Information Complexity: A Paradigm for Proving Lower Bounds

Amit Chakrabarti

Dartmouth College

Hanover, NH, USA

STOC 2013, Palo Alto

History

The applications came first; theory built in service of applications

History

The <u>applications</u> came first; theory built in service of applications

- Ablayev (*Theor. Comp. Sci., 1996*)

  Proved lower bound for communication problem INDEX

  Used technique that we now recognize as "information complexity"

History

The <u>applications</u> came first; theory built in service of applications

- Ablayev (*Theor. Comp. Sci., 1996*)

  Proved lower bound for communication problem INDEX

  Used technique that we now recognize as "information complexity"

- Chakrabarti, Shi, Wirth, Yao (*FOCS, 2001*)

  Proved direct sum results for simultaneous message complexity

  Formally defined "information cost" and "information complexity"

  Introduced notations $\mathsf{icost}, \mathsf{IC}$

  Anticipated wider applicability of paradigm

History

The <u>applications</u> came first; theory built in service of applications

- Ablayev (*Theor. Comp. Sci., 1996*)

  Proved lower bound for communication problem INDEX

  Used technique that we now recognize as “information complexity”

- Chakrabarti, Shi, Wirth, Yao (*FOCS, 2001*)

  Proved direct sum results for simultaneous message complexity

  Formally defined “information cost” and “information complexity”

  Introduced notations $\mathsf{icost}, \mathsf{IC}$

  Anticipated wider applicability of paradigm

> *“We introduce a new notion of <u>informational complexity</u> which is related to SM complexity and has nice direct sum properties. This notion is used as a tool to prove the above results; it appears to be quite powerful and may be of independent interest.”*

History

The applications came first; theory built in service of applications

- Ablayev (*Theor. Comp. Sci., 1996*)

  Proved lower bound for communication problem INDEX

  Used technique that we now recognize as "information complexity"

- Chakrabarti, Shi, Wirth, Yao (*FOCS, 2001*)

  Proved direct sum results for simultaneous message complexity

  Formally defined "information cost" and "information complexity"

  Introduced notations $\mathsf{icost}, \mathsf{IC}$

  Anticipated wider applicability of paradigm

History

The applications came first; theory built in service of applications

- Ablayev (*Theor. Comp. Sci., 1996*)

  Proved lower bound for communication problem INDEX

  Used technique that we now recognize as "information complexity"

- Chakrabarti, Shi, Wirth, Yao (*FOCS, 2001*)

  Proved direct sum results for simultaneous message complexity

  Formally defined "information cost" and "information complexity"

  Introduced notations icost, IC

  Anticipated wider applicability of paradigm

- Bar-Yossef, Jayram, Kumar, Sivakumar (*FOCS, 2002*)

  Gave extension to interactive communication

  Cleverly handled non-product distributions: "conditional icost"

  Improved some communication (hence data stream) lower bounds

This Talk

Goals:

- Tutorial style
- Diversity of results
- Extract common patterns in applying IC

Not goals:

- Be comprehensive
- Present latest results (but see Woodruff's talk next)

(Generalized) Direct Sum Theorems

Situation:

- Task $\mathscr{A}$: a simple computational task
- Task $\mathscr{B}$: combines N independent copies of task $\mathscr{A}$

Direct sum theorem:

$$\text{Complexity}(\mathscr{B}) \;=\; \Omega(N) \cdot \text{Complexity}(\mathscr{A})$$

The Information Complexity Paradigm

Situation:

- Task $\mathscr{A}$: a simple computational task
- Task $\mathscr{B}$: combines N independent copies of task $\mathscr{A}$

The paradigm:

1. Define information cost
2. Simulation Argument
3. Basic IC lower bound

The Information Complexity Paradigm

Situation:

- Task $\mathscr{A}$: a simple computational task
- Task $\mathscr{B}$: combines N independent copies of task $\mathscr{A}$

The paradigm:

1. Define information cost, hence information complexity (IC)

   Get $\text{Complexity}(\mathscr{B}) \geq \text{IC}(\mathscr{B})$

2. Simulation Argument

3. Basic IC lower bound

The Information Complexity Paradigm

Situation:

- Task $\mathscr{A}$: a simple computational task
- Task $\mathscr{B}$: combines N independent copies of task $\mathscr{A}$

The paradigm:

1. Define information cost, hence information complexity (IC)

   Get $\text{Complexity}(\mathscr{B}) \geq \text{IC}(\mathscr{B})$

2. Simulation Argument: solving $\mathscr{B} \Rightarrow$ solving each copy of $\mathscr{A}$

   Get $\text{IC}(\mathscr{B}) \geq N \cdot \text{IC}(\mathscr{A})$

3. Basic IC lower bound

The Information Complexity Paradigm

Situation:

- Task $\mathscr{A}$: a simple computational task
- Task $\mathscr{B}$: combines N independent copies of task $\mathscr{A}$

The paradigm:

1. Define information cost, hence information complexity (IC)

   Get $\text{Complexity}(\mathscr{B}) \geq \text{IC}(\mathscr{B})$

2. Simulation Argument: solving $\mathscr{B} \Rightarrow$ solving each copy of $\mathscr{A}$

   Get $\text{IC}(\mathscr{B}) \geq N \cdot \text{IC}(\mathscr{A})$

3. Basic IC lower bound: apply to simple task $\mathscr{A}$

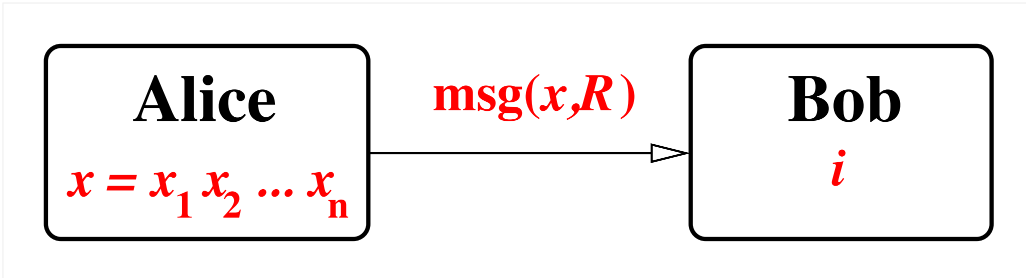
   Get $\text{IC}(\mathscr{A}) \gtrsim \text{Complexity}(\mathscr{A})$

Part One:
No Interaction

The INDEX Problem

Definition:

Alice holds $x \in \{0,1\}^n$, Bob holds $i \in [n]$; find x_i (error $\leq \varepsilon$)



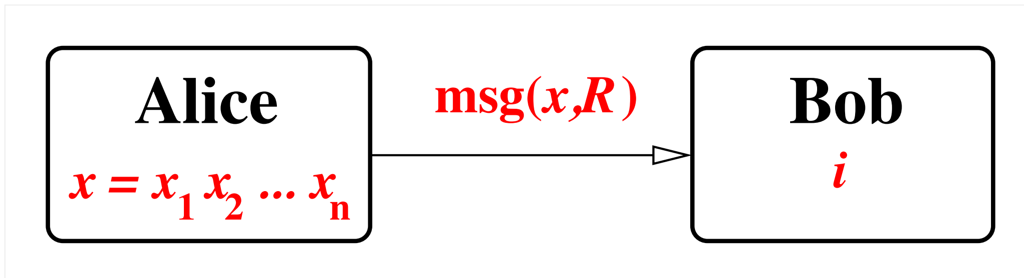
Correctness requirement:

$$\forall\, x, i \quad \Pr[\text{output} \neq x_i] \leq \varepsilon$$

The INDEX Problem

Definition:

Alice holds $x \in \{0,1\}^n$, Bob holds $i \in [n]$; find x_i (error $\leq \varepsilon$)



Correctness requirement:

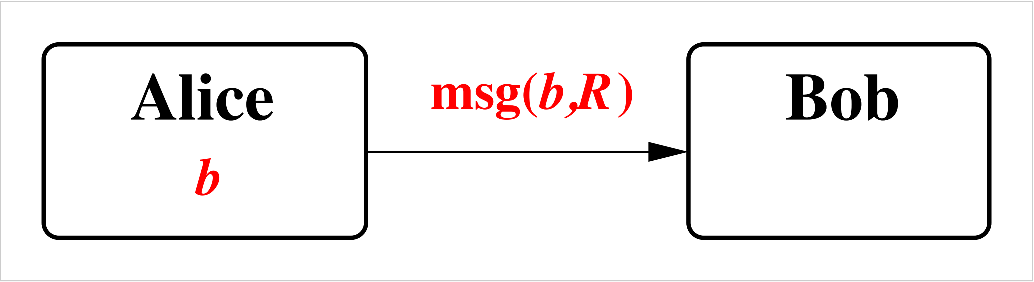
$$\forall\, x, i \quad \Pr[\text{output} \neq x_i] \leq \varepsilon$$

Theorem: Alice needs to send $\Omega(n)$ bits. [Ablayev'96]

Information Complexity Paradigm Demo

The ECHO problem:

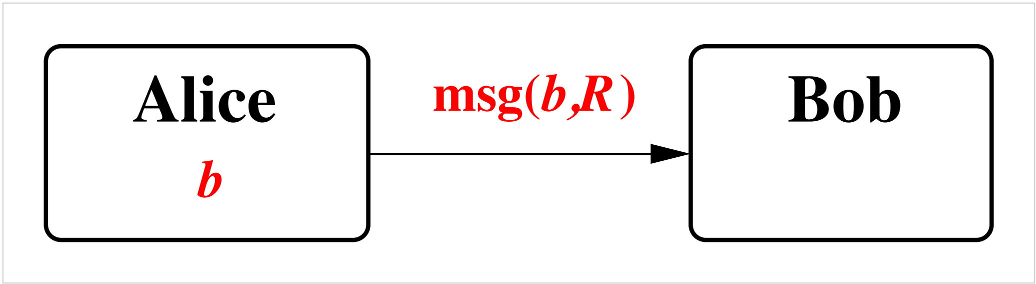
Alice holds $b \in \{0,1\}$, Bob to output b with error $\leq \varepsilon$



Information Complexity Paradigm Demo

The ECHO problem:

Alice holds $b \in \{0,1\}$, Bob to output b with error $\le \varepsilon$



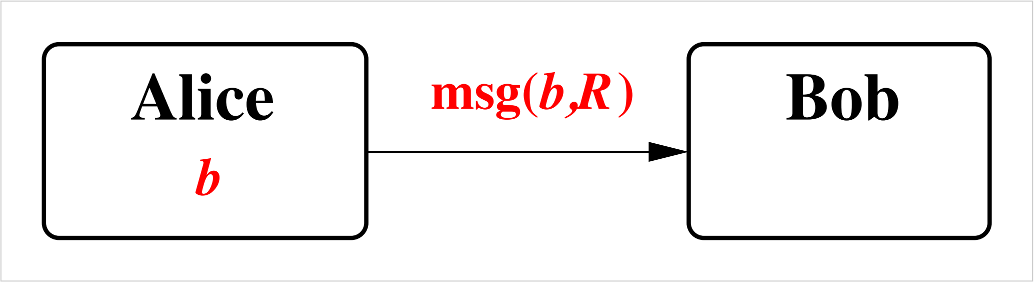
Simple task ($\mathscr{A}$): ECHO

Complex task ($\mathscr{B}$): INDEX

Information Complexity Paradigm Demo

The ECHO problem:

Alice holds $b \in \{0,1\}$, Bob to output b with error $\leq \varepsilon$



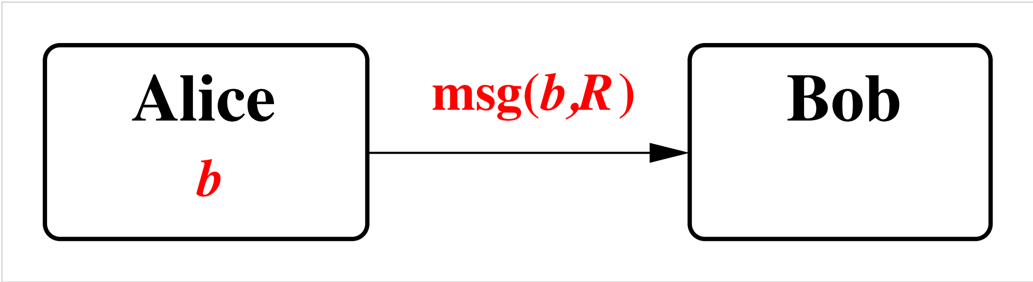
Simple task ($\mathscr{A}$): ECHO

Complex task ($\mathscr{B}$): INDEX $\approx$ combines n independent copies of ECHO

Information Complexity Paradigm Demo

The ECHO problem:

Alice holds $b \in \{0,1\}$, Bob to output b with error $\leq \varepsilon$



Simple task ($\mathscr{A}$): ECHO

Complex task ($\mathscr{B}$): INDEX $\approx$ combines n independent copies of ECHO

The paradigm

1. Define information cost
2. Simulation Argument
3. Basic IC lower bound (for ECHO)

Step 1: Define Information Cost

Generic notion, for a communication protocol Π:

$$\text{icost}(\Pi) = \begin{array}{l}\text{amount of info about (part of) the input to } \Pi \\ \text{revealed by (some of) the messages in } \Pi \\ \text{(possibly conditioned on some prior knowledge)}\end{array}$$

Step 1: Define Information Cost

Generic notion, for a communication protocol Π:

$$\text{icost}(\Pi) = \text{amount of info about (part of) the input to } \Pi$$
$$\text{revealed by (some of) the messages in } \Pi$$
$$\text{(possibly conditioned on some prior knowledge)}$$

In this case...

- Let Π_B be a protocol for task $\mathscr{B}$ (i.e., $\textsc{index}$)
- Let X = random input (distrib μ) for Alice
  R = random coins of Alice
  $M = \text{msg}(X, R)$; then

$$\text{icost}^{\mu}(\Pi_B) := \mathrm{I}(X : M)$$

- Notice:

$$\text{icost}^{\mu}(\Pi_B) \le \mathrm{H}(M) \le \text{length}(M) = \text{cost}(\Pi_B)$$

Step 2: Simulation Argument

Take $X = X_1 \ldots X_n \sim \mu_1 \otimes \cdots \otimes \mu_n =: \mu$; then $X_1, \ldots, X_n$ independent

$$
\begin{aligned}
\mathsf{cost}(\Pi_B) &\geq \mathsf{icost}^{\mu}(\Pi_B) \\
&= \mathrm{I}(X_1 X_2 \ldots X_n : M) \\
&\geq \mathrm{I}(X_1 : M) + \mathrm{I}(X_2 : M) + \cdots + \mathrm{I}(X_n : M) \qquad \text{[superadditivity]}
\end{aligned}
$$

Step 2: Simulation Argument

Take $X = X_1 \dots X_n \sim \mu_1 \otimes \cdots \otimes \mu_n =: \mu$; then $X_1, \dots, X_n$ independent

$$\begin{aligned} \mathsf{cost}(\Pi_B) &\geq \mathsf{icost}^{\mu}(\Pi_B) \\ &= \mathrm{I}(X_1 X_2 \dots X_n : M) \\ &\geq \mathrm{I}(X_1 : M) + \mathrm{I}(X_2 : M) + \cdots + \mathrm{I}(X_n : M) && \text{[superadditivity]} \\ &= \mathsf{icost}^{\mu_1}(\Pi_{A,1}) + \mathsf{icost}^{\mu_2}(\Pi_{A,2}) + \cdots + \mathsf{icost}^{\mu_n}(\Pi_{A,n}) \end{aligned}$$

To make this work, want protocols $\Pi_{A,j}$ s.t.

$M \equiv$ Alice's message in $\Pi_{A,j}$ on input $X_j \sim \mu_j$

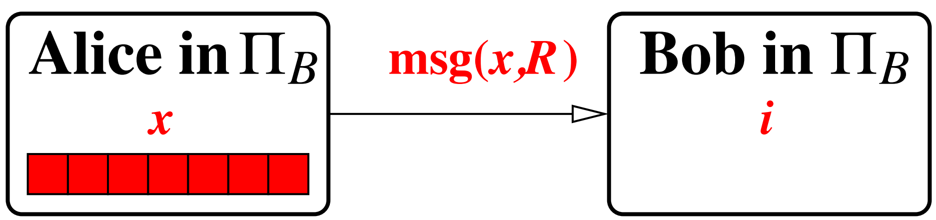
Step 2: Simulation Argument

Take $X = X_1 \ldots X_n \sim \mu_1 \otimes \cdots \otimes \mu_n =: \mu$; then $X_1, \ldots, X_n$ independent

$$
\begin{aligned}
\mathsf{cost}(\Pi_B) &\geq \mathsf{icost}^{\mu}(\Pi_B) \\
&= \mathrm{I}(X_1 X_2 \ldots X_n : M) \\
&\geq \mathrm{I}(X_1 : M) + \mathrm{I}(X_2 : M) + \cdots + \mathrm{I}(X_n : M) \qquad \text{[superadditivity]} \\
&= \mathsf{icost}^{\mu_1}(\Pi_{A,1}) + \mathsf{icost}^{\mu_2}(\Pi_{A,2}) + \cdots + \mathsf{icost}^{\mu_n}(\Pi_{A,n})
\end{aligned}
$$

To make this work, want protocols $\Pi_{A,j}$ s.t.

$M \equiv$ Alice's message in $\Pi_{A,j}$ on input $X_j \sim \mu_j$



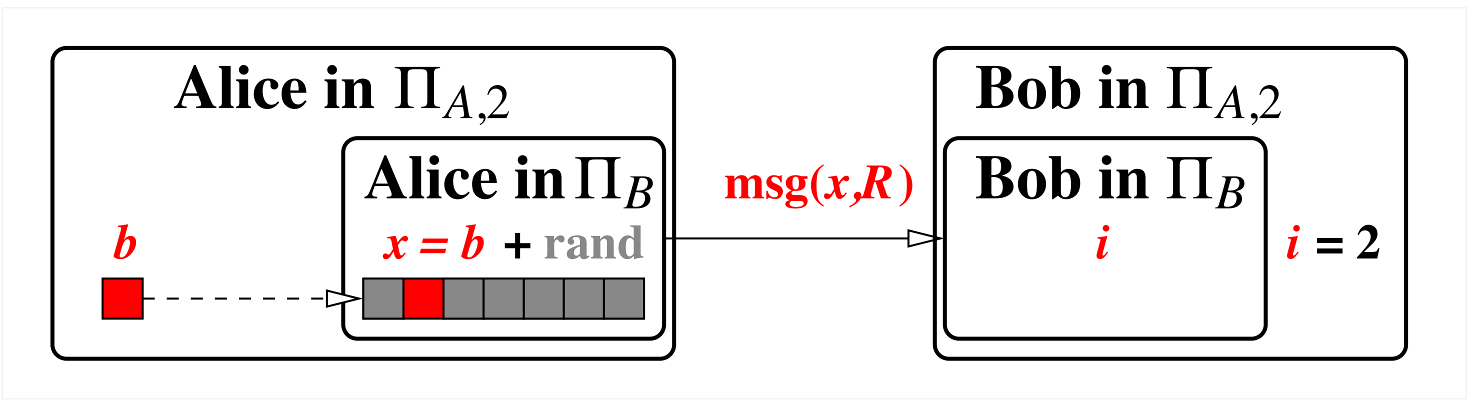
Step 2: Simulation Argument

Take $X = X_1 \ldots X_n \sim \mu_1 \otimes \cdots \otimes \mu_n =: \mu$;   then $X_1, \ldots, X_n$ independent

$$
\begin{aligned}
\mathsf{cost}(\Pi_B) &\geq \mathsf{icost}^{\mu}(\Pi_B) \\
&= \mathrm{I}(X_1 X_2 \ldots X_n : M) \\
&\geq \mathrm{I}(X_1 : M) + \mathrm{I}(X_2 : M) + \cdots + \mathrm{I}(X_n : M) \qquad \text{[superadditivity]} \\
&= \mathsf{icost}^{\mu_1}(\Pi_{A,1}) + \mathsf{icost}^{\mu_2}(\Pi_{A,2}) + \cdots + \mathsf{icost}^{\mu_n}(\Pi_{A,n})
\end{aligned}
$$

To make this work, want protocols $\Pi_{A,j}$ s.t.

$M \equiv$ Alice's message in $\Pi_{A,j}$ on input $X_j \sim \mu_j$



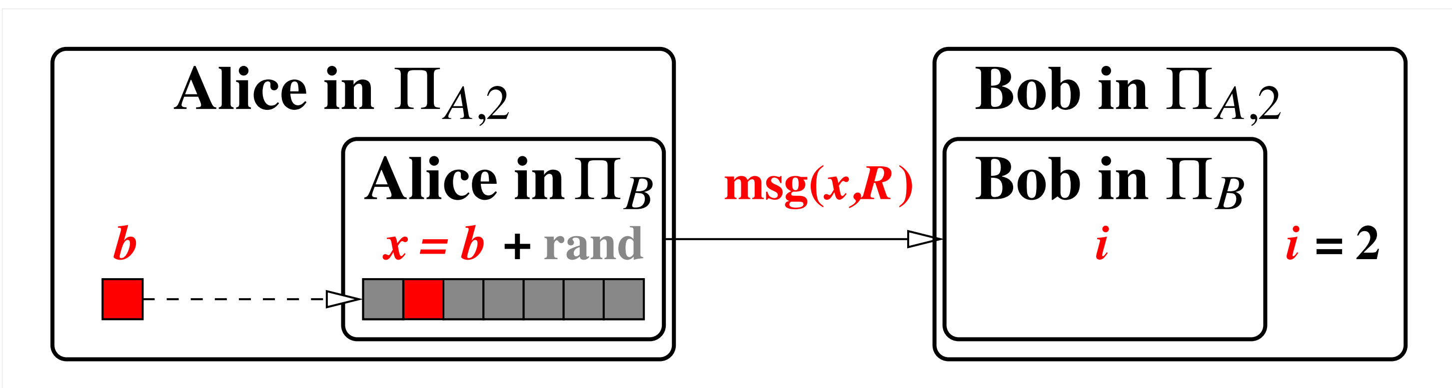
Step 2: Simulation Argument

Take $X = X_1 \ldots X_n \sim \mu_1 \otimes \cdots \otimes \mu_n =: \mu$; then $X_1, \ldots, X_n$ independent

$$\begin{aligned}
\mathsf{cost}(\Pi_B) &\geq \mathsf{icost}^{\mu}(\Pi_B) \\
&= \mathrm{I}(X_1 X_2 \ldots X_n : M) \\
&\geq \mathrm{I}(X_1 : M) + \mathrm{I}(X_2 : M) + \cdots + \mathrm{I}(X_n : M) \qquad \text{[superadditivity]} \\
&= \mathsf{icost}^{\mu_1}(\Pi_{A,1}) + \mathsf{icost}^{\mu_2}(\Pi_{A,2}) + \cdots + \mathsf{icost}^{\mu_n}(\Pi_{A,n})
\end{aligned}$$

To make this work, want protocols $\Pi_{A,j}$ s.t.

$M \equiv$ Alice's message in $\Pi_{A,j}$ on input $X_j \sim \mu_j$



Notice: each $\Pi_{A,j}$ solves ECHO

Step 3: Basic IC Lower Bound

Comm complexity: $\mathrm{R}_{\varepsilon}^{\to}(\mathscr{B}) = \min \{\mathsf{cost}(\Pi) : \Pi \text{ solves } \mathscr{B} \text{ with error } \varepsilon\}$

Info complexity: $\mathsf{IC}_{\varepsilon}^{\mu,\to}(\mathscr{B}) = \inf \{\mathsf{icost}^{\mu}(\Pi) : \Pi \text{ solves } \mathscr{B} \text{ with error } \varepsilon\}$

Pick $\mu =$ uniform distrib, ξ; so far

$$\mathrm{R}_{\varepsilon}^{\to}(\mathscr{B}) \;\geq\; \mathsf{IC}_{\varepsilon}^{\xi,\to}(\mathscr{B})$$

$$\mathsf{IC}_{\varepsilon}^{\xi,\to}(\mathscr{B}) \;\geq\; n \cdot \mathsf{IC}^{\xi,\to}(\mathscr{A})$$

Step 3: Basic IC Lower Bound

Comm complexity: $\mathrm{R}_{\varepsilon}^{\rightarrow}(\mathscr{B}) = \min\{\textsf{cost}(\Pi) : \Pi \textsf{ solves } \mathscr{B} \textsf{ with error } \varepsilon\}$

Info complexity: $\mathsf{IC}_{\varepsilon}^{\mu,\rightarrow}(\mathscr{B}) = \inf\{\textsf{icost}^{\mu}(\Pi) : \Pi \textsf{ solves } \mathscr{B} \textsf{ with error } \varepsilon\}$

Pick $\mu =$ uniform distrib, ξ; so far

$$\mathrm{R}_{\varepsilon}^{\rightarrow}(\mathscr{B}) \;\geq\; \mathsf{IC}_{\varepsilon}^{\xi,\rightarrow}(\mathscr{B})$$

$$\mathsf{IC}_{\varepsilon}^{\xi,\rightarrow}(\mathscr{B}) \;\geq\; n \cdot \mathsf{IC}^{\xi,\rightarrow}(\mathscr{A})$$

Intuitively clear that $\mathsf{IC}_{\varepsilon}^{\mu,\rightarrow}(\mathscr{A}) \neq 0$

Implication: $\mathrm{R}_{\varepsilon}^{\rightarrow}(\mathscr{B}) = \Omega(n)$. QED

Step 3: Basic IC Lower Bound

Comm complexity: $\mathrm{R}_\varepsilon^{\to}(\mathscr{B}) = \min \{\textsf{cost}(\Pi) : \Pi \textsf{ solves } \mathscr{B} \textsf{ with error } \varepsilon\}$

Info complexity: $\textsf{IC}_\varepsilon^{\mu,\to}(\mathscr{B}) = \inf \{\textsf{icost}^\mu(\Pi) : \Pi \textsf{ solves } \mathscr{B} \textsf{ with error } \varepsilon\}$

Pick $\mu =$ uniform distrib, ξ; so far

$$\mathrm{R}_\varepsilon^{\to}(\mathscr{B}) \geq \textsf{IC}_\varepsilon^{\xi,\to}(\mathscr{B})$$

$$\textsf{IC}_\varepsilon^{\xi,\to}(\mathscr{B}) \geq n \cdot \textsf{IC}^{\xi,\to}(\mathscr{A})$$

Intuitively clear that $\textsf{IC}_\varepsilon^{\mu,\to}(\mathscr{A}) \neq 0$

Implication: $\mathrm{R}_\varepsilon^{\to}(\mathscr{B}) = \Omega(n)$. QED

In fact we can work out the constant precisely...

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., ECHO)

- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$

$$\mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M)$$

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., ECHO)
- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$

$$\mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M)$$

$\mathrm{D}_{\mathrm{TV}}(P, Q)$	:	total variation distance
$\mathrm{D}_{\mathrm{KL}}(P \| Q)$	:	Kullback-Leibler divergence
$\mathrm{D}_{\mathrm{JS}}(P, Q)$	:	Jensen-Shannon divergence
$\mathrm{H}_b(x)$	:	binary entropy function
		$= -x \log x - (1 - x) \log(1 - x)$

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., ECHO)
- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$

$$\mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M)$$

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., ECHO)
- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$

- Basic information theory:

$$\mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M) = \tfrac{1}{2}(\mathrm{D_{KL}}(M^{(0)} \| M) + \mathrm{D_{KL}}(M^{(1)} \| M))$$

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., ECHO)
- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$
- Basic information theory:

$$\begin{aligned} \mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M) &= \tfrac{1}{2}(\mathrm{D_{KL}}(M^{(0)} \| M) + \mathrm{D_{KL}}(M^{(1)} \| M)) \\ &= \mathrm{D_{JS}}(M^{(0)}, M^{(1)}) \\ &\geq 1 - \mathrm{H}_b\left(\frac{1 - \mathrm{D_{TV}}(M^{(0)}, M^{(1)})}{2}\right) \end{aligned}$$

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., ECHO)
- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$
- Basic information theory:

$$
\begin{aligned}
\mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M) &= \tfrac{1}{2}(\mathrm{D}_{\mathrm{KL}}(M^{(0)} \| M) + \mathrm{D}_{\mathrm{KL}}(M^{(1)} \| M)) \\
&= \mathrm{D}_{\mathrm{JS}}(M^{(0)}, M^{(1)}) \\
&\geq 1 - \mathrm{H}_b\left(\frac{1 - \mathrm{D}_{\mathrm{TV}}(M^{(0)}, M^{(1)})}{2}\right)
\end{aligned}
$$

- Error $\leq \varepsilon$ implies $\mathrm{D}_{\mathrm{TV}}(M^{(0)}, M^{(1)}) \geq 1 - 2\varepsilon$

Step 3: Basic IC Lower Bound: Details

- Let Π_A be a protocol for task $\mathscr{A}$ (i.e., $\textsc{echo}$)
- Let Z = random input (uniform distrib ξ) for Alice

  $M = \mathsf{msg}(Z, R)$

  $M^{(z)} = \mathsf{msg}(z, R)$ for $z \in \{0, 1\}$
- Basic information theory:

$$\begin{aligned} \mathsf{icost}^{\xi}(\Pi_A) = \mathrm{I}(Z : M) &= \tfrac{1}{2}(\mathrm{D_{KL}}(M^{(0)} \| M) + \mathrm{D_{KL}}(M^{(1)} \| M)) \\ &= \mathrm{D_{JS}}(M^{(0)}, M^{(1)}) \\ &\geq 1 - \mathrm{H}_b\left(\frac{1 - \mathrm{D_{TV}}(M^{(0)}, M^{(1)})}{2}\right) \end{aligned}$$

- Error $\leq \varepsilon$ implies $\mathrm{D_{TV}}(M^{(0)}, M^{(1)}) \geq 1 - 2\varepsilon$
- Thus $\mathsf{icost}^{\xi}(\Pi_A) \geq 1 - \mathrm{H}_b(\varepsilon)$

  and so $\mathrm{R}_{\varepsilon}^{\rightarrow}(\textsc{index}) \geq (1 - \mathrm{H}_b(\varepsilon))n$ ... a tight bound!

The INDEX Problem: Applications

A humble lower bound, but with many applications!

- Complexity of sampling procedures
- Lower bounds for succinct data structures
- Space lower bounds for (one-pass) data stream algorithms
  - Median of n numbers: $\Omega(n)$
  - Mode of n numbers: $\Omega(n)$
  - Connectivity of n-vertex graph, given edges: $\Omega(n)$
  - Triangle-freeness of n-vertex graph: $\Omega(n^2)$

$$\vdots$$

The INDEX Problem: Applications

A humble lower bound, but with many applications!

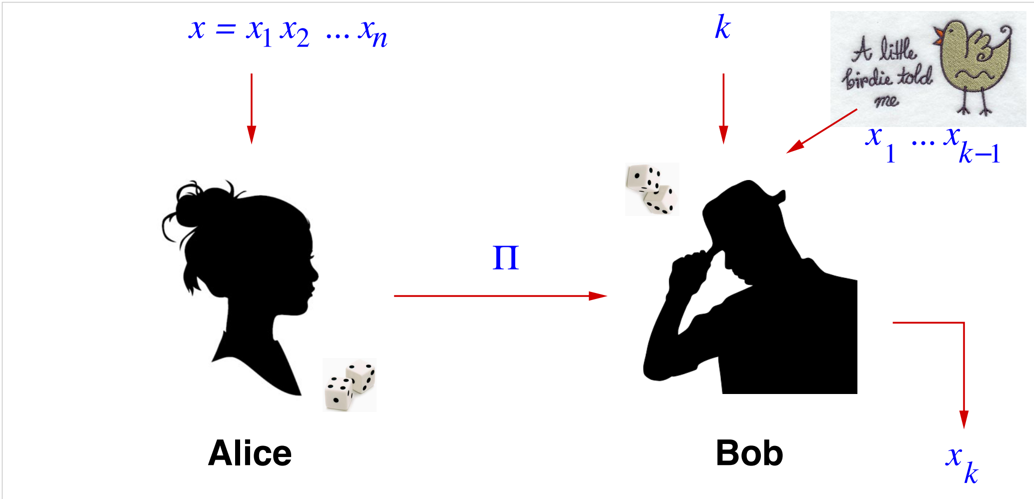
- Complexity of sampling procedures
- Lower bounds for succinct data structures
- Space lower bounds for (one-pass) data stream algorithms
  - Median of n numbers: $\Omega(n)$
  - Mode of n numbers: $\Omega(n)$
  - Connectivity of n-vertex graph, given edges: $\Omega(n)$
  - Triangle-freeness of n-vertex graph: $\Omega(n^2)$

$$\vdots$$

  - Diameter of n-vertex graph, k-approx: $\Omega(n^{1+1/k})$
    A very sophisticated reduction [Feigenbaum-K-M-S-Z'05]

The INDEX Problem: Extensions

- Generalize to AUGMENTED-INDEX

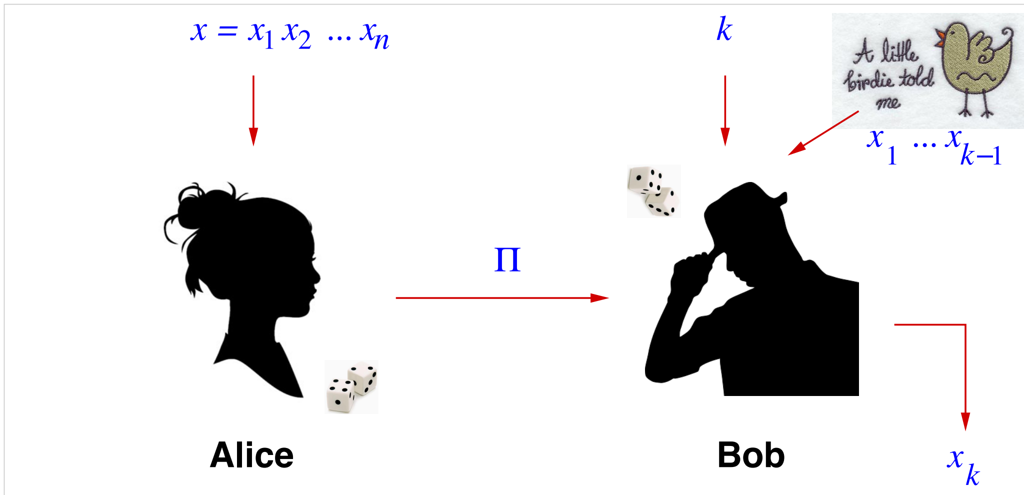


- Still $\Omega(n)$; replace superadditivity step with chain rule:

$$\mathrm{I}(X_1 X_2 \ldots X_n : M) = \sum_{i=1}^{n} \mathrm{I}(X_i : M \mid X_1 \ldots X_{i-1})$$

The INDEX Problem: Extensions

- Generalize to AUGMENTED-INDEX

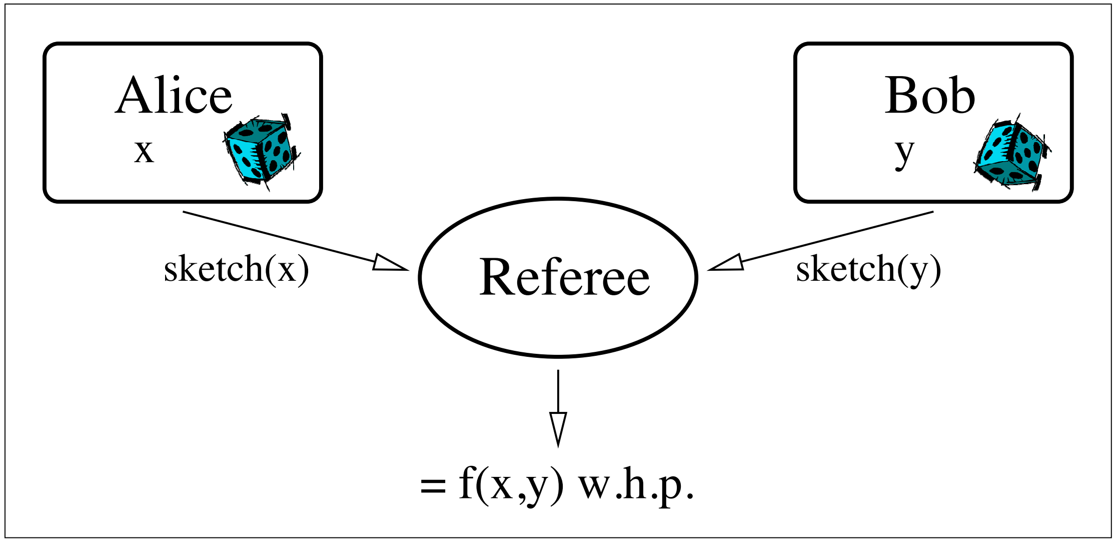


- Still $\Omega(n)$; replace superadditivity step with chain rule:

$$\mathrm{I}(X_1 X_2 \ldots X_n : M) = \sum_{i=1}^{n} \mathrm{I}(X_i : M \mid X_1 \ldots X_{i-1})$$

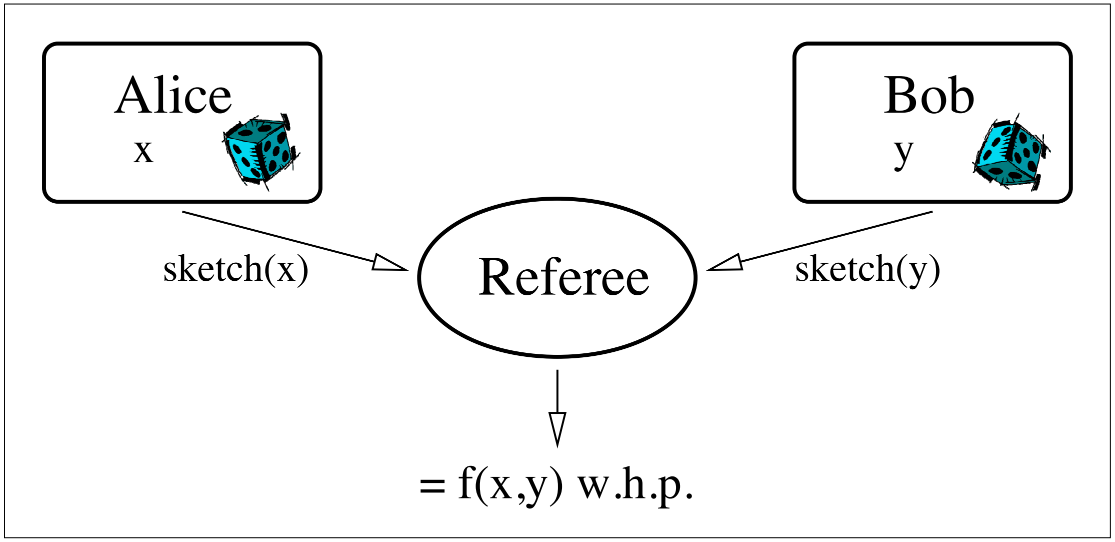
- Generalize to **interactive** communication
  - Communication complexity drops to $O(\log n)$
  - Seek tradeoffs [Magniez-Mathieu-Nayak'10], [C.-Kondapally'11]

Simultaneous Message Communication



Lower bound method: $\mathrm{R}^{\|}(f) := \mathrm{R}^{\|}_{1/3}(f) = \Omega(\sqrt{\mathrm{D}^{\|}(f)})$ [Babai-Kimmel'97]

Simultaneous Message Communication

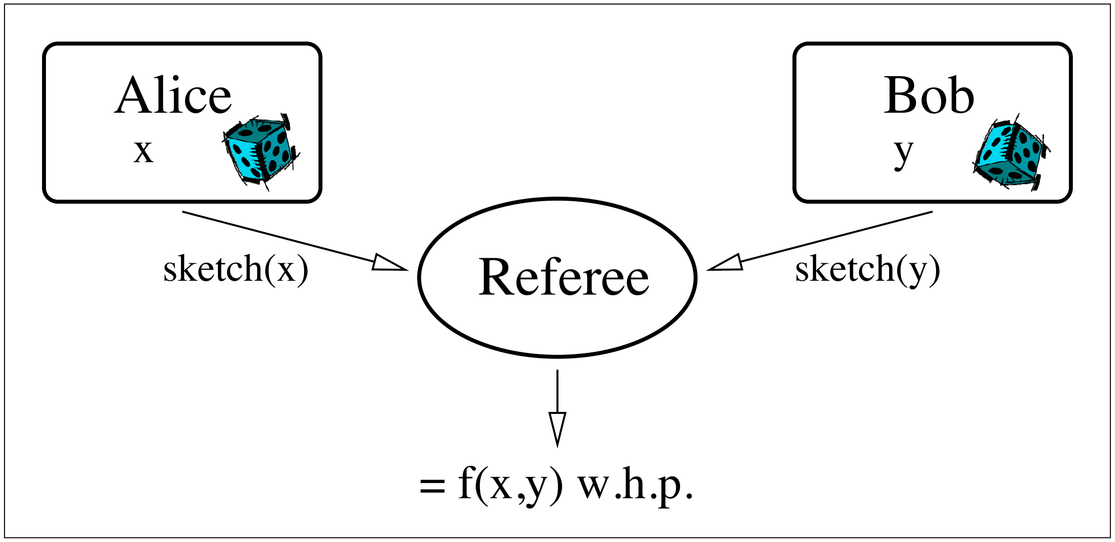


Lower bound method: $\mathrm{R}^{\|}(f) := \mathrm{R}^{\|}_{1/3}(f) = \Omega(\sqrt{\mathrm{D}^{\|}(f)})$ [Babai-Kimmel'97]

Equality function $\mathrm{EQ}_n(x, y) = 1 \iff x = y$ ... $x, y \in \{0, 1\}^n$

A neat result: $\mathrm{R}^{\|}(\mathrm{EQ}_n) = \Theta(\sqrt{n})$ [Ambainis'96]

Simultaneous Message Communication



Lower bound method: $\mathrm{R}^{\|}(f) := \mathrm{R}^{\|}_{1/3}(f) = \Omega(\sqrt{\mathrm{D}^{\|}(f)})$ [Babai-Kimmel'97]

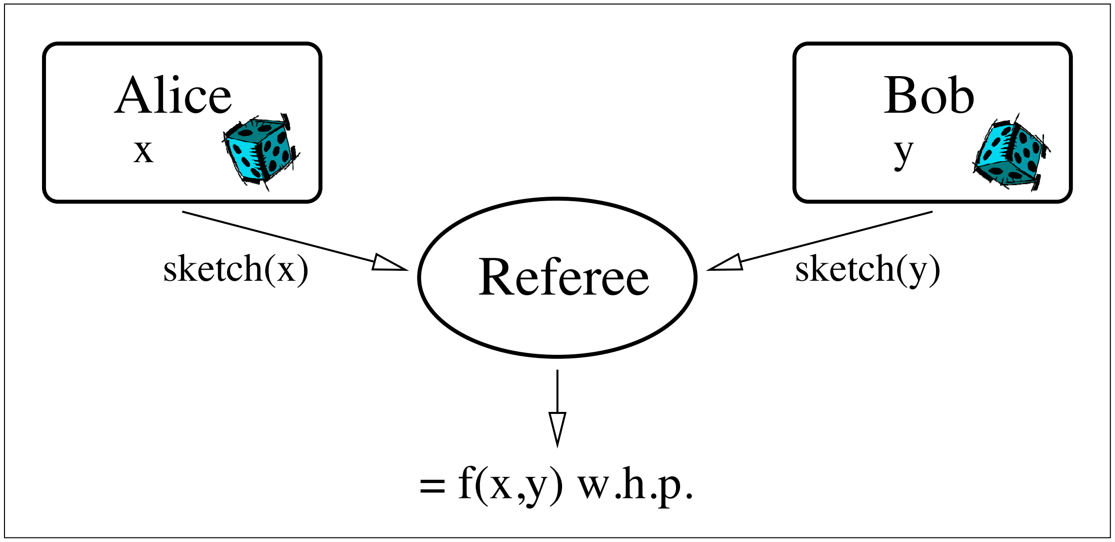
Equality function $\mathrm{EQ}_n(x,y) = 1 \iff x = y$ ... $x, y \in \{0,1\}^n$

A neat result: $\mathrm{R}^{\|}(\mathrm{EQ}_n) = \Theta(\sqrt{n})$ [Ambainis'96]

Direct sum: $\mathrm{OREQ}_{n,m}(x_1 \ldots x_m, y_1 \ldots y_m) = \bigvee_{i=1}^{m} \mathrm{EQ}_n(x_i, y_i)$

What is $\mathrm{R}^{\|}(\mathrm{OREQ}_{n,m})$? Above method only shows $\Omega(\sqrt{mn})$

Simultaneous Message Communication



Lower bound method: $\mathrm{R}^{\|}(f) := \mathrm{R}^{\|}_{1/3}(f) = \Omega(\sqrt{\mathrm{D}^{\|}(f)})$ [Babai-Kimmel'97]

Equality function $\mathrm{EQ}_n(x,y) = 1 \iff x = y$ ... $x, y \in \{0,1\}^n$

A neat result: $\mathrm{R}^{\|}(\mathrm{EQ}_n) = \Theta(\sqrt{n})$ [Ambainis'96]

Direct sum: $\mathrm{OREQ}_{n,m}(x_1 \ldots x_m, y_1 \ldots y_m) = \bigvee_{i=1}^{m} \mathrm{EQ}_n(x_i, y_i)$

What is $\mathrm{R}^{\|}(\mathrm{OREQ}_{n,m})$? Above method only shows $\Omega(\sqrt{mn})$

Theorem: (via IC) $\mathrm{R}^{\|}(\mathrm{OREQ}_{n,m}) = \Omega(m\sqrt{n})$ [C.-Shi-Wirth-Yao'01]

IC Paradigm: Second Demo

Alice: input $X \sim \xi$, message U; Bob: input $Y \sim \xi$, message V

1. Define information cost

2. Simulation Argument

3. Basic IC lower bound

IC Paradigm: Second Demo

Alice: input $X \sim \xi$, message U; Bob: input $Y \sim \xi$, message V

1. Define information cost

   $\text{icost}^{\xi}(\Pi) = \mathrm{I}(X : U) + \mathrm{I}(Y : V)$

2. Simulation Argument

   To solve EQ_n by simulating protocol for $\mathrm{OREQ}_{n,m}$

   Alice, Bob plug input into ith position, fill rest at random $\sim \xi$

   May change answer from 0 to 1 w.p. $\leq m/2^n = o(1)$

3. Basic IC lower bound

IC Paradigm: Second Demo

Alice: input $X \sim \xi$, message U; Bob: input $Y \sim \xi$, message V

1. Define information cost

   $$\text{icost}^{\xi}(\Pi) = \mathrm{I}(X : U) + \mathrm{I}(Y : V)$$

2. Simulation Argument

   To solve $\textsc{eq}_n$ by simulating protocol for $\textsc{oreq}_{n,m}$

   Alice, Bob plug input into ith position, fill rest at random $\sim \xi$

   May change answer from 0 to 1 w.p. $\leq m/2^n = o(1)$

3. Basic IC lower bound (for $\textsc{eq}$)

   So far: $\mathrm{R}^{\parallel}(\textsc{oreq}_{n,m}) \geq \text{IC}^{\xi,\parallel}(\textsc{oreq}_{n,m}) \geq m \cdot \text{IC}^{\xi,\parallel}(\textsc{eq}_n)$

   Must show $\text{IC}^{\xi,\parallel}(\textsc{eq}_n) \gtrsim \mathrm{R}^{\parallel}(\textsc{eq}_n)$

IC Paradigm: Second Demo

Alice: input $X \sim \xi$, message U; Bob: input $Y \sim \xi$, message V

1. Define information cost

   $\text{icost}^{\xi}(\Pi) = \mathrm{I}(X : U) + \mathrm{I}(Y : V)$

2. Simulation Argument

   To solve $\textsc{eq}_n$ by simulating protocol for $\textsc{oreq}_{n,m}$

   Alice, Bob plug input into ith position, fill rest at random $\sim \xi$

   May change answer from 0 to 1 w.p. $\leq m/2^n = o(1)$

3. Basic IC lower bound (for $\textsc{eq}$)

   So far: $\mathrm{R}^{\|}(\textsc{oreq}_{n,m}) \geq \mathsf{IC}^{\xi,\|}(\textsc{oreq}_{n,m}) \geq m \cdot \mathsf{IC}^{\xi,\|}(\textsc{eq}_n)$

   Must show $\mathsf{IC}^{\xi,\|}(\textsc{eq}_n) \gtrsim \mathrm{R}^{\|}(\textsc{eq}_n)$

For last step: compress Alice's/Bob's messages down to their info content

Main idea: Whittle down message space via rejection sampling [CSWY'01]

IC Paradigm: Second Demo

Alice: input $X \sim \xi$, message U; Bob: input $Y \sim \xi$, message V

1. Define information cost

   $\text{icost}^{\xi}(\Pi) = \mathrm{I}(X : U) + \mathrm{I}(Y : V)$

2. Simulation Argument

   To solve EQ_n by simulating protocol for $\text{OREQ}_{n,m}$

   Alice, Bob plug input into ith position, fill rest at random $\sim \xi$

   May change answer from 0 to 1 w.p. $\leq m/2^n = o(1)$

3. Basic IC lower bound (for EQ)

   So far: $\mathrm{R}^{\|}(\text{OREQ}_{n,m}) \geq \text{IC}^{\xi,\|}(\text{OREQ}_{n,m}) \geq m \cdot \text{IC}^{\xi,\|}(\text{EQ}_n)$

   Must show $\text{IC}^{\xi,\|}(\text{EQ}_n) \gtrsim \mathrm{R}^{\|}(\text{EQ}_n)$

For last step: compress Alice's/Bob's messages down to their info content

Main idea: Whittle down message space via rejection sampling [CSWY'01]

Deeper version of idea: comm complexity of correlation [Harsha-J-M-R'07]

Part Two:
Interaction, But Not Really

Lower Bounds for Data Structures

Preprocess data Y $\rightarrow$ data structure $T = T(Y)$ ... low storage space

Query x $\rightarrow$ algorithm $\mathcal{A}(x, T)$ $\rightarrow$ output z ... low query time

Satisfying some relation $R(x, Y, z)$.

Examples: take $x \in \{0, 1\}^d$, $Y \subseteq \{0, 1\}^d$ with $|Y| = n$

- Predecessor Search

  Treat data as d-bit integers

  $R(x, Y, z)$ iff $z \in Y$ is the predecessor of x in Y.

Lower Bounds for Data Structures

Preprocess data Y $\rightarrow$ data structure $T = T(Y)$ ... low storage space

Query x $\rightarrow$ algorithm $\mathcal{A}(x, T)$ $\rightarrow$ output z ... low query time

Satisfying some relation $R(x, Y, z)$.

Examples: take $x \in \{0,1\}^d$, $Y \subseteq \{0,1\}^d$ with $|Y| = n$

- Predecessor Search

  Treat data as d-bit integers

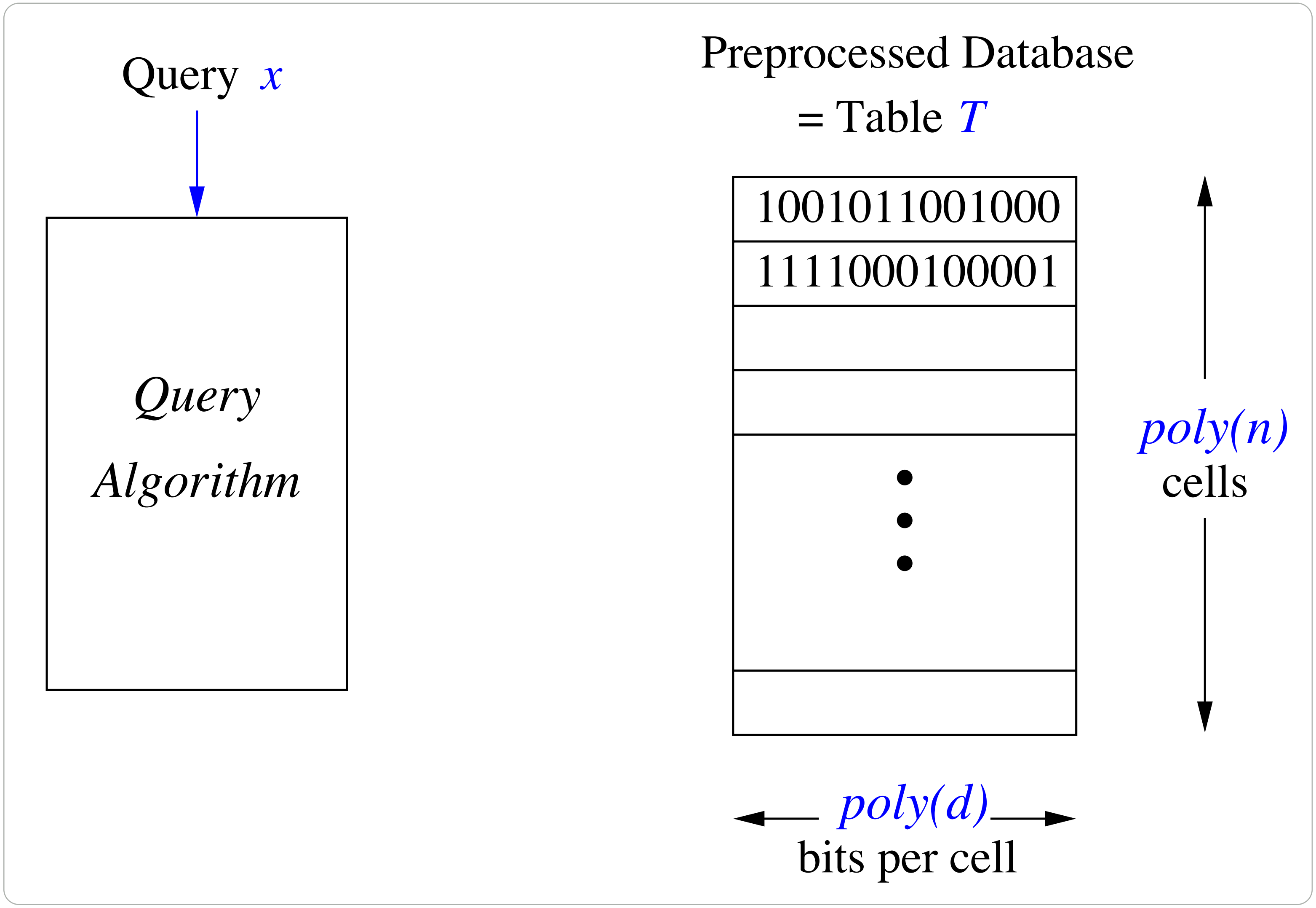
  $R(x, Y, z)$ iff $z \in Y$ is the predecessor of x in Y.

- Approx Nearest Neighbor (ANN) Search

  Treat data as points in Hamming cube

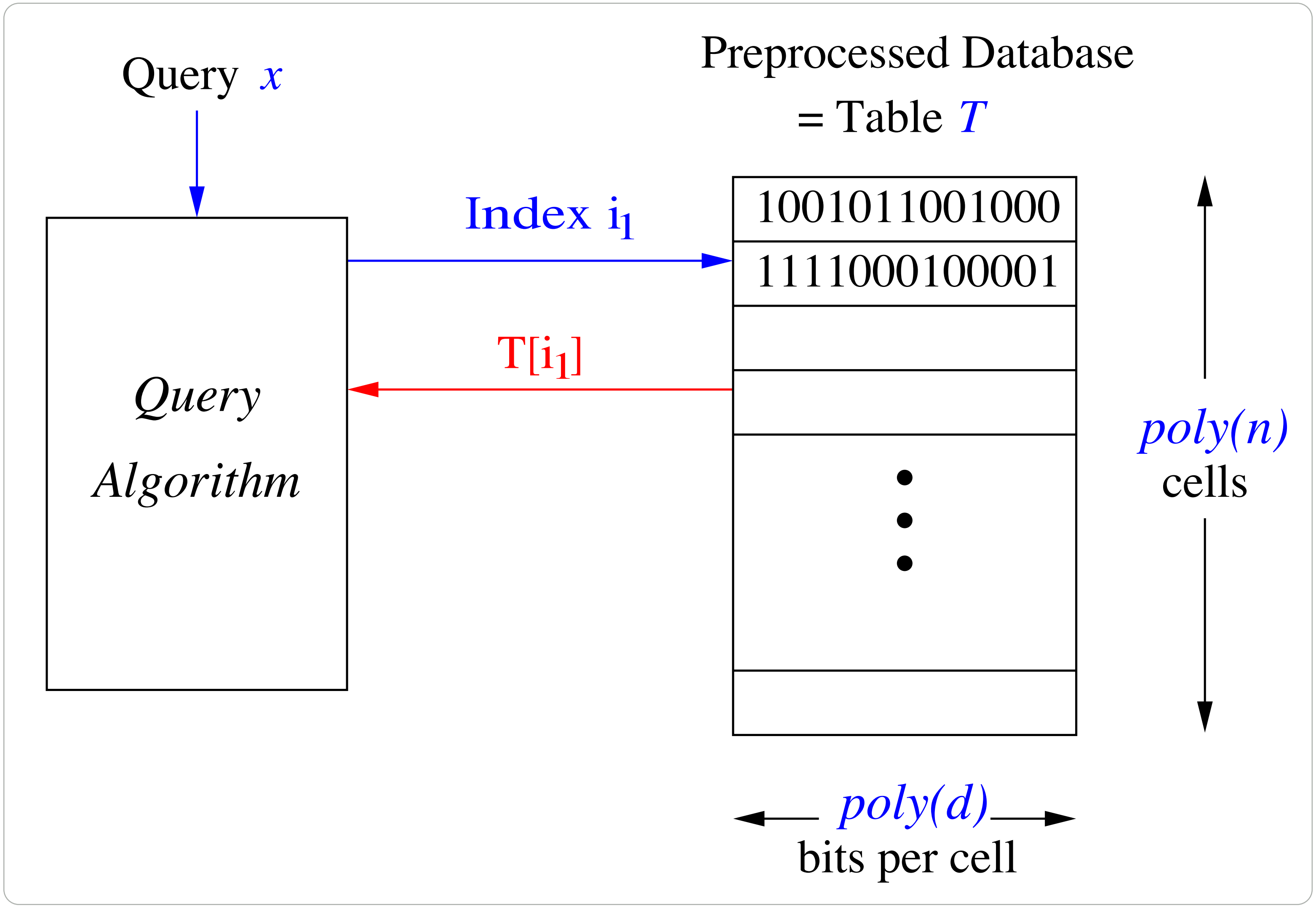
  $R(x, Y, z)$ iff $z \in Y$ is a β-ANN of x w.r.t. Y.

Cell-Probe Model



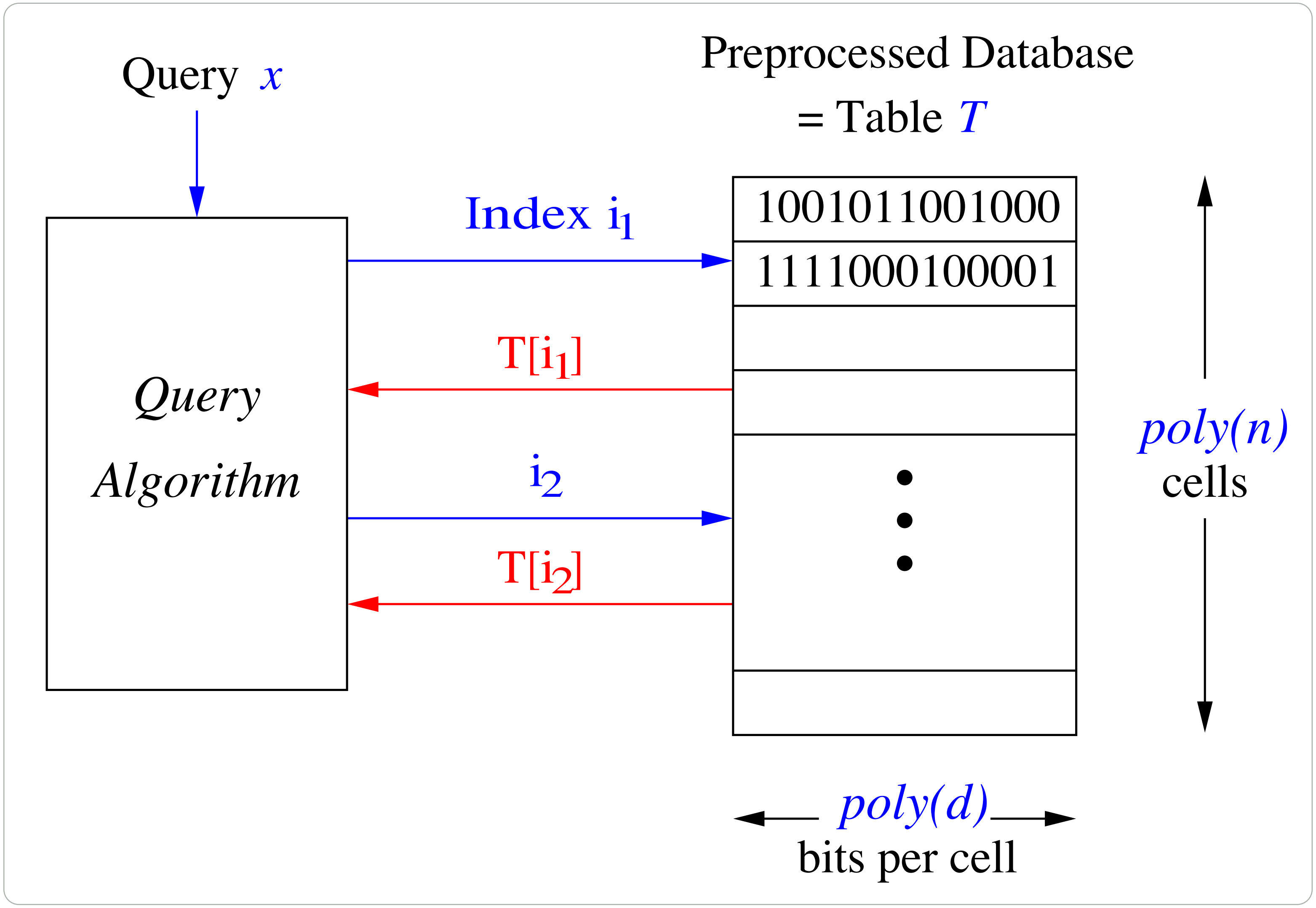
[Yao'81]

Cell-Probe Model



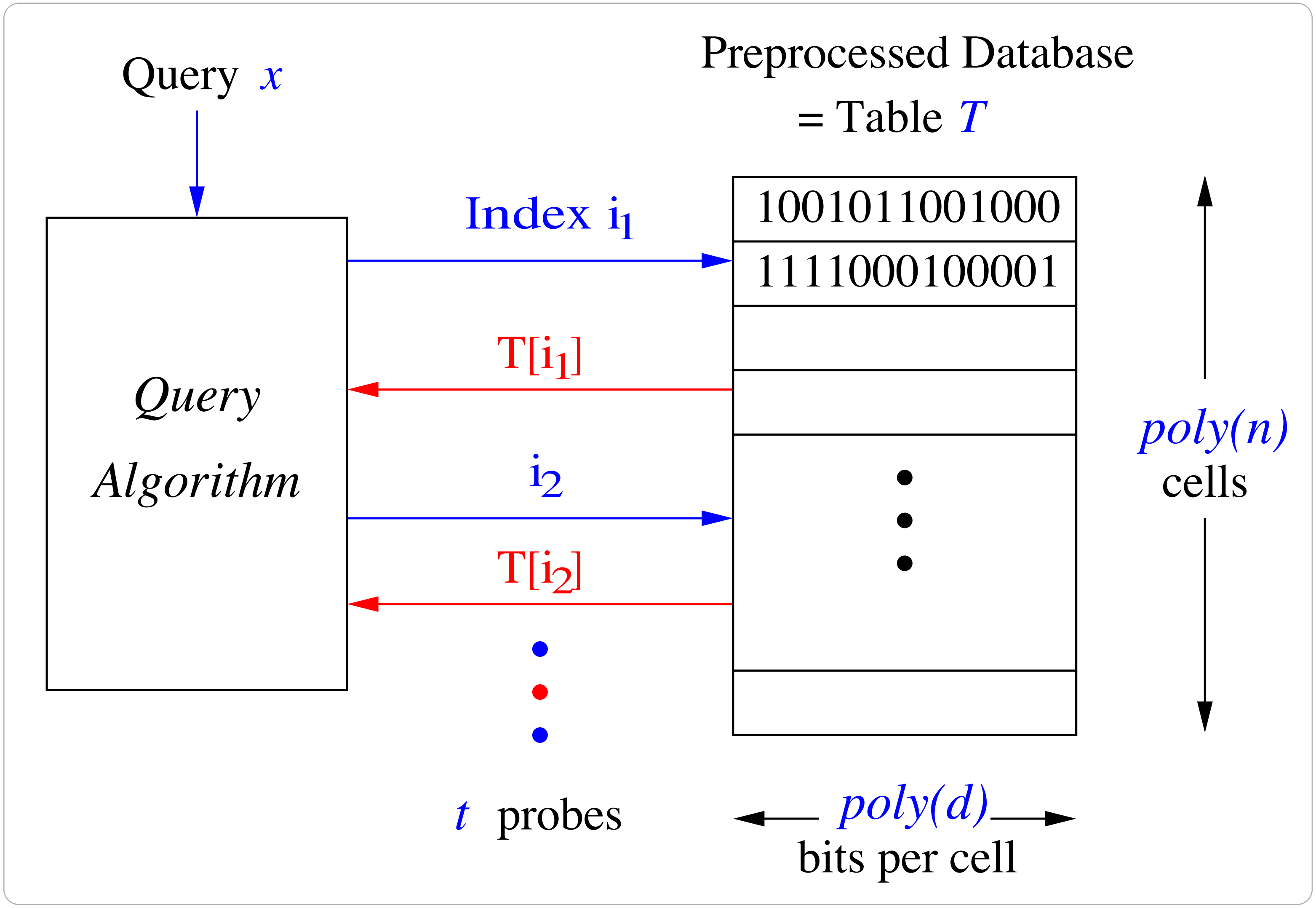
[Yao'81]

Cell-Probe Model



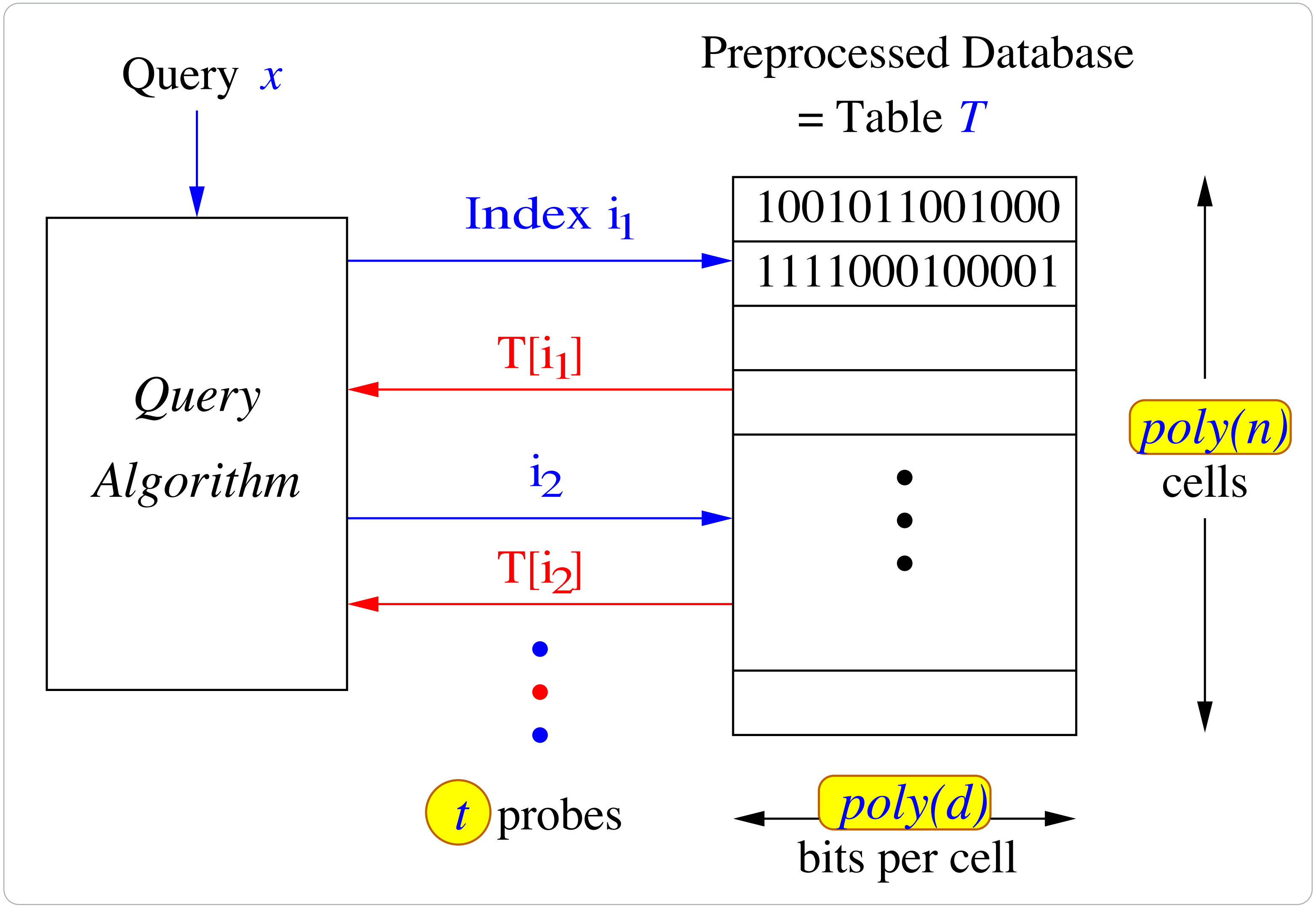
[Yao'81]

Cell-Probe Model



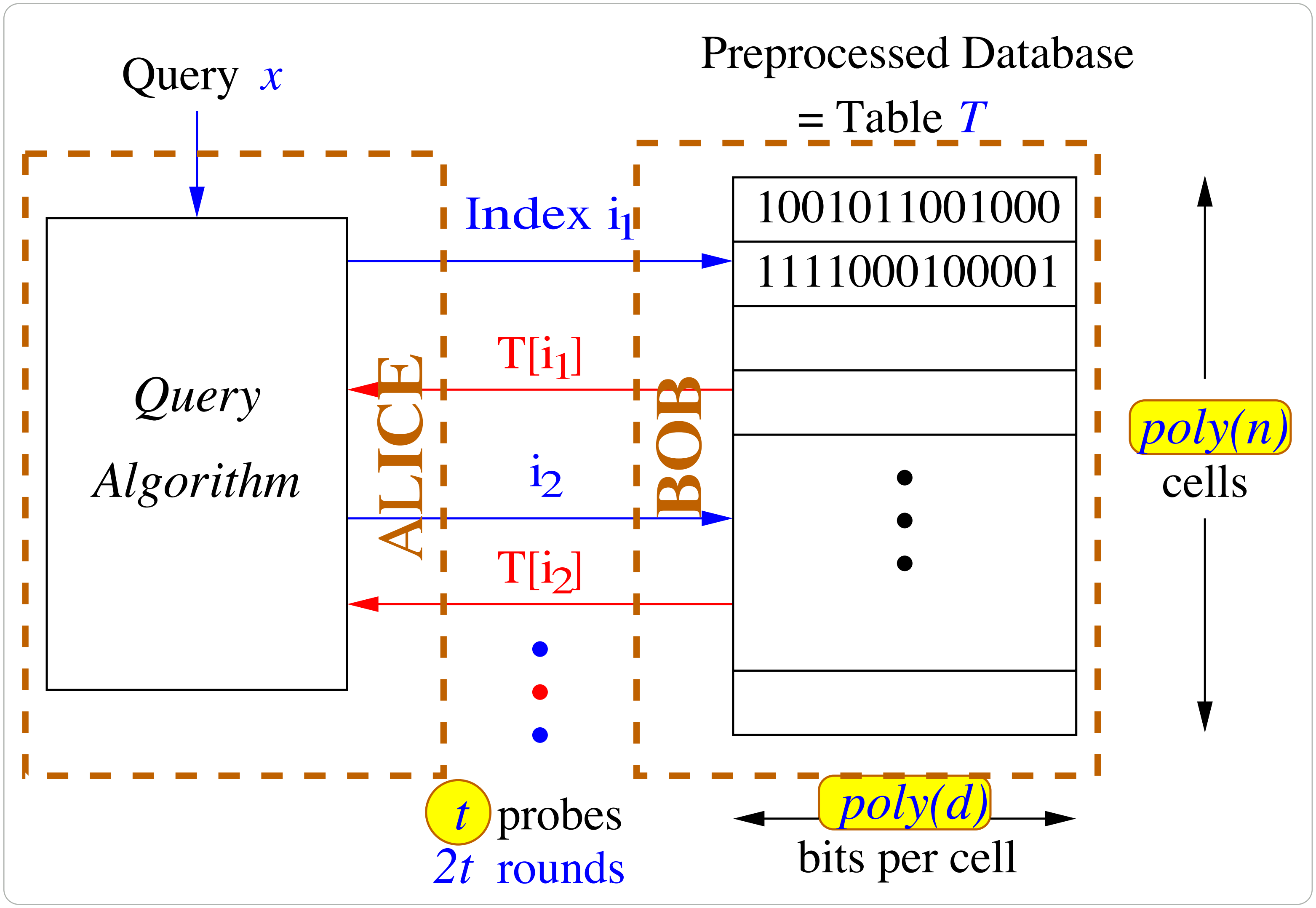
[Yao'81]

Cell-Probe Model



[Yao'81]

Cell-Probe Model

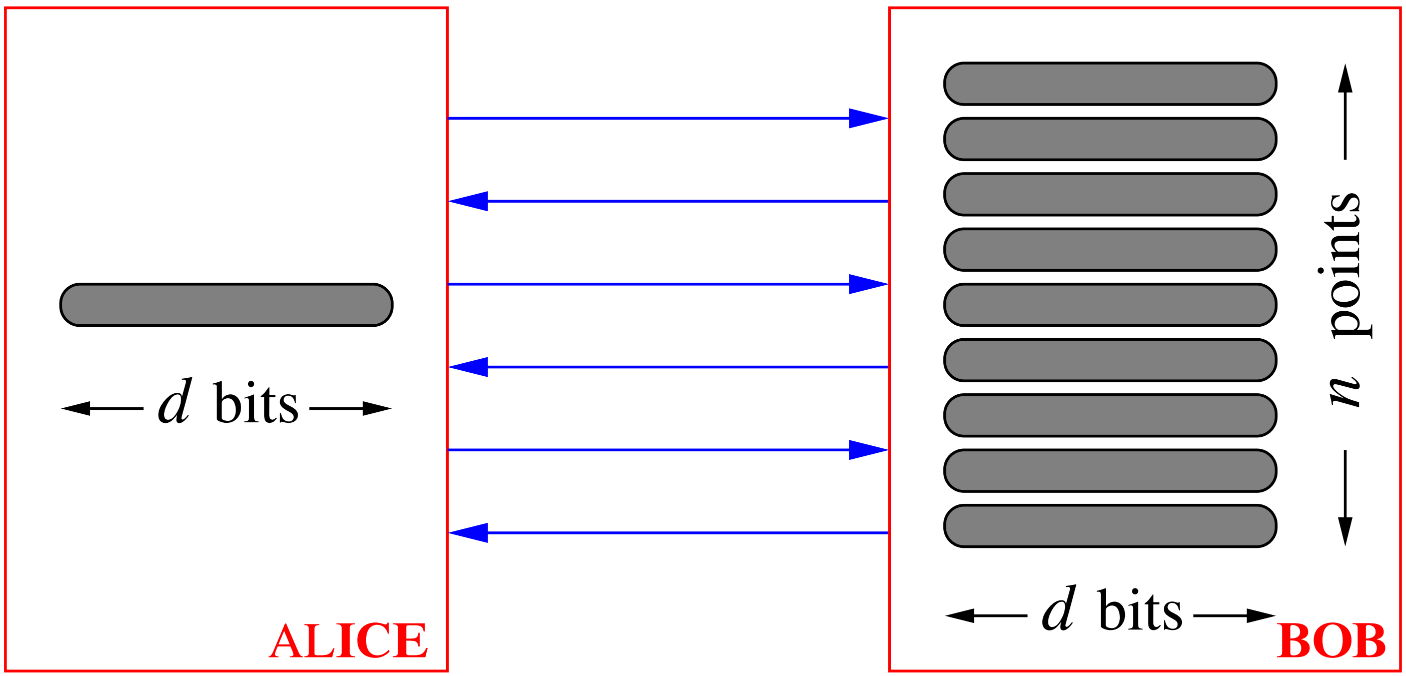


Alice: $O(\log n)$-bit messages; Bob: poly(d)-bit messages

Round Elimination: Predecessor Search

Repeatedly remove first round, shrink instance size

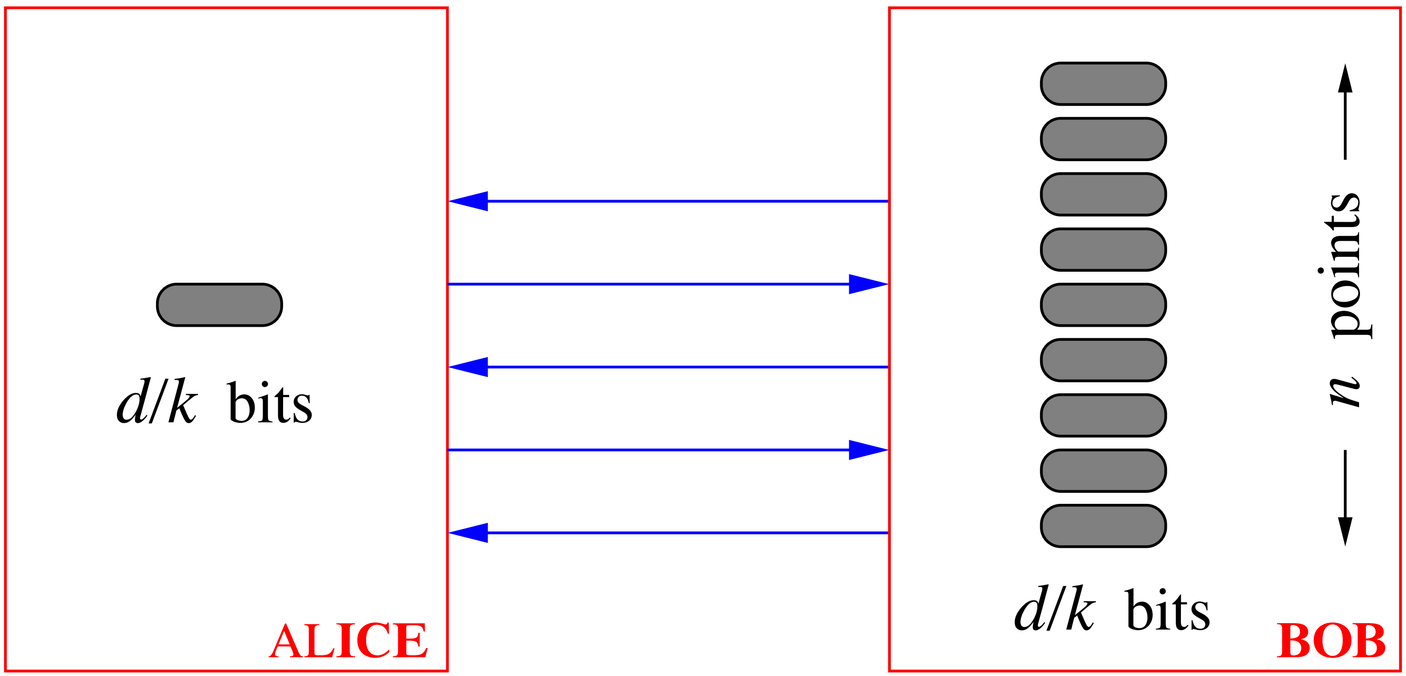
[Miltersen-Nisan-Safra-Wigderson'95], [Sen'03]



Round Elimination: Predecessor Search

Repeatedly remove first round, shrink instance size

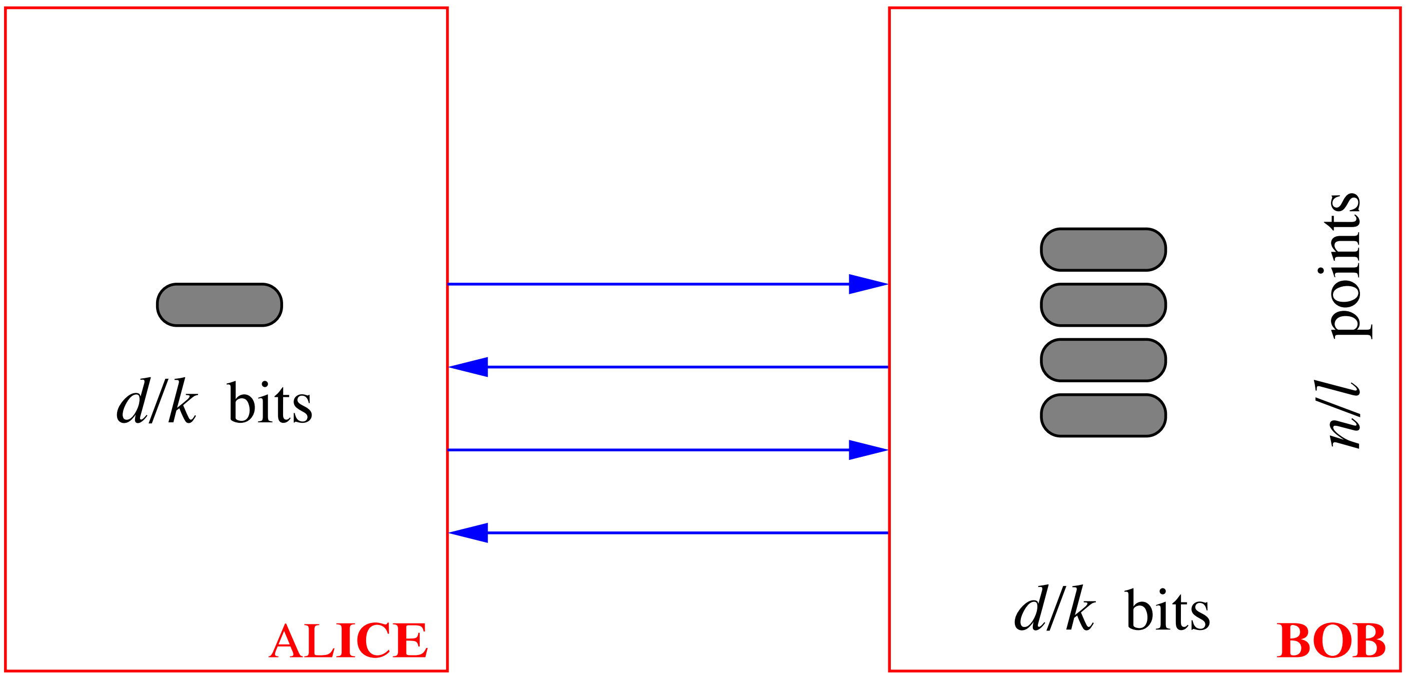
[Miltersen-Nisan-Safra-Wigderson'95], [Sen'03]



Round Elimination: Predecessor Search

Repeatedly remove first round, shrink instance size

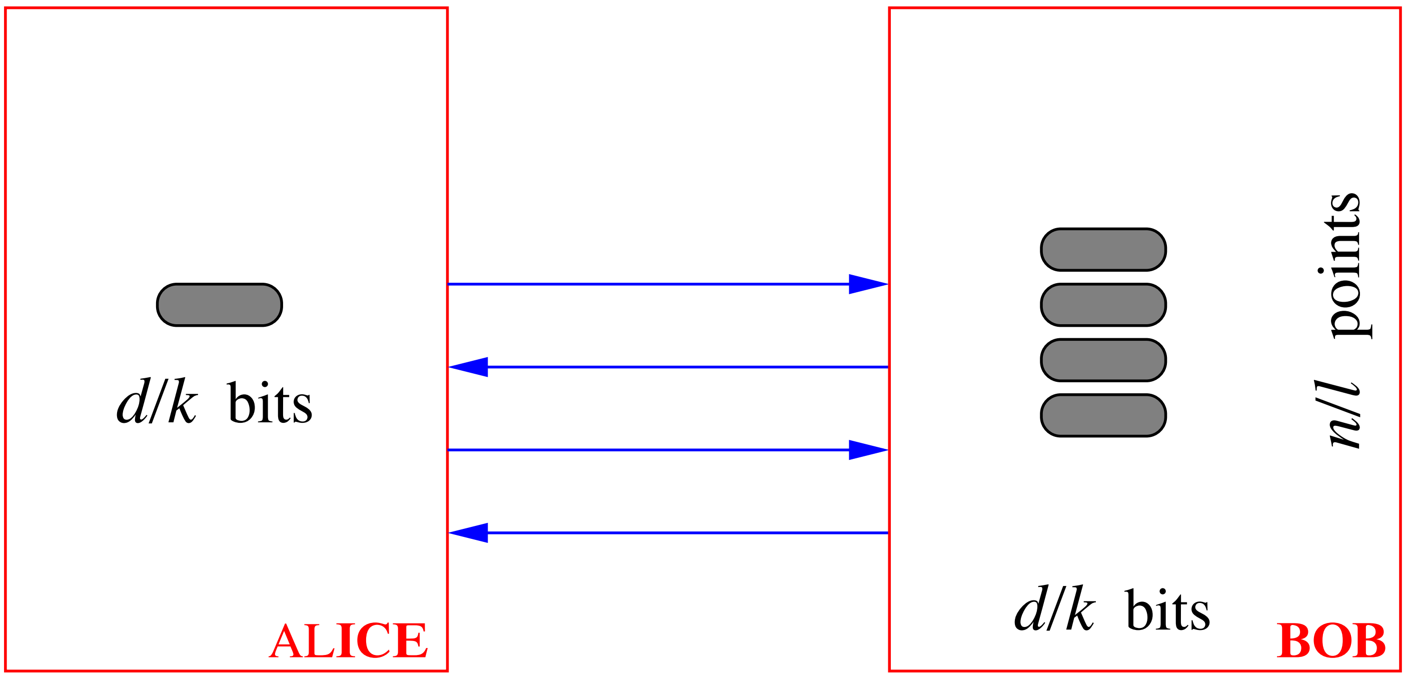
[Miltersen-Nisan-Safra-Wigderson'95], [Sen'03]



Round Elimination: Predecessor Search

Repeatedly remove first round, shrink instance size

[Miltersen-Nisan-Safra-Wigderson'95], [Sen'03]

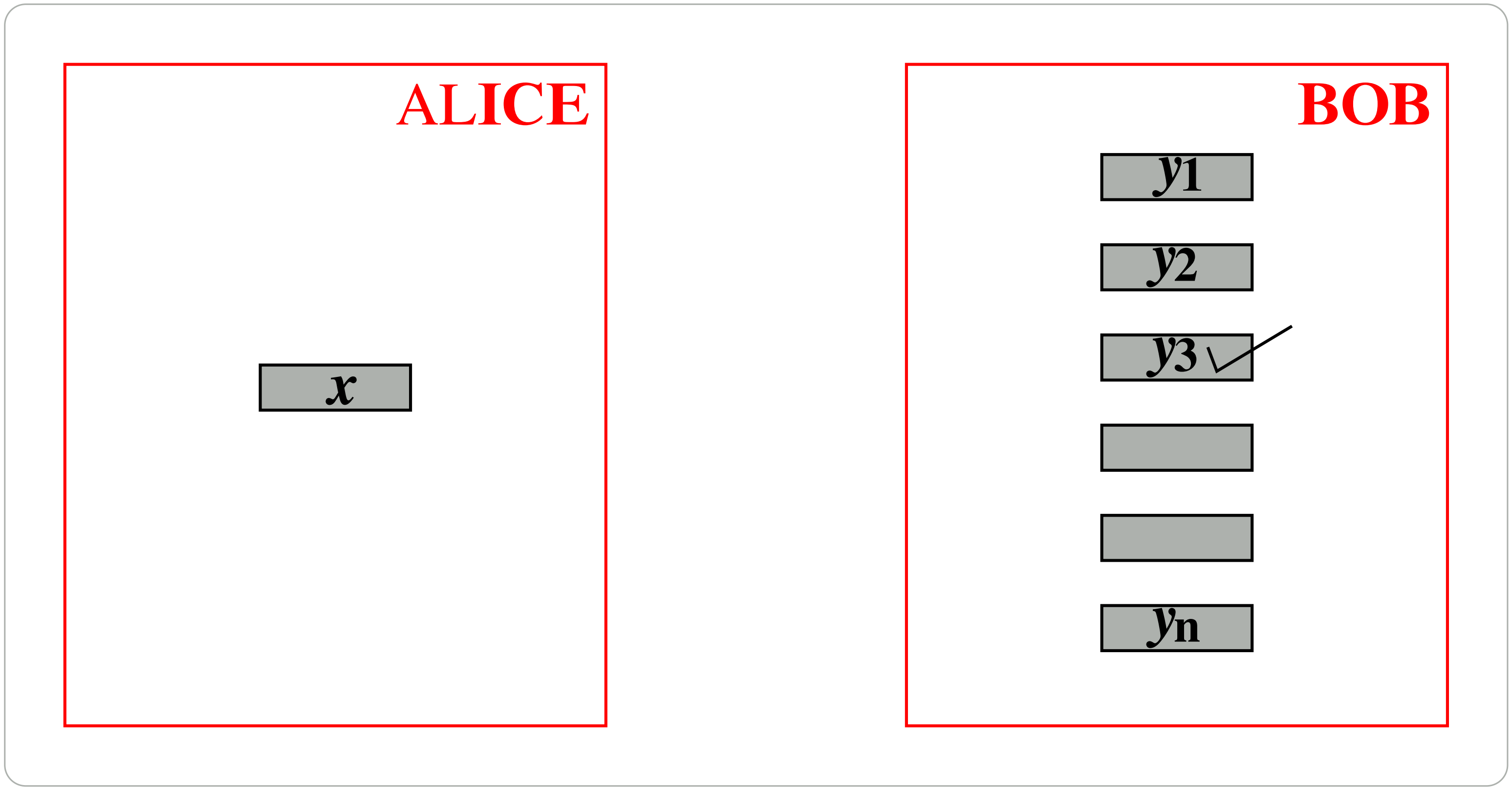


Eventually: zero communication protocol for instance size $(d/k^t, n/\ell^t)$

Implying... **Theorem:** Query time $t = \Omega\left(\frac{\log d}{\log\log d}\right)$

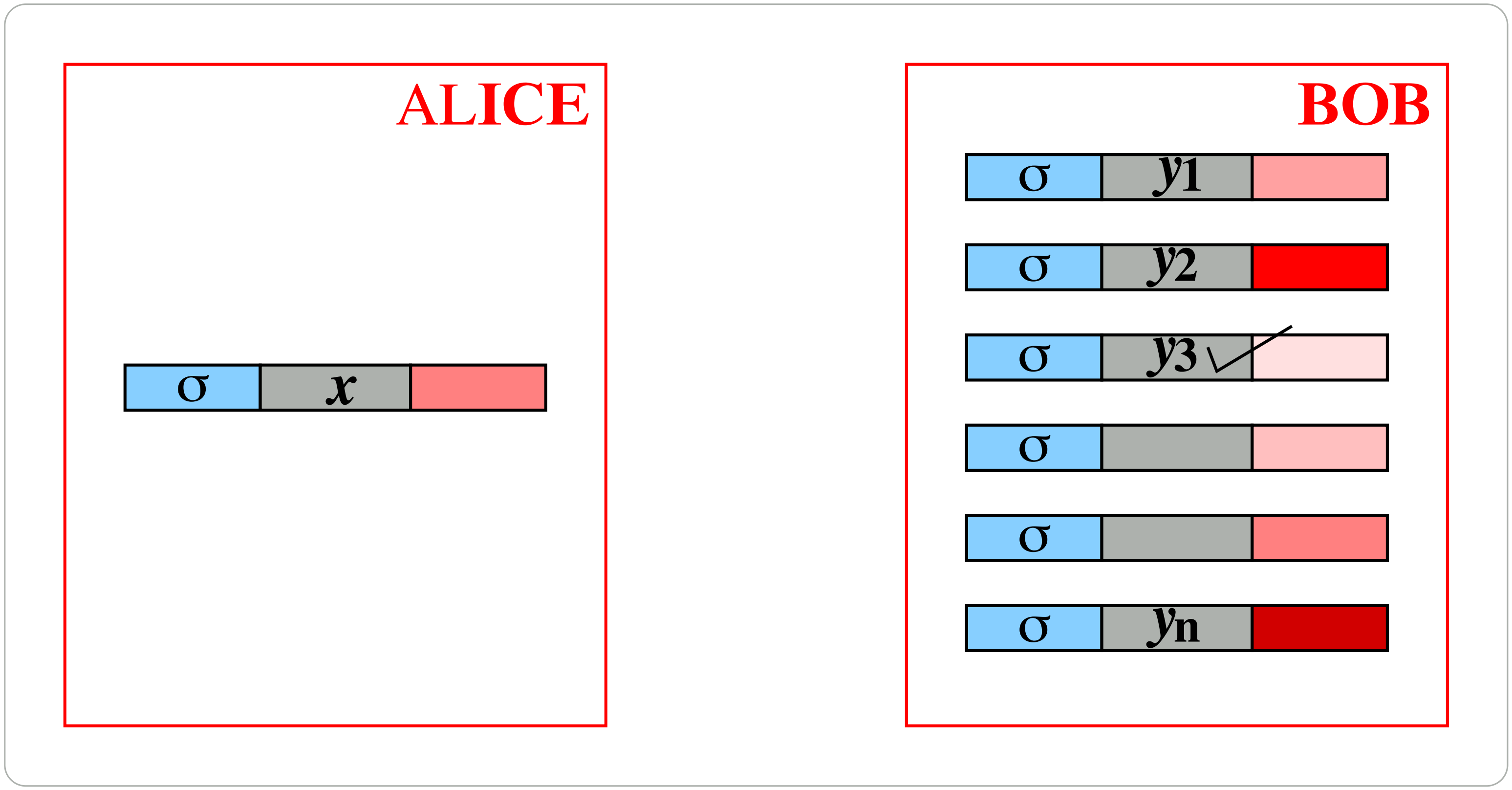
Predecessor Search: Embeddability Property

Input (x, Y) solvable using input $(\sigma \circ x \circ \textsc{rand},\ \sigma \circ Y \circ \textsc{rand})$.



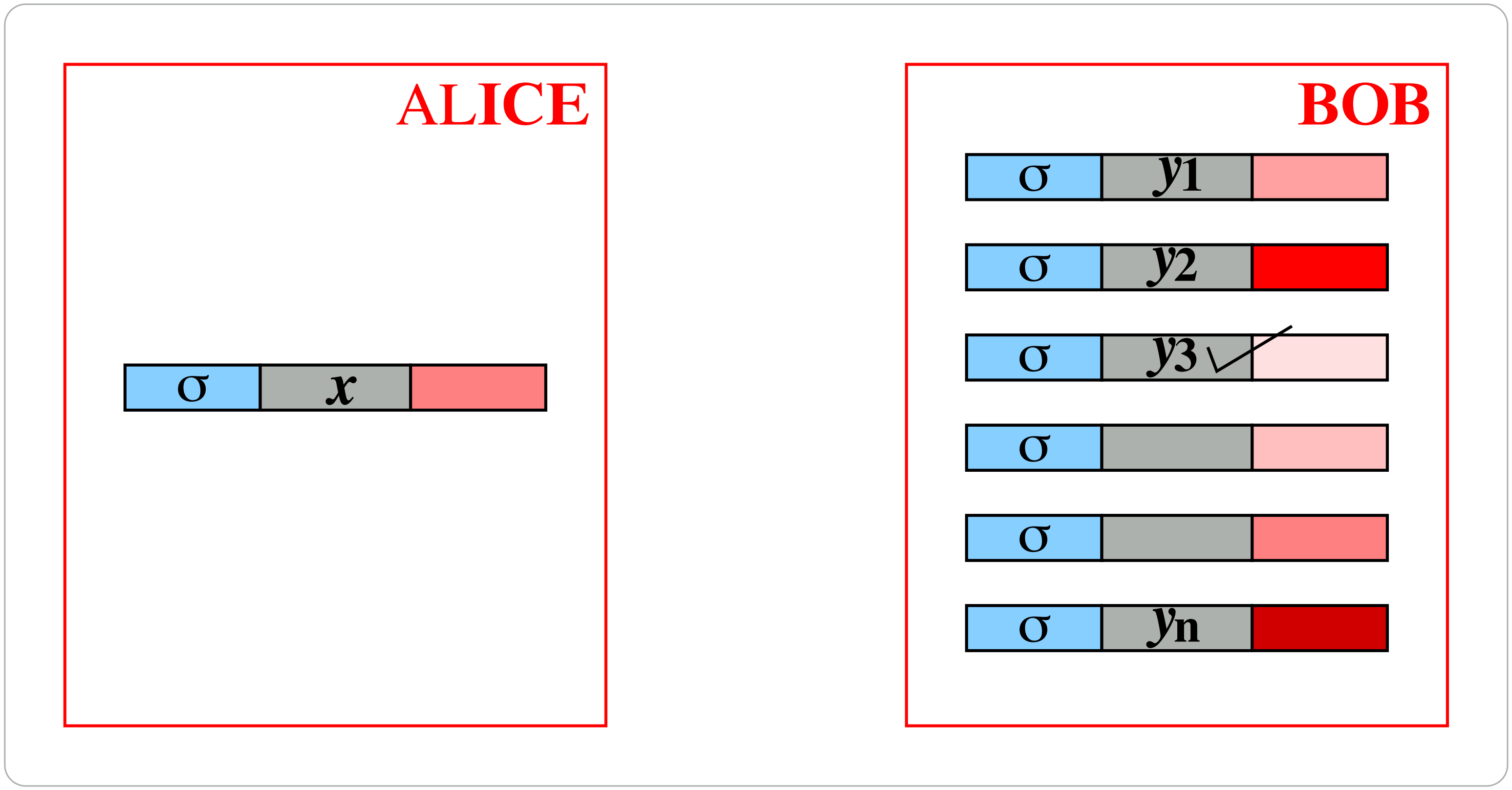
Predecessor Search: Embeddability Property

Input (x, Y) solvable using input $(\sigma \circ x \circ \textsc{rand},\ \sigma \circ Y \circ \textsc{rand})$.



Predecessor Search: Embeddability Property

Input (x, Y) solvable using input $(\sigma \circ x \circ \text{RAND},\ \sigma \circ Y \circ \text{RAND})$.



Simple task ($\mathscr{A}$): instance size $(n/k, d)$, using $2t$ rounds

Complex task ($\mathscr{B}$): instance size (n, d), using $2t$ rounds $\approx k$ times $\mathscr{A}$

IC Paradigm For Round Elimination

1. Define information cost

   $\text{icost}_1^{\mu}(\Pi) = \mathrm{I}(X : M_1)$ Alice's input $X \sim \mu$, $M_1 = \text{msg}_1(X, R)$

2. Simulation Argument

3. Compression Argument

IC Paradigm For Round Elimination

1. Define information cost

   $\mathsf{icost}_1^{\mu}(\Pi) = \mathrm{I}(X : M_1)$ Alice's input $X \sim \mu$, $M_1 = \mathsf{msg}_1(X, R)$

2. Simulation Argument

   Put $X = X_1 X_2 \ldots X_k$, each X_i: a (d/k)-bit chunk

   Protocol $\Pi_{A,\sigma}$: pad instance using prefix σ of length $(i-1)d/k$

$$
\begin{aligned}
O(\log n) \geq \mathsf{icost}_1^{\mu^{\otimes k}}(\Pi_B) &= \mathrm{I}(X_1 X_2 \ldots X_k : M_1) \\
&= \textstyle\sum_{i=1}^{k} \mathrm{I}(X_i : M_1 \mid X_1 \ldots X_{i-1}) \\
&= \sum_{i=1}^{k} \mathbb{E}_\sigma [\mathrm{I}(X_i : M_1 \mid X_1 \ldots X_{i-1} = \sigma)] = \sum_{i=1}^{k} \mathbb{E}_\sigma [\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma})]
\end{aligned}
$$

3. Compression Argument

IC Paradigm For Round Elimination

1. Define information cost

   $\mathsf{icost}_1^{\mu}(\Pi) = \mathrm{I}(X : M_1)$ Alice's input $X \sim \mu$, $M_1 = \mathsf{msg}_1(X, R)$

2. Simulation Argument

   Put $X = X_1 X_2 \ldots X_k$, each X_i: a (d/k)-bit chunk

   Protocol $\Pi_{A,\sigma}$: pad instance using prefix σ of length $(i-1)d/k$

$$O(\log n) \geq \mathsf{icost}_1^{\mu^{\otimes k}}(\Pi_B) = \mathrm{I}(X_1 X_2 \ldots X_k : M_1)$$
$$= \textstyle\sum_{i=1}^{k} \mathrm{I}(X_i : M_1 \mid X_1 \ldots X_{i-1})$$
$$= \sum_{i=1}^{k} \mathbb{E}_\sigma[\mathrm{I}(X_i : M_1 \mid X_1 \ldots X_{i-1} = \sigma)] = \sum_{i=1}^{k} \mathbb{E}_\sigma[\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma})]$$

3. Compression Argument

   So far: exists $\Pi_{A,\sigma}$ with $\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma}) \leq O((\log n)/k) = o(1)$

   Pretend $\mathsf{msg}_1(X', R)$ sent as first message, $X' \equiv X$ but indep.

IC Paradigm For Round Elimination

1. Define information cost

   $\mathsf{icost}_1^{\mu}(\Pi) = \mathrm{I}(X : M_1)$ Alice's input $X \sim \mu$, $M_1 = \mathsf{msg}_1(X, R)$

2. Simulation Argument

   Put $X = X_1 X_2 \ldots X_k$, each X_i: a (d/k)-bit chunk

   Protocol $\Pi_{A,\sigma}$: pad instance using prefix σ of length $(i-1)d/k$

$$O(\log n) \geq \mathsf{icost}_1^{\mu^{\otimes k}}(\Pi_B) = \mathrm{I}(X_1 X_2 \ldots X_k : M_1)$$

$$= \sum_{i=1}^{k} \mathrm{I}(X_i : M_1 \mid X_1 \ldots X_{i-1})$$

$$= \sum_{i=1}^{k} \mathbb{E}_\sigma [\mathrm{I}(X_i : M_1 \mid X_1 \ldots X_{i-1} = \sigma)] = \sum_{i=1}^{k} \mathbb{E}_\sigma [\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma})]$$

3. Compression Argument

   So far: exists $\Pi_{A,\sigma}$ with $\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma}) \leq O((\log n)/k) = o(1)$

   Pretend $\mathsf{msg}_1(X', R)$ sent as first message, $X' \equiv X$ but indep.

   Error $+= O(\sqrt{\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma})}) = o(1)$ [Pinsker's inequality]

Round Elimination: ANN Search

1. Define information cost

   As before, $\text{icost}_1(\Pi) = \mathrm{I}(X : M_1)$

2. Simulation Argument

   ANN does not have embeddability property

3. Compression Argument

Round Elimination: ANN Search

1. Define information cost

   As before, $\mathsf{icost}_1(\Pi) = \mathrm{I}(X : M_1)$

2. Simulation Argument

   ANN does not have embeddability property

   Reduce from Longest Prefix Match (LPM), which does

   Get $\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma}) \leq O((\log n)/k)$ but this bound $= \omega(1)$

3. Compression Argument

Round Elimination: ANN Search

1. Define information cost

   As before, $\mathsf{icost}_1(\Pi) = \mathrm{I}(X : M_1)$

2. Simulation Argument

   ANN does not have embeddability property

   Reduce from Longest Prefix Match (LPM), which does

   Get $\mathsf{icost}_1^{\mu}(\Pi_{A,\sigma}) \le O((\log n)/k)$ but this bound $= \omega(1)$

3. Compression Argument

   Use comm complexity of correlation [Harsha-J-M-R'07]

   Compress first message down to its info content

   Now it's short enough: easy (combinatorial) round eliminiation

Eventually... **Theorem:** Query time $t = \Omega\left(\dfrac{\log\log d}{\log\log\log d}\right)$ [C.-Regev'10]

Pointer Jumping Problems

Input: One pointer per level in layered graph; plus one bit per leaf

Task: output bit at leaf reached by following pointers from root

Multilayer Ptr Jumping, $\mathbf{MPJ}_{n,p}$

Full layered DAG, n nodes/layer

Number-on-Forehead (NOF)

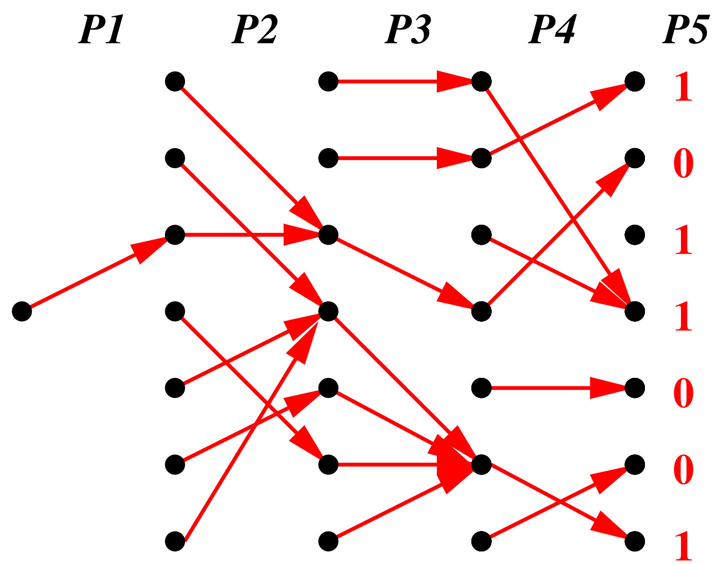
Speaking order $P_1, \ldots, P_p$

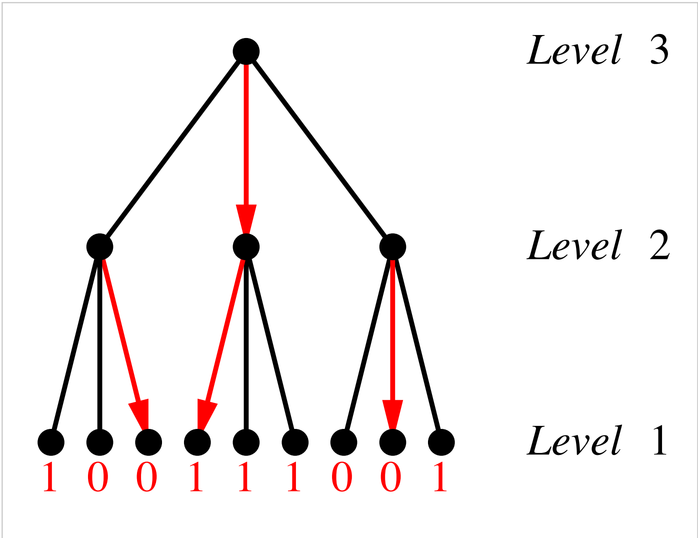
Tree Pointer Jumping, $\mathbf{TPJ}_{n,p+1}$

Complete $(p+1)$-level n-ary tree

Number-In-Hand (NIH)

Use p rounds, going up tree: ↑↑↑

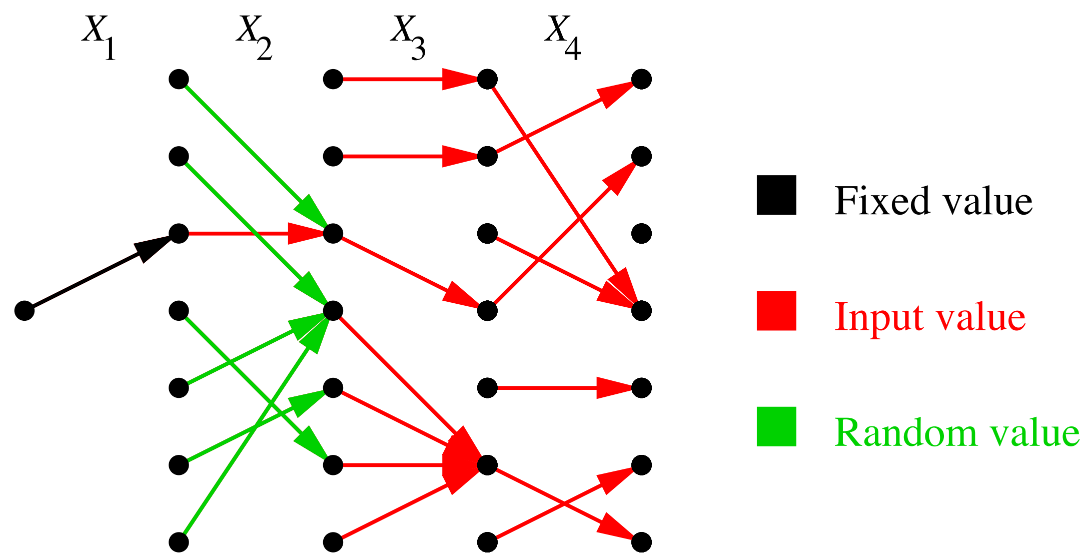




Theorems: $R^{\rightarrow}(\text{MPJ}_{n,p}) = \Omega(n/p)^{\star}$ $\qquad R^{p}(\text{TPJ}_{n,p+1}) = \Omega(n/p^2)$

The Importance of a Careful Definition

For the NOF problem $\text{MPJ}_{n,p}$ [Chakrabarti'07]



Protocol P, input $(X_1, \ldots, X_p) \sim \mu$, $M_1 =$ message of player P_1:

$$\text{icost}^{\mu}(\Pi) := \mathrm{I}(M_1 : X_2 \mid X_3, \ldots, X_k)$$

Simulation argument works only under one of these protocol restrictions:

- Myopic: each player sees only one layer ahead ... $\Omega(n/p)$
- Conservative: don't see behind except know current node ... $\Omega(n/p^2)$

Pointer Jumping: Applications

- Layered DAG version, NOF
  - Strong NOF lower bounds imply circuit lower bounds [Yao'90]
  - If myopic/conservative restriction removed, $\mathsf{ACC}^0 \neq \mathsf{LOGSPACE}$

Pointer Jumping: Applications

- Layered DAG version, NOF
  - Strong NOF lower bounds imply circuit lower bounds [Yao'90]
  - If myopic/conservative restriction removed, $\mathsf{ACC}^0 \neq \mathsf{LOGSPACE}$
- Tree version, NIH
  - Multi-pass data stream lower bounds
  - Classic example: median of n numbers, p passes: $\Omega(n^{1/p})$ space
  - Modern: median, randomly-ordered, p passes: $\Omega(n^{2^{-\mathsf{p}}})$ space
    A sophisticated reduction [C.-Cormode-McGregor'08]

Part Three:
Full Interaction

The DISJOINTNESS problem

- Input: $2 \times n$ Boolean matrix
- Task: distinguish between the following two cases
  - Case 0: Every column has weight ≤ 1

0	1	0	0	0	1	1	0
1	0	1	1	0	0	0	0

  - Case 1: One column has weight 2, rest have weight ≤ 1 (i.e., the subsets of $[n]$ represented by the rows intersect)

0	1	1	0	0	1	1	0
1	0	1	1	0	0	0	0

The DISJOINTNESS problem

- Input: $2 \times n$ Boolean matrix
- Task: distinguish between the following two cases
  - Case 0: Every column has weight ≤ 1

0	1	0	0	0	1	1	0	$\longleftarrow$ Alice
1	0	1	1	0	0	0	0	$\longleftarrow$ Bob

  - Case 1: One column has weight 2, rest have weight ≤ 1 (i.e., the subsets of $[n]$ represented by the rows intersect)

0	1	1	0	0	1	1	0	$\longleftarrow$ Alice
1	0	1	1	0	0	0	0	$\longleftarrow$ Bob

- This problem is called $\text{DISJ}_{n,2}$
- Later: t-player generalization $\text{DISJ}_{n,t}$

New Twist in Applying IC Paradigm

Problem $\textsc{and}_2$:

- Alice holds $a \in \{0,1\}$, Bob holds $b \in \{0,1\}$
- Bob to output $a \wedge b$

Simple task ($\mathscr{A}$): $\textsc{and}_2$

Complex task ($\mathscr{B}$): $\textsc{disj}_{n,2} \approx$ combines n copies of $\textsc{and}_2$

New Twist in Applying IC Paradigm

Problem $\textsc{and}_2$:

- Alice holds $a \in \{0,1\}$, Bob holds $b \in \{0,1\}$
- Bob to output $a \wedge b$

Simple task ($\mathscr{A}$): $\textsc{and}_2$

Complex task ($\mathscr{B}$): $\textsc{disj}_{n,2} \approx$ combines n copies of $\textsc{and}_2$

The key twist:

- Step 2 (Simulation): pad $\textsc{and}_2$ instance to get $\textsc{disj}_{n,2}$ instance
- Careless random padding will drown out the answer!
- Must ensure "padding" columns j have $\textsc{and}_2(X_j, Y_j) = 0$
- Need shared coins to do this ... so must condition on these coins

IC Paradigm Applied to $\textsc{disjointness}$ (1/3)

Inputs: $\vec{X} = X_1 \ldots X_n$, $\vec{Y} = Y_1 \ldots Y_n$; Auxiliary coins: $D_1 \ldots D_n$

Each $(X_i, Y_i, D_i) \sim \mu$; $\quad M = \text{transcript}^{\Pi}(X, Y, R_{\text{pub}}, R_{\text{priv}})$

μ =

$XY \rightarrow$	00	01	10	11
D = 0	1/2	0	1/2	0
D = 1	1/2	1/2	0	0

1. Define information cost

   $$\text{icost}^{\mu}(\Pi) = \mathrm{I}(\vec{X}\vec{Y} : M \mid \vec{D}, R_{\text{pub}})$$

   (Note: external icost)

2. Simulation Argument

3. Basic IC lower bound (for $\textsc{and}_2$)

IC Paradigm Applied to $\textsc{disjointness}$ (1/3)

Inputs: $\vec{X} = X_1 \ldots X_n$, $\vec{Y} = Y_1 \ldots Y_n$; Auxiliary coins: $D_1 \ldots D_n$

Each $(X_i, Y_i, D_i) \sim \mu$; $\quad M = \text{transcript}^{\Pi}(X, Y, R_{\text{pub}}, R_{\text{priv}})$

μ =

$XY \rightarrow$	00	01	10	11
$D = 0$	1/2	0	1/2	0
$D = 1$	1/2	1/2	0	0

1. Define information cost

   $\text{icost}^{\mu}(\Pi) = \mathrm{I}(\vec{X}\vec{Y} : M \mid \vec{D}, R_{\text{pub}})$ (Note: external icost)

2. Simulation Argument

   Protocols $\Pi_{A,i}$ for $\textsc{and}_2(X, Y)$ simulating Π_B for $\textsc{disj}_{n,2}$

   Public coins for $\vec{D}$, private for $\vec{X}, \vec{Y}$ s.t. each $(X_j, Y_j, D_j) \sim \mu$

3. Basic IC lower bound (for $\textsc{and}_2$)

IC Paradigm Applied to DISJOINTNESS (1/3)

Inputs: $\vec{X} = X_1 \ldots X_n$, $\vec{Y} = Y_1 \ldots Y_n$; Auxiliary coins: $D_1 \ldots D_n$

Each $(X_i, Y_i, D_i) \sim \mu$; $M = \text{transcript}^{\Pi}(X, Y, R_{\text{pub}}, R_{\text{priv}})$

μ =

$XY \rightarrow$	00	01	10	11
$D = 0$	1/2	0	1/2	0
$D = 1$	1/2	1/2	0	0

1. Define information cost

   $\text{icost}^{\mu}(\Pi) = \mathrm{I}(\vec{X}\vec{Y} : M \mid \vec{D}, R_{\text{pub}})$ (Note: external icost)

2. Simulation Argument

   Protocols $\Pi_{A,i}$ for $\text{AND}_2(X, Y)$ simulating Π_B for $\text{DISJ}_{n,2}$

   Public coins for $\vec{D}$, private for $\vec{X}, \vec{Y}$ s.t. each $(X_j, Y_j, D_j) \sim \mu$

   Crucial property: D_j **factorizes** (X_j, Y_j)

   Plug in $X_i \leftarrow X$ and $Y_i \leftarrow Y$

3. Basic IC lower bound (for AND_2)

IC Paradigm Applied to DISJOINTNESS (2/3)

Each $(X_i, Y_i, D_i) \sim \mu$; $\quad M = \text{transcript}^{\Pi}(X, Y, R_{\text{pub}}, R_{\text{priv}})$

μ =

$XY \rightarrow$	00	01	10	11
D = 0	1/2	0	1/2	0
D = 1	1/2	1/2	0	0

2. Simulation Argument

Protocols $\Pi_{A,i}$ for $\text{AND}_2(X, Y)$ simulating Π_B for $\text{DISJ}_{n,2}$

$$\text{icost}^{\mu^{\otimes n}}(\Pi_B) = \mathrm{I}(\vec{X}\vec{Y} : M \mid \vec{D}) \geq \sum_{i=1}^{n} \mathrm{I}(X_i Y_i : M \mid \vec{D})$$

$$= \sum_{i=1}^{n} \mathrm{I}(X_i Y_i : M \mid D_i, \vec{D}_{-i}) = \sum_{i=1}^{n} \text{icost}^{\mu}(\Pi_{A,i})$$

3. Basic IC lower bound (for AND_2)

IC Paradigm Applied to DISJOINTNESS (2/3)

Each $(X_i, Y_i, D_i) \sim \mu$; $\quad M = \text{transcript}^{\Pi}(X, Y, R_{\text{pub}}, R_{\text{priv}})$

μ =

$XY \rightarrow$	00	01	10	11
D = 0	1/2	0	1/2	0
D = 1	1/2	1/2	0	0

2. Simulation Argument

   Protocols $\Pi_{A,i}$ for $\text{AND}_2(X, Y)$ simulating Π_B for $\text{DISJ}_{n,2}$

$$\text{icost}^{\mu^{\otimes n}}(\Pi_B) = \mathrm{I}(\vec{X}\vec{Y} : M \mid \vec{D}) \geq \sum_{i=1}^{n} \mathrm{I}(X_i Y_i : M \mid \vec{D})$$

$$= \sum_{i=1}^{n} \mathrm{I}(X_i Y_i : M \mid D_i, \vec{D}_{-i}) = \sum_{i=1}^{n} \text{icost}^{\mu}(\Pi_{A,i})$$

3. Basic IC lower bound (for AND_2)

   To prove: $\forall\ \Pi_A$ solving AND_2, have $\text{icost}^{\mu}(\Pi_A) = \Omega(1)$

   Twist: distrib μ not hard for AND_2: $\quad \mathbb{E}_{(X,Y)\sim\mu}[\text{AND}_2(X, Y)] = 0$

IC Paradigm Applied to DISJOINTNESS (3/3)

- Protocol Π_A for $\textsc{and}_2$; $(X, Y, D) \sim \mu$; $M = \mathsf{trans}^{\Pi}(X, Y, R_{\mathsf{priv}})$
- Let $M^{(wz)} = \mathsf{transcript}^{\Pi}(w, z, R_{\mathsf{priv}})$ for $w, z \in \{0, 1\}$; then

$$\begin{aligned}
\mathsf{icost}^{\mu}(\Pi_A) = \mathrm{I}(XY : M \mid D) &= \tfrac{1}{2}(\mathrm{I}(X : M \mid D = 0) + \mathrm{I}(Y : M \mid D = 1)) \\
&= \tfrac{1}{2}(\mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(10)}) + \mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{2}(\mathrm{h}^2(M^{(00)}, M^{(10)}) + \mathrm{h}^2(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{4}\mathrm{h}^2(M^{(10)}, M^{(01)})
\end{aligned}$$

IC Paradigm Applied to DISJOINTNESS (3/3)

- Protocol Π_A for $\textsc{and}_2$; $(X, Y, D) \sim \mu$; $M = \mathsf{trans}^{\Pi}(X, Y, R_{\mathsf{priv}})$
- Let $M^{(wz)} = \mathsf{transcript}^{\Pi}(w, z, R_{\mathsf{priv}})$ for $w, z \in \{0, 1\}$; then

$$\begin{aligned}
\mathsf{icost}^{\mu}(\Pi_A) = \mathrm{I}(XY : M \mid D) &= \tfrac{1}{2}(\mathrm{I}(X : M \mid D = 0) + \mathrm{I}(Y : M \mid D = 1)) \\
&= \tfrac{1}{2}(\mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(10)}) + \mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{2}(\mathrm{h}^2(M^{(00)}, M^{(10)}) + \mathrm{h}^2(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{4}\mathrm{h}^2(M^{(10)}, M^{(01)}) \;=\; \tfrac{1}{4}\mathrm{h}^2(M^{(00)}, M^{(11)})
\end{aligned}$$

by cut-and-paste property, conseq of rectangle property for protocols

Digression: Cut-And-Paste

- Protocol Π; inputs w, z
- Let $M^{(wz)} = \text{transcript}^{\Pi}(w, z, R_{\text{priv}})$ for $w, z \in \{0, 1\}$
- Rectangle property:

  Can "factorize" distrib of $M^{(wz)}$ as $f^{(w)} \odot g^{(z)}$

  This means $\forall\, a :\quad \Pr[M^{(wz)} = a] = f^{(w)}(a) g^{(z)}(a)$
- Hellinger distance: $\mathrm{h}^2((, P), Q) = 1 - \sum_a \sqrt{P(a)Q(a)}$
- Put these together:

$$\mathrm{h}^2(M^{(bc)}, M^{(wz)}) = \mathrm{h}^2(M^{(bz)}, M^{(cw)})$$

IC Paradigm Applied to DISJOINTNESS (3/3)

- Protocol Π_A for AND_2; $(X, Y, D) \sim \mu$; $M = \mathsf{trans}^{\Pi}(X, Y, R_{\mathsf{priv}})$
- Let $M^{(wz)} = \mathsf{transcript}^{\Pi}(w, z, R_{\mathsf{priv}})$ for $w, z \in \{0, 1\}$; then

$$\begin{aligned}
\mathsf{icost}^{\mu}(\Pi_A) = \mathrm{I}(XY : M \mid D) &= \tfrac{1}{2}(\mathrm{I}(X : M \mid D = 0) + \mathrm{I}(Y : M \mid D = 1)) \\
&= \tfrac{1}{2}(\mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(10)}) + \mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{2}(\mathrm{h}^2(M^{(00)}, M^{(10)}) + \mathrm{h}^2(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{4}\mathrm{h}^2(M^{(10)}, M^{(01)}) \;=\; \tfrac{1}{4}\mathrm{h}^2(M^{(00)}, M^{(11)})
\end{aligned}$$

by cut-and-paste property, conseq of rectangle property for protocols

IC Paradigm Applied to DISJOINTNESS (3/3)

- Protocol Π_A for AND_2; $(X, Y, D) \sim \mu$; $M = \text{trans}^{\Pi}(X, Y, R_{\text{priv}})$
- Let $M^{(wz)} = \text{transcript}^{\Pi}(w, z, R_{\text{priv}})$ for $w, z \in \{0, 1\}$; then

$$\begin{aligned}
\text{icost}^{\mu}(\Pi_A) = \mathrm{I}(XY : M \mid D) &= \tfrac{1}{2}(\mathrm{I}(X : M \mid D = 0) + \mathrm{I}(Y : M \mid D = 1)) \\
&= \tfrac{1}{2}(\mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(10)}) + \mathrm{D}_{\mathrm{JS}}(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{2}(\mathrm{h}^2(M^{(00)}, M^{(10)}) + \mathrm{h}^2(M^{(00)}, M^{(01)})) \\
&\geq \tfrac{1}{4}\mathrm{h}^2(M^{(10)}, M^{(01)}) \;=\; \tfrac{1}{4}\mathrm{h}^2(M^{(00)}, M^{(11)})
\end{aligned}$$

by cut-and-paste property, conseq of rectangle property for protocols

- Error $\leq \varepsilon$ implies

$$\mathrm{D}_{\mathrm{TV}}(M^{(00)}, M^{(11)}) \geq 1 - 2\varepsilon \implies \mathrm{h}^2(M^{(00)}, M^{(11)}) \geq 1 - 2\sqrt{\varepsilon}$$

- Overall: $\mathrm{R}_{\varepsilon}(\text{DISJ}_{n,2}) \geq \frac{1}{4}(1 - 2\sqrt{\varepsilon})n$ [BarYossef-J-K-S'04]

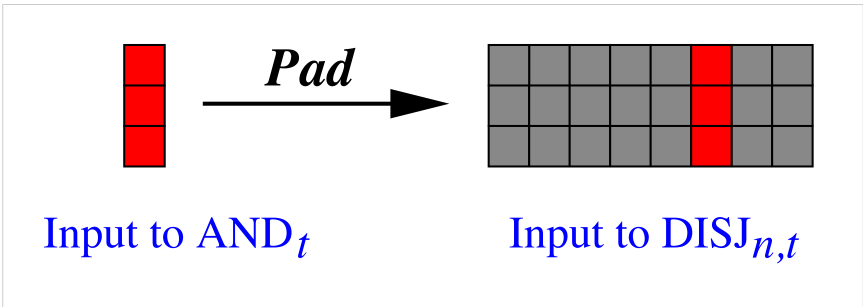
Applications of DISJOINTNESS

Multi-pass lower bounds for many data stream problems

- Connectivity of n-vertex graphs: $\Omega(n)$ space

Generalization (t players, NIH)

Theorem: $\mathrm{R}(\text{DISJ}_{n,t}) = \Omega(n/t)$ [C.-Khot-Sun'03], [Gronemeier'09]



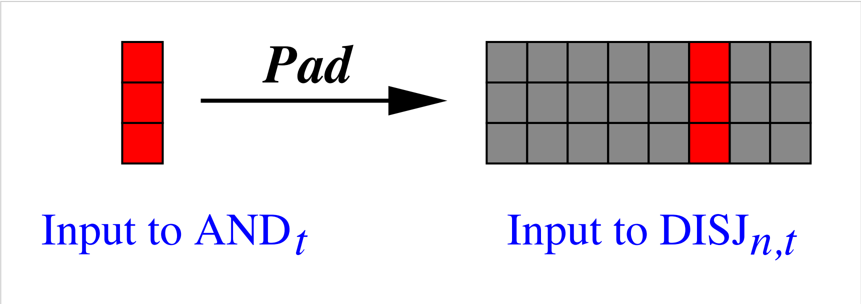
Applications of DISJOINTNESS

Multi-pass lower bounds for many data stream problems

- Connectivity of n-vertex graphs: $\Omega(n)$ space

Generalization (t players, NIH)

Theorem: $\mathrm{R}(\mathrm{DISJ}_{n,t}) = \Omega(n/t)$ [C.-Khot-Sun'03], [Gronemeier'09]

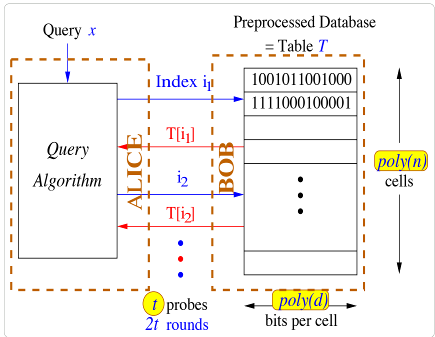


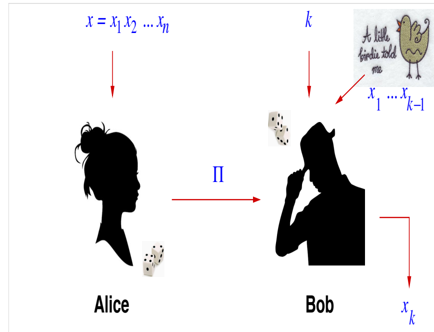
- Classic data stream problem: frequency moments $F_k = \sum_j f_j^k$ where $f_j :=$ number of occurrences of 'j' in stream
- Approximating F_k: space $\widetilde{\Theta}(n^{1-2/k})$ lower bound via $\mathrm{DISJ}_{n,t}$ [Alon-Matias-Szegedy'96]

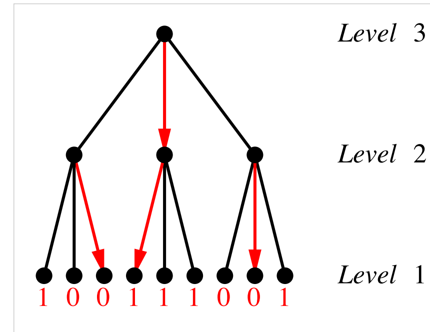
Yet More Applications of IC

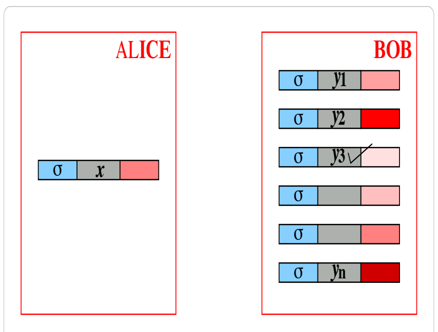
Sadly, left on cutting room floor ...

- Separation of nondet and randomized CC [Jayram-Kumar-Sivakumar'03]
- CC of read-once-formula problems [Saks-Leonardos'09] [Jayram-Kopparty-Raghavendra'09]
- Det vs rand decision trees [Jayram-Kumar-Sivakumar'03]
- Increasingly complex data stream lower bounds [Jayram-Woodruff'09], [Magniez-Mathieu-Nayak'10] [C.-Cormode-Kondapally-McGregor'10], [Magniez'13]
- Data structure query/update time lower bounds [Patrascu'10]
- Quantum communication... [Jain-Radhakrishnan-Sen]

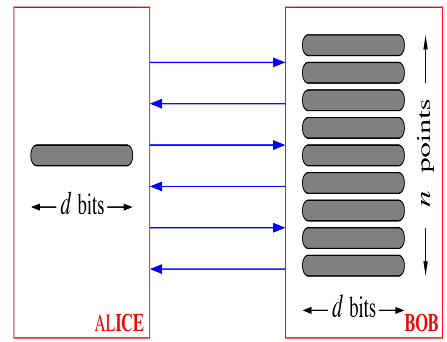

Query x
Preprocessed Database
= Table T
Index i_1
1001011001000
1111000100001
Query Algorithm
ALICE
BOB
$T[i_1]$
i_2
$T[i_2]$
poly(n)
cells
t probes
$2t$ rounds
poly(d)
bits per cell

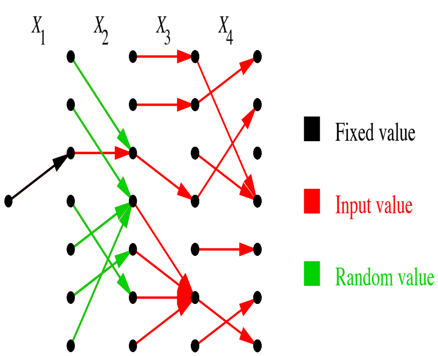

$x = x_1 x_2 \dots x_n$
k
$x_1 \dots x_{k-1}$
Π
Alice
Bob
x_k

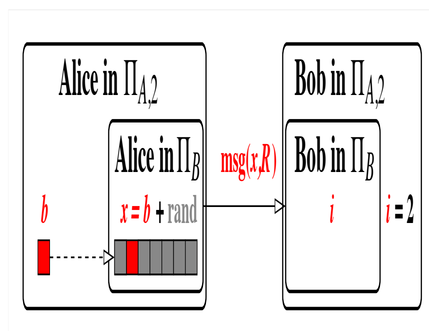

Level 3
Level 2
Level 1
1 0 0 1 1 1 0 0 1

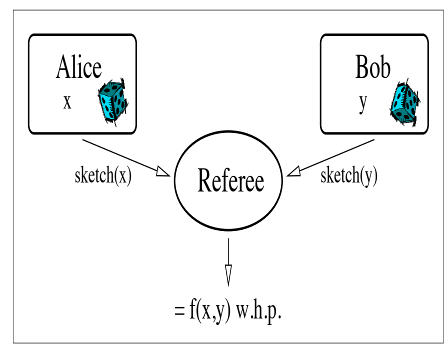

ALICE
BOB
σ x
σ y1
σ y2
σ y3
σ
σ
σ yn

THANKS!


d bits
ALICE
n points
d bits
BOB


X_1
X_2
X_3
X_4
Fixed value
Input value
Random value


Alice in $\Pi_{A,2}$
Alice in Π_B
b
$x = b + \text{rand}$
msg(x,R)
Bob in $\Pi_{A,2}$
Bob in Π_B
i
$i = 2$


Alice
x
Bob
y
sketch(x)
Referee
sketch(y)
= f(x,y) w.h.p.