Patricia's bad distributions

Ralph Neininger
(Goethe University Frankfurt)

joint work with **Louigi Addario-Berry** and **Pat Morin**

AofA 2024
University of Bath

Trie

Trie

Alphabet $\Sigma = \{0, 1\}$

Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$

Trie

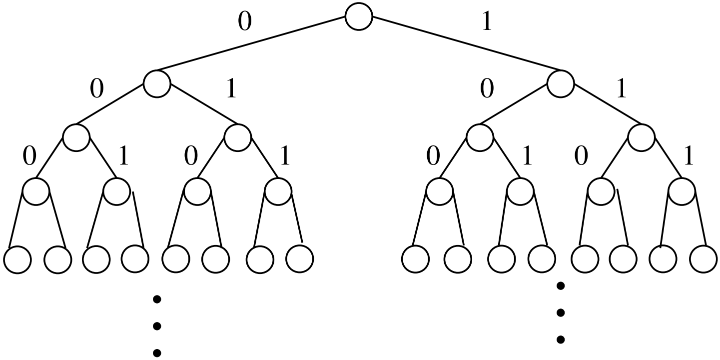
Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.

Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:

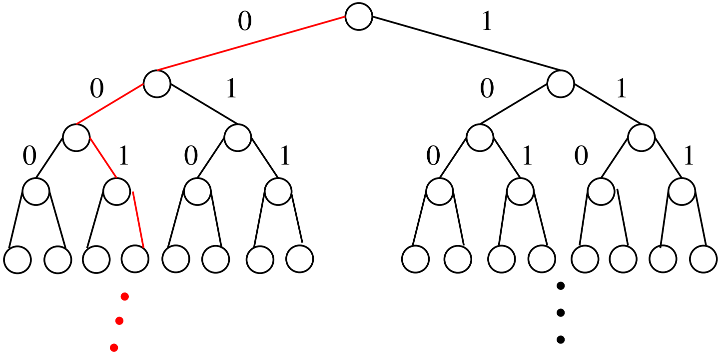
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



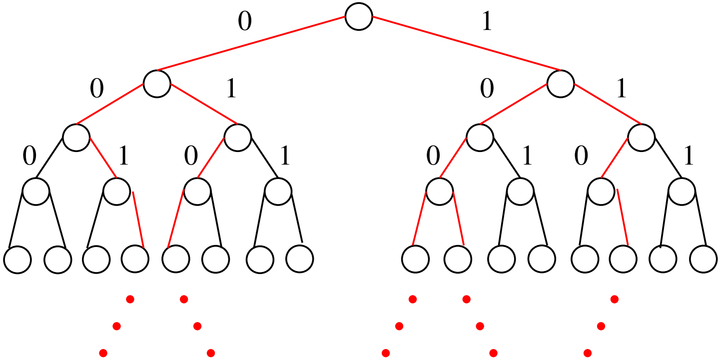
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



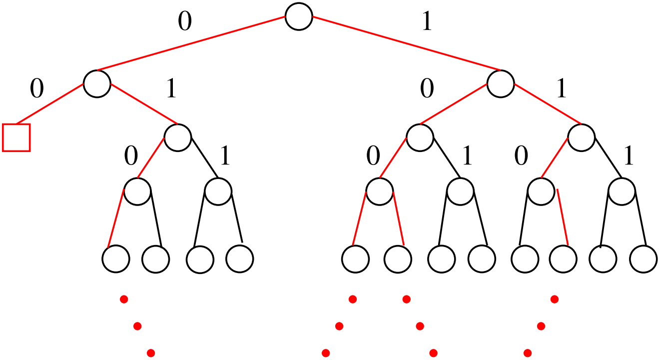
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



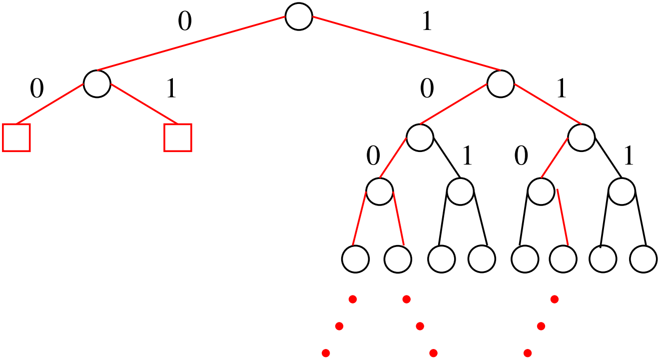
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



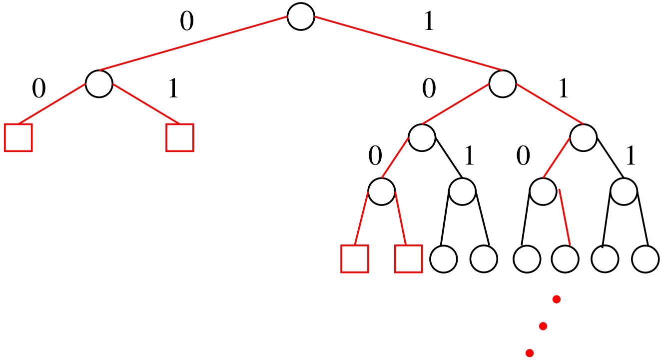
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



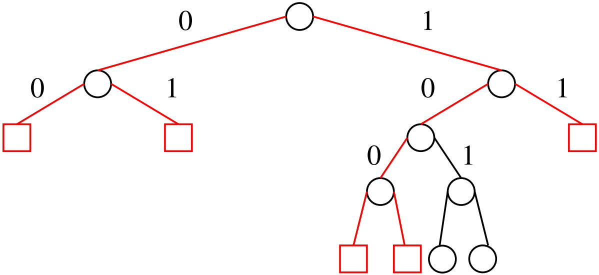
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



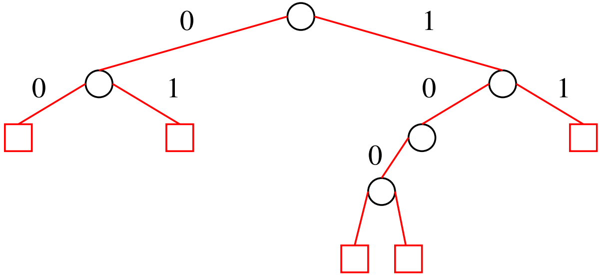
Trie

Alphabet $\Sigma = \{0, 1\}$

Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$

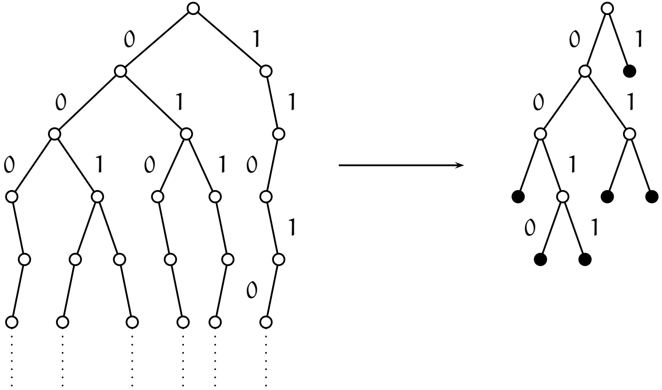
Infinite sequences of Bits called strings.

Construction of the trie:



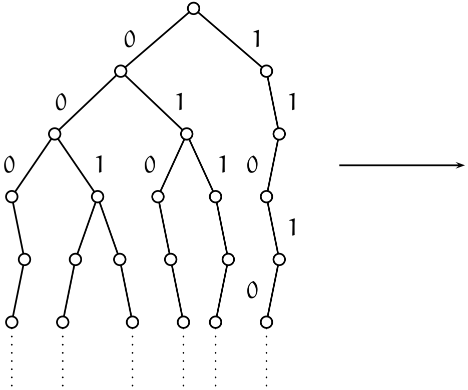
Trie

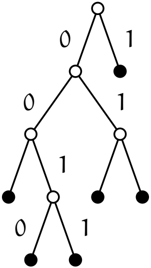
Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie:



Trie

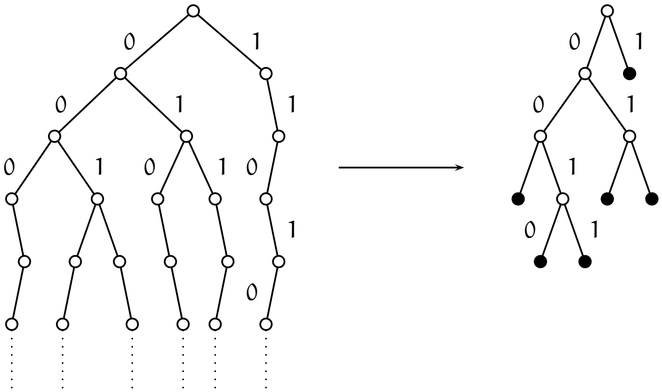
Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \dots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie: **(retrieval)**





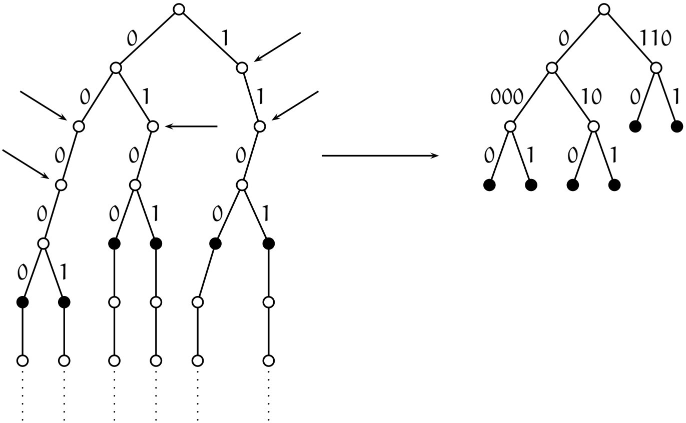
Trie

Alphabet $\Sigma = \{0, 1\}$
Data: $x^{(1)}, \ldots, x^{(n)} \in \Sigma^{\mathbb{N}}$
Infinite sequences of Bits called strings.
Construction of the trie: **(retrieval)** **Height**



PATRICIA tree

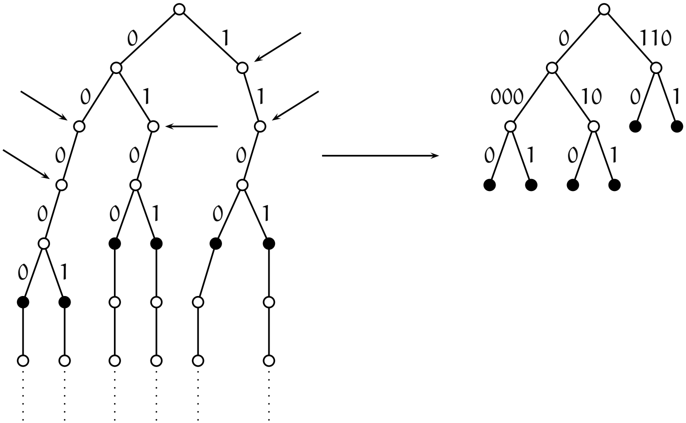
From Trie to PATRICIA tree:



PATRICIA tree

From Trie to PATRICIA tree:

Height: H_n

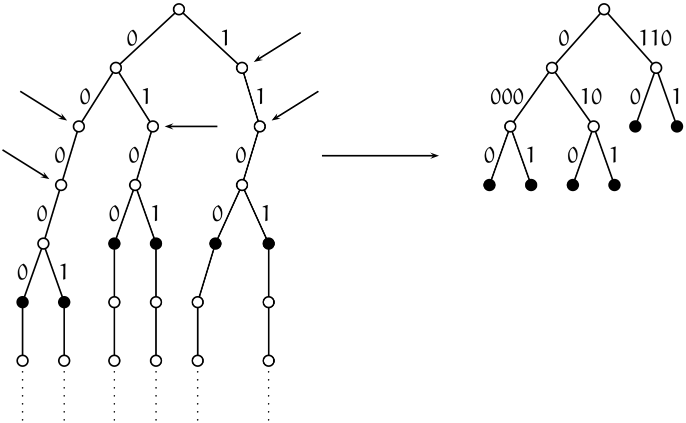


PATRICIA tree

From Trie to PATRICIA tree:

Height: H_n

PATRICIA: Practical Algorithm To Retrieve Information Coded In Alphanumeric (D. Morrison 1968, G. Gwehenberger 1968)



Probabilistic models

Data random: $X^{(1)}, \ldots, X^{(n)}$ i.i.d. with a distribution μ on $\Sigma^{\mathbb{N}}$.
General assumption: μ diffuse (non-atomic)

Probabilistic models

Data random: $X^{(1)}, \ldots, X^{(n)}$ i.i.d. with a distribution μ on $\Sigma^{\mathbb{N}}$.

General assumption: μ diffuse (non-atomic)

Special cases of $\mu = \mathcal{L}((X_j^{(1)})_{j\geq 1})$:

Probabilistic models

Data random: $X^{(1)}, \ldots, X^{(n)}$ i.i.d. with a distribution μ on $\Sigma^{\mathbb{N}}$.
General assumption: μ diffuse (non-atomic)
Special cases of $\mu = \mathcal{L}((X_j^{(1)})_{j \geq 1})$:

Bernoulli model: $X_1^{(1)}, X_2^{(1)}, \ldots$ i.i.d. Bernoulli(p)
Books of Knuth, Szpankowski, Mahmoud.
Devroye (2002), Janson(2012, 2022), Ischebeck (2023+)

Probabilistic models

Data random: $X^{(1)}, \dots, X^{(n)}$ i.i.d. with a distribution μ on $\Sigma^{\mathbb{N}}$.
General assumption: μ diffuse (non-atomic)
Special cases of $\mu = \mathcal{L}((X_j^{(1)})_{j\geq 1})$:

Bernoulli model: $X_1^{(1)}, X_2^{(1)}, \dots$ i.i.d. Bernoulli(p)
Books of Knuth, Szpankowski, Mahmoud.
Devroye (2002), Janson(2012, 2022), Ischebeck (2023+)

Markov model: $(X_j^{(1)})_{j\geq 1}$ is a homogenous Markov chain
Szpankowski & Jacquet (1991) Leckey, N. & Szpankowski (2013),
Leckey, N. & Sulzbach (2014)

Probabilistic models

Data random: $X^{(1)}, \dots, X^{(n)}$ i.i.d. with a distribution μ on $\Sigma^{\mathbb{N}}$.
General assumption: μ diffuse (non-atomic)
Special cases of $\mu = \mathcal{L}((X_j^{(1)})_{j \geq 1})$:

Bernoulli model: $X_1^{(1)}, X_2^{(1)}, \dots$ i.i.d. Bernoulli(p)
Books of Knuth, Szpankowski, Mahmoud.
Devroye (2002), Janson(2012, 2022), Ischebeck (2023+)

Markov model: $(X_j^{(1)})_{j \geq 1}$ is a homogenous Markov chain
Szpankowski & Jacquet (1991) Leckey, N. & Szpankowski (2013),
Leckey, N. & Sulzbach (2014)

Density model: Z with density f on $[0, 1]$.
Then binary expansion $Z = 0.X_1^{(1)} X_2^{(1)} \dots$
Devroye (1992)

Probabilistic models

Data random: $X^{(1)}, \dots, X^{(n)}$ i.i.d. with a distribution μ on $\Sigma^{\mathbb{N}}$.
General assumption: μ diffuse (non-atomic)
Special cases of $\mu = \mathcal{L}((X_j^{(1)})_{j \geq 1})$:

Bernoulli model: $X_1^{(1)}, X_2^{(1)}, \dots$ i.i.d. Bernoulli(p)
Books of Knuth, Szpankowski, Mahmoud.
Devroye (2002), Janson(2012, 2022), Ischebeck (2023+)

Markov model: $(X_j^{(1)})_{j \geq 1}$ is a homogenous Markov chain
Szpankowski & Jacquet (1991) Leckey, N. & Szpankowski (2013), Leckey, N. & Sulzbach (2014)

Density model: Z with density f on $[0, 1]$.
Then binary expansion $Z = 0.X_1^{(1)} X_2^{(1)} \dots$
Devroye (1992)

Dynamical sources: Clément, Flajolet & Vallée (2001)

Selected results for the height of PATRICIA trees

Symmetric Bernoulli model ($p = \frac{1}{2}$):

Selected results for the height of PATRICIA trees

Symmetric Bernoulli model ($p = \frac{1}{2}$):

Theorem: (Pittel, 1989)

$$\lim_{n\to\infty} \frac{H_n}{\log_2 n} = 1 \quad \text{almost surely}$$

Selected results for the height of PATRICIA trees

Symmetric Bernoulli model ($p = \frac{1}{2}$):

Theorem: (Pittel, 1989)

$$\lim_{n\to\infty} \frac{H_n}{\log_2 n} = 1 \quad \text{almost surely}$$

Arbitrary diffuse distribution

Selected results for the height of PATRICIA trees

Symmetric Bernoulli model ($p = \frac{1}{2}$):

Theorem: (Pittel, 1989)

$$\lim_{n\to\infty} \frac{H_n}{\log_2 n} = 1 \quad \text{almost surely}$$

Arbitrary diffuse distribution

Theorem: (Devroye, 2005)
For all $t > 0$ we have

$$\mathbb{P}(H_n \geq \mathbb{E}[H_n] + t) \leq \exp\left(-\frac{t^2}{2(\mathbb{E}[H_n]+1) + 2t/3}\right)$$

$$\mathbb{P}(H_n \leq \mathbb{E}[H_n] - t) \leq \exp\left(-\frac{t^2}{2(\mathbb{E}[H_n]+1)}\right)$$

Question of S. Evans and A. Wakolbinger

Question of S. Evans and A. Wakolbinger

How high can a random PATRICIA tree grow?

Question of S. Evans and A. Wakolbinger

How high can a random PATRICIA tree grow?

(For arbitrary diffuse μ on $\Sigma^{\mathbb{N}}$.)

Results (with L. Addario-Berry and P. Morin)

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

$$\frac{\mathbb{E}[H_n]}{n} \to 0, \qquad \frac{H_n}{n} \to 0 \quad \text{almost surely.}$$

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

$$\frac{\mathbb{E}[H_n]}{n} \to 0, \qquad \frac{H_n}{n} \to 0 \quad \text{almost surely.}$$

Theorem 2: (Any sublinear height is possible)

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

$$\frac{\mathbb{E}[H_n]}{n} \to 0, \qquad \frac{H_n}{n} \to 0 \quad \text{almost surely.}$$

Theorem 2: (Any sublinear height is possible)
For any sequence $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

$$\frac{\mathbb{E}[H_n]}{n} \to 0, \qquad \frac{H_n}{n} \to 0 \quad \text{almost surely.}$$

Theorem 2: (Any sublinear height is possible)
For any sequence $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$
there exists a diffuse μ_α on $\{0,1\}^{\mathbb{N}}$, such that

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

$$\frac{\mathbb{E}[H_n]}{n} \to 0, \qquad \frac{H_n}{n} \to 0 \quad \text{almost surely.}$$

Theorem 2: (Any sublinear height is possible)
For any sequence $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$
there exists a diffuse μ_α on $\{0,1\}^{\mathbb{N}}$, such that

$$\frac{\mathbb{E}[H_n]}{n/\alpha_n} \to \infty, \qquad \frac{H_n}{n/\alpha_n} \to \infty \quad \text{almost surely.}$$

Results (with L. Addario-Berry and P. Morin)

Theorem 1: (The height is sublinear)
For any diffuse probability distribution μ on $\{0,1\}^{\mathbb{N}}$,

$$\frac{\mathbb{E}[H_n]}{n} \to 0, \qquad \frac{H_n}{n} \to 0 \quad \text{almost surely.}$$

Theorem 2: (Any sublinear height is possible)
For any sequence $\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$
there exists a diffuse μ_α on $\{0,1\}^{\mathbb{N}}$, such that

$$\frac{\mathbb{E}[H_n]}{n/\alpha_n} \to \infty, \qquad \frac{H_n}{n/\alpha_n} \to \infty \quad \text{almost surely.}$$

Remark: Devroye (1992) has Theorem 1 for the density model and Theorem 2 for $\alpha_n = n^\varepsilon$.

Idea for Theorem 2

Auxiliary distribution μ_N on $\{0,1\}^{\mathbb{N}}$ for $N \in \mathbb{N}$:

Idea for Theorem 2

Auxiliary distribution μ_N on $\{0,1\}^{\mathbb{N}}$ for $N \in \mathbb{N}$:

T uniformly on $\{0, \dots, N^2 - 1\}$

Idea for Theorem 2

Auxiliary distribution μ_N on $\{0,1\}^{\mathbb{N}}$ for $N \in \mathbb{N}$:

T uniformly on $\{0, \ldots, N^2 - 1\}$

$(B_i)_{i \geq 1}$ i.i.d. Bernoulli($\frac{1}{2}$) (indep. of T)

Idea for Theorem 2

Auxiliary distribution μ_N on $\{0,1\}^{\mathbb{N}}$ for $N \in \mathbb{N}$:

T uniformly on $\{0, \ldots, N^2 - 1\}$

$(B_i)_{i \geq 1}$ i.i.d. Bernoulli($\frac{1}{2}$) (indep. of T)

Conditional on $\{T = k\}$,

$$\xi_i := \begin{cases} 0, & \text{if } i < k, \\ 1, & \text{if } i = k, \\ B_{i-k}, & \text{if } i > k. \end{cases}$$

Idea for Theorem 2

Auxiliary distribution μ_N on $\{0,1\}^{\mathbb{N}}$ for $N \in \mathbb{N}$:

T uniformly on $\{0, \ldots, N^2 - 1\}$

$(B_i)_{i \geq 1}$ i.i.d. Bernoulli($\frac{1}{2}$) (indep. of T)

Conditional on $\{T = k\}$,

$$\xi_i := \begin{cases} 0, & \text{if } i < k, \\ 1, & \text{if } i = k, \\ B_{i-k}, & \text{if } i > k. \end{cases}$$

μ_N is defined as the distribution of $\Xi = (\xi_i)_{i \in \mathbb{N}}$.

Idea for Theorem 2

Auxiliary distribution μ_N on $\{0,1\}^{\mathbb{N}}$ for $N \in \mathbb{N}$:

T uniformly on $\{0, \dots, N^2 - 1\}$

$(B_i)_{i \geq 1}$ i.i.d. Bernoulli($\frac{1}{2}$) (indep. of T)

Conditional on $\{T = k\}$,

$$\xi_i := \begin{cases} 0, & \text{if } i < k, \\ 1, & \text{if } i = k, \\ B_{i-k}, & \text{if } i > k. \end{cases}$$

μ_N is defined as the distribution of $\Xi = (\xi_i)_{i \in \mathbb{N}}$.

Lemma: For all $n \in \{1, \dots, N\}$ and μ_N we have

$$\mathbb{E}[H_n] \geq n - 2.$$

Idea for Theorem 2

Idea for Theorem 2

$\alpha = (\alpha_n)_{n \in \mathbb{N}}$ with $\alpha_n \to \infty$

Idea for Theorem 2

$\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$

$\beta_n := \lfloor \log_2 \alpha_n \rfloor - 2.$

$$B(n) := \max\{m \in \mathbb{N} \mid \beta_m \leq n\}, \quad n \in \mathbb{N}.$$

Idea for Theorem 2

$\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$

$\beta_n := \lfloor \log_2 \alpha_n \rfloor - 2.$

$$B(n) := \max\{m \in \mathbb{N} \mid \beta_m \leq n\}, \quad n \in \mathbb{N}.$$

G geometric($\frac{1}{2}$) distributed.

Idea for Theorem 2

$\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$

$\beta_n := \lfloor \log_2 \alpha_n \rfloor - 2.$

$$B(n) := \max\{m \in \mathbb{N} \mid \beta_m \leq n\}, \quad n \in \mathbb{N}.$$

G geometric($\frac{1}{2}$) distributed.

Conditional on $\{G = k\}$

$$\phi_i = \begin{cases} 0, & \text{if } i < k, \\ 1, & \text{if } i = k, \\ \xi_{i-k}, & \text{if } i > k, \end{cases}$$

where $\Xi = (\xi_i)_{i\in\mathbb{N}}$ is independent of G and has distribution $\mu_{B(k)}$.

Idea for Theorem 2

$\alpha = (\alpha_n)_{n\in\mathbb{N}}$ with $\alpha_n \to \infty$

$\beta_n := \lfloor \log_2 \alpha_n \rfloor - 2.$

$$B(n) := \max\{m \in \mathbb{N} \,|\, \beta_m \le n\}, \quad n \in \mathbb{N}.$$

G geometric($\frac{1}{2}$) distributed.

Conditional on $\{G = k\}$

$$\phi_i = \begin{cases} 0, & \text{if } i < k, \\ 1, & \text{if } i = k, \\ \xi_{i-k}, & \text{if } i > k, \end{cases}$$

where $\Xi = (\xi_i)_{i\in\mathbb{N}}$ is independent of G and has distribution $\mu_{B(k)}$.

$$\mu_\alpha := \mathcal{L}((\phi_i)_{i\in\mathbb{N}}).$$

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left|\left\{1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)}\right) = (0, \ldots, 0, 1)\right\}\right|$$

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left|\left\{1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)}\right) = (0, \ldots, 0, 1)\right\}\right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$.

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left|\left\{1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)}\right) = (0, \ldots, 0, 1)\right\}\right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$. Note $2^{-\beta_n} \geq 4/\alpha_n$.

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left| \left\{ 1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)}\right) = (0, \ldots, 0, 1) \right\} \right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$. Note $2^{-\beta_n} \geq 4/\alpha_n$.

$$\mathbb{P}\left(X_n < \frac{2n}{\alpha_n}\right) \leq \exp\left(-\frac{n}{\alpha_n}\right) \qquad \text{(Okamoto)}$$

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left| \left\{ 1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)} \right) = (0, \ldots, 0, 1) \right\} \right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$. Note $2^{-\beta_n} \geq 4/\alpha_n$.

$$\mathbb{P}\left(X_n < \frac{2n}{\alpha_n} \right) \leq \exp\left(-\frac{n}{\alpha_n} \right) \qquad \text{(Okamoto)}$$

W.h.p. $2n/\alpha_n$ of the n strings start with $(0^{\beta_n - 1}, 1)$.

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left|\left\{1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \dots, \Phi_{\beta_n}^{(j)}\right) = (0, \dots, 0, 1)\right\}\right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$. Note $2^{-\beta_n} \geq 4/\alpha_n$.

$$\mathbb{P}\left(X_n < \frac{2n}{\alpha_n}\right) \leq \exp\left(-\frac{n}{\alpha_n}\right) \qquad \text{(Okamoto)}$$

W.h.p. $2n/\alpha_n$ of the n strings start with $(0^{\beta_n - 1}, 1)$.

They have suffixes $(\Xi_1^{(j)}, \Xi_2^{(j)}, \dots)$ drawn ind. from $\mu_{B(\beta_n)}$

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left|\left\{1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)}\right) = (0, \ldots, 0, 1)\right\}\right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$. Note $2^{-\beta_n} \geq 4/\alpha_n$.

$$\mathbb{P}\left(X_n < \frac{2n}{\alpha_n}\right) \leq \exp\left(-\frac{n}{\alpha_n}\right) \qquad \text{(Okamoto)}$$

W.h.p. $2n/\alpha_n$ of the n strings start with $(0^{\beta_n - 1}, 1)$.

They have suffixes $(\Xi_1^{(j)}, \Xi_2^{(j)}, \ldots)$ drawn ind. from $\mu_{B(\beta_n)}$

For all large n we have $\lceil 2n/\alpha_n \rceil \leq n \leq B(\beta_n)$.

Idea for Theorem 2

Aim: $\mathbb{E}[H_n] \geq n/\alpha_n$ for all $n \geq n_0$.

$$X_n := \left|\left\{1 \leq j \leq n \,:\, \left(\Phi_1^{(j)}, \ldots, \Phi_{\beta_n}^{(j)}\right) = (0, \ldots, 0, 1)\right\}\right|$$

$X_n \sim \mathrm{B}(n, 2^{-\beta_n})$. Note $2^{-\beta_n} \geq 4/\alpha_n$.

$$\mathbb{P}\left(X_n < \frac{2n}{\alpha_n}\right) \leq \exp\left(-\frac{n}{\alpha_n}\right) \qquad \text{(Okamoto)}$$

W.h.p. $2n/\alpha_n$ of the n strings start with $(0^{\beta_n - 1}, 1)$.

They have suffixes $(\Xi_1^{(j)}, \Xi_2^{(j)}, \ldots)$ drawn ind. from $\mu_{B(\beta_n)}$

For all large n we have $\lceil 2n/\alpha_n \rceil \leq n \leq B(\beta_n)$.

The Lemma implies the assertion. $\square$

Almost sure behavior

Consider bad ν with $\mathbb{E}[H_n] \geq n/\log \alpha_n$.

Almost sure behavior

Consider bad ν with $\mathbb{E}[H_n] \geq n/\log \alpha_n$.

Devroye (2005)

$$\mathbb{P}\left(H_n \leq \mathbb{E}[H_n] - t\right) \leq \exp\left(-\frac{t^2}{2n}\right).$$

Almost sure behavior

Consider bad ν with $\mathbb{E}[H_n] \geq n / \log \alpha_n$.

Devroye (2005)

$$\mathbb{P}\left(H_n \leq \mathbb{E}[H_n] - t\right) \leq \exp\left(-\frac{t^2}{2n}\right).$$

This implies

$$\mathbb{P}\left(H_n \leq \frac{n}{\log^2 \alpha_n}\right) \leq \exp\left(-\frac{n}{2\log^4 n}\right)$$

Almost sure behavior

Consider bad ν with $\mathbb{E}[H_n] \geq n / \log \alpha_n$.

Devroye (2005)

$$\mathbb{P}\left(H_n \leq \mathbb{E}[H_n] - t\right) \leq \exp\left(-\frac{t^2}{2n}\right).$$

This implies

$$\mathbb{P}\left(H_n \leq \frac{n}{\log^2 \alpha_n}\right) \leq \exp\left(-\frac{n}{2\log^4 n}\right)$$

Borel–Cantelli Lemma implies

$$\frac{H_n}{n/\alpha_n} \to \infty \text{ almost surely.}$$

THANK YOU!

Idea for Theorem 1

$\Xi = \xi_1\xi_2\xi_3\dots$ with $\mathcal{L}(\Xi) = \mu$. **Claim:**

Idea for Theorem 1

$\Xi = \xi_1 \xi_2 \xi_3 \ldots$ with $\mathcal{L}(\Xi) = \mu$. **Claim:**

$$\forall \varepsilon > 0 \, \exists \, k \in \mathbb{N} \, \forall \, v = v_1 \ldots v_k : \mathbb{P}(\xi_1 \ldots \xi_k = v_1 \ldots v_k) < \varepsilon.$$

Idea for Theorem 1

$\Xi = \xi_1\xi_2\xi_3\ldots$ with $\mathcal{L}(\Xi) = \mu$. **Claim:**

$$\forall\, \varepsilon > 0\, \exists\, k \in \mathbb{N}\, \forall\, v = v_1 \ldots v_k : \mathbb{P}(\xi_1 \ldots \xi_k = v_1 \ldots v_k) < \varepsilon.$$

Assume

$$\exists\, \varepsilon > 0\, \forall\, k \in \mathbb{N}\, \exists\, v = v_1 \ldots v_k : \mathbb{P}(\xi_1 \ldots \xi_k = v_1 \ldots v_k) \geq \varepsilon.$$

Idea for Theorem 1

$\Xi = \xi_1\xi_2\xi_3\ldots$ with $\mathcal{L}(\Xi) = \mu$. **Claim:**

$$\forall\, \varepsilon > 0\, \exists\, k \in \mathbb{N}\, \forall\, v = v_1 \ldots v_k : \mathbb{P}(\xi_1 \ldots \xi_k = v_1 \ldots v_k) < \varepsilon.$$

Assume

$$\exists\, \varepsilon > 0\, \forall\, k \in \mathbb{N}\, \exists\, v = v_1 \ldots v_k : \mathbb{P}(\xi_1 \ldots \xi_k = v_1 \ldots v_k) \geq \varepsilon.$$

By a diagonal argument exists $v = v_1 v_2 v_3 \ldots \in \{0,1\}^{\mathbb{N}}$:

$$\mathbb{P}(\xi_1 \ldots \xi_k = v_1 \ldots v_k) \geq \varepsilon, \qquad \text{for all } k.$$

Idea for Theorem 1

$\Xi = \xi_1\xi_2\xi_3\ldots$ with $\mathcal{L}(\Xi) = \mu$. **Claim:**

$$\forall\,\varepsilon > 0\,\exists\,k \in \mathbb{N}\,\forall\,v = v_1\ldots v_k : \mathbb{P}(\xi_1\ldots\xi_k = v_1\ldots v_k) < \varepsilon.$$

Assume

$$\exists\,\varepsilon > 0\,\forall\,k \in \mathbb{N}\,\exists\,v = v_1\ldots v_k : \mathbb{P}(\xi_1\ldots\xi_k = v_1\ldots v_k) \geq \varepsilon.$$

By a diagonal argument exists $v = v_1 v_2 v_3 \ldots \in \{0,1\}^{\mathbb{N}}$:

$$\mathbb{P}(\xi_1\ldots\xi_k = v_1\ldots v_k) \geq \varepsilon, \qquad \text{for all } k.$$

Then, by continuity of measure

$$\mathbb{P}(\Xi = v) = \lim_{k\to\infty} \mathbb{P}(\xi_1\ldots\xi_k = v_1\ldots v_k) \geq \varepsilon.$$

Contradiction to diffuse.

Idea for Theorem 1

Consider $\varepsilon \in (0, 1/4)$ and $k = k(\varepsilon)$ as above.

Idea for Theorem 1

Consider $\varepsilon \in (0, 1/4)$ and $k = k(\varepsilon)$ as above. Sufficient is

$$\mathbb{P}(\exists\, n_0\, \forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n) = 1.$$

Idea for Theorem 1

Consider $\varepsilon \in (0, 1/4)$ and $k = k(\varepsilon)$ as above. Sufficient is

$$\mathbb{P}(\exists\, n_0 \,\forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n) = 1.$$

For $v = v_1 \dots v_k$ set

$$N_v := |\{1 \leq j \leq n : \xi_1^{(j)} \dots \xi_k^{(j)} = v_1 \dots v_k\}|,$$

Idea for Theorem 1

Consider $\varepsilon \in (0, 1/4)$ and $k = k(\varepsilon)$ as above. Sufficient is

$$\mathbb{P}(\exists\, n_0\, \forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n) = 1.$$

For $v = v_1 \dots v_k$ set

$$N_v := |\{1 \leq j \leq n : \xi_1^{(j)} \dots \xi_k^{(j)} = v_1 \dots v_k\}|,$$

$$E_{n,k} := \left\{ \max_{v \in \{0,1\}^k} N_v \geq 2\varepsilon n \right\}.$$

Idea for Theorem 1

Consider $\varepsilon \in (0, 1/4)$ and $k = k(\varepsilon)$ as above. Sufficient is

$$\mathbb{P}(\exists\, n_0\, \forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n) = 1.$$

For $v = v_1 \dots v_k$ set

$$N_v := |\{1 \leq j \leq n : \xi_1^{(j)} \dots \xi_k^{(j)} = v_1 \dots v_k\}|,$$
$$E_{n,k} := \left\{ \max_{v \in \{0,1\}^k} N_v \geq 2\varepsilon n \right\}.$$

Then

$$\{\exists\, n_0 : \forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n\} \supset \{E_{n,k} \text{ for finitely many } n\}$$
$$= \left(\limsup_{n\to\infty} E_{n,k} \right)^c$$

Idea for Theorem 1

Consider $\varepsilon \in (0, 1/4)$ and $k = k(\varepsilon)$ as above. Sufficient is

$$\mathbb{P}(\exists\, n_0 \,\forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n) = 1.$$

For $v = v_1 \ldots v_k$ set

$$N_v := |\{1 \leq j \leq n : \xi_1^{(j)} \ldots \xi_k^{(j)} = v_1 \ldots v_k\}|,$$

$$E_{n,k} := \left\{ \max_{v \in \{0,1\}^k} N_v \geq 2\varepsilon n \right\}.$$

Then

$$\begin{aligned} \{\exists\, n_0 : \,\forall\, n \geq n_0 : H_n \leq k + 2\varepsilon n\} &\supset \{E_{n,k} \text{ for finitely many } n\} \\ &= \left(\limsup_{n \to \infty} E_{n,k} \right)^c \end{aligned}$$

Tail bound for Binomial plus Borel–Cantelli.