Lecture 14

Predecessor lower bounds (static, cell-probe model)

$\sim$ a story

- Ajtai – Combinatorica 1988:
  - first $\omega(1)$ bound, complicated
  - $\forall w \; \exists n$ s.t. $\Omega(\sqrt{\lg w})$ [accidental $\Omega(\lg w)$ claim]
- Miltersen – STOC 1994:
  - better understanding of "same" proof
  - connection to communication complexity
  - $\forall w \; \exists n$ s.t. $\Omega(\sqrt{\lg w})$ (again – slightly more general)
  - $\forall n \; \exists w$ s.t. $\Omega(\sqrt[3]{\lg n})$
- Miltersen, Nisan, Safra, Wigderson – STOC 1995 & JCSS 1998:
  - clean proofs of same bounds
  - round elimination idea & lemma

  citation
- Beame & Fich – STOC 1999 & JCSS 2002 & manuscript 1994 (!):
  - $\forall w \; \exists n$ s.t. $\Omega\left(\frac{\lg w}{\lg \lg w}\right)$
  - $\forall n \; \exists w$ s.t. $\Omega\left(\sqrt{\frac{\lg n}{\lg \lg n}}\right)$

  messy
  - static DS with $O\left(\min\left\{\frac{\lg w}{\lg \lg w}, \sqrt{\frac{\lg n}{\lg \lg n}}\right\}\right)$ $\boxed{\text{L15}}$

  $\Rightarrow$ best "pure" bounds in n & w
- Xiao – Ph.D. thesis 1992 @ Stanford
  - same lower bounds! (still messy)

  $\Rightarrow$ Beame & Fich was independent discovery

- Sen – CCC 2003 & arXiv:cs.CC/0309033 with Venkatesh:
  - clean proofs of same bounds
  - uses <u>round elimination</u> & new lemma

TODAY

- Pǎtraşcu & Thorup 2005 & 2006:
  - complete n vs. w vs. space trade-off

$\boxed{L15}$

Colored predecessor problem:

- each element has a <u>color</u> of red or blue
- predecessor/succ. query just needs to return color
- easier problem ⇒ stronger lower bound
- useful for reductions later

Communication complexity viewpoint:

$$\begin{array}{ccc} \text{Alice} & \xrightarrow[\text{msgs.} \le a \text{ bits}]{} & \text{Bob} \\ x & & y \\ & \xleftarrow[\text{msgs.} \le b \text{ bits}]{} & \end{array}$$

goal: $f(x,y)$

- Alice knows input x
  ↳ query algorithm ↳ query
- Bob knows input y
  ↳ DS/memory ↳ contents of DS/memory
- a = # address bits = lg(space) = $\underbrace{O(\lg n) \text{ if space} = n^{O(1)}}_{\text{typical assumption}}$
- b = word size w
- # messages = 2 · # cell probes

Predecessor lower bound: $\Omega(\min\{\log_a w, \log_b n\})$

Beame-Fich-Xiao pure bound:

- assume $a = O(\lg n)$
- LB largest (strongest) when w & n satisfy:

$$\log_a w = \log_b n$$

i.e. $$\frac{\lg w}{\lg\lg n} = \frac{\lg n}{\lg w} \qquad (\Theta)$$

i.e. $$\lg^2 w = \lg n \cdot \lg\lg n$$

i.e. $$\lg w = \sqrt{\lg n \cdot \lg\lg n}$$

$$\Rightarrow \lg\lg n = \lg\lg w$$

LB: $$\frac{\lg w}{\lg\lg n} = \sqrt{\frac{\lg n}{\lg\lg n}} = \frac{\lg w}{\lg\lg w}$$

Round elimination warmup:

$f^{(k)}$: variation on problem f

- Alice has k inputs $x_1, x_2, \ldots, x_k$
- Bob has 2 inputs y, $i \in \{1, 2, \ldots, k\}$
  & already knows $x_1, x_2, \ldots, x_{i-1}$
- goal: compute $f(x_i, y)$

Intuition: first message sent by Alice is $\approx$useless
if a bits $\ll k$ inputs

- unlikely for Alice to send anything useful about x_i

$\Rightarrow$ can start communication protocol @second msg.
i.e. eliminate first message

- repeat in Bob $\to$ Alice direction $\Rightarrow$ eliminate round

Round elimination lemma:

if there is a protocol for $f^{(k)}$
where Alice speaks first
using m messages & error probability δ
then there is a protocol for f
where Bob speaks first
using $m-1$ messages & error probability $\delta + O(\sqrt{a/k})$

Intuition: (not a proof)

- if i were chosen uniformly at random, then expect a/k bits to be "about" x_i
- Bob can guess these bits randomly
- $\Pr\{\text{correct guess}\} = 1/2^{a/k}$

$\Rightarrow$ error increase $= 1 - 1/2^{a/k}$ (union bound)
$\approx a/k$ $(1 - 1/e^x \approx x)$
$< \sqrt{a/k}$ — safer & "more correct"

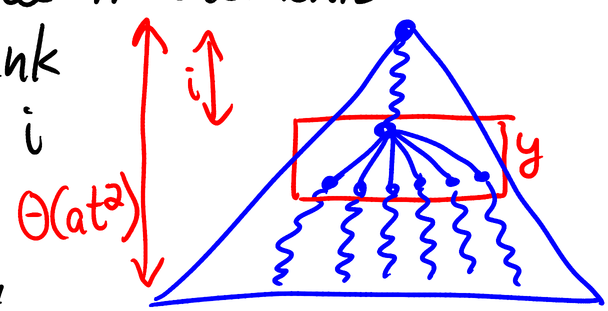
- real proof uses information theory (see below)

Proof of predecessor lower bound:

- let t = # cell probes (rounds) for predecessor
- goal: t round eliminations
  - ⇒ remaining protocol has ∅ messages
  - ⇒ answer must be guessed — assuming $n' \geq 2$
  - ⇒ $\Pr\{\text{success}\} \leq 1/2$
  - get contradiction when t small (error $< 1/2$)

① eliminating message from Alice to Bob:

- Alice's input x has w' bits (initially w)
- break into $k = \Theta(a t^2)$ equal-size chunks $x_1, \dots, x_k$
  - ⇒ error increase from elimination $= O(\sqrt{a/at^2}) = O(1/t)$
- tell Alice & Bob that all n' elements first differ in ith chunk
- but only Bob knows i
- Bob also knows $x_1, x_2, \dots, x_{i-1}$ of query (common prefix of all elements)

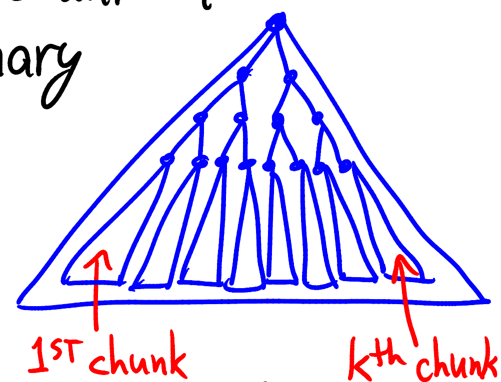


- goal: query x_i in DS y for ith chunk
- ⇒ elimination reduces $w' \to \Theta(w'/at^2)$

ANALOGY: van Emde Boas binary searches on levels to find longest prefix match, reducing w as you go

② eliminating message from Bob to Alice:

- Bob's input is n' integers of w' bits each
- divide integers into $k = \Theta(bt^2)$ equal chunks
  $\Rightarrow$ error increase again $O(1/t)$
- tell Alice & Bob that ith chunk x_i starts with prefix "i" in binary
- goal: search for query in ith chunk x_i containing query y
- only Alice knows i



$\Rightarrow$ elimination reduces $n' \to \Theta(n'/bt^2)$

$w' \to w' - \lg(\underbrace{bt^2}_{\le \lg^2 w})$

negligible: at most $\times 2$ reduction if $w' \ge c \cdot \lg b = c \cdot \lg w'$

$\boxed{\text{ANALOGY: fusion trees branch by polynomial factor in } w\text{, reducing } n}$

- round elimination reduces $w' \to \Theta(w'/at^2)$
  $n' \to \Theta(n'/bt^2)$
- t-round error $\le 1/3$ if set constants right
- stop when w' hits $\lg b$ or when n' hits 2

$\Rightarrow t = \Omega(\min\{\log_{\underbrace{at^2}_{O(a^3)}} w, \log_{\underbrace{bt^2}_{O(b^3)}} n\})$

because $t = O(\lg n)$ & $a \ge \lg n$ | because $t = O(\lg w)$ & $b = w$

$= \Omega(\min\{\log_a w, \log_b n\})$. □

Information-theory basics:

- $H(x)$ = entropy of x
  = # bits to represent x as sample from distrib.
  $= \sum_{x_0} \Pr\{x=x_0\} \cdot \lg \frac{1}{\Pr\{x=x_0\}}$
- $H(x|y)$ = entropy of x given y
  = # bits to represent x if you know y
  $= E_{y_0}[H(x|y=y_0)]$ – propagate into Pr's
- $I(x:y)$ = shared information between x & y
  $= H(x) + H(y) - H((x,y))$
- $I(x:y|z) = E_{z_0}[I(x:y|z=z_0)]$ – prop. into H's

Proof sketch of round-elimination lemma:

- call Alice's first message $m = m(x_1, \dots, x_k)$
- $a = |m| \geq H(m) = \sum_{i=1}^{k} I(x_i : m \mid x_1, \dots, x_{i-1})$

  chain rule for information [info. theory]
- if i is distributed uniformly,

  then $E_i[I(x_i : m \mid x_1, \dots, x_{i-1})]$ = average term in sum $\leq H(m)/k \leq a/k$
- intuition: Bob knows $x_1, \dots, x_{i-1}$ & receives m

  $\Rightarrow$ learns $I(x_i : m)$ about x_i
- build protocol for $f(x)$ as follows:
  - fix $x_1, \dots, x_{i-1}$ & i randomly in advance
  - now query x comes along
  - set $x_i = x$
  - run $f^{(k)}$ protocol, starting at second message, assuming first message $m = m(x_1, \dots, x_{i-1}, \tilde{x}_i, \dots, \tilde{x}_k)$ for $\tilde{x}_i, \dots, \tilde{x}_k$ chosen uniformly by Bob (who doesn't know x_i)
  - guess $I(x_i : m)$ correctly with probability $\approx a/k$
  - claim: with probability $\sqrt{a/k}$,

    $\exists x_{i+1}, \dots, x_k$ such that $m(x_1, \dots, x_{i-1}, \tilde{x}_i, \dots, \tilde{x}_k) = m(x_1, \dots, x_k)$

    distributed roughly the same

    ($\Rightarrow$ error probability δ preserved)

    $\rightarrow$ Average Encoding Theorem