TRIES

Wojciech Szpankowski*
Department of Computer Science
Purdue University, W. Lafayette, IN
U.S.A.

November 25, 2008

AofA and **IT** logos

This one is for you, PHILIPPE!

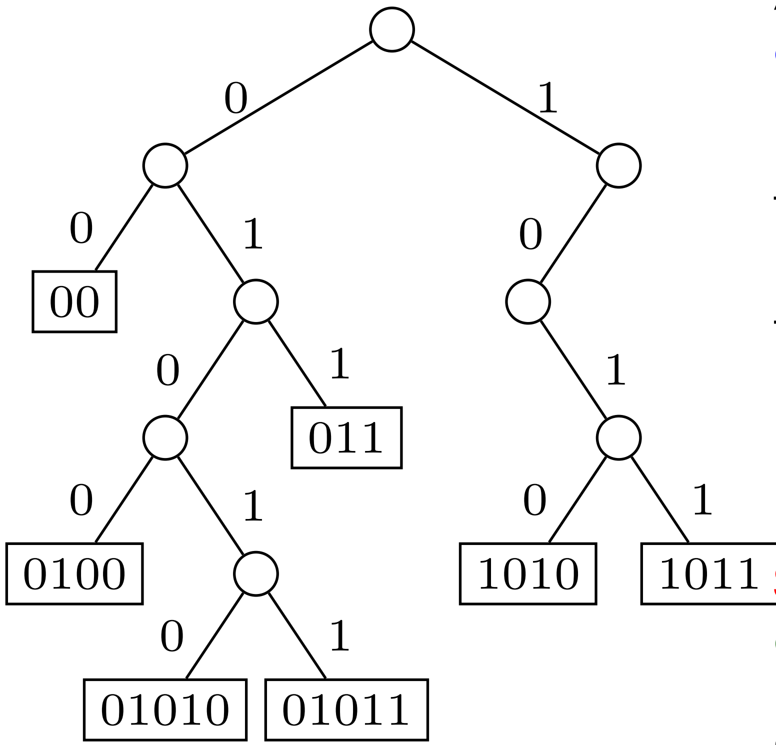
---

*Joint work with M. Drmota, H-K. Hwang, P. Nicodeme, and G. Park.

Outline of the Presentation

1. Tries and Suffix Trees
2. Flajolet and Tries – A Look Back
3. Profiles of Tries
   - Trie Parameters
   - Main Results
   - Sketch of Proofs
   - Consequences (height, fill-up, shortest path)
4. Profile of Digital Search Trees (announcement)

Tries and Suffix Trees



A **trie**, or prefix tree, is an ordered tree data structure that stores keys usually represented by strings.

Tries were introduced by de la Briandais (1959) and Fredkin (1960) who introduced the name:

"tries" derived from **retrieval**.

Suffix tree is a trie built form suffixes of one string.

Other digital trees are: PATRICIA and digital search trees.

Typical Tries: In this talk we mostly discuss random tries built from n (independent) sequences generated by a binary **memoryless source** with p denoting the probability of generating a "0" ($q = 1 - p \leq p$).

Usefulness of Tries

Tries and suffix tress are widely used in diverse applications:

- automatically correcting words in texts; Kukich (1992);
- taxonomies of regular language; Watson (1995);
- event history in datarace detection for multi-threaded object-oriented programs; Choi et al. (2002);
- internet IP addresses lookup; Nilsson and Tikkanen (2002);
- data compression, Lempel-Ziv, . . . ; W.S. (2001);
- **distributed hash tables**, Malkhi et al. (2002) and Adler et al. (2003).
- compression of graphical structures, Choi and W.S. (2008).

Fundamental, prototype data structures:

- variations and extensions: Patricia, DST, bucket digital search trees, k-d tries, quadtries, LC-tries, multiple-tries, etc.;
- splitting procedures using coin-flipping: collision resolution in multi-access (or broadcast) communication models, loser selection or leader election, etc.
- combinatorial interpretations in terms of words and urn models.

Outline Update

1. Tries and Suffix Trees
2. Flajolet and Tries – A Look Back
3. Profiles of Tries
4. Profile of Digital Search Trees

Flajolet, Devroye, and Tries – A Look Back

Flajolet, Devroye, and Tries – A Look Back

A Look Back:

January 1983 –conference in Paris from gloomy Poland. "Among many good talks one stood out for me. It was on approximate counting, by Philippe Flajolet. The precision of the analysis and the brightness of the speaker made a lasting impression on me".

Flajolet, Devroye, and Tries – A Look Back

A Look Back:

January 1983 –conference in Paris from gloomy Poland. "Among many good talks one stood out for me. It was on approximate counting, by Philippe Flajolet. The precision of the analysis and the brightness of the speaker made a lasting impression on me".

January 1984 – moved to McGill, worked on a recurrence about conflict resolution algorithms, and Luc Devroye told me about D.E. Knuth's three volume opus, and reminded me about Philippe Flajolet.

Flajolet, Devroye, and Tries – A Look Back

A Look Back:

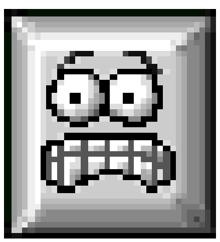
January 1983 –conference in Paris from gloomy Poland. "Among many good talks one stood out for me. It was on approximate counting, by Philippe Flajolet. The precision of the analysis and the brightness of the speaker made a lasting impression on me".

January 1984 – moved to McGill, worked on a recurrence about conflict resolution algorithms, and Luc Devroye told me about D.E. Knuth's three volume opus, and reminded me about Philippe Flajolet.

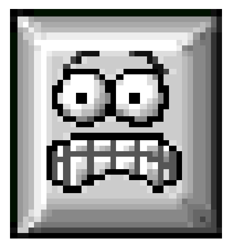
January 1985 – moved to Purdue, and discovered **tries**. I contacted **Flajolet** who sent me tons of papers and young P. Jacquet.
1985 - – my Flajolet's number is one.

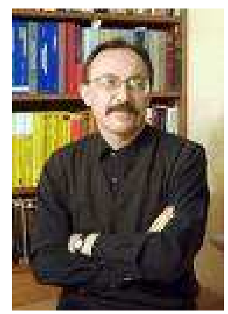
Outline Update

1. Tries and Suffix Trees
2. Flajolet and Tries – A Look Back
3. Profiles of Tries
   - Trie Parameters
   - Main Results
   - Sketch of Proofs
   - Consequences
4. Profile of Digital Search Trees



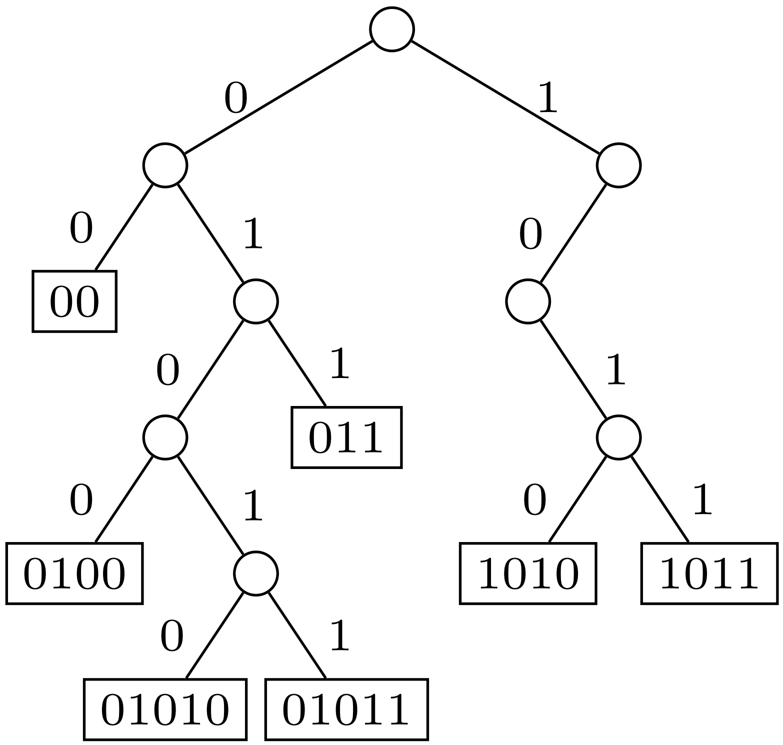
(G. Park)





(P. Nicodeme)

External and Internal Profiles



$B_n^0 = 0, \quad I_n^0 = 1$

$B_n^1 = 0, \quad I_n^1 = 2$

$B_n^2 = 1, \quad I_n^2 = 2$

$B_n^3 = 1, \quad I_n^3 = 2$

$B_n^4 = 3, \quad I_n^4 = 1$

$B_n^5 = 2 \quad I_n^5 = 0$

External profile and internal profile:

$B_n^k = \#$ external nodes at distance k from the root;

$I_n^k = \#$ internal nodes at distance k from the root.

Why to Study Profiles?

- Fine, informative shape characteristic;
- Related to path length, depth, height, shortest path, width, etc.;
- Breadth-first search;
- Compression algorithms.
- Mathematically challenging, phenomenally interesting!

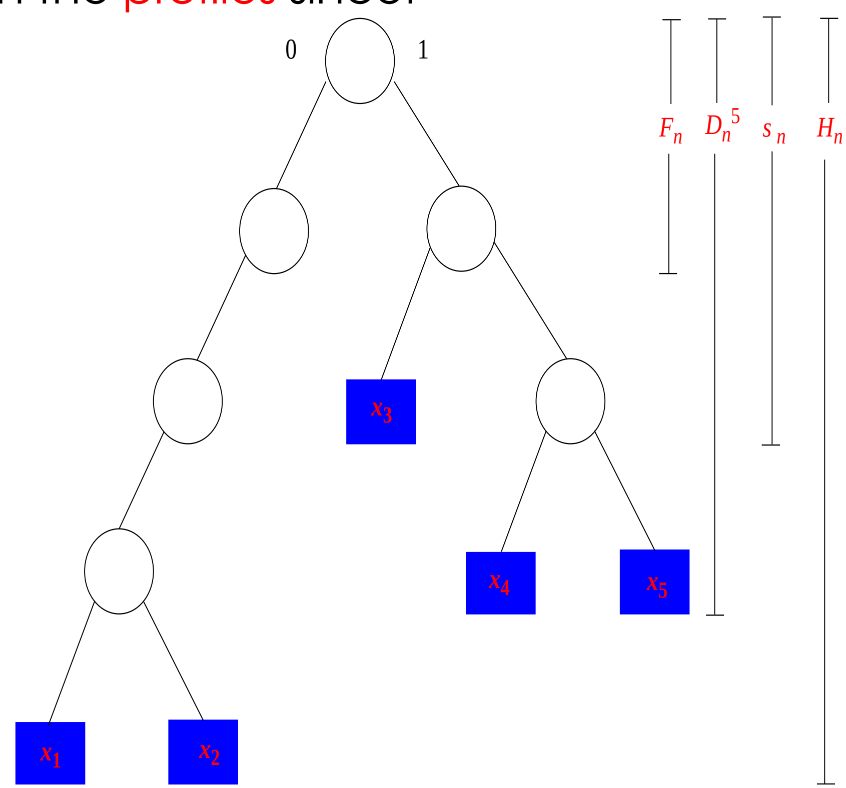
Example: Parameters such height H_n, shortest path, s_n, fill-up level F_n, and depth, D_n can be studied through the profiles since:

$$H_n = \max\{k : \ B_n^k > 0\},$$

$$s_n = \min\{k : \ B_n^k > 0\},$$

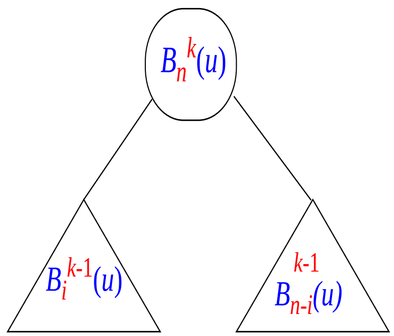
$$F_n = \max\{k : \ I_n^k = 2^k\},$$

$$\Pr(D_n = k) = \frac{\mathbb{E}[B_n^k]}{n}.$$



Recurrence for the Profiles

External Profile B_n^k:
Define the probability generating function as



$$B_n^k(u) = \mathbb{E}[u^{B_n^k}] = \sum_{\ell \geq 0} P(B_n^k = l)u^l.$$

Then

$$B_n^k(u) = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} B_i^{k-1}(u) B_{n-i}^{k-1}(u)$$

with $B_n^0 = 1$ for $n \neq 1$ and $B_1^0 = u$

Internal Profile probability generating function $I_n^k(u) = \mathbb{E}[I_n^k]$ satisfies the same recurrence with $U_n^0(u) = u$ for $n > 1$ and $U_0^0(u) = U_1^0(u) = 1$.

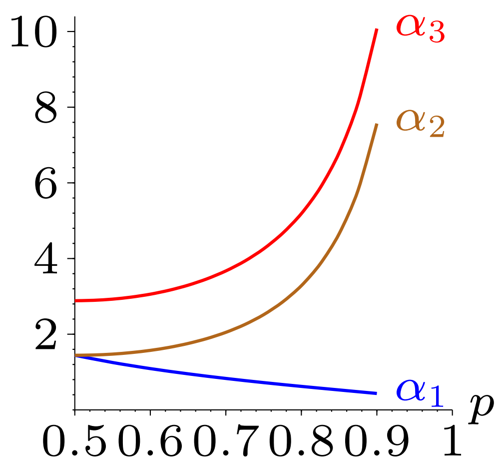
Average External Profile:

$$\boxed{\mathbb{E}[B_n^k] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^{k-1}] + \mathbb{E}[B_{n-i}^{k-1}]), \quad n \geq 2, k \geq 1,}$$

under some initial conditions (e.g., $\mathbb{E}[B_0^k] = 0$ for all k).

Main Results

Notation: $r = p/q = p/(1-p) > 1$, and $\alpha := \alpha_{n,k} = \frac{k}{\log n}$. Also:



$$\alpha_1 := \frac{1}{\log(1/q)},$$

$$\alpha_2 := \frac{p^2+q^2}{p^2\log(1/p)+q^2\log(1/q)},$$

$$\alpha_3 := \frac{2}{\log(1/(p^2+q^2))}.$$

1: **Exponential Growth** $(0 < \alpha < \alpha_1)$:

Let $1 \leq k \leq \frac{1}{\log q^{-1}}(\log n - \log\log\log n + \log(r-1) - \varepsilon)$:

$$\mathbb{E}[B_n^k] = nq^k(1-q^k)^{n-1}\left(1 + O\left((\log n)^{-\delta}\right)\right) = O(2^{-n\nu})$$

2: **Logarithmic Growth** $(0 < \alpha < \alpha_1)$:

Let $1 \leq k \leq \frac{1}{\log q^{-1}}(\log n - \log\log\log n + m\log(r-1) - \varepsilon)$:

$$\mathbb{E}[B_n^k] = O(\log\log n \cdot \log^{m-\beta} n).$$

where m and β are constants (smaller or greater than m).

Phase Transitions

3: **Polynomial Growth**: $\alpha_1 \cdot \log n < k < \alpha_2 \cdot \log n$: $(\alpha_1 < \alpha < \alpha_2)$

$$\mathbb{E}[B_n^k] \sim G_1(\log n)\frac{p^\rho q^\rho(p^{-\rho}+q^{-\rho})}{\sqrt{2\pi\alpha}\,\log(p/q)} \cdot \frac{n^{\upsilon_1}}{\sqrt{\log n}},$$

where $G_1(x)$ is a periodic function and

$$\upsilon_1 = -\rho + \alpha\log(p^{-\rho}+q^{-\rho}), \quad \rho = -\frac{1}{\log(p/q)}\log\left(\frac{-1-\alpha\log q}{1+\alpha\log p}\right).$$

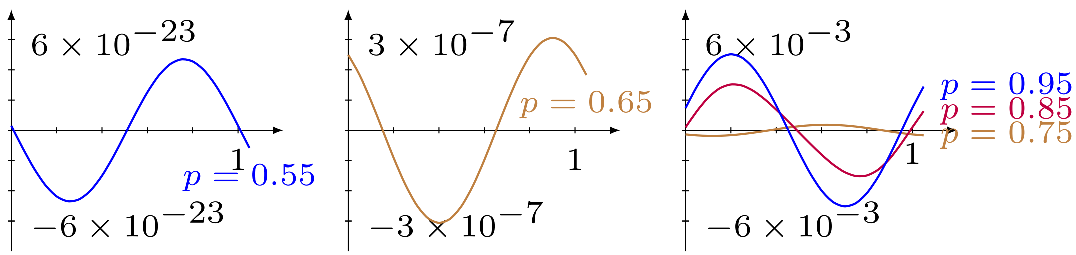


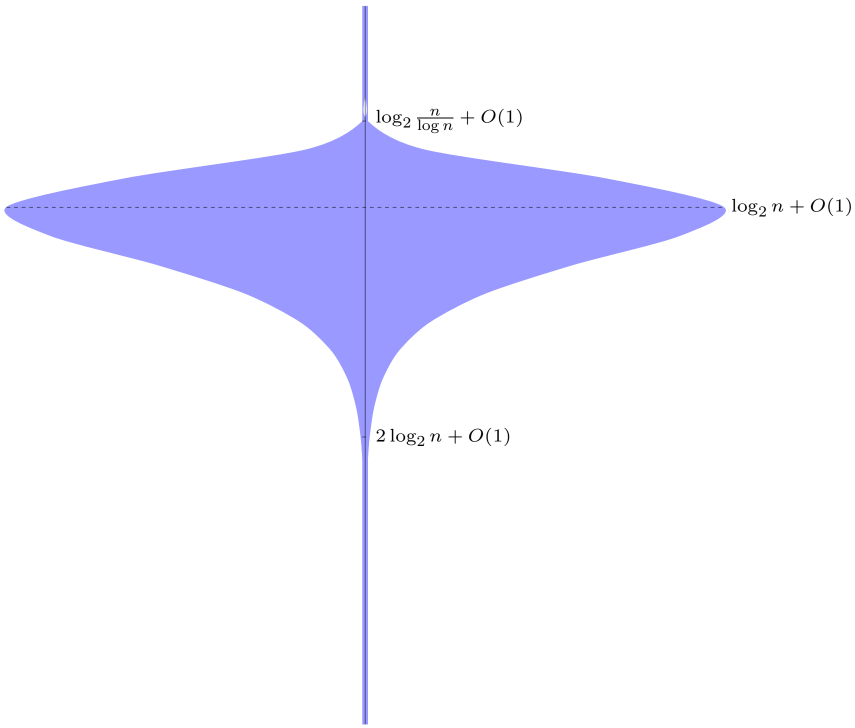
Figure 1: *The fluctuating part of the periodic function $G_1(x)$ for $p = 0.55, 0.65, \ldots, 0.95$.*

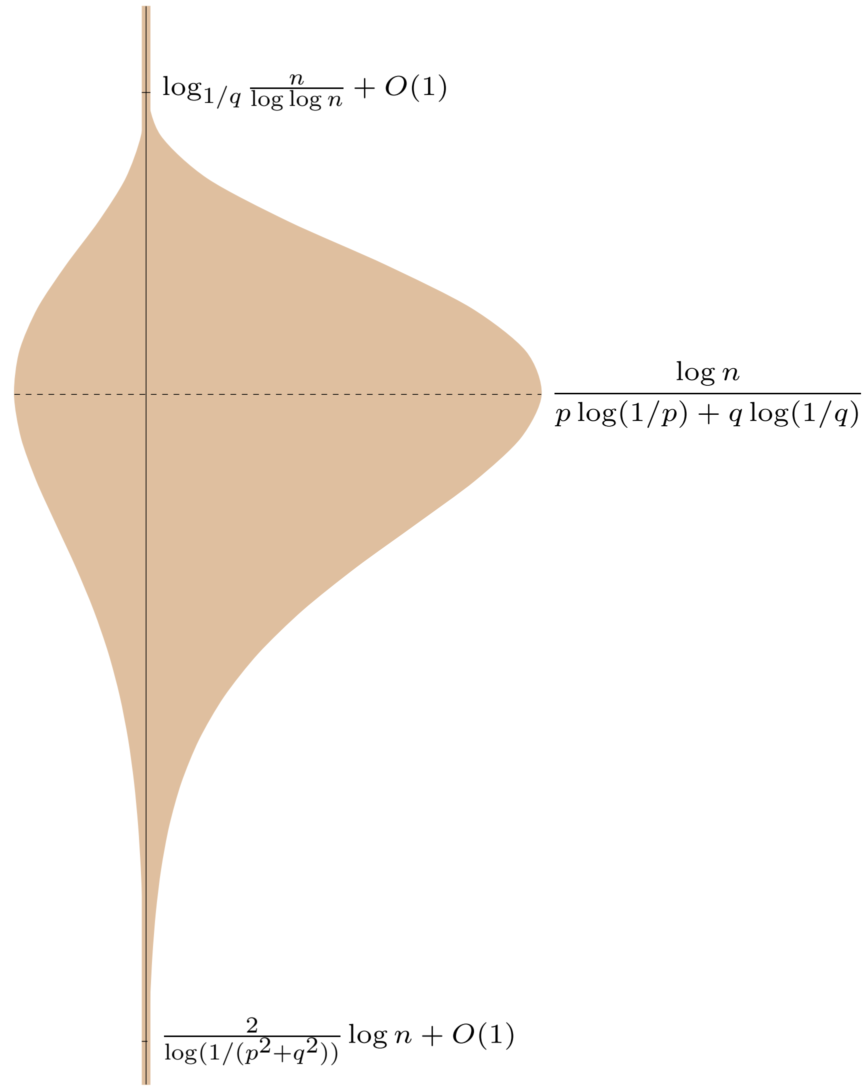
4: **Polynomial Growth/Decay**: $\alpha_2 \cdot \log n < k$: $(\alpha_2 < \alpha)$

$$\mathbb{E}[B_n^k] = \frac{2pq}{p^2+q^2}n^{\nu_2} + O\left(n^{\nu_3}\right)$$

where $\nu_2 = 2 + \alpha\log(p^2+q^2)$ for some $\nu_3 < \nu_2$.

External Shapes





$(p = 0.5, \alpha_1 = \alpha_2 = 1/\log 2)$ $(p = 0.75)$

Average Internal Profile

1: **Almost Full Tree**: $k < \alpha_1 \cdot \log n$

$$\mathbb{E}(I_n^k) = 2^k - \mathbb{E}(B_{n,k})(1 + o(1)).$$

2: **Phase Transition I**: $\alpha_1 \cdot \log n < k < \alpha_0 \cdot \log n$, where $\alpha_0 = \frac{2}{\log(1/p)+\log(1/q)}$

$$\mathbb{E}[I_n^k] = 2^k - G_2(\log n)\mathbb{E}(B_{n,k})(1 + o(1))$$

where $G_2(x)$ is a periodic function.

3: **Phase Transition II**: $\alpha_0 \cdot \log n < k < \alpha_2 \cdot \log n$

$$\mathbb{E}[I_n^k] = G_2(\log n)\mathbb{E}(B_{n,k})(1 + o(1))$$

where $G_2(x)$ is a periodic function.

4: **Polynomial Growth/Decay**: $\alpha_2 \cdot \log n < k$

$$\mathbb{E}[I_n^k] = \frac{1}{2}n^{\nu_2}(1 + o(1))$$

where $\nu_2 = 2 - \alpha \log(p^2 + q^2)$.

Variance and Limiting Distributions of the External Profile

Variance:

1: $k < \alpha_1 \cdot \log n$: $\mathbb{V}[B_n^k] \sim \mathbb{E}[B_n^k]$.

2: $\alpha_1 \cdot \log n < k < \alpha_2 \cdot \log n$: $\mathbb{V}[B_n^k] \sim G_3(\log n)\mathbb{E}[B_n^k]$.
where $G_3(\log n)$ is a periodic function.

3: $\alpha_2 \cdot \log n < k$: $\mathbb{V}[B_n^k] \sim 2\mathbb{E}[B_n^k]$.

Limiting Distributions:
Central Limit Theorem: For $\alpha_1 \cdot \log n < k < \alpha_3 \cdot \log n$:

$$\frac{B_n^k - \mathbb{E}[B_n^k]}{\sqrt{\mathbb{V}[B_n^k]}} \to N(0,1),$$

where $N(0,1)$ is the standard normal distribution.

Poisson Distribution: For $\alpha_3 \cdot \log n < k$:

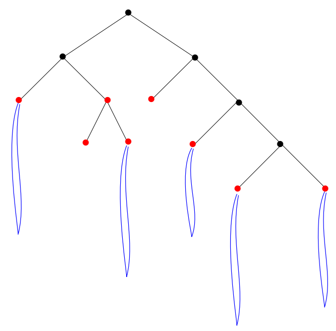
$$P\left(B_{n,k} = 2m\right) = \frac{{\lambda_0}^m}{m!} e^{-\lambda_0} + o(1), \quad \text{and} \quad P\left(B_{n,k} = 2m+1\right) = o(1),$$

where $\lambda_0 := pqn^2(p^2+q^2)^{k-1}$.

Consequences

Height: For large n (cf. Flajolet, 1980, Pittel, 1985, W.S., 1988, Devroye, 1992)

$$H_n = \frac{2}{\log(p^2+q^2)^{-1}} \log n = \alpha_3 \log n := k_H, \quad (\text{whp}).$$



Upper Bound: $P(H_n > (1+\epsilon)k_H) \leq P(B_n^k \geq 1) \leq \mathbb{E}[B_n^k] \to 0.$

Lower Bound: $P(H_n < (1-\epsilon)k_H) \leq P(B_n^{\lceil(1-\varepsilon)k_H\rceil)} = 0)$

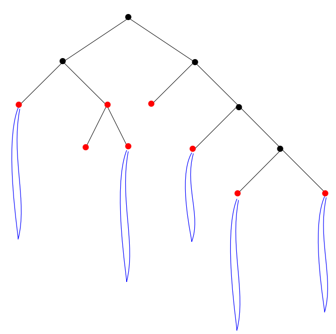
$$\leq \frac{\mathbb{V}[B_n^{\lceil(1-\epsilon)k_H\rceil}]}{(\mathbb{E}[B_n^{\lceil(1-\epsilon)k_H\rceil}])^2} = O\left(\frac{1}{\mathbb{E}[B_n^{\lceil(1-\epsilon)k_H\rceil}]}\right) \to 0.$$

Define: $k_S := \lfloor \frac{1}{\log q^{-1}} (\log n - \log\log\log n + \log(e \log r)) \rfloor.$

Consequences

Height: For large n (cf. Flajolet, 1980, Pittel, 1985, W.S., 1988, Devroye, 1992)

$$H_n = \frac{2}{\log(p^2+q^2)^{-1}} \log n = \alpha_3 \log n := k_H, \quad (\text{whp}).$$



Upper Bound: $P(H_n > (1+\epsilon)k_H) \le P(B_n^k \ge 1) \le \mathbb{E}[B_n^k] \to 0.$

Lower Bound: $P(H_n < (1-\epsilon)k_H) \le P(B_n^{\lceil (1-\varepsilon)k_H \rceil)} = 0)$

$$\le \frac{\mathbb{V}[B_n^{\lceil (1-\epsilon)k_H \rceil}]}{(\mathbb{E}[B_n^{\lceil (1-\epsilon)k_H \rceil}])^2} = O\left(\frac{1}{\mathbb{E}[B_n^{\lceil (1-\epsilon)k_H \rceil}]}\right) \to 0.$$

Define: $k_S := \lfloor \frac{1}{\log q^{-1}} (\log n - \log\log\log n + \log(e \log r)) \rfloor.$

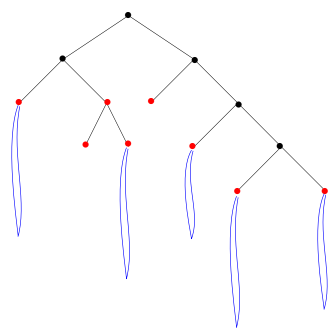
Shortest Path: For large n (cf. Knessl and W.S., 2005)

$$P(s_n = k_S \ \text{ or } \ s_n = k_S + 1) \to 1.$$

Consequences

Height: For large n (cf. Flajolet, 1980, Pittel, 1985, W.S., 1988, Devroye, 1992)

$$H_n = \frac{2}{\log(p^2+q^2)^{-1}} \log n = \alpha_3 \log n := k_H, \quad \text{(whp)}.$$



Upper Bound: $P(H_n > (1+\epsilon)k_H) \le P(B_n^k \ge 1) \le \mathbb{E}[B_n^k] \to 0.$

Lower Bound: $P(H_n < (1-\epsilon)k_H) \le P(B_n^{\lceil (1-\varepsilon)k_H \rceil}) = 0)$

$$\le \frac{\mathbb{V}[B_n^{\lceil (1-\epsilon)k_H \rceil}]}{(\mathbb{E}[B_n^{\lceil (1-\epsilon)k_H \rceil}])^2} = O\left(\frac{1}{\mathbb{E}[B_n^{\lceil (1-\epsilon)k_H \rceil}]}\right) \to 0.$$

Define: $k_S := \lfloor \frac{1}{\log q^{-1}}(\log n - \log\log\log n + \log(e \log r)) \rfloor.$

Shortest Path: For large n (cf. Knessl and W.S., 2005)

$$P(s_n = k_S \text{ or } s_n = k_S + 1) \to 1.$$

Fill-up: For large n (cf. Pittel, 1986, Devroye, 1992, Knessl & W.S., 2005)

$$P(F_n = k_S - 1 \text{ or } F_n = k_S) \to 1.$$

Sketch of the Proof

1. **Recurrence**: $\mathbb{E}[B_n^k] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^{k-1}] + \mathbb{E}[B_{n-i}^{k-1}])$, $n \geq 2, k \geq 1.$

Sketch of the Proof

1. **Recurrence**: $\mathbb{E}[B_n^k] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^{k-1}] + \mathbb{E}[B_{n-i}^{k-1}])$, $n \geq 2, k \geq 1$.

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}_k(z) = \tilde{E}_{k-1}(zp) + \tilde{E}_{k-1}(zq), \;\; k \geq 2,$$

Sketch of the Proof

1. **Recurrence**: $\mathbb{E}[B_n^k] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^{k-1}] + \mathbb{E}[B_{n-i}^{k-1}]),\, n \geq 2, k \geq 1.$

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}_k(z) = \tilde{E}_{k-1}(zp) + \tilde{E}_{k-1}(zq), \;\; k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = (p^{-s} + q^{-s}) \tilde{E}_{k-1}^*(s)$:

$$\tilde{E}_k^*(s) = (p^{-s} + q^{-s})^{k-1} \cdot s \cdot (p^{-s} + q^{-s} - 1)\Gamma(s)$$

for $\Re(s) \in (-2, \infty)$, where $\Gamma(s)$ is the Euler Gamma function.

Sketch of the Proof

1. **Recurrence**: $\mathbb{E}[B_n^k] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^{k-1}] + \mathbb{E}[B_{n-i}^{k-1}])$, $n \geq 2$, $k \geq 1$.

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}_k(z) = \tilde{E}_{k-1}(zp) + \tilde{E}_{k-1}(zq), \ \ k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = (p^{-s} + q^{-s}) \tilde{E}_{k-1}^*(s)$:

$$\tilde{E}_k^*(s) = (p^{-s} + q^{-s})^{k-1} \cdot s \cdot (p^{-s} + q^{-s} - 1)\Gamma(s)$$

for $\Re(s) \in (-2, \infty)$, where $\Gamma(s)$ is the Euler Gamma function.

4. **Inverse Mellin Transform**: $\tilde{E}_k(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} z^{-s} \tilde{E}_k^*(s) ds$:

$$\tilde{E}_k(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} s(p^{-s} + q^{-s} - 1)\Gamma(s) z^{-s} (p^{-s} + q^{-s})^{k-1} ds$$

through the **saddle point method**.

Sketch of the Proof

1. **Recurrence**: $\mathbb{E}[B_n^k] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^{k-1}] + \mathbb{E}[B_{n-i}^{k-1}])$, $n \geq 2, k \geq 1$.

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}_k(z) = \tilde{E}_{k-1}(zp) + \tilde{E}_{k-1}(zq), \;\; k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = (p^{-s} + q^{-s}) \tilde{E}_{k-1}^*(s)$:

$$\tilde{E}_k^*(s) = (p^{-s} + q^{-s})^{k-1} \cdot s \cdot (p^{-s} + q^{-s} - 1)\Gamma(s)$$

for $\Re(s) \in (-2, \infty)$, where $\Gamma(s)$ is the Euler Gamma function.

4. **Inverse Mellin Transform**: $\tilde{E}_k(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} z^{-s} \tilde{E}_k^*(s) ds$:

$$\tilde{E}_k(z) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} s(p^{-s} + q^{-s} - 1)\Gamma(s) z^{-s} (p^{-s} + q^{-s})^{k-1} ds$$

through the **saddle point method**.

5. **Depoissonization**: From the Poisson transform $\tilde{E}_k(z)$ to $\mathbb{E}[B_n^k]$.

Saddle Point Method: Phase Transitions

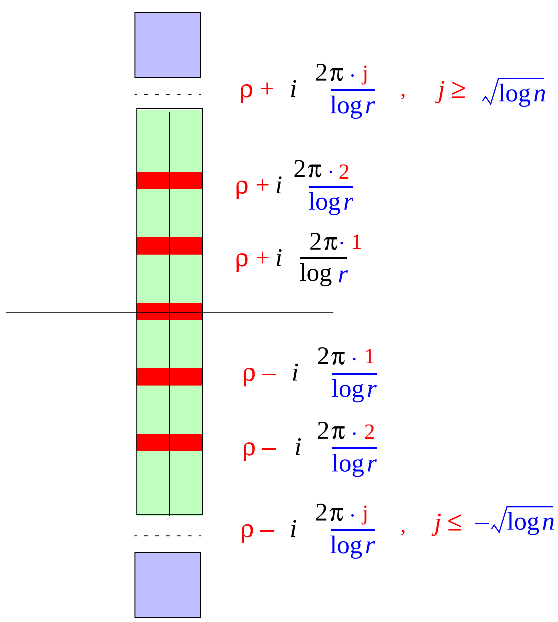
By depoisonization we have $\tilde{E}_k(n) \sim \tilde{E}_k(z)$, where recall

$$\begin{aligned}\tilde{E}_k(n) &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} g(s)\Gamma(s+1)n^{-s}(p^{-s}+q^{-s})^k ds \\ &= \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} g(s)\Gamma(s+1)\exp(h(s)\log n)ds, \qquad k=\alpha\log n.\end{aligned}$$

The saddle point equation $h'(s)=0$ has a unique real root:

$$\rho = \frac{-1}{\log r}\log\left(\frac{\alpha\log q^{-1}-1}{1-\alpha\log p^{-1}}\right), \quad \frac{1}{\log q^{-1}} < \alpha < \frac{1}{\log p^{-1}}.$$

There are **infinitely many** saddle points $\rho + it_j$ for $t_j = 2\pi j/\log r$, $j \in \mathbb{Z}$.



Phase Transitions:

1. $\rho \to \infty$ as $\alpha \downarrow 1/\log q^{-1} = \alpha_1$.

2 $\rho \to -\infty$ when $\alpha \uparrow 1/\log p^{-1}$.

3. **Saddle points coalesce** with poles of the $\Gamma(s+1)$ function at $s=-2,-3,\ldots$.
Pole $s=-2$ leads to α_2.

Depoissonization

Theorem 1. ***(Jacquet and W.S., 1998)*** *Let $\tilde{G}(z)$ be the Poisson transform of a sequence g_n, that is,*

$$\tilde{G}(z) = \sum_n g_n \frac{z}{n!} e^{-z}.$$

$\tilde{G}(z)$ is assumed to be an entire function of z.
Two conditions simultaneously to hold for a cone S_θ:
(I) *For $z \in S_\theta$ and some reals $B, R > 0, \nu$*

$$|z| > R \Rightarrow |\tilde{G}(z)| \leq B|z|^\nu \Psi(|z|),$$

where $\Psi(x)$ is a slowly varying function.
(O) *For $z \notin S_\theta$ and $A, \alpha < 1$*

$$|z| > R \Rightarrow |\tilde{G}(z)e^z| \leq A\exp(\alpha|z|).$$

Then

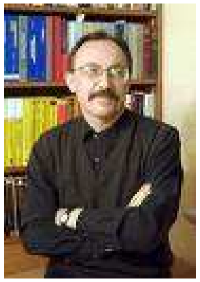
$$g_n = \tilde{G}(n) + O(n^{\nu-1}\Psi(n)).$$

Back to Profile:

Using the depoissonization theorem, we find $\mathbb{E}[B_n^k] = \tilde{E}_k(n) + O\left(\frac{n^{\nu-1}}{\sqrt{\log n}}\right)$.

Outline Update

1. Tries and Suffix Trees
2. Flajolet and Tries – A Look Back
3. Profiles of Tries
4. Profile of Digital Search Trees (announcement)



Profile of Digital Search Trees

1. **Recurrence**: $\mathbb{E}[B_{n+1}^{k+1}] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^k] + \mathbb{E}[B_{n-i}^k]),\ n \geq 2,\ k \geq 0.$

Profile of Digital Search Trees

1. **Recurrence**: $\mathbb{E}[B_{n+1}^{k+1}] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^k] + \mathbb{E}[B_{n-i}^k])$, $n \geq 2, k \geq 0$.

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}'_{k+1}(z) + \tilde{E}_{k+1}(z) = \tilde{E}_k(zp) + \tilde{E}_k(zq), \;\; k \geq 2,$$

Profile of Digital Search Trees

1. **Recurrence**: $\mathbb{E}[B_{n+1}^{k+1}] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^k] + \mathbb{E}[B_{n-i}^k])$, $n \geq 2, k \geq 0$.

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}'_{k+1}(z) + \tilde{E}_{k+1}(z) = \tilde{E}_k(zp) + \tilde{E}_k(zq), \;\; k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = -\Gamma(s) F_k(s)$:

$$F_{k+1}^*(s) - F_{k+1}^*(s-1) = (p^{-s} + q^{-s}) \cdot F_k^*(s)$$

for $\Re(s) \in (-k-1, 0)$, and $F_0^*(s) = 1$.

Profile of Digital Search Trees

1. **Recurrence**: $\mathbb{E}[B_{n+1}^{k+1}] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^k] + \mathbb{E}[B_{n-i}^k])$, $n \geq 2, k \geq 0$.

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}'_{k+1}(z) + \tilde{E}_{k+1}(z) = \tilde{E}_k(zp) + \tilde{E}_k(zq), \;\; k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = -\Gamma(s) F_k(s)$:

$$F_{k+1}^*(s) - F_{k+1}^*(s-1) = (p^{-s} + q^{-s}) \cdot F_k^*(s)$$

for $\Re(s) \in (-k-1, 0)$, and $F_0^*(s) = 1$.

4. **The power Series**: $f(x, s) = \sum_{k \geq 0} F_k(s) x^s$ becomes

$$f(x, s) = \frac{g(x, s)}{g(x, -1)},$$

where $g(x, s) = 1 + x \sum_{j \geq 0} g(x, s-j)(p^{-s+j} + q^{-s+j})$, and asymptotically

$$F_k^*(s) \sim f(s)(p^{-s} + q^{-s})^k.$$

where $f(s)$ is analytic with $f(-r) = 0$ for $r = 1, 2 \ldots$.

Profile of Digital Search Trees

1. **Recurrence**: $\mathbb{E}[B_{n+1}^{k+1}] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^k] + \mathbb{E}[B_{n-i}^k])$, $n \geq 2, k \geq 0.$

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}'_{k+1}(z) + \tilde{E}_{k+1}(z) = \tilde{E}_k(zp) + \tilde{E}_k(zq), \;\; k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = -\Gamma(s) F_k(s)$:

$$F_{k+1}^*(s) - F_{k+1}^*(s-1) = (p^{-s} + q^{-s}) \cdot F_k^*(s)$$

for $\Re(s) \in (-k-1, 0)$, and $F_0^*(s) = 1.$

4. **The power Series**: $f(x, s) = \sum_{k \geq 0} F_k(s) x^s$ becomes

$$f(x, s) = \frac{g(x, s)}{g(x, -1)},$$

where $g(x, s) = 1 + x \sum_{j \geq 0} g(x, s-j)(p^{-s+j} + q^{-s+j})$, and asymptotically

$$F_k^*(s) \sim f(s)(p^{-s} + q^{-s})^k.$$

where $f(s)$ is analytic with $f(-r) = 0$ for $r = 1, 2 \ldots.$

5. **Inverse Mellin Transform** by **saddle point** and **depoissonization**.

Profile of Digital Search Trees

1. **Recurrence**: $\mathbb{E}[B_{n+1}^{k+1}] = \sum_{i=0}^{n} \binom{n}{i} p^i q^{n-i} (\mathbb{E}[B_i^k] + \mathbb{E}[B_{n-i}^k]),\ n \geq 2, k \geq 0.$

2. **Poisson Transform**: $\tilde{E}_k(z) = \sum_{n=0}^{\infty} \mathbb{E}[B_n^k] \frac{z^n}{n!} e^{-z}$:

$$\tilde{E}'_{k+1}(z) + \tilde{E}_{k+1}(z) = \tilde{E}_k(zp) + \tilde{E}_k(zq),\ \ k \geq 2,$$

3. **Mellin Transform**: $\tilde{E}_k^*(s) := \int_0^\infty z^{s-1} \tilde{E}_k(z) dz = -\Gamma(s) F_k(s)$:

$$F_{k+1}^*(s) - F_{k+1}^*(s-1) = (p^{-s} + q^{-s}) \cdot F_k^*(s)$$

for $\Re(s) \in (-k-1, 0)$, and $F_0^*(s) = 1$.

4. **The power Series**: $f(x, s) = \sum_{k \geq 0} F_k(s) x^s$ becomes

$$f(x, s) = \frac{g(x, s)}{g(x, -1)},$$

where $g(x, s) = 1 + x \sum_{j \geq 0} g(x, s-j)(p^{-s+j} + q^{-s+j})$, and asymptotically

$$F_k^*(s) \sim f(s)(p^{-s} + q^{-s})^k.$$

where $f(s)$ is analytic with $f(-r) = 0$ for $r = 1, 2 \ldots$.

5. **Inverse Mellin Transform** by **saddle point** and **depoissonization**.

6. Asymptotically DST profile behaves as the profile of tries.

Analysis of Algorithms (AofA): Analytic Algorithmics

- **Analysis of Algorithms** is concerned with **precise** estimates of complexity parameters of algorithms and aims at predicting algorithms' behaviour. It develops general methods for obtaining closed-form formulae, asymptotic estimates, and probability distributions for combinatorial or probabilistic quantities. Properties of discrete structures such as strings, trees, tries, dags, graphs are investigated.

- The area of analysis of algorithms was born on July 27, 1963, when D. E. Knuth wrote his "Notes on Open Addressing".

- Following **Hadamard's precept**[1], we study algorithmic problems using techniques of **complex analysis** such as generating functions, combinatorial calculus, Rice's formula, Mellin transform, Fourier series, sequences distributed modulo 1, saddle point methods, analytic poissonization and depoissonization, and singularity analysis.

- This program, which applies complex-analytic tools to analysis of algorithm, constitutes **analytic algorithmics**.

[1] The shortest path between two truths on the real line passes through the complex plane.

Thank you, Philippe . . .

. . . for long standing support, friendship, and sharing your knowledge!

Merci au bon docteur Flajolet!