The complexity of distributions

Emanuele Viola
Northeastern University
August 2010
Local functions

- **f**: \(\{0,1\}^n \rightarrow \{0,1\}\) **d-local**: output depends on \(d\) input bits

- **Fact**: Parity(x) = 1 ⇔ \(\sum x_i = 1 \mod 2\) is not \(n-1\) local

- **Proof**: Flip any input bit \(\Rightarrow\) output flips ♦
Local generation of \( (Y, \text{parity}(Y)) \)

- **Theorem** [Babai ‘87; Boppana Lagarias '87]

  There is \( f : \{0,1\}^n \rightarrow \{0,1\}^{n+1} \), each bit is 2-local.

  Distribution \( f(X) \equiv (Y, \text{parity}(Y)) \) \( (X, Y \in \{0,1\}^n \text{ uniform}) \)

\[
\begin{align*}
y_1 &= x_1 \\
y_2 &= x_1 \oplus x_2 \\
y_3 &= x_3 \oplus x_2 \\
... & \\
y_n &= x_{n-1} \oplus x_n \\
\text{parity}(y) &= x_n
\end{align*}
\]
• Complexity theory of distributions (as opposed to functions)

How hard is it to generate distribution $D$ given random bits?

E.g., $D = (Y, \text{parity}(Y))$, $D = W_k := \text{uniform n-bit with k 1's}$
Rest of this talk

- Connection with succinct data structures
- Lower bound for locally generating $W_{n/2} = n$-bit with $n/2$ 1's
- Decision tree model
- Bounded-depth circuit model (with Shachar Lovett)
Succinct data structures for sets

• Store $S \subseteq \{1, 2, \ldots, n\}$ of size $|S| = k$

In $u$ bits $b_1, \ldots, b_u \in \{0,1\}$

• Want:
  Small space $u$ (optimal $= \lceil \lg_2 \binom{n}{k} \rceil$)
  Answer “$i \in S$?” by probing few bits (optimal = 1)

• In combinatorics: Nešetřil Pultr, …, Körner Monti
Previous results

• Store $S \subseteq \{1, 2, \ldots, n\}$, $|S| = k$, in bits, answer “$i \in S$?”

• [Minsky Papert '69, Buhrman Miltersen Radhakrishnan Venkatesh; Pagh; ...; Pătraşcu; V. '09]

• Surprising upper bounds
  space = optimal + $o(n)$, probe $O(\log n)$

• No lower bounds for $k = n / 2^a$
General connection

- **Claim:** If store $S \subseteq \{1, 2, \ldots, n\}$, $|S| = k$ in $u = \text{optimal} + r$ bits answer “$i \in S$?” by (non-adaptively) probing $d$ bits.

Then $\exists f : \{0,1\}^u \rightarrow \{0,1\}^n$, $d$-local
Distance($f(X)$, $W_k = \text{uniform set of size } k$) $< 1 - 2^{-r}$

$$
\left( \text{Distance}(A, B) := \max_T \left| \Pr[A \in T] - \Pr[B \in T] \right| \right)
$$

- **Proof:** $f_i := “i \in S?”$

$f(X) = W_k$ with probability $(n \text{ choose } k) / 2^u = 2^{-r}$
Our result

- **Theorem (V.)** \( f : \{0, 1\}^{\text{optimal} + n^{O(1)}} \rightarrow \{0, 1\}^n \), \((d < \varepsilon \log n)\)-local. 
  \[ \text{Distance}(f(X), W_k = \text{uniform set of size } k = \Theta(n)) > 1 - n^{-\Omega(1)} \]

- Tight up to \( \Omega() \) if \( k = n/2 \): \( f(x) = x \), \((\text{binomial choose } n/2) = O(2^{n/\sqrt{n}})\)

- **Corollary:** To store \( S \subseteq \{1, 2, \ldots, n\} \), \( |S| = k = n / 2^a \) 
  answer “\( i \in S? \)” probing \( d < \varepsilon \log(n) \) bits: 
  Need space > optimal + \( \Omega(\log n) \)
Rest of this talk

- Connection with succinct data structures

- Lower bound for locally generating $W_{n/2} = n$-bit with $n/2$ 1's

- Decision tree model

- Bounded-depth circuit model
Our result

- **Theorem [V.]:** Let \( f : \{0,1\}^n \rightarrow \{0,1\}^n \) be \((d=O(1))\)-local.

  There is \( T \subseteq \{0,1\}^n \) such that
  \[
  \left| \Pr[f(x) \in T] - \Pr[W_{n/2} \in T] \right| > 1 - n^{-\Omega(1)}
  \]

- **Warm-up scenarios:**

  - \( f(x) = 000111 \) **Low-entropy** \( T := \{000111\} \)
    \[
    \left| \Pr[ f(x) \in T] - \Pr[W_{n/2} \in T] \right| = \left| 1 - |T| / \binom{n}{n/2} \right|
    \]

  - \( f(x) = x \) **"Anti-concentration"** \( T := \{ z : \sum_i z_i = n/2 \} \)
    \[
    \left| \Pr[ f(x) \in T] - \Pr[W_{n/2} \in T] \right| = \left| 1/\sqrt{n} - 1 \right|
    \]
Proof

- Partition input bits $X = (X_1, X_2, \ldots, X_s, H)$

- Fix $H$. Output block $B_i$ depends only on bit $X_i$

- Many $B_i$ constant ( $B_i(0,H) = B_i(1,H)$ ) $\Rightarrow$ low-entropy

- Many $B_i$ depend on $X_i$ ( $B_i(0,H) \neq B_i(1,H)$ )

Idea: Independent $\Rightarrow$ anti-concentration: can't sum to $n/2$
If many $B_i(0,H)$, $B_i(1,H)$ have different sum of bits, use Anti-concentration Lemma [Littlewood Offord]

For $a_1, a_2, ..., a_s \neq 0$, any $c$, $\Pr_{x \in \{0,1\}^s} \left[ \sum_i a_i x_i = c \right] < 1/\sqrt{n}$

Problem: $B_i(0,H) = 100$, $B_i(1,H) = 010$
high entropy but no anti-concentration

Fix: want many blocks 000, so high entropy $\Rightarrow$ different sum
Test $T \subseteq \{0,1\}^n$ : $\Pr[f(X_1,\ldots,X_s,H) \in T] \approx 1$ ; $\Pr[W_{n/2} \in T] \approx 0$

$z \in T \iff$

$\exists H : \exists X_1,\ldots,X_s$ w/ many blocks $B_i$ fixed : $f(X_1,\ldots,X_s,H) = z$

OR

Few blocks $z|_{B_i}$ are 000

OR

$\sum_i z_i \neq n/2$
Rest of this talk

- Connection with succinct data structures
- Lower bound for locally generating $W_{n/2} = \text{n-bit with n/2 1's}$
- Decision tree model
- Bounded-depth circuit model
Decision tree model

- $f : \{0,1\}^m \rightarrow \{0,1\}^n$ depth-$d$
  - each output bit $f_i$
  - is depth-$d$ decision tree

- Depth $d \subseteq 2^d$ local
Our result for decision trees

- **Theorem[V.]** $f : \{0,1\}^* \rightarrow \{0,1\}^n : \text{each bit depth} < 0.1 \log n$
  \[
  \text{Distance}(f(X), W_{n/2}) > n^{-\Omega(1)}
  \]

- Worse than $1 - n^{-\Omega(1)}$ bound for $O(1)$-local functions

- **Theorem**[Czumaj Kanarek Lorys Kutyłowski, V.]
  \[
  \exists f : \{0,1\}^* \rightarrow \{0,1\}^n : \text{each bit depth} O(\log n)
  \]
  \[
  \text{Distance}(f(X), W_{n/2}) < 1/n
  \]
Tool for lower bound proof

- Central limit theorem:
  \[ x_1, x_2, \ldots, x_n \text{ independent} \Rightarrow \sum x_i \approx \text{normal} \]

- Bounded-independence central limit theorem
  \[ \text{[Diakonikolas Gopalan Jaiswal Servedio V.]} \]
  \[ x_1, x_2, \ldots, x_n \text{ k-wise independent} \Rightarrow \sum x_i \approx \text{normal} \]

- Note: For next result, Paley–Zygmund inequality enough
Proof

- **Theorem**: \( f : \{0,1\}^* \rightarrow \{0,1\}^n \): each bit depth < 0.1 \( \log n \)
  \[
  \text{Distance}( f(X), W_{n/2} ) > n^{-\Omega(1)}
  \]

- **Proof**: Is output distribution \( f(X) \) \((k = 10)\)-wise independent?

  NO \( \Rightarrow \) \( W_{n/2} \approx k \)-wise independent

  Distance(those \( k \) bits, uniform on \( \{0,1\}^k \)) > 2^{-k(0.1 \log n)}
  (granularity of decision tree probability)

  YES \( \Rightarrow \) by prev. theorem \( \sum f(X)_i \approx \) normal

  so often \( \sum f(X)_i \neq n/2 \)
Rest of this talk

- Connection with succinct data structures
- Lower bound for locally generating $W_{n/2} = n$-bit with $n/2$ 1's
- Decision tree model
- Bounded-depth circuit model
Bounded-depth circuits

- More general model: small bounded-depth circuits (AC$^0$)

- **Challenge**: $\exists$ explicit boolean $f : \text{cannot generate } (Y, f(Y))$ ?

- **Theorem** [Matias Vishkin, Hagerup, Czumaj Kanarek Lorys Kutyłowski, V.]
  Can generate $(Y, \text{majority}(Y)) \quad \text{(exp. small error)}$

- **Theorem** [Lovett V.] Cannot generate error-correcting code
Lower bound for codes

- Code $C \subseteq \{0,1\}^n$ of size $|C| = 2^k = \Omega(n)$
  
  $x \neq y \in C \Rightarrow x, y$ far : hamming distance $\Omega(n)$

- Theorem [Lovett V.] $f : \{0,1\}^* \rightarrow \{0,1\}^n$, $f \in AC^0$
  
  Distance($f(X)$, uniform over $C$) $> 1 - n^{-\Omega(1)}$

- Consequences for data structures for codewords, complexity of pseudorand. generators against $AC^0$ [Nisan]
Warm-up

- **Fact**: $f : \{0,1\}^k \to \{0,1\}^n$, $f \in AC^0$
  f cannot **compute encoding** function of $C$,
  mapping message $m \in \{0,1\}^k$ to codeword

- **Proof**: 

- [Linial Mansour Nisan, Boppana] **low sensitivity of $AC^0$**: 
  $m, m'$ random at hamming distance 1
  $\Rightarrow f(m), f(m')$ **close** in hamming distance.

  - But $f(m) \neq f(m') \in C \Rightarrow$ **far** in hamming distance
Lower bound for codes

- **Theorem** [Lovett V.]: \( f : \{0,1\}^L \gg k \rightarrow \{0,1\}^n, f \in \text{AC}^0 \)
  \[
  \text{Distance}(f(X), \text{uniform over } C) > 1 - n^{-\Omega(1)}
  \]

**Problem**: \( f \) needs not compute encoding function.
Input length \( \gg \) message length

- **Idea**: Input \( \{0,1\}^L \) to \( f \) partitioned in \( |C| \) sets

- **Isoperimetric inequality** [Harper, Hart]:
  Random \( m, m' \) at distance 1 often in \( \neq \) sets \( \Rightarrow \) low sensitivity
Theorem [Lovett V.] \( f : \{0,1\}^L \rightarrow \{0,1\}^n \), \( f \in AC^0 \)

Distance(f(X), uniform over C) > 1 - \( n^{-\Omega(1)} \)

Note: to get
Need isoperimetric inequality for m, m' at distance >> 1

Fact [thanks to Samorodnitsky] \( \forall A \subseteq \{0,1\}^L \) of density \( \alpha \)
random m, m' obtained flipping bits w/ probability p:
\( \alpha^2 \leq \Pr[\text{both } m \in A \text{ and } m' \in A] \leq \alpha^{1/(1-p)} \)
Complexity of generators against $\text{AC}^0$

- Pseudorandom generator against circuit of depth $d$ (want: reduce randomness w/ minimum overhead)

- **Direct implementation** of Nisan's generator takes depth $\geq d$ (circuit + generator $\rightarrow$ depth $2d$)

- [Lovett V.] Generating output distribution of Nisan's generator takes depth $\geq d$ (for some choice of designs)

- [V.] Generator in depth 2 (circuit + generator $\rightarrow$ depth $d+1$) [Braverman] + [Guruswami Umans Vadhan]
Conclusion

- Complexity of distributions = uncharted territory

- Lower bound for generating $W_k$ locally
  $\Rightarrow$ lower bound for succinct data structures for storing sets of size $n / 2^a$

- Lower bound for decision trees

- Lower bound for bounded-depth circuits ($AC^0$)
Recall: edit style changes ALL settings.

Click on “line” for just the one you highlight
More connections

- More uses of generating $W_k := \text{uniform n-bit string with k 1's}$
- McEliece cryptosystem
- Switching networks, …
Previous results

- Store $S \subseteq \{1, 2, \ldots, n\}$, $|S| = k$, in bits, answer “$i \in S$?”

- [Minsky Papert '69] Average-case study

- [Buhrman Miltersen Radhakrishnan Venkatesh; Pagh '00]
  
  Space $O(\text{optimal})$, probe $O(1)$ when $k = \Theta(n)$

  Lower bounds for $k < n^{1-\varepsilon}$

- [..., Pagh, Pătrașcu] space = optimal $+ o(n)$, probe $O(\log n)$

- [V. '09] lower bounds for $k = \Omega(n)$, except $k = n / 2^a$