Concentration bounds for RIC^2 (Randomized Incremental Construction)

Sandeep Sen

IIT Delhi, India

Feb 8, 2019

[2]to appear in STACS'19

Once upon a time ..

Conceived : 1986 *Forgotten: 1993*

Copyright:(c) Nilikha | Dreamstime.com

Randomized Incremental Construction (RIC)

Starting from an empty set

Repeat:

1. *Insert the next object*
2. *Update the partial construction (data-structures)*

Randomized Incremental Construction (RIC)

Starting from an empty set

Repeat:

1. *Insert the next object*
2. *Update the partial construction (data-structures)*

Total Time $= \sum_i$ Time to insert the i-th object.

[6] to appear in STACS'19

Randomized Incremental Construction (RIC)

Starting from an empty set

Repeat:

1. *Insert the next object*
2. *Update the partial construction (data-structures)*

Total Time $= \sum_i$ Time to insert the i-th object.

$T_s(N) =$ Total time to insert a sequence s. (s is **good** if total time is less).

[6] to appear in STACS'19

Randomized Incremental Construction (RIC)

Starting from an empty set

Repeat:

1. *Insert the next object*
2. *Update the partial construction (data-structures)*

Total Time $= \sum_i$ Time to insert the i-th object.

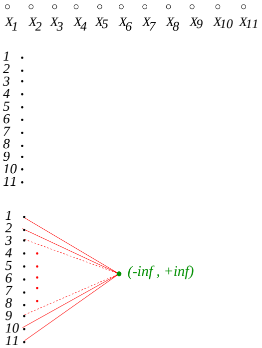
$T_s(N) =$ Total time to insert a sequence s. (s is **good** if total time is less).

Expected total time $= \max_{input\ I} \mathbb{E}[T_s(I)]$
(worst case for any input of size n).

[6] to appear in STACS'19

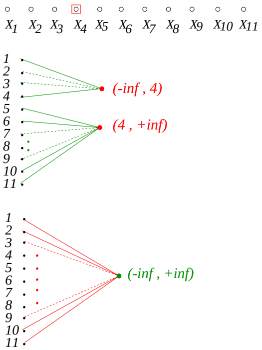
Quicksort as R.I.C.

Gradual refinement of partition



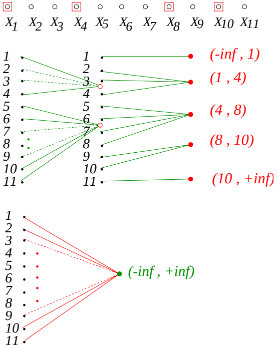
Quicksort as R.I.C.

Conflict graph



Quicksort as R.I.C.

Conflict graph



Bounding the maximum sub-problem size

$\Pi(n)$: set of subproblems defined by n objects
σ: a subproblem is defined by $D(\sigma)$ elements
$I(\sigma)$: size of the subproblem (unchosen elements in σ)

Bounding the maximum sub-problem size

$\Pi(n)$: set of subproblems defined by n objects
σ: a subproblem is defined by $D(\sigma)$ elements
$I(\sigma)$: size of the subproblem (unchosen elements in σ)

Example: Quicksort
$\Pi(n) = \binom{n}{2}$ pairs of points

A subproblem is defined by a pair of sample points
$D(\sigma)$ end-points of σ
$I(\sigma)$: unsampled points in σ

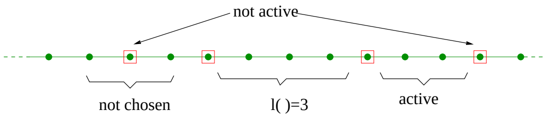
[11] to appear in STACS'19

Bounding the maximum sub-problem size

$\Pi(n)$: set of subproblems defined by n objects
σ: a subproblem is defined by $D(\sigma)$ elements
$l(\sigma)$: size of the subproblem (unchosen elements in σ)

Example: Quicksort
$\Pi(n) = \binom{n}{2}$ pairs of points

A subproblem is defined by a pair of sample points
$D(\sigma)$ end-points of σ
$l(\sigma)$: unsampled points in σ



$\Pi^0(n)$ special significance : $l(\sigma) = 0$

[11] to appear in STACS'19

Bounding the maximum sub-problem

Claim :

$$\Pr\{\max_{active\ \sigma} l(\sigma) \geq c\frac{n}{r}\log r\} \leq \frac{1}{2}$$

R chosen by Bernoulli sampling $p = r/n$

Bounding the maximum sub-problem

Claim :

$$\Pr\{\max_{active\ \sigma} l(\sigma) \geq c\frac{n}{r}\log r\} \leq \frac{1}{2}$$

R chosen by Bernoulli sampling $p = r/n$

$p(\sigma, r)$: conditional probability that none of the k $(= l(\sigma))$ elements are selected given $D(\sigma)$ chosen

$\leq (1 - r/n)^k$

$\leq e^{-c\log r} = 1/r^c$ for $\boxed{k \geq cn/r \ln r}$

BAD σ

[13] to appear in STACS'19

$q(\sigma, r)$: Prob. that $D(\sigma) \subset R$

$q(\sigma, r)$: Prob. that $D(\sigma) \subset R$

Prob. that σ is active $= p(\sigma, r) \times \Pr[D(\sigma) \subset R]$

[15] to appear in STACS'19

$q(\sigma, r)$: Prob. that $D(\sigma) \subset R$

Prob. that σ is active $= p(\sigma, r) \times \Pr[D(\sigma) \subset R]$

Prob. that some σ is active and **BAD** ($l(\sigma) \geq c(n \ln r)/r$) :

[15] to appear in STACS'19

$q(\sigma, r)$: Prob. that $D(\sigma) \subset R$

Prob. that σ is active $= p(\sigma, r) \times \Pr[D(\sigma) \subset R]$

Prob. that some σ is active and **BAD** $(I(\sigma) \geq c(n \ln r)/r)$:

$$\leq \frac{1}{r^c} \sum_{\sigma \in \Pi(n)} \Pr[D(\sigma) \subset R] = \frac{1}{r^c} \sum_{\sigma \in \Pi(n)} E[D(\sigma) \subset R]$$

$$= \frac{1}{r^c} \quad E[\text{ number sub-problems for which } D(\sigma) \subset R]$$

(linearity of Expectation)

$= r^{O(1)}$ for $D(\sigma) = O(1)$

[15] to appear in STACS'19

$q(\sigma, r)$: Prob. that $D(\sigma) \subset R$

Prob. that σ is active $= p(\sigma, r) \times \Pr[D(\sigma) \subset R]$

Prob. that some σ is active and **BAD** $(l(\sigma) \geq c(n \ln r)/r)$:

$$\leq \frac{1}{r^c} \sum_{\sigma \in \Pi(n)} \Pr[D(\sigma) \subset R] = \frac{1}{r^c} \sum_{\sigma \in \Pi(n)} E[D(\sigma) \subset R]$$

$$= \frac{1}{r^c} \ E[\text{ number sub-problems for which } D(\sigma) \subset R]$$

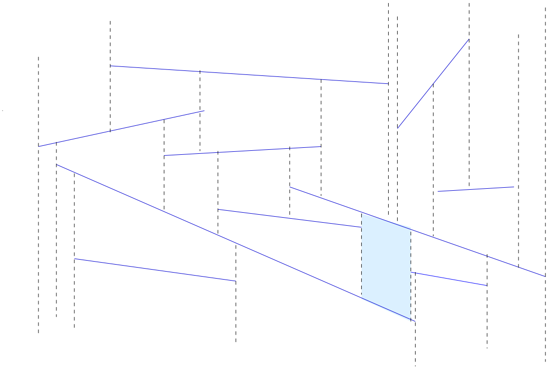
(linearity of Expectation)

$= r^{O(1)}$ for $D(\sigma) = O(1) \leq$

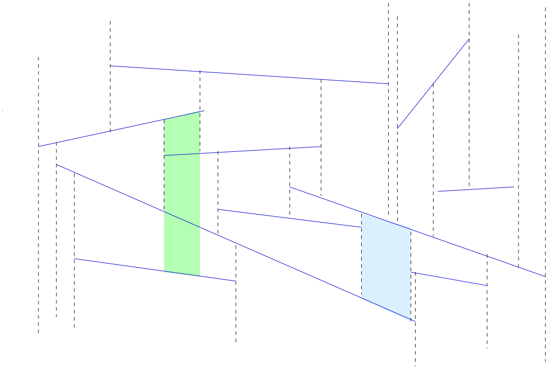
$\leq 1/2$ for appropriate c

[15] to appear in STACS'19

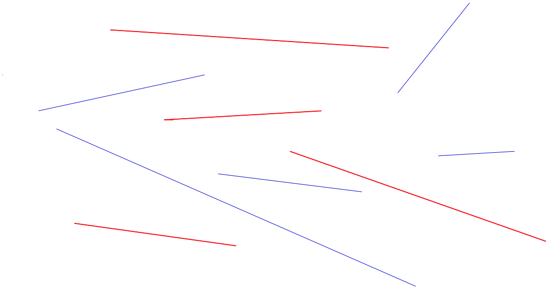
More general : Trapezoidal Map



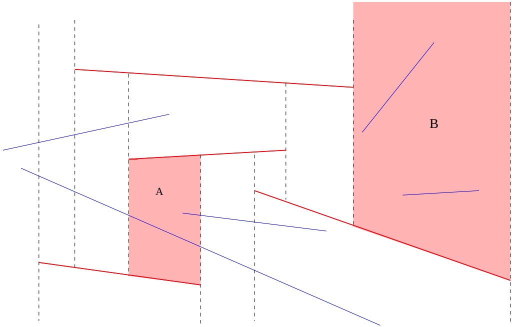
Trapezoidal Map : Ranges



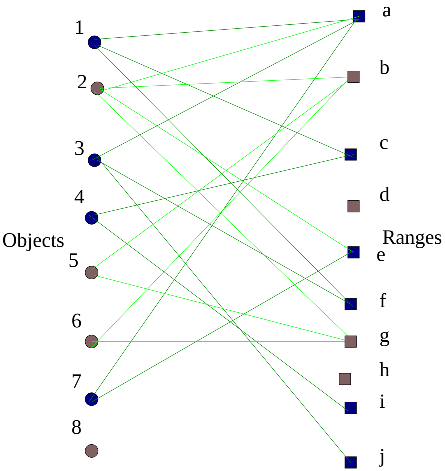
Randomized Incremental Construction



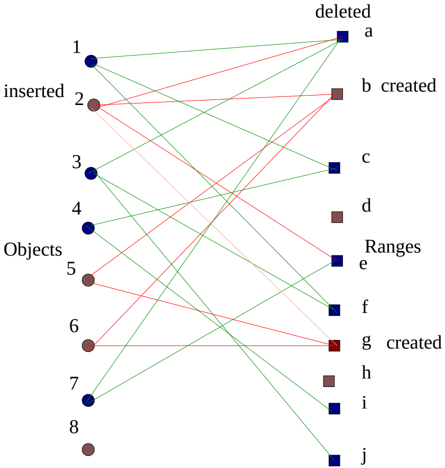
Objects (segments) and ranges (trapezoids)



A more general scenario



Modifications caused by insertion of an object



A general bound for RIC [Cl-Sh]

Total (amortised) cost = O(edges created in conflict graph)

Edges can be deleted at most once

General Step: $R \leftarrow R \cup s$ (both random subsets)

Expected work (#edges created in the conflict graph)=

$$\sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma) \cdot \Pr\{\sigma \in \Pi^0(R \cup s) - \Pi^0(R)\}$$

A general bound for RIC [Cl-Sh]

Total (amortised) cost $= O$(edges created in conflict graph)
Edges can be deleted at most once
General Step: $R \leftarrow R \cup s$ (both random subsets)
Expected work (#edges created in the conflict graph)=

$$\sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma) \cdot \Pr\{\sigma \in \Pi^0(R \cup s) - \Pi^0(R)\}$$

From *backward analysis* this probability is the same as deleting a random element from $R \cup s$ which is $\frac{d(\sigma)}{r+1}$. Substituting

$$\sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma) \cdot \frac{d(\sigma)}{r+1} = \frac{d(\sigma)}{r+1} \sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma)$$

[23] to appear in STACS'19

A general bound for RIC [Cl-Sh]

Total (amortised) cost $= O$(edges created in conflict graph)
Edges can be deleted at most once
General Step: $R \leftarrow R \cup s$ (both random subsets)
Expected work (#edges created in the conflict graph)=

$$\sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma) \cdot \Pr\{\sigma \in \Pi^0(R \cup s) - \Pi^0(R)\}$$

From *backward analysis* this probability is the same as deleting a random element from $R \cup s$ which is $\frac{d(\sigma)}{r+1}$. Substituting

$$\sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma) \cdot \frac{d(\sigma)}{r+1} = \frac{d(\sigma)}{r+1} \sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma)$$

Sum bound

$$\sum_{\sigma \in \Pi^0(R \cup s)} l(\sigma) = \frac{n}{r} \cdot E[\Pi^0(R \cup s)]$$

[23] to appear in STACS'19

A general bound on Expected running time of RIC

$$= O(\frac{d(\sigma)}{r} \cdot \frac{n}{r} E[\Pi^0(R \cup s)])$$

A common scenario $E[\Pi^0(R) = O(r)$.

$$\text{Total expected cost of RIC } = \sum_{r=1}^{r=n} O\left(\frac{d}{r} \cdot n\right)$$

$= O(n \log n)$ (also applicable to convex hulls)

A general bound on Expected running time of RIC

$$= O(\frac{d(\sigma)}{r} \cdot \frac{n}{r} E[\Pi^0(R \cup s)])$$

<u>A common scenario</u> $E[\Pi^0(R) = O(r)$.

$$\text{Total expected cost of RIC } = \sum_{r=1}^{r=n} O\left(\frac{d}{r} \cdot n\right)$$

$= O(n \log n)$ (also applicable to convex hulls)

OPEN PROBLEM

Tail estimates in the general case

[25] to appear in STACS'19

A general bound on Expected running time of RIC

$$= O(\frac{d(\sigma)}{r} \cdot \frac{n}{r} E[\Pi^0(R \cup s)])$$

A common scenario $E[\Pi^0(R) = O(r)$.

$$\text{Total expected cost of RIC } = \sum_{r=1}^{r=n} O\left(\frac{d}{r} \cdot n\right)$$

$= O(n \log n)$ (also applicable to convex hulls)

OPEN PROBLEM

Tail estimates in the general case without independent repetitions

[25] to appear in STACS'19

Sum of subproblem sizes

Def: c-order conflict $\binom{l(\sigma)}{c}$, for some $c \geq 0$

Let $T_c = \sum_{\sigma \in \Pi^0(R)} \binom{l(\sigma)}{c}$

Remark For technical reasons it is not $l(\sigma)^c$. $T_0 = |\Pi^0(R)|$. $T_1 =$ sum of subproblems.

Sum of subproblem sizes

Def: c-order conflict $\binom{l(\sigma)}{c}$, for some $c \geq 0$

Let $T_c = \sum_{\sigma \in \Pi^0(R)} \binom{l(\sigma)}{c}$

Remark For technical reasons it is not $l(\sigma)^c$. $T_0 = |\Pi^0(R)|$. $T_1 =$ sum of subproblems.

$$\textbf{Claim } E[T_c] = O(\left(\tfrac{n}{r}\right)^c E[\Pi^c(R)])$$

For constant c, $E[\Pi^c(R) = O(E[\Pi^0(R)]$ implying that average conflict size is very close to $\frac{n}{r}$

[27] to appear in STACS'19

Sum of subproblem sizes

$T_c = \sum_{\sigma \in \Pi(N)(R)} \binom{l(\sigma)}{c} I_{\sigma,R}$ where $I_{\sigma,R} = 1$ if $\sigma \in \Pi^0(R)$.

$E[T_c] = \sum_{\sigma \in \Pi(N)} \binom{l(\sigma)}{c} p^{d(\sigma)} \cdot (1-p)^{l(\sigma)}$ for $l(\sigma) \geq c$.

$= \sum_{\sigma \in \Pi(N)} \binom{l(\sigma)}{c} p^{d(\sigma)+c} \cdot (1-p)^{l(\sigma)-c} \cdot \left(\frac{1-p}{p}\right)^c$

$\leq \left(\frac{1-p}{p}\right)^c \cdot E[\Pi^c(R)]$

Sum of subproblem sizes

$T_c = \sum_{\sigma \in \Pi(N)(R)} \binom{l(\sigma)}{c} I_{\sigma,R}$ where $I_{\sigma,R} = 1$ if $\sigma \in \Pi^0(R)$.

$E[T_c] = \sum_{\sigma \in \Pi(N)} \binom{l(\sigma)}{c} p^{d(\sigma)} \cdot (1-p)^{l(\sigma)}$ for $l(\sigma) \geq c$.

$= \sum_{\sigma \in \Pi(N)} \binom{l(\sigma)}{c} p^{d(\sigma)+c} \cdot (1-p)^{l(\sigma)-c} \cdot \left(\frac{1-p}{p}\right)^c$

$\leq \left(\frac{1-p}{p}\right)^c \cdot E[\Pi^c(R)]$

since $\Pr\{\sigma \in \Pi^c(R)\} = \Pr\{d(\sigma)$ objects chosen and c out of $l(\sigma)$ not chosen

[29] to appear in STACS'19

Sum of subproblem sizes

$$T_c = \sum_{\sigma \in \Pi(N)(R)} \binom{l(\sigma)}{c} I_{\sigma,R} \text{ where } I_{\sigma,R} = 1 \text{ if } \sigma \in \Pi^0(R).$$

$$E[T_c] = \sum_{\sigma \in \Pi(N)} \binom{l(\sigma)}{c} p^{d(\sigma)} \cdot (1-p)^{l(\sigma)} \text{ for } l(\sigma) \geq c.$$

$$= \sum_{\sigma \in \Pi(N)} \binom{l(\sigma)}{c} p^{d(\sigma)+c} \cdot (1-p)^{l(\sigma)-c} \cdot \left(\frac{1-p}{p}\right)^c$$

$$\leq \left(\frac{1-p}{p}\right)^c \cdot E[\Pi^c(R)]$$

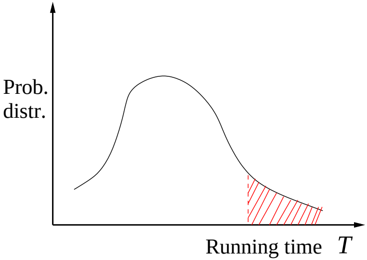
since $\Pr\{\sigma \in \Pi^c(R)\} = \Pr\{d(\sigma)$ objects chosen and c out of $l(\sigma)$ not chosen

$$\leq \left(\frac{1}{p}\right)^c \cdot E[\Pi^c(R)] \text{ where } p = \frac{r}{n}.$$

[29] to appear in STACS'19

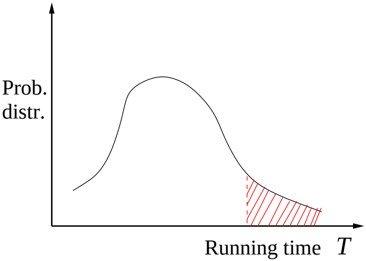
Standard Tools

Running time is bounded by some probability distribution



Standard Tools

Running time is bounded by some probability distribution



Probabilistic Inequalities :

- Markov $\Pr[X > kE[X]] < 1/k$
- Chernoff $\Pr[X > A] \leq G_X(s) \cdot s^{-A}$
- Chebychev $\Pr[|X - E[X]| > r] \leq \sigma^2/r^2$

Notation : $\widetilde{O}(\cdot) \stackrel{def}{=} O(\cdot)$ with prob. $1 - \frac{1}{n}$ Inv polynomial

[31] to appear in STACS'19

A martingale framework

$(\Omega, \mathcal{U})$: all permutations of n objects $\mathcal{U}$ uniform probability distribution.
$\mathcal{B}_i$ *blocks* of permutations, to the i-prefixes $\bar{X}^{(i)}$

A martingale framework

$(\Omega, \mathcal{U})$: all permutations of n objects $\mathcal{U}$ uniform probability distribution.
$\mathcal{B}_i$ *blocks* of permutations, to the i-prefixes $\bar{X}^{(i)}$
Example: Objects $\{1, 2, 3\}$.

B_0	[1, 2, 3], [1, 3, 2], [2, 1, 3], [2, 3, 1], [3, 1, 2], [3, 2, 1]
B_1	[1, 2, 3], [1, 3, 2] , [2, 1, 3], [2, 3, 1] , [3, 1, 2], [3, 2, 1]
B_2	[1, 2, 3], [1, 3, 2] ...

[33] to appear in STACS'19

A martingale framework

$(\Omega, \mathcal{U})$: all permutations of n objects $\mathcal{U}$ uniform probability distribution.
$\mathcal{B}_i$ *blocks* of permutations, to the i-prefixes $\bar{X}^{(i)}$
Example: Objects $\{1, 2, 3\}$.

B_0	[1, 2, 3], [1, 3, 2], [2, 1, 3], [2, 3, 1], [3, 1, 2], [3, 2, 1]
B_1	[1, 2, 3], [1, 3, 2] , [2, 1, 3], [2, 3, 1] , [3, 1, 2], [3, 2, 1]
B_2	[1, 2, 3], [1, 3, 2] ...

$Z_i = \mathbb{E}[Z|\bar{X}^{(i)}]$: random variable Z over the probability space $(\Omega, \mathcal{U})$

[33] to appear in STACS'19

A martingale framework

$(\Omega, \mathcal{U})$: all permutations of n objects $\mathcal{U}$ uniform probability distribution.
$\mathcal{B}_i$ *blocks* of permutations, to the i-prefixes $\bar{X}^{(i)}$
Example: Objects $\{1, 2, 3\}$.

B_0	[1, 2, 3], [1, 3, 2], [2, 1, 3], [2, 3, 1], [3, 1, 2], [3, 2, 1]
B_1	[1, 2, 3], [1, 3, 2] , [2, 1, 3], [2, 3, 1] , [3, 1, 2], [3, 2, 1]
B_2	[1, 2, 3], [1, 3, 2] ...

$Z_i = \mathbb{E}[Z|\bar{X}^{(i)}]$: random variable Z over the probability space $(\Omega, \mathcal{U})$

Claim: $Z_1, Z_2 \ldots$ satisfies $\mathbb{E}[Z_i|Z_0, Z_2 \ldots Z_{i-1}] = Z_{i-1}$

[33] to appear in STACS'19

A martingale framework

$(\Omega, \mathcal{U})$: all permutations of n objects $\mathcal{U}$ uniform probability distribution.
$\mathcal{B}_i$ *blocks* of permutations, to the i-prefixes $\bar{X}^{(i)}$
Example: Objects $\{1, 2, 3\}$.

B_0	$[1, 2, 3], [1, 3, 2], [2, 1, 3], [2, 3, 1], [3, 1, 2], [3, 2, 1]$
B_1	$[1, 2, 3], [1, 3, 2]$, $[2, 1, 3], [2, 3, 1]$, $[3, 1, 2], [3, 2, 1]$
B_2	$[1, 2, 3]$, $[1, 3, 2]$ ...

$Z_i = \mathbb{E}[Z|\bar{X}^{(i)}]$: random variable Z over the probability space $(\Omega, \mathcal{U})$

Claim: $Z_1, Z_2 \ldots$ satisfies $\mathbb{E}[Z_i|Z_0, Z_2 \ldots Z_{i-1}] = Z_{i-1}$ **Doob's martingale**

[33] to appear in STACS'19

A martingale framework

$(\Omega, \mathcal{U})$: all permutations of n objects $\mathcal{U}$ uniform probability distribution.
$\mathcal{B}_i$ *blocks* of permutations, to the i-prefixes $\bar{X}^{(i)}$
Example: Objects $\{1, 2, 3\}$.

B_0	[1, 2, 3], [1, 3, 2], [2, 1, 3], [2, 3, 1], [3, 1, 2], [3, 2, 1]
B_1	[1, 2, 3], [1, 3, 2] , [2, 1, 3], [2, 3, 1] , [3, 1, 2], [3, 2, 1]
B_2	[1, 2, 3], [1, 3, 2] ...

$Z_i = \mathbb{E}[Z|\bar{X}^{(i)}]$: random variable Z over the probability space $(\Omega, \mathcal{U})$

Claim: $Z_1, Z_2 \ldots$ satisfies $\mathbb{E}[Z_i | Z_0, Z_2 \ldots Z_{i-1}] = Z_{i-1}$ **Doob's martingale**

Deviation from Expectation

Z_0 : expected runnning time of ric and Z_n: actual runnning time
Concentration bound : $\Pr[|Z_0 - Z_n| \geq \lambda] < ???$

[33] to appear in STACS'19

A martingale inequality

Theorem [Freedman 75]

Let $X_1, X_2 \ldots X_n$ be a sequence of random variables and let Y_k, a function of $X_1 \ldots X_k$ be a martingale sequence, i.e., $\mathbb{E}[Y_k | X_1 \ldots X_k] = Y_{k-1}$ such that $\max_{1 \leq k \leq n}\{|Y_k - Y_{k-1}|\} \leq M_n$. Let

$$W_k = \sum_{j=1}^{k} \mathbb{E}[(Y_j - Y_{j-1})^2 | X_1 \ldots X_{j-1}] = \sum_{j=1}^{k} Var(Y_j | X_1 \ldots X_{j-1})$$

where *Var* is the variance using $\mathbb{E}[Y_j] = Y_{j-1}$. Then for all λ and $W_n \leq \Delta^2, \ \ \Delta^2 > 0$,

$$\Pr[|Y_n - Y_0| \geq \lambda] \leq 2\exp\left(-\frac{\lambda^2}{2(\Delta^2 + M_n \cdot \lambda/3)}\right)$$

Martingale inequality

Extended Freedman inequality

Let $\Pr[\max_{1 \le k \le n}\{|Y_k - Y_{k-1}|\} \ge M_n, W_n \ge \Delta^2] \le \frac{1}{f(n)}$, then,

$$\Pr[|Y_n - Y_0| \ge \lambda] \le 2\exp\left(-\frac{\lambda^2}{2(\Delta^2 + M_n \cdot \lambda/3)}\right) + O(1/f(n))$$

Martingale inequality

Extended Freedman inequality

Let $\Pr[\max_{1 \le k \le n}\{|Y_k - Y_{k-1}|\} \ge M_n, W_n \ge \Delta^2] \le \frac{1}{f(n)}$, then,

$$\Pr[|Y_n - Y_0| \ge \lambda] \le 2\exp\left(-\frac{\lambda^2}{2(\Delta^2 + M_n \cdot \lambda/3)}\right) + O(1/f(n))$$

Azuma-Hoeffding

$$\Pr[|Y_n - Y_0| \ge t] \le \exp\left(\frac{-t^2}{\sum_{i=1}^{n} M_i^2}\right)$$

[36] to appear in STACS'19

Martingale inequality

Extended Freedman inequality

Let $\Pr[\max_{1\leq k\leq n}\{|Y_k - Y_{k-1}|\} \geq M_n, W_n \geq \Delta^2] \leq \frac{1}{f(n)}$, then,

$$\Pr[|Y_n - Y_0| \geq \lambda] \leq 2\exp\left(-\frac{\lambda^2}{2(\Delta^2 + M_n\cdot\lambda/3)}\right) + O(1/f(n))$$

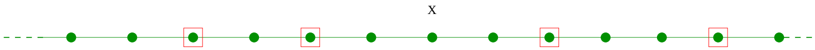
Azuma-Hoeffding

$$\Pr[|Y_n - Y_0| \geq t] \leq \exp\left(\frac{-t^2}{\sum_{i=1}^{n} M_i^2}\right)$$

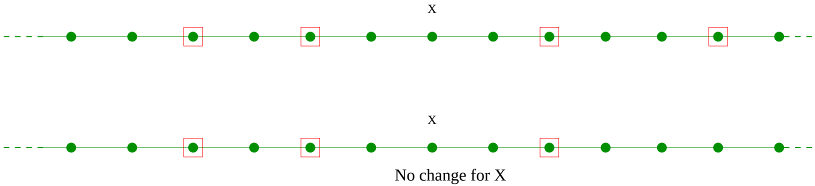
Mehlhorn-Sharir-Welzl [93] obtained a special kind of Martingale concentration bound that was effective for some cases of line segment intersection RIC.

[36] to appear in STACS'19

Application to quicksort

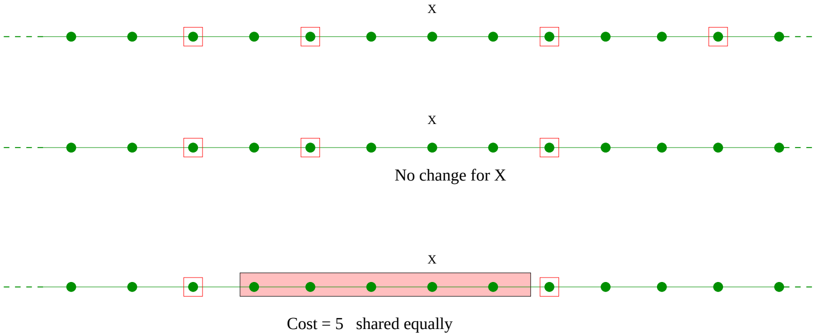


Application to quicksort



[38] to appear in STACS'19

Application to quicksort



$$I_j^x = \begin{cases} 1 & \text{if interval containing } x \text{ changes in step } j \\ 0 & \text{otherwise.} \end{cases}$$

[38] to appear in STACS'19

Analysis

$$\mathbb{E}[\sum_j I_j^x] = \sum_j \Pr[I_j^x = 1] = \sum_j \frac{2}{j} = \log n$$

Analysis

$$\mathbb{E}[\sum_j I_j^x] = \sum_j \Pr[I_j^x = 1] = \sum_j \frac{2}{j} = \log n$$

$M_n = 1$ since maximum charge on X is 1 for any pivot.

Analysis

$$\mathbb{E}[\sum_j I_j^x] = \sum_j \Pr[I_j^x = 1] = \sum_j \frac{2}{j} = \log n$$

$M_n = 1$ since maximum charge on X is 1 for any pivot.

$$Y_j - Y_{j-1} = w(v_{j-1}, v_j) + \left(\sum_{k=j+1}^{n} \mathbb{E}[I'_k]\right) - \left(\sum_{k=j}^{n} \mathbb{E}[I_k]\right)$$

$$= I_j - E[I_j] \text{ assuming } I_j, I'_j\text{'s have the same distribution}$$

Analysis

$$\mathbb{E}[\sum_j I_j^x] = \sum_j \Pr[I_j^x = 1] = \sum_j \frac{2}{j} = \log n$$

$M_n = 1$ since maximum charge on X is 1 for any pivot.

$$Y_j - Y_{j-1} = w(v_{j-1}, v_j) + \left(\sum_{k=j+1}^{n} \mathbb{E}[I'_k] \right) - \left(\sum_{k=j}^{n} \mathbb{E}[I_k] \right)$$

$$= I_j - E[I_j] \text{ assuming } I_j, I'_j\text{'s have the same distribution}$$

$$\mathbb{E}_j[(I_j - E[I_j])^2] = \mathbb{E}_j[I_j^2] - \mathbb{E}_j^2[I_j] \ \le \mathbb{E}_j[I_j^2] - \frac{4}{(n-j)^2}$$

$$\le \frac{2}{n-j} \quad I_j^2 \text{ also 0-1 indicator rv}$$

$$\sum_{j=1}^{n} \mathbb{E}_{j-1}[(Y_j - Y_{j-1})^2] \le \sum_{j=1}^{n} \frac{2}{n-j} \le 2\log n = W_n$$

Inverse polynomial bound for quicksort

$$\Pr[|Y_n - Y_0| \geq c \log n] \leq \exp\left(-\frac{4c^2 \log^2 n}{2(\log n + c \log n/3)}\right) \leq \frac{1}{n^c}.$$

Inverse polynomial bound for quicksort

$$\Pr[|Y_n - Y_0| \geq c \log n] \leq \exp\left(-\frac{4c^2 \log^2 n}{2(\log n + c \log n/3)}\right) \leq \frac{1}{n^c}.$$

Since I_j^x are independent, one can apply Chernoff bounds directly to get similar concentration bounds [Seidel89]

[42] to appear in STACS'19

Inverse polynomial bound for quicksort

$$\Pr[|Y_n - Y_0| \geq c \log n] \leq \exp\left(-\frac{4c^2 \log^2 n}{2(\log n + c \log n/3)}\right) \leq \frac{1}{n^c}.$$

Since I_j^x are independent, one can apply Chernoff bounds directly to get similar concentration bounds [Seidel89]

Use of Martingales can reduce the number of random bits significantly.

[42] to appear in STACS'19

Inverse polynomial bound for quicksort

$$\Pr[|Y_n - Y_0| \geq c \log n] \leq \exp\left(-\frac{4c^2 \log^2 n}{2(\log n + c \log n/3)}\right) \leq \frac{1}{n^c}.$$

Since I_j^x are independent, one can apply Chernoff bounds directly to get similar concentration bounds [Seidel89]

Use of Martingales can reduce the number of random bits significantly.

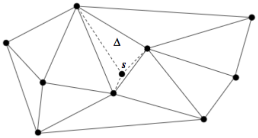
Azuma-Hoeffding

$$\Pr[|Y_n - Y_0| \geq t] \leq \exp\left(\frac{-t^2}{\sum_{i=1}^{n} M_i^2}\right)$$

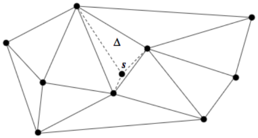
For $M_i = 1$, there is no meaningful bound for this setting.

[42] to appear in STACS'19

Incremental Delaunay triangulation



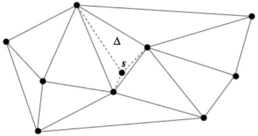
Incremental Delaunay triangulation



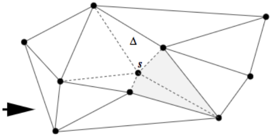
Total size is $O(i)$ but the degree of a vertex can be large

[44] to appear in STACS'19

Incremental Delaunay triangulation



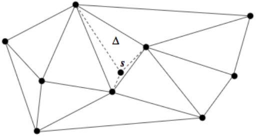
Total size is $O(i)$ but the degree of a vertex can be large



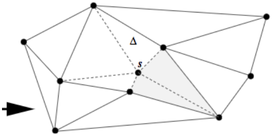
Number of new Δ's is 5 and can be as large as i but ..

[44] to appear in STACS'19

Incremental Delaunay triangulation



Total size is $O(i)$ but the degree of a vertex can be large

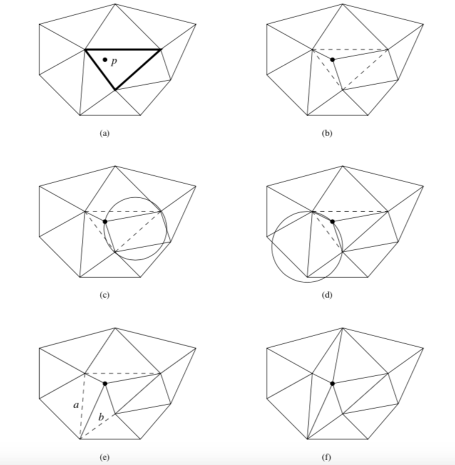


Number of new Δ's is 5 and can be as large as i but ..
the average degree of a planar graph is $O(1)$.

~~Expected number of triangles that appear~~ over the course of RIC $= O(n)$

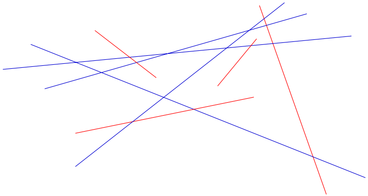
[44] to appear in STACS'19

Flip algorithm [Green and Sibson]



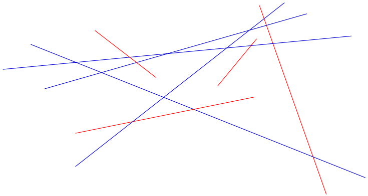
Courtesy: Lischinski[93]

Line segment intersections

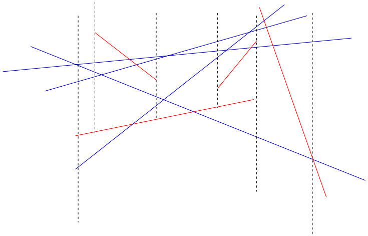


No intersections between red. All blue segments intersect.

Line segment intersections



No intersections between red. All blue segments intersect.



Intermediate structures (trapezoidal maps) can have huge variance

[47] ... in STACS'10

Expected number of intersections at stage i

Probability intersection (s_j, s_k) appear in stage i = Probabilty s_j, s_k have been chosen.

Expected number of intersections at stage i

Probability intersection (s_j, s_k) appear in stage i = Probabilty s_j, s_k have been chosen.

$$\mathbb{E}[M_i] = m \cdot \frac{i}{n} \cdot \frac{i}{n} = \frac{m \cdot i^2}{n^2}$$

[49] to appear in STACS'19

Expected number of intersections at stage i

Probability intersection (s_j, s_k) appear in stage i = Probabilty s_j, s_k have been chosen.

$$\mathbb{E}[M_i] = m \cdot \frac{i}{n} \cdot \frac{i}{n} = \frac{m \cdot i^2}{n^2}$$

Ideal bound for segment intersections : $O(m + n \log n)$

[49] to appear in STACS'19

Summary of results

Problem	Exp Run time	Tail estimates
Quicksort	$O(n \log n)$	$O(c\gamma n \log n)$ w.p. $\geq 1 - n^{-c}$
Delaunay Triangulation	$O(n \log n)$	$O(c\gamma n \log n)$ w.p. $\geq 1 - 2^{-c}$
Segment intersections/ Trapezoidal maps	$O(n \log n + m)$ *no conflict list*	$O(n \log n + m)$ w.p. $1 - \exp -(\frac{m+n \log n}{n\alpha(n)})$
	*using conflict list**	w.p. $\geq 1 - \exp -(\frac{m}{n \log^2 n})$

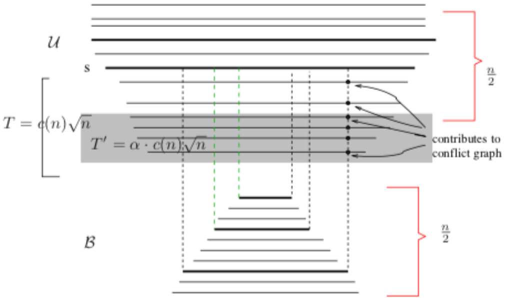
Summary of results

Problem	Exp Run time	Tail estimates
Quicksort	$O(n \log n)$	$O(c\gamma n \log n)$ w.p. $\geq 1 - n^{-c}$
Delaunay Triangulation	$O(n \log n)$	$O(c\gamma n \log n)$ w.p. $\geq 1 - 2^{-c}$
Segment intersections/ Trapezoidal maps	$O(n \log n + m)$ *no conflict list*	$O(n \log n + m)$ w.p. $1 - \exp -(\frac{m+n\log n}{n\alpha(n)})$
	*using conflict list**	w.p. $\geq 1 - \exp -(\frac{m}{n\log^2 n})$

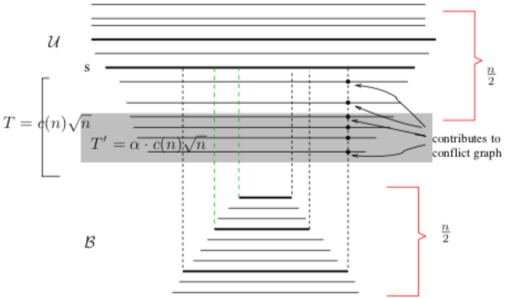
The non-conflict bound is nearly inverse polynomial for $m = 0$, i.e., $\exp -(\log n/\alpha(n))$

[51] to appear in STACS'19

A lower bound construction



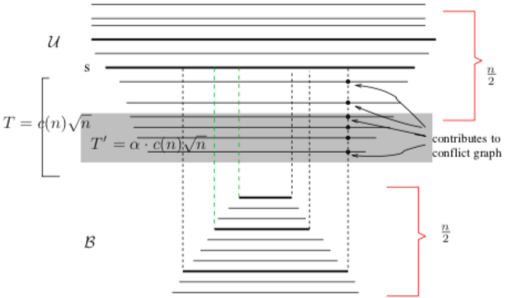
A lower bound construction



$|\mathcal{U}| = |\mathcal{B}| = \frac{n}{2}$
Among first $3\sqrt{n}$ insertions, $\Pr[\mathcal{B} \geq \sqrt{n}] \geq 1 - 2^{\sqrt{n}}$.

[53] to appear in STACS'19

A lower bound construction

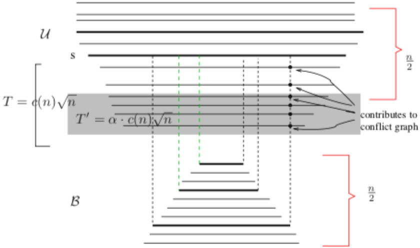


$|\mathcal{U}| = |\mathcal{B}| = \frac{n}{2}$
Among first $3\sqrt{n}$ insertions, $\Pr[\mathcal{B} \geq \sqrt{n}] \geq 1 - 2^{\sqrt{n}}$.

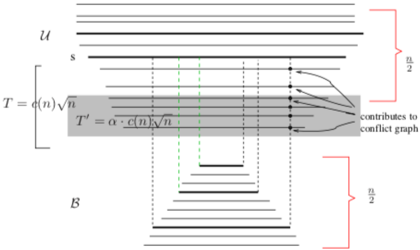
$$\Pr[|T| \geq c(n)\sqrt{n}] \geq 4^{-3c(n)}$$

[53] to appear in STACS'10

Construction cont.



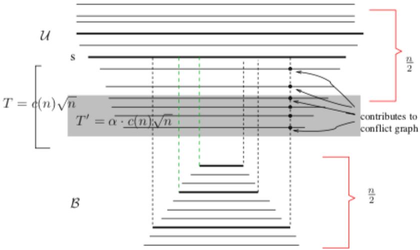
Construction cont.



Number of times s moves down without touching T' (shaded region) $= r$
$\Rightarrow$ RIC incurs $r \cdot c(n)\sqrt{(n)}\sqrt{(n)}$ cost

[55] to appear in STACS'19

Construction cont.



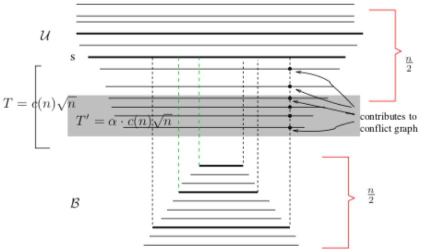
Number of times s moves down without touching T' (shaded region) $= r$

$\Rightarrow$ RIC incurs $r \cdot c(n)\sqrt{(n)}\sqrt{(n)}$ cost

$$\Pr[r \geq \Omega(\log\log n)] \geq \frac{1}{\log n} \quad \text{say } \alpha = 1/2$$

[55] to appear in STACS'19

Construction cont.



Number of times s moves down without touching T' (shaded region) $= r$
$\Rightarrow$ RIC incurs $r \cdot c(n)\sqrt{(n)}\sqrt{(n)}$ cost

$$\Pr[r \geq \Omega(\log \log n)] \geq \frac{1}{\log n} \quad \text{say } \alpha = 1/2$$

$$\Pr[X \geq n \log n \log \log n] \geq \frac{1}{\sqrt{n}} \quad \text{for } c(n) = \Omega(\log n)$$

[55] to appear in STACS'19

Future directions

Future directions

- Inverse polynomial bounds open for Delaunay, Line segments and many others

[57] to appear in STACS'19

Future directions

- Inverse polynomial bounds open for Delaunay, Line segments and many others
- More lower bound constructions

[57] to appear in STACS'19

Future directions

- Inverse polynomial bounds open for Delaunay, Line segments and many others
- More lower bound constructions
- Distinct variations of RICs (rebuild ?) may have different performance including data structures like conflict lists (without).

[57] to appear in STACS'19