequence is accepted, so that there are no other unique sequences in $L(\mathcal{A}'_1) \cap L(\mathcal{A}'_2)$.



Last part of the work

What we have here is the fact that whenever $p(h,k,m) \leq P(m)$ then for any $q < p(h,k,m)$ we can establish the critical base for such an m.
But when and how it is possible to determine h for a given m ?
The existence proof for such h, k and m is proved in the second part of the work:

Theorem

The condition $p(k,h,m) \leq P(m)$ of Theorem 11 is satisfied for some $k, h \geq 0$ and $m \geq 2$ if and only if $m \in [m_{k,h}, M_{k,h}]$, where $m_{k,h}$ and $M_{k,h}$ are the unique solutions of the equations

$$\pi_{P(m)}\left((S_{k+1,m}S_{k,m}^h)^\infty\right) = m - 1$$

$$\pi_{P(m)}\left((S_{k,m}^h S_{k+1,m})^\infty\right) = \frac{m}{P(m)-1} - 1,$$

respectively.

Golden number

Let $A = \{0, 1\}$ and consider the sequence $(c_i) = 1(10)^\infty$. This sequence is nontrivial and eventually periodic with a nontrivial period, thus there exists a critical base q_c such that $q > q_c$ if and only if (c_i) is unique in base q. By applying the algorithm of the proof of Corollary 5 we obtain that the sequence $\{q_n\}$ is composed of the following elements:

- if $n = 1$ and $c_n = 1$, then $c_{n+i} = (10)^\infty$, so that the equation

  $$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q_1^i} = \frac{1}{q-1} - 1$$

  is equivalent to $q^2 - 2 = 0$ whence $q_1 = \sqrt{2}$.
- if $n = 2$ and $c_n = 1$, then $c_{n+i} = (10)^\infty$, so that the same equation is equivalent to $q^2 - q - 1 = 0$ whence $q_2 = \frac{1+\sqrt{5}}{2}$.
- if $n = 3$ and $c_n = 0$, then $c_{n+i} = (10)^\infty$, and the equation

  $$\sum_{i=1}^{\infty} \frac{c_{n+i}}{q_1^i} = 1$$

A_3

Fix $m = 3$ and consider the sequence $c := (S_1)^\infty = (31)^\infty$.
We recall from Example $A_3 = \{0, 1, 3\}$ that the critical base of this sequence is equal to $q_c \approx 2.18614$.
Since $q_c \leq P_m \approx 2.224744$, applying Theorem 11 we conclude that q_c is the critical base of the alphabet $\{0, 1, 3\}$.

Numerical computation of the intervals of Theorem 16

In the following table we indicate the first five intervals $I_{k,h} := [m_{k,h}, M_{k,h}]$ of Theorem 16,

k	h	$(S_{k+1,m}S^h_{k,m})$	$I_{k,h}$
0	0	$(1)^\infty$	$[1.61803, 2.32471]$
1	3	$(m1111)^\infty$	$[2.34687, 2.37782]$
1	2	$(m111)^\infty$	$[2.37897, 2.46001]$
1	1	$(m11)^\infty$	$[2.46788, 2.72274]$
1	0	$(m1)^\infty$	$[2.80194, 3.56811]$