Optimal Static Range Reporting in One Dimension

Gerth Stølting Brodal

BRICS, University of Aarhus

Joint work with

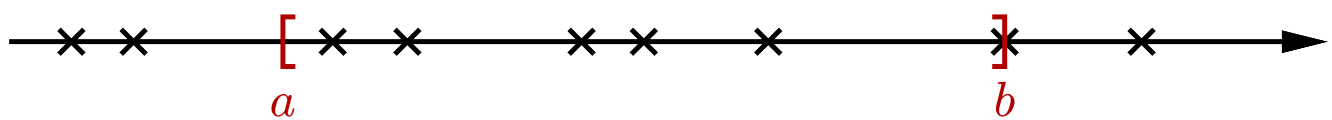
Stephen Alstrup   Theis Rauhe

The IT University of Copenhagen

Range Queries in One Dimension

Preprocess a set S of n elements from a totally ordered universe U to support online range queries

Member(a)	$a \in S$?
Pred(a)	$\max\{x \in S \mid x \leq a\}$ or $\perp$
FindAny(a, b)	any $e \in S \cap [a, b]$ or $\perp$
Report(a, b)	$S \cap [a, b]$
Count(a, b)	$\lvert S \cap [a, b] \rvert$
Count$_\varepsilon(a, b)$	Count$(a, b) \leq$ Count$_\varepsilon(a, b) \leq (1 + \varepsilon) \cdot$ Count(a, b)



Fact Comparison based all queries require time $\Omega(\log n)$

RAM Model

Unit cost RAM with word size w, $U = \{0, 1, \ldots, 2^w - 1\}$

Instructions: $+$, $*$, $\neg$, $\vee$, shifting, random bits, ...

Theorem

Supporting Member requires space $\Omega(n)$ words, for $n \leq 2^{(1-\varepsilon)w}$

Theorem

Member, Pred, FindAny, Count, Count$_\varepsilon$ in constant time and Report in time $O(1 + |\text{output}|)$ using space $O(2^w)$ words

Pred	0	1	2	3	3	3	3	7	7	9	10	10	10	13	13	13	16	16	16	16	16	21	22	22	22	22	22	27	28	29	29	31
Rank	0	1	2	3	0	0	0	4	0	5	6	0	0	7	0	0	8	0	0	0	0	9	10	0	0	0	0	11	12	13	0	14
S		X	X	X				X		X	X			X			X					X	X					X	X	X		X
	0	1	2	3							10										20										30	31

Previous Results for the RAM

	Query time	Space
Member		
Fredman *et al.* 1984 †	1	n
Pred, Count		
Willard 1983	$\log w$	n
Beame, Fich 1999 ⋆	$\min\left(\frac{\log w}{\log\log w}, \sqrt{\frac{\log n}{\log\log n}}\right)$	$n^{1+\delta}$
Andersson, Thorup 2000	$\min\left(\frac{\log w \cdot \log\log n}{\log\log w}, \sqrt{\frac{\log n}{\log\log n}}\right)$	n
FindAny		
Miltersen *et al.* 1998	1	$n \cdot w$
Report		
Miltersen *et al.* 1998	$1 + \lvert\text{output}\rvert$	$n \cdot w$

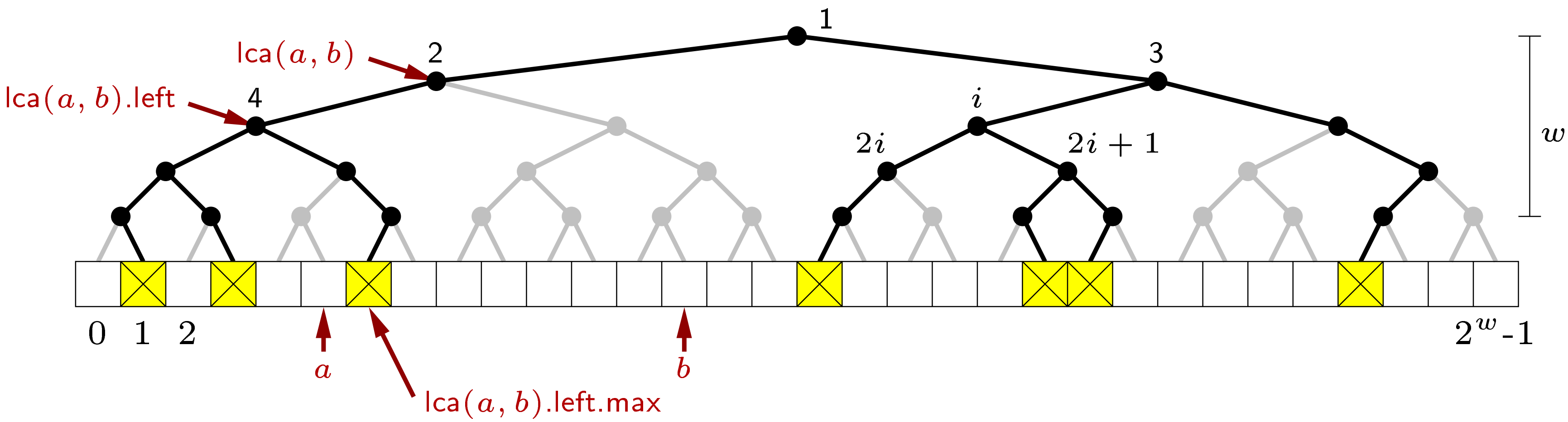
† "FKS" ⋆ Optimal query time

Our Result

	Query time	Space
FindAny	1	n
Report	$1 + \lvert\text{output}\rvert$	n
Count_ε	$\log \frac{1}{\varepsilon}$	n

Preprocessing expected time $O(n\sqrt{w})$

The Basic Idea



$$\text{FindAny}(a,b) \in [a,b] \cap \{\text{lca}(a,b).\text{left.max}\,,\ \text{lca}(a,b).\text{right.min}\}$$

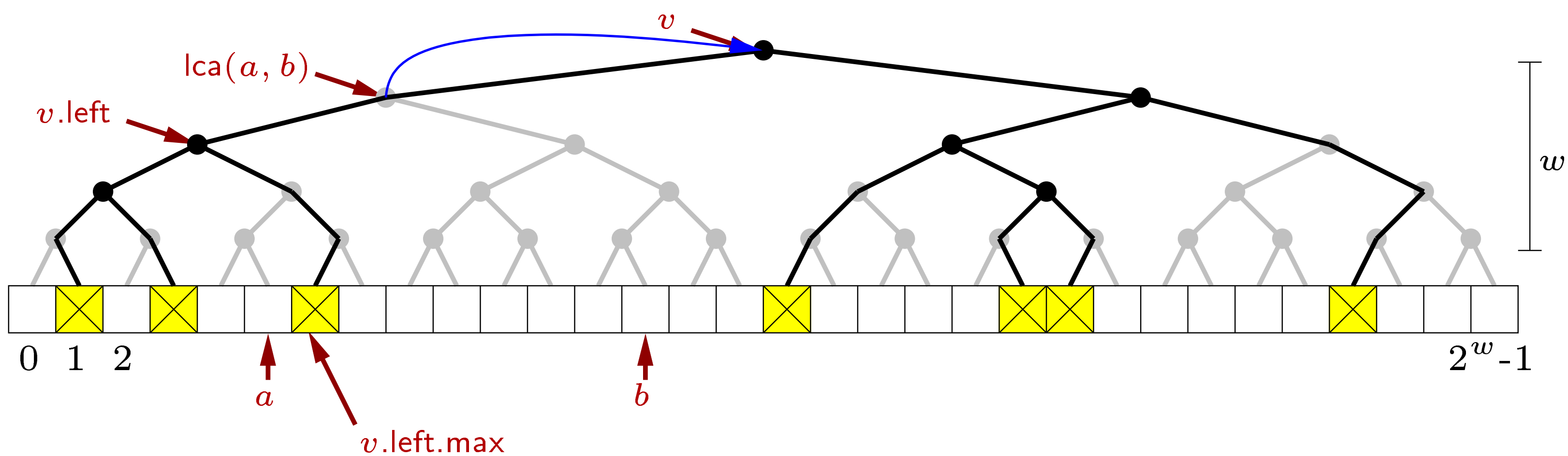
where

$$\text{lca}(a,b) = (2^{w-1} + \lfloor a/2 \rfloor) \downarrow \text{msb}(a \oplus b)$$

Lemma

FKS on ● nodes ⇒ space $O(n \cdot w)$ words and FindAny constant time

The Second Idea



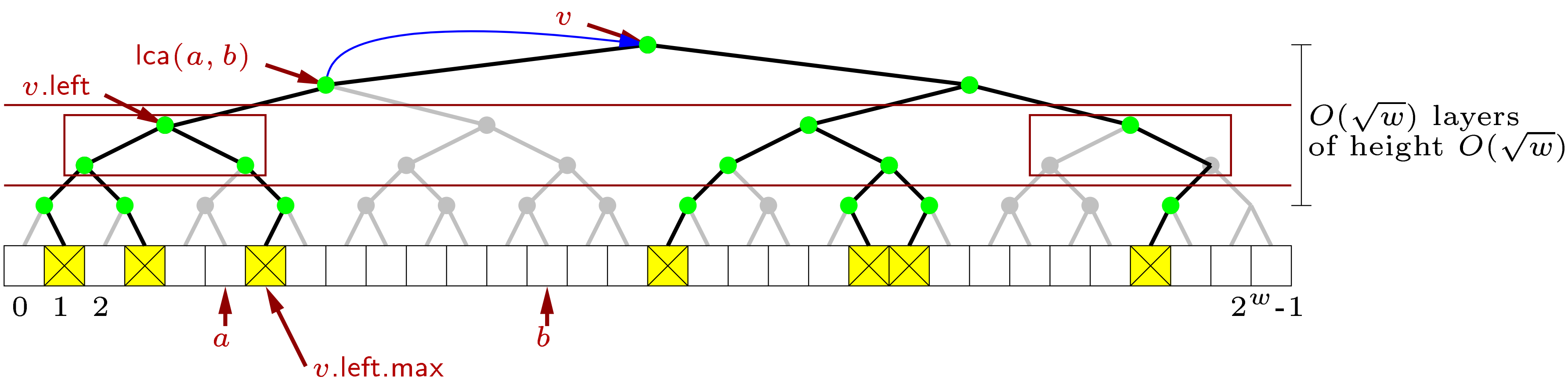
v is the nearest ancestor of lca(a, b) where both subtrees are nonempty

$$\text{FindAny}(a, b) \in [a, b] \cap \{v.\text{left}.\{\min, \max\}, v.\text{right}.\{\min, \max\}\}$$

Lemma

Store ● nodes, and depth(v) − depth(lca(a, b)) for all lca(a, b) with nonempty subtrees ($\leq nw$ nodes) using a perfect hash function [Schmidt, Siegel 1990] ⇒ space $O(n + \frac{n \cdot w \cdot \log w}{w}) = O(n \log w)$

The Last Idea



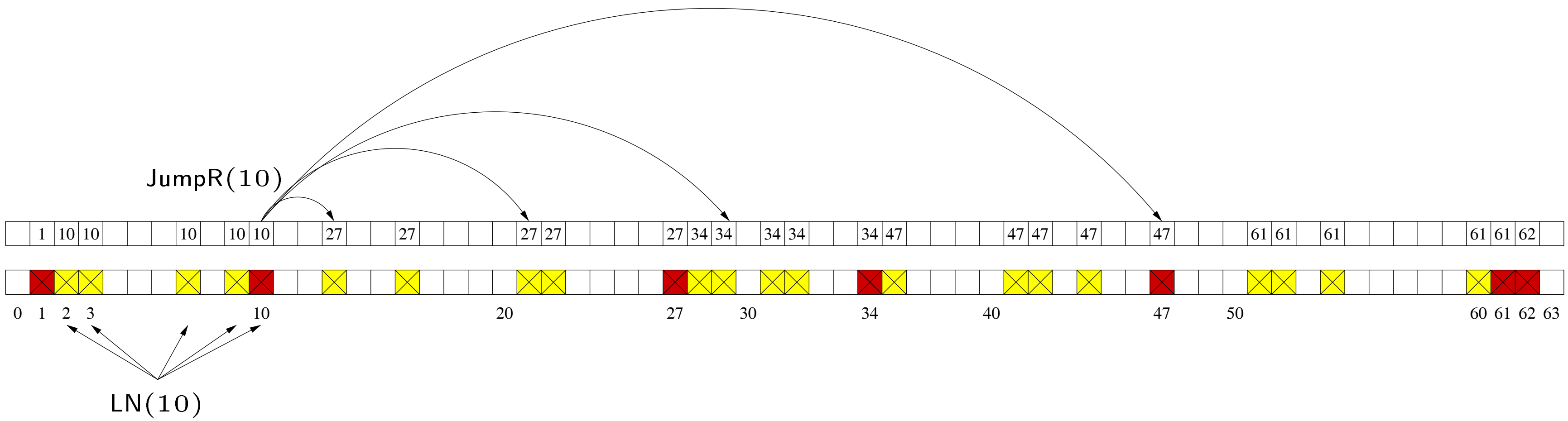
Lemma

Only store a subset of depth(v) − depth(lca(a, b)) (● nodes)
[Schmidt, Siegel 1990] ⇒ space $O(n + \frac{n \cdot \sqrt{w} \cdot \log w}{w}) = O(n)$

Theorem

FindAny can be supported in constant time and space $O(n)$

Approximate Range Counting



Partition elements into blocks of size $\log n$

For each block keep two Q-heaps [Fredman, Willard, 1994]

- Linear space
- Rank queries in constant time (for logarithmic sized problems)
- One Q-heap for the $O(\log n)$ elements in the block
- One Q-heap for the $1, 2, \ldots, 2^i, \ldots$ neighbors

Theorem Count$_\varepsilon$ can be supported in time $O(\log \frac{1}{\varepsilon})$ and space $O(n)$

Note ε is not used in the preprocessing

Summary

	Query time	Space
FindAny	1	n
Report	$1 + \lvert\text{output}\rvert$	n
Count$_\varepsilon$	$\log \frac{1}{\varepsilon}$	n

Preprocessing expected time O$(n\sqrt{w})$

FindAny

Proc FindAny(a, b)

if $a \leq b$ **then**

$H \ = 1 \uparrow (\mathsf{msb}(w) \downarrow 1)$

$d \ = \mathsf{msb}(a \oplus b)$

$u \ = ((1 \uparrow (w - 1)) + (a \downarrow 1)) \downarrow d$

$z \ = u \downarrow ((w - 1 - d) \wedge (H - 1))$

$v \ = B[z]\ ?\ V[u \downarrow D[u]]\ :\ V[z \downarrow D[z]]$

for $x \in \{\, v.\mathsf{left}.m,\ v.\mathsf{left}.M,\ v.\mathsf{right}.m,\ v.\mathsf{right}.M \,\}$

if $x \in [a, b]$ **then return** x

return $\perp$