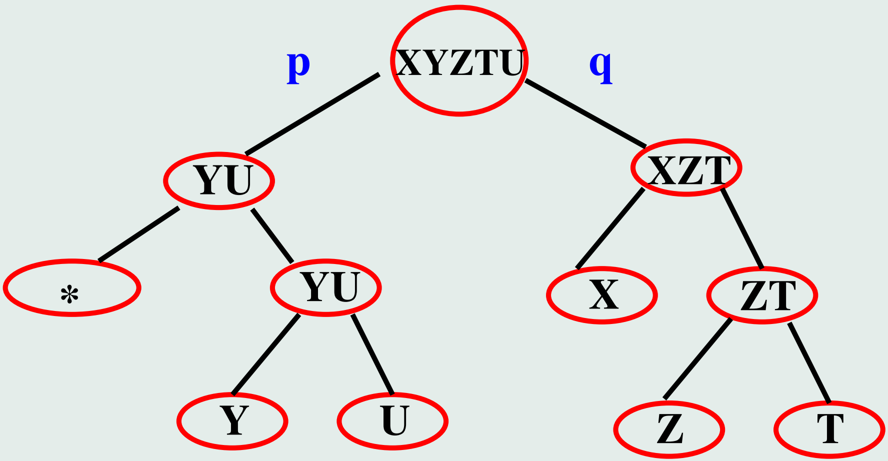
Probabilistic Methods for Algorithms and Stochastic Networks

Lecture 3

Philippe Robert
INRIA Paris-Rocquencourt

Previous Episode



p
XYZTU
q
YU
XZT
*
YU
X
ZT
Y
U
Z
T

Non-Symmetrical Case

n items at the root:
$\mathbb{E}(R_n)$ average size of tree,

$$\phi_R(x) = \sum_{n=0}^{+\infty} \mathbb{E}(R_n) \frac{x^n}{n!} e^{-x}.$$

Non-Symmetrical Case

n items at the root:
$\mathbb{E}(R_n)$ average size of tree,

$$\phi_R(x) = \sum_{n=0}^{+\infty} \mathbb{E}(R_n) \frac{x^n}{n!} e^{-x}.$$

If $\psi(x) = \phi_R(x) - 1$,

$$\psi(x) = \psi(px) + \psi(qx) + 2\mathbb{P}(t_2 \leq x).$$

t_2 second point of Poisson process with rate 1.

The Renewal Theorem

Some History

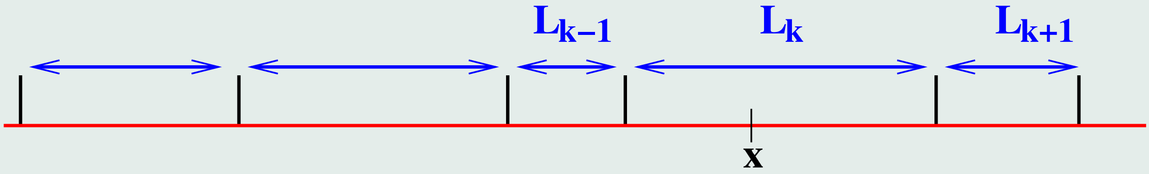
Blackwell (1948):

— A light bulb lasts two years in average.

— How many are necessary for ten years ?

Doob, Feller, Blackwell, . . .

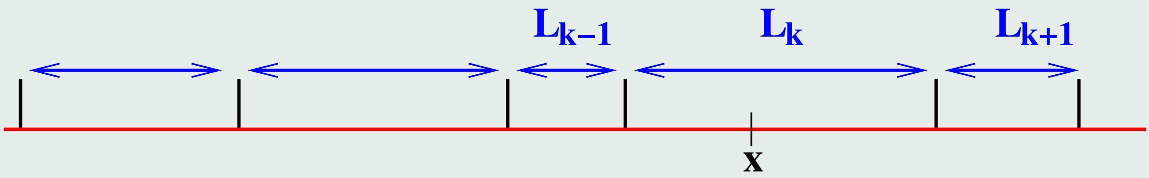
Random Walk (S_n) Associated to (L_n)



(L_n) i.i.d. non-negative random variables,

$$S_n = \sum_{i=1}^{n} L_i.$$

Random Walk (S_n) Associated to (L_n)



(L_n) i.i.d. non-negative random variables,

$$S_n = \sum_{i=1}^{n} L_i.$$

Does the landscape around x converge in dist. when $x \to +\infty$?

The Renewal Theorem

Potential Functional

f continuous compact support on $\mathbb{R}$,

$$G(f) \stackrel{\text{def.}}{=} \mathbb{E}\left(\sum_{n=0}^{+\infty} f(S_n)\right).$$

The Renewal Theorem

Translated Potential Functional

f continuous compact support on $\mathbb{R}$,

$$T_x G(f) \stackrel{\text{def.}}{=} \mathbb{E}\left(\sum_{n=0}^{+\infty} f(S_n - x)\right).$$

The Renewal Theorem

Translated Potential Functional

f continuous compact support on $\mathbb{R}$,

$$T_x G(f) \stackrel{\text{def.}}{=} \mathbb{E}\left(\sum_{n=0}^{+\infty} f(S_n - x)\right).$$

Theorem. If L_1 is non-lattice, then

$$\lim_{x\to+\infty} T_x G(f) = \frac{1}{\mathbb{E}(L_1)} \int_{\mathbb{R}_+} f(u)\, du.$$

Non-Lattice: $\mathbb{P}(L_1 \in \delta\mathbb{N}) < 1, \forall\delta > 0.$

Consequences

If L_1 is non-lattice and $M(A) = G(\mathbb{1}_{\{A\}})$,

$$\lim_{x \to +\infty} M[0, x]/x = \frac{1}{\mathbb{E}(L_1)}, a.s.$$

Consequences

If L_1 is non-lattice and $M(A) = G(\mathbb{1}_{\{A\}})$,

$$\lim_{x\to+\infty} M[0,x]/x = \frac{1}{\mathbb{E}(L_1)}, a.s.$$

$$\lim_{x\to+\infty} \mathbb{E}(M[x,x+a]) = \frac{|a|}{\mathbb{E}(L_1)}$$

Consequences

If L_1 is non-lattice and $M(A) = G(\mathbb{1}_{\{A\}})$,

$$\lim_{x \to +\infty} M[0, x]/x = \frac{1}{\mathbb{E}(L_1)}, a.s.$$

$$\lim_{x \to +\infty} \mathbb{E}(M[x, x+a]) = \frac{|a|}{\mathbb{E}(L_1)}$$

If F_x first element S_n after x,
then $F_x - x$ cv in dist. to F

$$\mathbb{P}(F \in A) = \frac{1}{\mathbb{E}(L_1)} \int_A \mathbb{P}(L_1 \geq x)\, dx, \qquad A \in \mathcal{B}(\mathbb{R}).$$

Lattice Case

If $\mathbb{P}(L_1 \in \delta\mathbb{N}) = 1$ **and** $\mathbb{P}(L_1 = \delta) > 0$**:**

Lattice Case

If $\mathbb{P}(L_1 \in \delta\mathbb{N}) = 1$ and $\mathbb{P}(L_1 = \delta) > 0$:

For $h > 0$,

$$\lim_{n\to+\infty} T_{h+n\delta} G(f) = \mathcal{L}_h(f) = \frac{1}{\mathbb{E}(L_1)} \sum_{j\in\mathbb{Z}} f(h+j\delta)$$

Lattice Case

If $\mathbb{P}(L_1 \in \delta\mathbb{N}) = 1$ and $\mathbb{P}(L_1 = \delta) > 0$:

For $h > 0$,

$$\lim_{n \to +\infty} T_{h+n\delta} G(f) = \mathcal{L}_h(f) = \frac{1}{\mathbb{E}(L_1)} \sum_{j \in \mathbb{Z}} f(h+j\delta)$$

Periodic Behavior: $\mathcal{L}_h(f) = \mathcal{L}_{h+\delta}(f)$.

Back to Tree Algorithm

Poisson Transform

$$\psi(x) = \phi_R(x) - 1,$$

$$\psi(x) = \psi(px) + \psi(qx) + 2\mathbb{P}(t_2 < x)$$

Poisson Transform

$\psi(x) = \phi_R(x) - 1,$

$$\psi(x) = \psi(px) + \psi(qx) + 2\mathbb{P}(t_2 < x)$$

Take $W_1 \in \{p, q\}$ random variable
such that $\mathbb{P}(W_1 = p) = p$ and $\mathbb{P}(W_1 = q) = q$

Poisson Transform

$$\psi(x) = \phi_R(x) - 1,$$

$$\psi(x) = \psi(px) + \psi(qx) + 2\mathbb{P}(t_2 < x)$$

Take $W_1 \in \{p, q\}$ random variable
such that $\mathbb{P}(W_1 = p) = p$ and $\mathbb{P}(W_1 = q) = q$

$$\psi(x) = \mathbb{E}\left(\frac{\psi(W_1 x)}{W_1}\right) + 2\mathbb{E}\left(\mathbb{1}_{\{t_2 < x\}}\right)$$

Poisson Transform

$$\psi(x) = \mathbb{E}\left(\frac{\psi(W_1 x)}{W_1}\right) + 2\mathbb{E}\left(\mathbb{1}_{\{t_2 < x\}}\right)$$

Poisson Transform

$$\psi(x) = \mathbb{E}\left(\frac{\psi(W_1 x)}{W_1}\right) + 2\mathbb{E}\left(\mathbb{1}_{\{t_2 < x\}}\right)$$

$$\psi(x) = \mathbb{E}\left(\frac{\psi(W_2 W_1 x)}{W_2 W_1}\right) + 2\mathbb{E}\left(\frac{1}{W_1}\mathbb{1}_{\{t_2 < x W_1\}}\right) + 2\mathbb{E}\left(\mathbb{1}_{\{t_2 < x\}}\right)$$

Poisson Transform

$$\psi(x) = \mathbb{E}\left(\frac{\psi(W_1 x)}{W_1}\right) + 2\mathbb{E}\left(\mathbb{1}_{\{t_2 < x\}}\right)$$

$$\psi(x) = 2\sum_{k=0}^{+\infty} \mathbb{E}\left(\frac{1}{\prod_1^k W_i} \mathbb{1}_{\{t_2 < x \prod_1^k W_i\}}\right)$$

(W_i) i.i.d. distributed as W_1.

Poisson Transform (III)

$$\mathbb{E}\left(R_{N[0,x]}\right) = \sum_{n=0}^{+\infty} \mathbb{E}(R_n)\frac{x^n}{n!}\, e^{-x}$$

$$= 1 + 2\sum_{k=0}^{+\infty} \mathbb{E}\left(\frac{1}{\prod_1^k W_i} \mathbb{1}_{\{t_2 < x \prod_1^k W_i\}}\right)$$

An Associated Random Walk

$U_{2,n}$: 2th min. of n ind. uniform r.v. on $[0,1]$

$$\mathbb{E}(R_n) = 1 + 2\mathbb{E}\left(\sum_{k\geq 0} \frac{1}{\prod_1^k W_i} \mathbb{1}_{\{U_{2,n} < \prod_1^k W_i\}}\right), \; n \geq 2.$$

An Associated Random Walk

$U_{2,n}$: 2th min. of n ind. uniform r.v. on $[0,1]$

$$\mathbb{E}(R_n) = 1+2\mathbb{E}\left(\sum_{k\geq 0}\frac{1}{\prod_1^k W_i}\mathbb{1}_{\{U_{2,n}<\prod_1^k W_i\}}\right), \; n\geq 2.$$

$$\Delta_n \overset{\text{def.}}{=} \frac{\mathbb{E}(R_n)-1}{2n}$$

$$=\mathbb{E}\left(\sum_{k} e^{\sum_{i=1}^{k}-\log(W_i)-\log n}\;\mathbb{1}_{\{-\sum_1^k \log(W_i)\leq -\log U_{2,n}\}}\right)$$

Renewal Theorem Context

$$L_i = -\log W_i$$

$$\Delta_n = \mathbb{E}\left(\sum_k e^{-\left(\log n - \sum_{i=1}^k L_i\right)} \mathbb{1}_{\{\sum_1^k L_i \le -\log U_{2,n}\}}\right)$$

Renewal Theorem Context

$$L_i = -\log W_i \qquad U_{2,n} \backsim t_2/n$$

$$\Delta_n = \mathbb{E}\left(\sum_k e^{-\left(\log n - \sum_{i=1}^k L_i\right)} \mathbb{1}_{\{\sum_1^k L_i \leq -\log U_{2,n}\}}\right)$$

$$\Delta_n \sim \mathbb{E}\left(\sum_k \frac{1}{t_2} e^{-\left(\log(n/t_2) - \sum_{i=1}^k L_i\right)} \mathbb{1}_{\{\sum_1^k L_i \leq \log(n/t_2)\}}\right)$$

The Use of Renewal Theorem

If $L_1 = -\log W_1$

$$\boldsymbol{\Delta}_n \sim \mathbb{E}\left(\sum_k \frac{1}{t_2} e^{-\left(\log(n/t_2) - \sum_{i=1}^k L_i\right)} \mathbf{1}_{\left\{\sum_1^k L_i \le \log(n/t_2)\right\}}\right)$$

The Use of Renewal Theorem

If $L_1 = -\log W_1$ **non-lattice**

$$\Delta_n \sim \mathbb{E}\left(\sum_k \frac{1}{t_2} e^{-\left(\log(n/t_2) - \sum_{i=1}^k L_i\right)} 1_{\left\{\sum_1^k L_i \le \log(n/t_2)\right\}}\right)$$

$$\longrightarrow \mathbb{E}\left(\frac{1}{t_2}\right) \frac{1}{\mathbb{E}(L_1)} \int_0^{+\infty} e^{-u}\, du$$

The Use of Renewal Theorem

$L_1 = -\log W_1$ non-lattice if $\log p / \log q \not\in \mathbb{Q}$,

$$\lim_{n \to +\infty} \frac{\mathbb{E}(R_n)}{n} = \frac{2}{H}$$

H entropy of W_1:

$H = \mathbb{E}(L_1) = -\mathbb{E}(\log W_1) = -p \log p - q \log q.$

The Use of Renewal Theorem

$L_1 = -\log W_1$ lattice if $\log p / \log q = m/n$.
Set $\lambda = \log(p)/m$,

$$\frac{\mathbb{E}(R_n)}{n} \sim F\left(\frac{\log n}{\lambda}\right)$$

$$F(x) = \frac{2\lambda}{H(1-e^{-\lambda})} \int_0^{+\infty} e^{-\lambda\left\{x - \frac{\log y}{\lambda}\right\}} e^{-y}\, dy$$

and $\{x\} = x - \lfloor x \rfloor$.

Renewal Theorems

— Distribution of $-\log W$ non-lattice:
$\log p/\log q \notin \mathbb{Q}$.

— Dist. $-\log W$ lattice
$\log p/\log q \in \mathbb{Q}$.

Renewal Theorems

— Distribution of $-\log W$ non-lattice:
$\log p/\log q \notin \mathbb{Q}$.
Continuous Renewal Theorem.
Convergence of $\mathbb{E}(R_n)/n$.

— Dist. $-\log W$ lattice
$\log p/\log q \in \mathbb{Q}$.
Discrete Renewal Theorem.
Periodic Fluctuations of $\mathbb{E}(R_n)/n$.

Ingredients of the Analysis

— Probabilistic de-Poissonnization

— Fubini's Theorem.

— Renewal Theorem

Ingredients of the Analysis

— Probabilistic de-Poissonnization

— Fubini's Theorem.

— Renewal Theorem

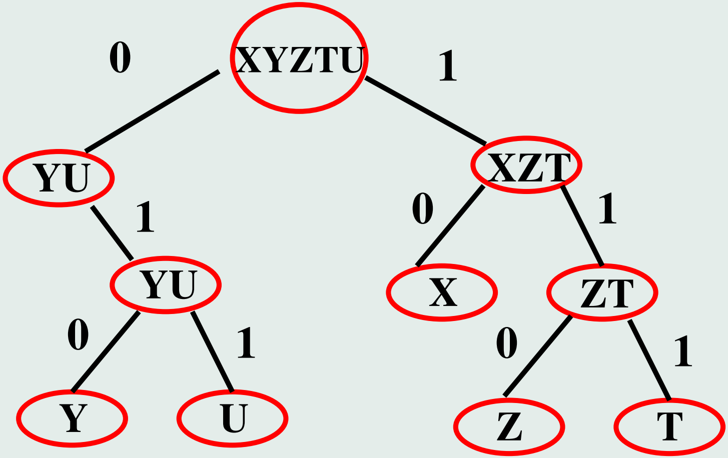
Works in a general branching setting.

An Alternative to Recursion

Initial Recursion

If $n \geq 2$, $n = n_1 + n_2$,

$$R_n = 1 + R_{n_1} + R_{n_2}$$



$\boldsymbol{X}$=10. . ., $\boldsymbol{Y}$=010 . . . , $\boldsymbol{Z}$=110 . . . ,
$\boldsymbol{T}$=111 . . . , $\boldsymbol{U}$=011 . . .

Nodes as Words

— An Alphabet $\mathcal{A} = \{0, 1\}$

— Set of finite and infinite words

$$\mathcal{M}^* = \bigcup_{n \geq 0} \mathcal{A}^n, \quad \mathcal{M} = \mathcal{A}^{\mathbb{N}}$$

— $x \in \mathcal{M}^*$, $|x|$ length of x

— $\pi_k : \mathcal{M} \to \mathcal{M}^*$ proj. on k first coordinates

— If $|x| = k$,
P_x proba that first k letters are $x_1, x_2, \ldots, x_k$.

A Direct Calculation

$$\mathbb{E}(R_n) = 1+2 \sum_{y \in \mathcal{M}^*} 1-(1-P_y)^n - nP_y(1-P_y)^{n-1}$$

A Direct Calculation

$$\mathbb{E}(R_n) = 1 + 2 \sum_{y \in \mathcal{M}^*} 1 - (1 - P_y)^n - n P_y (1 - P_y)^{n-1}$$

Valid even if P_y is not a product probability.

A Direct Calculation

$$\mathbb{E}(R_n) = 1+2 \sum_{y \in \mathcal{M}^*} 1-(1-P_y)^n - nP_y(1-P_y)^{n-1}$$

Valid even if P_y is not a product probability.

$$\mathbb{E}(R_n) = 1 + 2 \sum_{y \in \mathcal{M}^*} \mathbb{P}(U_{2,n} \leq P_y)$$

A Direct Calculation

$$\mathbb{E}(R_n) = 1+2 \sum_{y\in\mathcal{M}^*} 1-(1-P_y)^n-nP_y(1-P_y)^{n-1}$$

Valid even if P_y is not a product probability.

$$\mathbb{E}(R_n) = 1 + 2 \sum_{y\in\mathcal{M}^*} \mathbb{P}(U_{2,n} \leq P_y)$$

$nU_{2,n} \sim t_2$,

$$\mathbb{E}(R_n) \sim 1 + 2 \sum_{y\in\mathcal{M}^*} \mathbb{P}(t_2 \leq nP_y)$$

A Direct Calculation

$$\mathbb{E}(R_n) \sim 1 + 2 \sum_{k=0}^{+\infty} \sum_{\substack{y \in \mathcal{M}^* \\ |y|=k}} \frac{\mathbb{P}(t_2 \leq nP_y)}{P_y} P_y$$

A Direct Calculation

$$\mathbb{E}(R_n) \sim 1 + 2 \sum_{k=0}^{+\infty} \sum_{\substack{y \in \mathcal{M}^* \\ |y|=k}} \frac{\mathbb{P}(t_2 \leq nP_y)}{P_y} P_y$$

X **infinite word with distribution** $(P_y, y \in \mathcal{M}^*)$,

$$\mathbb{E}(R_n) \sim 1 + 2 \sum_{k=0}^{+\infty} \mathbb{E}\left(\frac{\mathbb{P}(t_2 \leq nP_{\pi_k(X)})}{P_{\pi_k(X)}} \right)$$

A Direct Calculation

$$\mathbb{E}(R_n) \sim 1 + 2 \sum_{k=0}^{+\infty} \sum_{\substack{y \in \mathcal{M}^* \\ |y|=k}} \frac{\mathbb{P}(t_2 \leq nP_y)}{P_y} P_y$$

X **infinite word with distribution** $(P_y, y \in \mathcal{M}^*)$,

$$\mathbb{E}(R_n) \sim 1 + 2 \sum_{k=0}^{+\infty} \mathbb{E}\left(\frac{\mathbb{P}(t_2 \leq nP_{\pi_k(X)})}{P_{\pi_k(X)}} \right)$$

$$\mathbb{E}(R_n) \sim 1 + 2\mathbb{E}\left(\sum_{k=0}^{+\infty} \frac{1}{P_{\pi_k(X)}} \mathbb{1}_{\{t_2 \leq nP_{\pi_k(X)}\}} \right)$$

Non-Additive quantities: Insertion Cost

n items stored:
nb I_n of operations to store the $(n+1)$th ?

X : the infinite word for the $(n+1)$th item.

Insertion Cost

$$\mathbb{P}(I_n \leq k) = \mathbb{E}\left(\left(1 - P_{\pi_k(X)}\right)^n\right)$$

Insertion Cost

$$\mathbb{P}(I_n \leq k) = \mathbb{E}\left((1 - P_{\pi_k(X)})^n\right)$$

$k \sim \log(n)/H$, H entropy.

Insertion Cost

$$\mathbb{P}(I_n \leq k) = \mathbb{E}\left((1 - P_{\pi_k(X)})^n\right)$$

$k \sim \log(n)/H$, H entropy.

$$\mathbb{E}(I_n) = \sum_{k=0}^{+\infty} \mathbb{E}\left(1 - (1 - P_{\pi_k(X)})^n\right)$$
$$= \sum_{k=0}^{+\infty} \mathbb{P}(U_{1,n} \leq P_{\pi_k(X)}).$$

$U_{1,n}$ min of n uniform r.v. on $[0, 1]$.

Insertion Cost

$$\mathbb{P}(I_n \leq k) = \mathbb{E}\left((1 - P_{\pi_k(X)})^n\right)$$

$k \sim \log(n)/H$, H entropy.

$$\mathbb{E}(I_n) = \sum_{k=0}^{+\infty} \mathbb{E}\left(1 - (1 - P_{\pi_k(X)})^n\right)$$

$$= \sum_{k=0}^{+\infty} \mathbb{P}(U_{1,n} \leq P_{\pi_k(X)}).$$

$U_{1,n}$ min of n uniform r.v. on $[0, 1]$.

$nU_{1,n} \sim E_1$+ Renewal Th. give asymptotics.

Exercise (Erdös et al. (1987))

Define $a_0 = 1$, $a_n = a_{\lfloor n/2 \rfloor} + a_{\lfloor n/3 \rfloor} + a_{\lfloor n/6 \rfloor}$, then

$$\lim_{n \to +\infty} \frac{a_n}{n} = \frac{12}{\log 432}.$$

Exercise (Erdös et al. (1987))

Define $a_0 = 1$, $a_n = a_{\lfloor n/2 \rfloor} + a_{\lfloor n/3 \rfloor} + a_{\lfloor n/6 \rfloor}$, **then**

$$\lim_{n \to +\infty} \frac{a_n}{n} = \frac{12}{\log 432}.$$

General: If $m_i \geq 2$, $r_i \in \mathbb{R}_+$, $i = 1, \ldots, d$, **and** $a_0 = 1$,

$$a_n = r_1 a_{\lfloor n/m_1 \rfloor} + r_2 a_{\lfloor n/m_2 \rfloor} + \cdots + r_d a_{\lfloor n/m_d \rfloor}$$

Exercise (Erdös et al. (1987))

Define $a_0 = 1$, $a_n = a_{\lfloor n/2 \rfloor} + a_{\lfloor n/3 \rfloor} + a_{\lfloor n/6 \rfloor}$, then

$$\lim_{n \to +\infty} \frac{a_n}{n} = \frac{12}{\log 432}.$$

General: If $m_i \geq 2$, $r_i \in \mathbb{R}_+$, $i = 1, \ldots, d$, and $a_0 = 1$,

$$a_n = r_1 a_{\lfloor n/m_1 \rfloor} + r_2 a_{\lfloor n/m_2 \rfloor} + \cdots + r_d a_{\lfloor n/m_d \rfloor}$$

then $\exists \nu \in \mathbb{R}$,

$$\lim_{n \to +\infty} \frac{a_n}{n^\nu} = \alpha \in (0, +\infty).$$