

Normality and Automata

Olivier Carton

LIAFA

Université Paris Diderot & CNRS

Join work with Verónica Becher and Pablo Heiber
(Universidad de Buenos Aires & CONICET)

Work supported by LIA Infinis

GDR-IM SDA2

Outline

Normality

Compressibility

One-way transducers

Two-way transducers

Selection

Prefix selection

Suffix selection

Expansion of real numbers

Fix an integer base $b \geq 2$. The alphabet is $A = \{0, 1, \dots, b-1\}$.

- ▶ if $b = 2$, $A = \{0, 1\}$,
- ▶ if $b = 10$, $A = \{0, 1, 2, \dots, 9\}$.

Each real number $\xi \in [0, 1)$ has an **expansion** in base b :
 $x = a_1 a_2 a_3 \dots$ where $a_i \in A$ and

$$\xi = \sum_{k \geq 1} \frac{a_k}{b^k}.$$

In the rest of this talk:

real number $\xi \in [0, 1)$ \longleftrightarrow infinite word $x \in A^\omega$

$1/3$ \longleftrightarrow $010101 \dots = (01)^\omega$

$\pi/4$ \longleftrightarrow $1100100100001111 \dots$

Normality (Borel 1909)

The number of **occurrences** of a word u in a word w is

$$\text{occ}(w, u) = |\{i : w[i..i + |u| - 1] = u\}|$$

An infinite word $x \in A^\omega$ (resp. a real number ξ) is **simply normal** (in base b) if for any $a \in A$,

$$\lim_{n \rightarrow \infty} \frac{\text{occ}(x[1..n], a)}{n} = \frac{1}{b}.$$

An infinite word $x \in A^\omega$ (resp. a real number ξ) is **normal** (in base b) if for any $u \in A^*$,

$$\lim_{n \rightarrow \infty} \frac{\text{occ}(x[1..n], u)}{n} = \frac{1}{b^{|u|}}.$$

In base $b = 2$, this means

- ▶ the frequencies in x of the 2 digits 0 and 1 are $1/2$,
- ▶ the frequencies in x of the 4 words 00, 01, 10, 11 are $1/4$,
- ▶ the frequencies in x of the 8 words 000, 001, ..., 111 are $1/8$,
- ▶ ...

Examples

Theorem (Borel 1909)

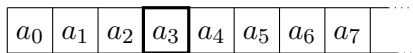
Almost all real numbers are normal, that is, the measure of the set of normal numbers in $[0, 1)$ is 1.

Examples

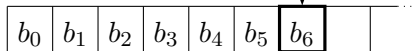
- ▶ the infinite word $(001)^\omega = 0010010 \dots$ is not simply normal in base 2,
- ▶ the infinite word $(01)^\omega = 01010 \dots$ is simply normal in base 2 but it is not normal,
- ▶ the Champernowne word $012345678910111213 \dots$ is normal in base 10.
- ▶ the Champernowne word $011011100101110111 \dots$ is normal in base 2.

Transducers

Input tape

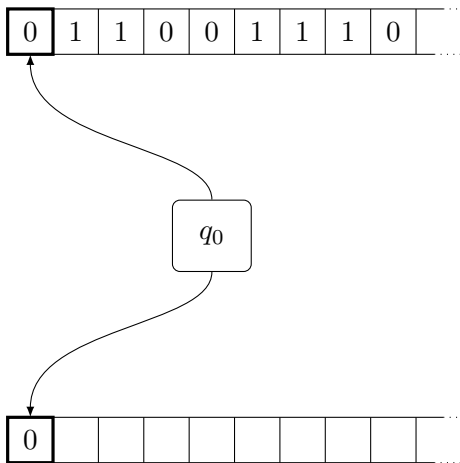
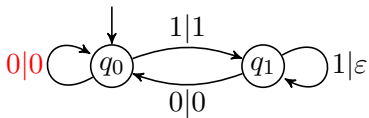


Output tape

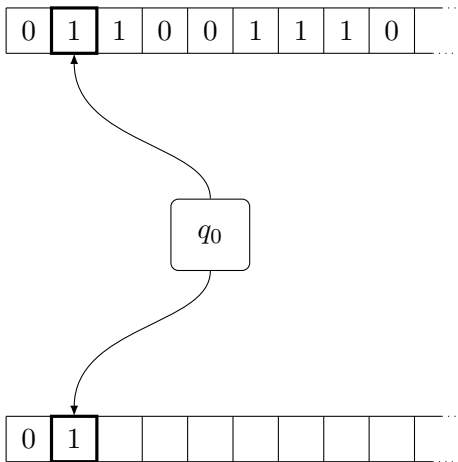
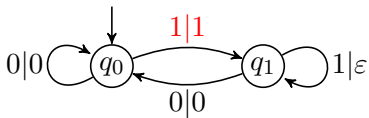


Transitions $p \xrightarrow{a|v} q$ for $a \in A, v \in B^*$.

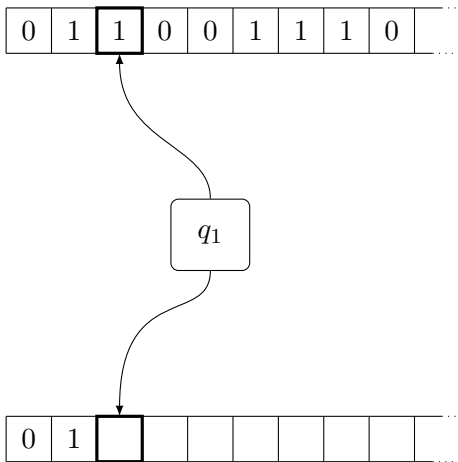
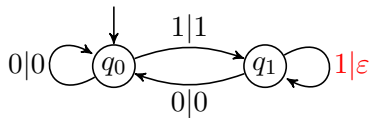
Example



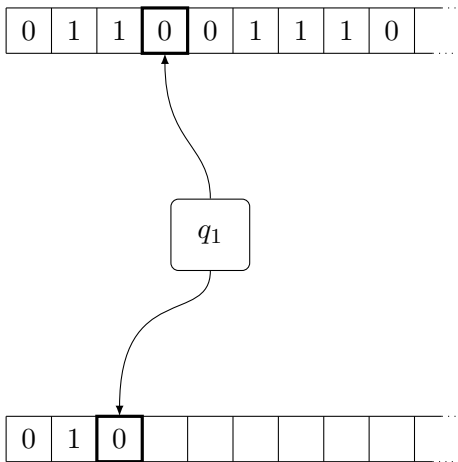
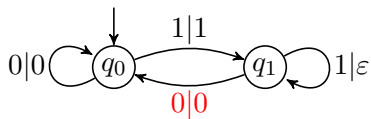
Example



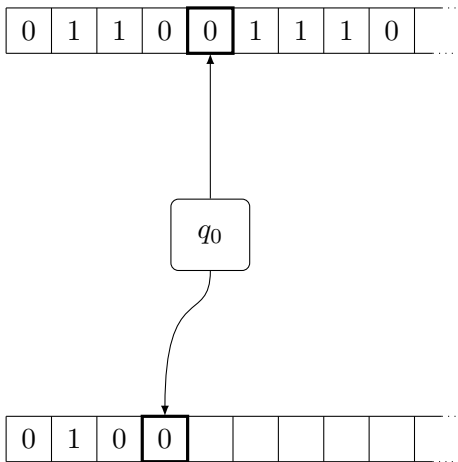
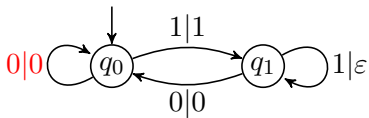
Example



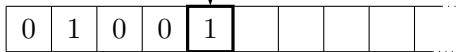
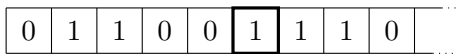
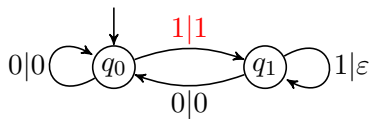
Example



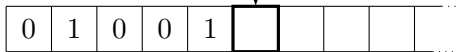
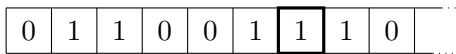
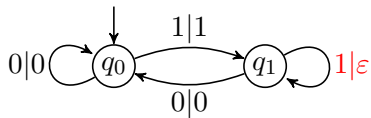
Example



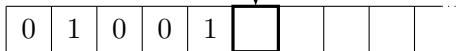
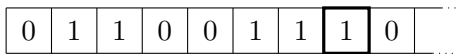
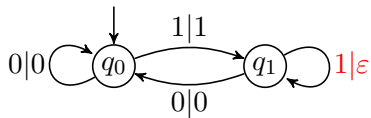
Example



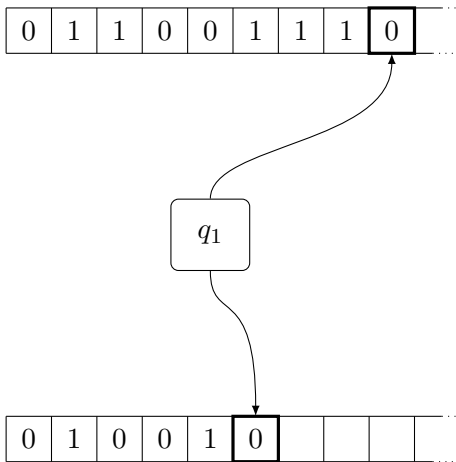
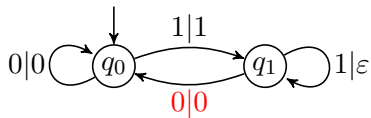
Example



Example



Example



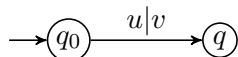
Transducers as compressors

An infinite word $x = a_1 a_2 a_3 \dots$ is *compressible* by a transducer if there is an accepting run $q_0 \xrightarrow{a_1|v_1} q_1 \xrightarrow{a_2|v_2} q_2 \xrightarrow{a_3|v_3} q_3 \dots$ satisfying

$$\liminf_{n \rightarrow \infty} \frac{|v_1 v_2 \dots v_n| \log |B|}{|a_1 a_2 \dots a_n| \log |A|} < 1.$$

Different notions of compressors

- ▶ the function $x \mapsto T(x)$ is one-to-one
- ▶ deterministic lossless: the map $u \mapsto (v, q)$ is one-to-one



- ▶ the function $x \mapsto T(x)$ is bounded-to-one
There is a constant K such that $|\{x : T(x) = y\}| \leq K$.

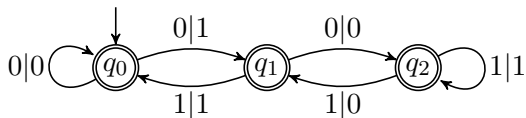
Characterization of normal words

Theorem (Many people)

An infinite word is normal if and only if it cannot be compressed by deterministic lossless transducers.

- ▶ Schnorr and Stimm (1971)
non-normality \Leftrightarrow finite-state martingale success
- ▶ Dai, Lathrop, Lutz and Mayordomo (2004)
compressibility \Leftrightarrow finite-state martingale success
normality \Rightarrow no martingale success
- ▶ Bourke, Hitchcock and Vinodchandran (2005)
non-normality \Rightarrow martingale success
- ▶ Becher and Heiber (2013)
non-normality \Leftrightarrow compressibility (direct)
- ▶ Becher, Carton and Heiber
generalized to bounded-to-one

Deterministic vs Non-Deterministic transducers



Multiplication by 3 in base 2

Theorem

Non-deterministic bounded-to-one transducers cannot compress normal infinite words.

Counter transducers

- ▶ the transducer uses k -counters with integer values that can be incremented, decremented and tested for zero
- ▶ real-time restriction: incrementation and decrementation can only occur when a input symbol is processed

Theorem

Bounded-to-one counter transducers cannot compress normal infinite words.

Non-real-time two-counter machines are Turing complete.

Summary of the results

| | det | non-det | non-rt |
|---------------------|-----|---------|--------|
| finite-state | N | N | N |
| 1 counter | N | N | N |
| ≥ 2 counters | N | N | T |
| 1 stack | ? | C | C |
| 1 stack + 1 counter | C | C | T |

where

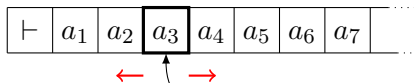
N means *cannot compress normal words*

C means *can compress some normal word*

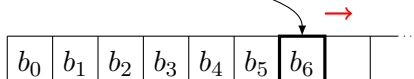
T means *is Turing complete* and thus can compress.

Two-way transducers

Two-way
input tape

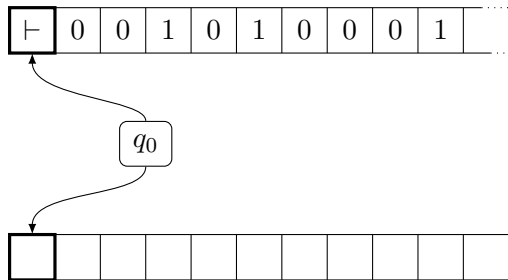
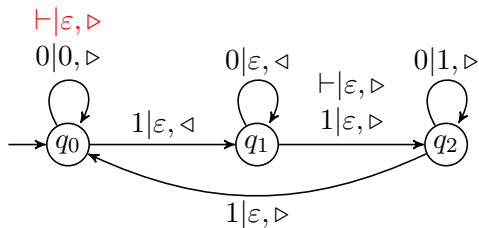


One-way
output tape

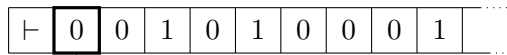
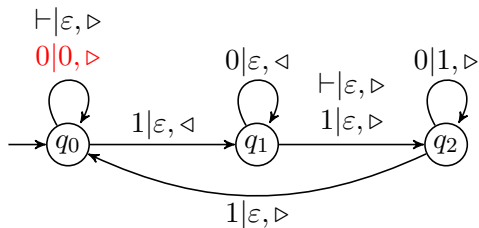


Transitions $p \xrightarrow{a|v,d} q$ for $a \in A$, $v \in B^*$ and $d \in \{\triangleleft, \triangleright\}$.

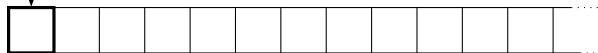
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



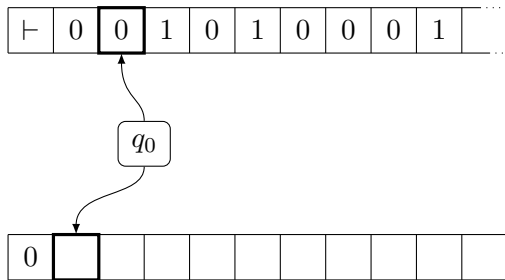
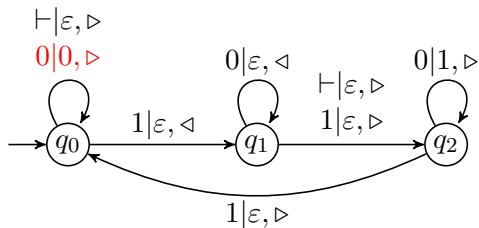
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



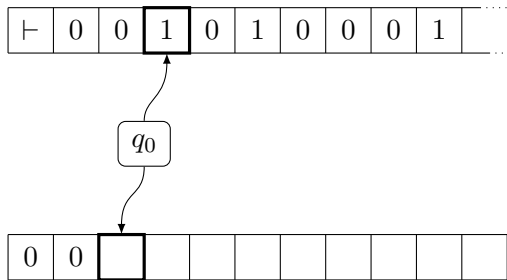
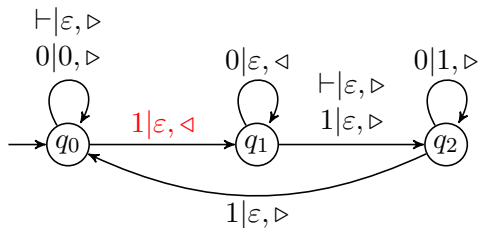
q_0



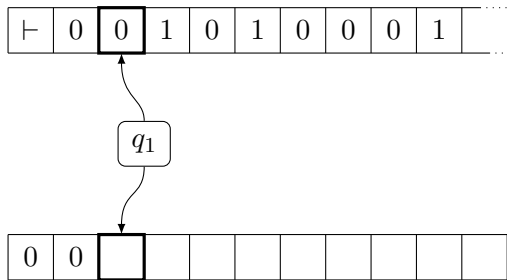
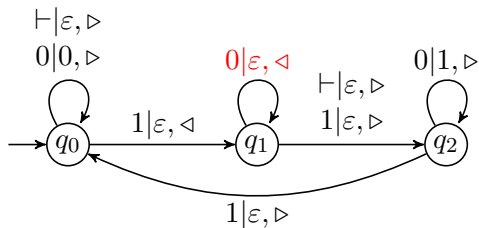
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



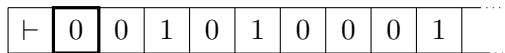
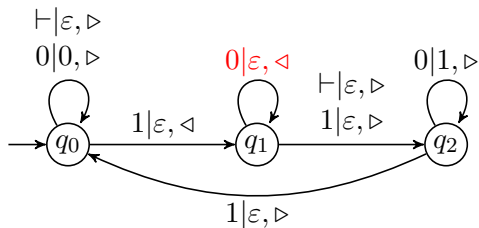
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



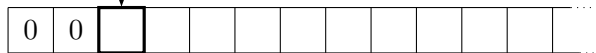
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



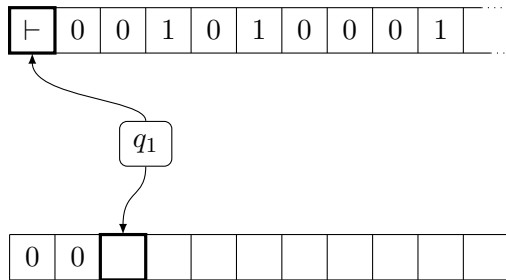
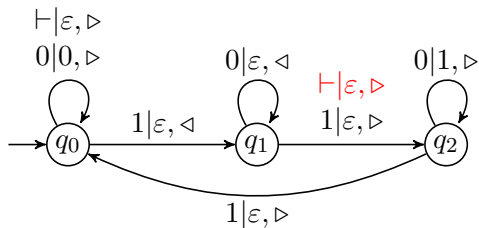
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



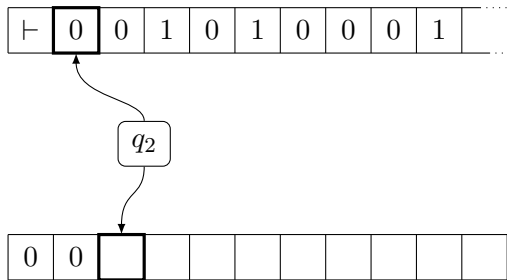
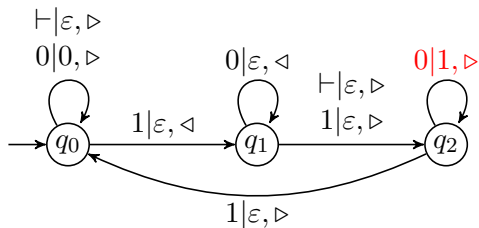
q_1



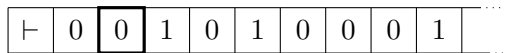
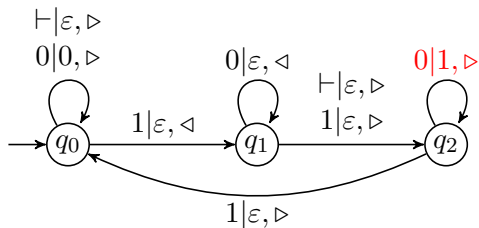
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



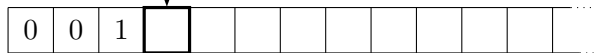
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



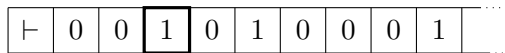
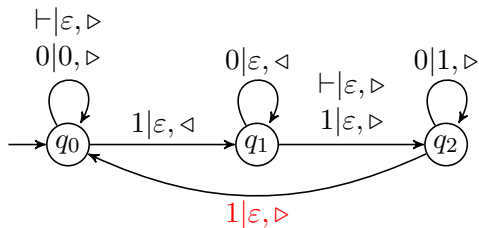
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



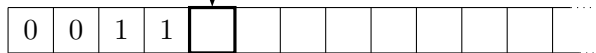
q_2



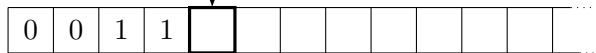
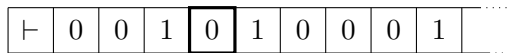
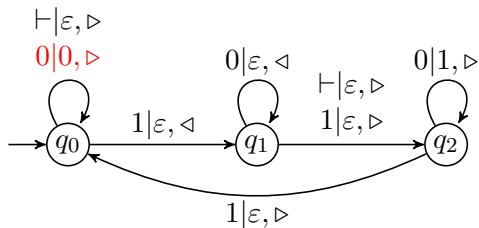
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



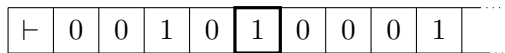
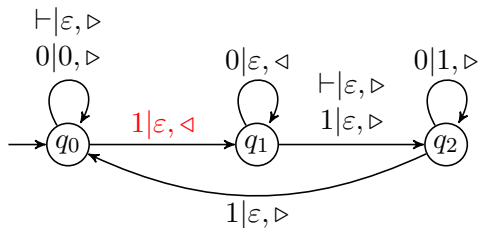
q_2



Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



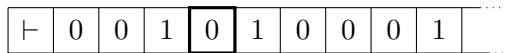
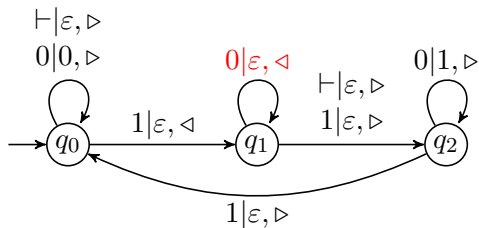
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



q_0



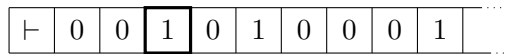
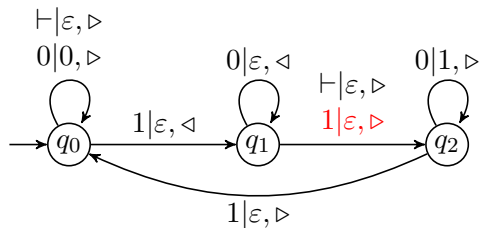
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



q_1



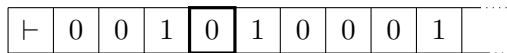
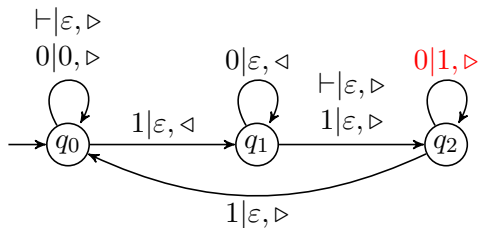
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



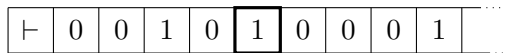
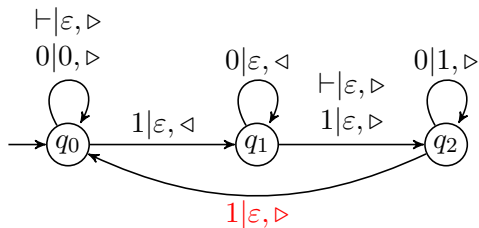
q_1



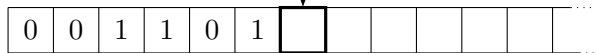
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



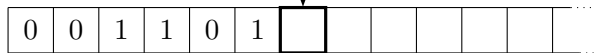
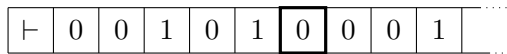
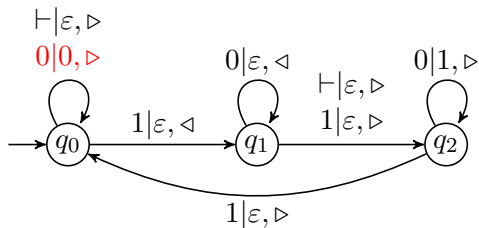
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



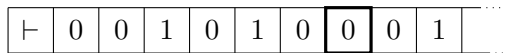
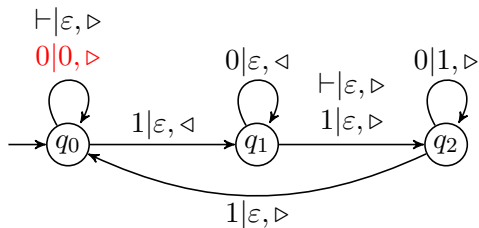
q_2



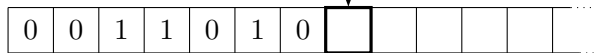
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



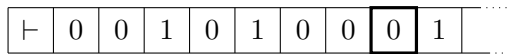
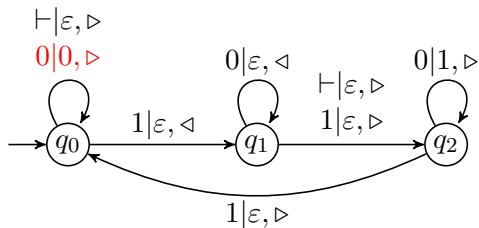
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



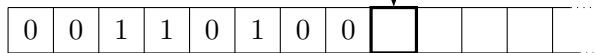
q_0



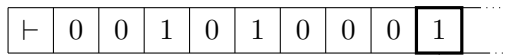
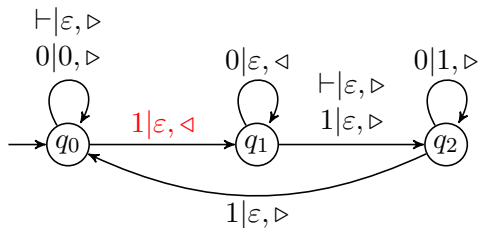
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



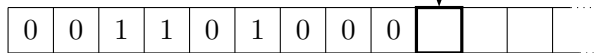
q_0



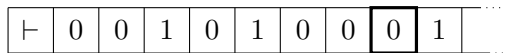
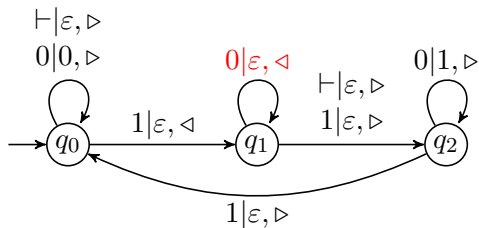
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



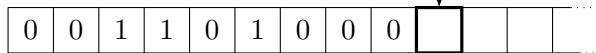
q_0



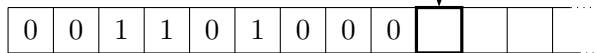
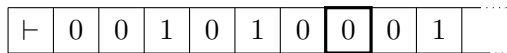
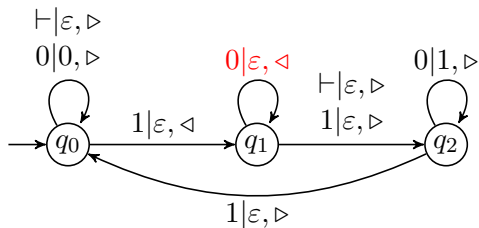
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



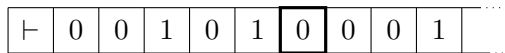
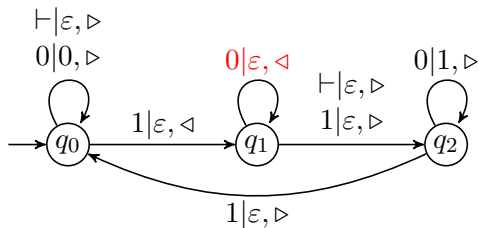
q_1



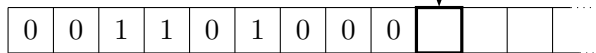
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



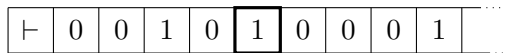
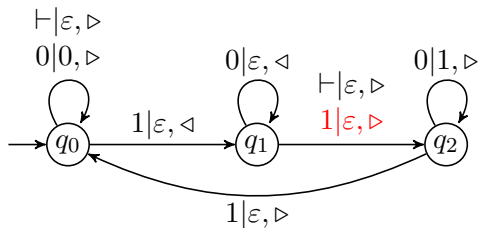
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



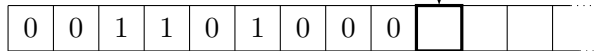
q_1



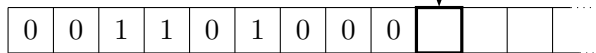
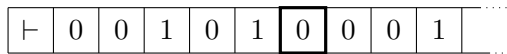
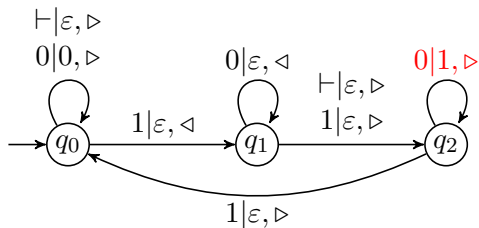
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



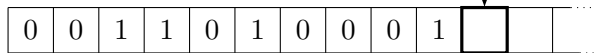
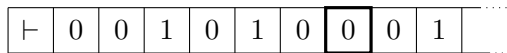
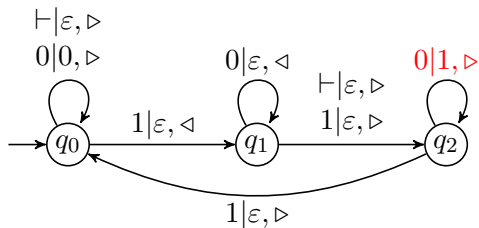
q_1



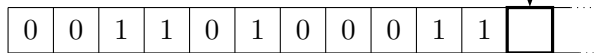
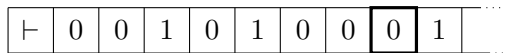
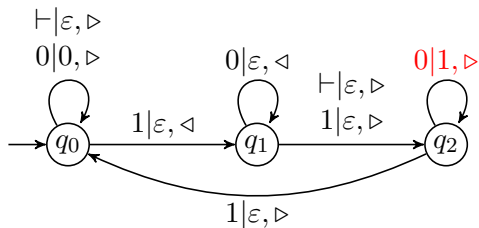
Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$



Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$

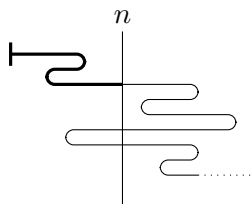


Example: $0^{n_0}10^{n_1}10^{n_2}1\dots \mapsto 0^{n_0}1^{n_0}0^{n_1}1^{n_1}0^{n_2}1^{n_2}\dots$

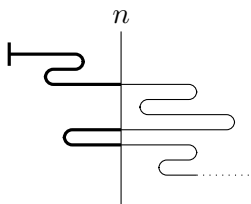


Ratios: first hit, last hit and in the middle

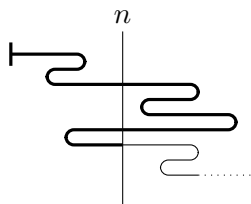
$$\liminf_{n \rightarrow \infty} \frac{|\text{?}|}{n} < 1.$$



First hit



Middle



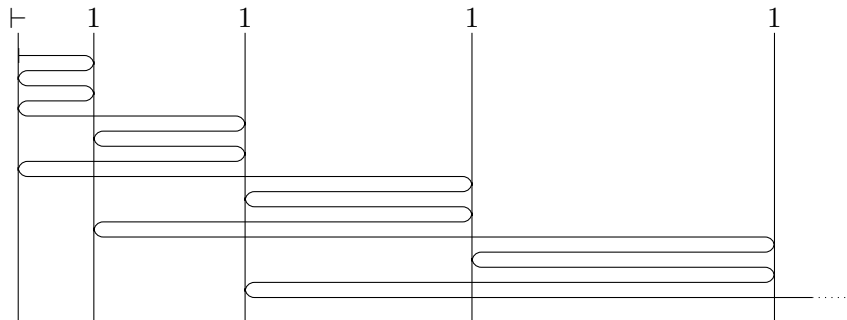
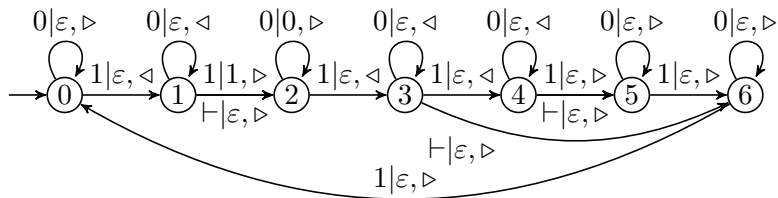
Last hit

First hit all output made up to the first hit of position n

Middle all output made at positions less than n

Last hit all output made up to the last hit of position n

Different ratios



Two-way transducers cannot compress normal words

Theorem

The first-hit, middle and last-hit ratios of the accepting run of a deterministic bounded-to-one two-way transducer over a normal infinite word coincide.

Theorem

For any run ρ of a non-deterministic two-way bounded-to-one transducer, there is another run ρ' with smaller ratios, such that first-hit, middle and last-hit ratios coincide.

Theorem

Deterministic and non-deterministic two-way bounded-to-one transducers cannot compress normal infinite words.

Selection rules

- ▶ If $x = a_1a_2a_3\cdots$ is a normal infinite word, then so is $x' = a_2a_3a_4\cdots$ made of symbols at all positions but the first one.
- ▶ If $x = a_1a_2a_3\cdots$ is normal infinite word, then so is $x' = a_2a_4a_6\cdots$ made of symbols at even positions.
- ▶ What about selecting symbols at positions 2^n ?
- ▶ What about selecting symbols at prime positions ?
- ▶ What about selecting symbols following a 1 ?
- ▶ What about selecting symbols followed by a 1 ?

Prefix selection

Let $L \subseteq A^*$ be a set of finite words and $x = a_1 a_2 a_3 \cdots \in A^\omega$.

The **prefix selection** of x by L is the word $x \upharpoonright L = a_{i_1} a_{i_2} a_{i_3} \cdots$ where $\{i_1 < i_2 < i_3 < \cdots\} = \{i : a_1 a_2 \cdots a_{i-1} \in L\}$.

Example (Symbols following a 1)

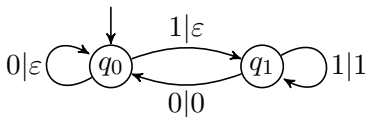
If $L = (0 + 1)^* 1$, then $i_1 - 1, i_2 - 1, i_3 - 1$ are the positions of 1 in x and $x \upharpoonright L$ is made of the symbols following a 1.

Theorem (Agafonov 1968)

Prefix selection by a rational set of finite words preserves normality.

The selection can be realized by a transducer.

Example (Selection of symbols following a 1)



Suffix selection

Let $X \subseteq A^\omega$ be a set of infinite words and $x = a_1a_2a_3 \cdots \in A^\omega$. The **suffix selection** of x by X is the word $x \upharpoonright X = a_{i_1}a_{i_2}a_{i_3} \cdots$ where $\{i_1 < i_2 < i_3 < \cdots\} = \{i : a_{i+1}a_{i+2}a_{i+3} \cdots \in X\}$.

Example (Symbols followed by a 1)

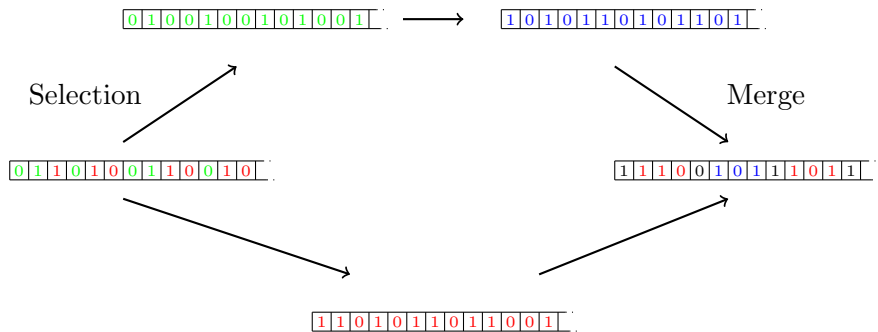
If $L = 1(0 + 1)^\omega$, then $i_1 + 1, i_2 + 1, i_3 + 1$ are the positions of 1 in x and $x \upharpoonright X$ is made of the symbols followed by a 1.

Theorem

Suffix selection by a rational set of infinite words preserves normality.

Picture

Compression



Combined prefix-suffix selection

Proposition

Let $x = a_1a_2a_3 \cdots \in A^\omega$ be an normal infinite word. The word $x' = a_{i_1}a_{i_2}a_{i_3} \cdots$ where $\{i_1 < i_2 < i_3 < \cdots\} = \{i : a_{i-1} = a_{i+1} = 1\}$ is not normal.

Merci

Randomness

Randomness can be characterized as non-compressibility:

$$\liminf_{n \rightarrow \infty} \mathcal{H}(x[1..n]) - n > -\infty$$

where \mathcal{H} is the prefix Kolmogorov complexity of the finite word w .

Normal infinite words are the **random words** for automata.

Turing may compress some normal words (Champernowne's).
What is the real power needed to compress a normal word ?

Ingredients

Shannon (1958)

- ▶ frequency of u different from $b^{-|u|}$ implies non maximum entropy
- ▶ non-maximum entropy implies compressibility

Huffman (1952)

- ▶ simple greedy implementation of Shannon's general idea
- ▶ implementation by a finite state transducer