The complexity of distributions: boolean average-case lower bounds

Emanuele Viola

Northeastern University

April 2018

The complexity of distributions:

boolean average-case lower bounds

Classical talk at quantum workshop

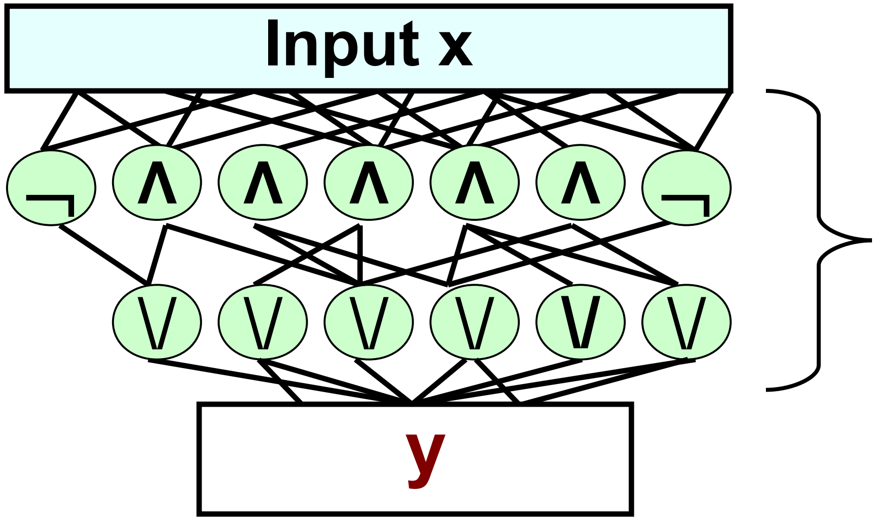
The complexity of distributions

- Leading goal of computational complexity: lower bounds for computing a function on a given input

- Since 2009 have advocated lower bounds for sampling distributions, given uniform bits

- Several papers, connections, still uncharted

Bounded-depth circuits (AC^0)



$\bigvee$ = or
$\bigwedge$ = and
$\neg$ = not

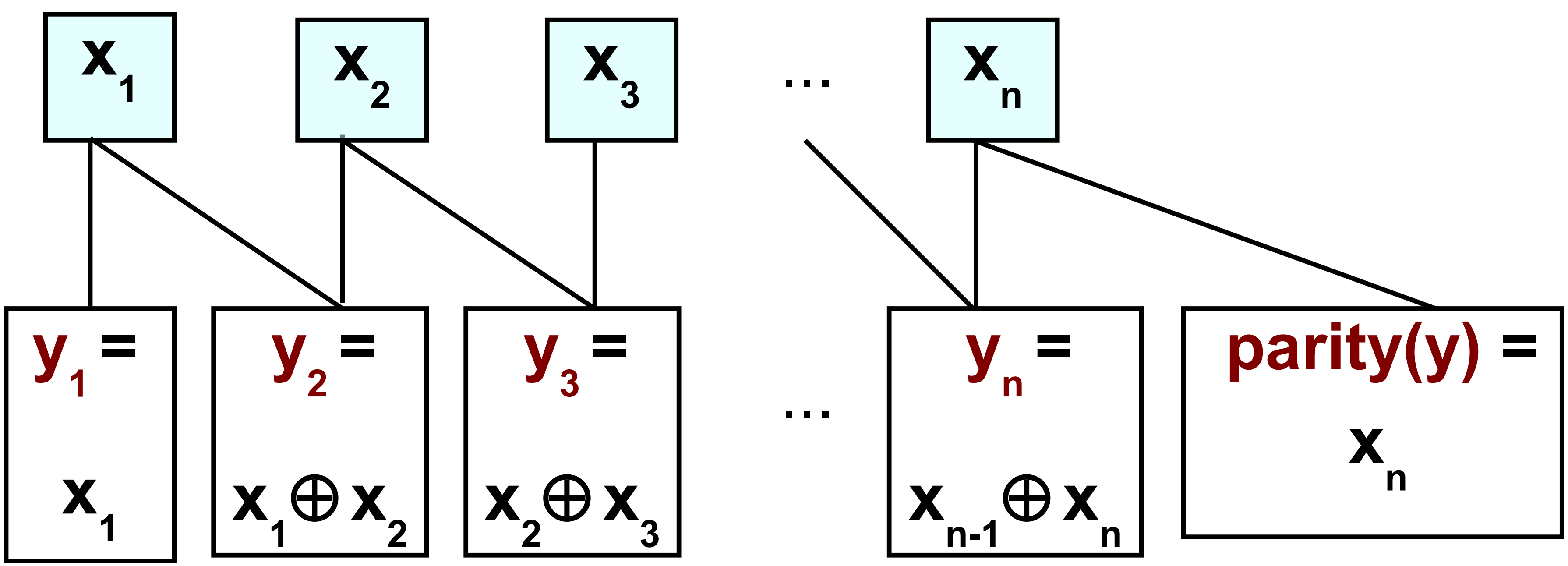
- AC^0 cannot compute parity
  [1980's: Furst Saxe Sipser, Ajtai, Yao, Hastad, ….]

Sampling (Y, parity(Y))

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

There is $f : \{0,1\}^n \rightarrow \{0,1\}^{n+1}$, in AC^0

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



AC^0 can sample

- (Y, Inner-Product(Y)) [Impagliazzo Naor]
- Permutations (error 2^{-n}) [Matias Vishkin, Hagerup]
- (Y, f(Y)), any symmetric f (error 2^{-n}) [V]

  e.g. f = Majority, Mod-3, ...

AC^0 cannot sample

AC^0 cannot sample

- Error-correcting codes [Lovett V 2011, Beck Impagliazzo Lovett]

  Z = uniform on good binary code $\subseteq \{0,1\}^n$

  AC^0 circuit $C : \{0,1\}^L \to \{0,1\}^n$

  ➔ Statistical-Distance(Z, C(X)) $\geq 1 - \exp(-n^{0.1})$

AC^0 cannot sample

- Error-correcting codes [Lovett V 2011, Beck Impagliazzo Lovett]

  Z = uniform on good binary code $\subseteq \{0,1\}^n$

  AC^0 circuit C : $\{0,1\}^L \to \{0,1\}^n$

  ➔ Statistical-Distance(Z, C(X)) $\geq 1 - \exp(-n^{0.1})$

- (Y, f(Y)) for bit-block extractor f : $\{0,1\}^n \to \{0,1\}$

  Statistical-Distance((Y, f(Y) ,C(X)) > 0 [V 2011]

AC^0 cannot sample

- Error-correcting codes [Lovett V 2011, Beck Impagliazzo Lovett]

  Z = uniform on good binary code $\subseteq \{0,1\}^n$

  AC^0 circuit $C : \{0,1\}^L \to \{0,1\}^n$

  ➔ Statistical-Distance(Z, $C(X)$) $\geq 1 - \exp(-n^{0.1})$

- $(Y, f(Y))$ for bit-block extractor $f : \{0,1\}^n \to \{0,1\}$

  Statistical-Distance($(Y, f(Y))$,$C(X)$) > 0 [V 2011]

  $> 1/2 - 1/n^{\omega(1)}$ [V now]

"Cannot compute f better than tossing a coin, even if you can sample the input yourself"

AC^0 cannot sample

- Error-correcting codes [Lovett V 2011, Beck Impagliazzo Lovett]

  Z = uniform on good binary code $\subseteq \{0,1\}^n$

  AC^0 circuit $C : \{0,1\}^L \to \{0,1\}^n$

  ➔ Statistical-Distance(Z, C(X)) $\geq 1 - \exp(-n^{0.1})$

Next

- (Y, f(Y)) for bit-block extractor $f : \{0,1\}^n \to \{0,1\}$

  Statistical-Distance((Y, f(Y) ,C(X)) > 0 [V 2011]

  $> 1/2 - 1/n^{\omega(1)}$ [V now]

"Cannot compute f better than tossing a coin, even if you can sample the input yourself"

- Theorem: AC^0 circuit C
  min-entropy $C(X) \geq k$ $(\forall\ a, \Pr[C(X) = a] \leq 2^{-k})$
  ➔ C(X) close to convex combination of bit-block sources
  with min-entropy $\geq k$ (k/n)

- Bit-block source: each bit is either constant or literal
  Example: $(0, 1, z_5, 1\text{-}z_3, z_3, z_3, 0, z_2)$

- Corollary: f bit-block extractor ➔ $C(X) \neq (Y, f(Y))$

- Proof:

- Theorem: AC^0 circuit C
  min-entropy $C(X) \geq k$ ($\forall$ a, $Pr[C(X) = a] \leq 2^{-k}$)
  ➔ $C(X)$ close to convex combination of bit-block sources
  with min-entropy $\geq k$ (k/n)

- Bit-block source: each bit is either constant or literal
  Example: $(0, 1, z_5 , 1\text{-}z_3 , z_3 , z_3 , 0 , z_2)$

- Corollary: f bit-block extractor ➔ $C(X) \neq (Y, f(Y))$

- Proof: $C(X) = (Y, f(Y))$ ➔ min-entropy $C(X) \geq |Y| = n$
  ➔ convex combination high min-entropy bit-block sources
  can fix "f(Y)" bit leaving high min-entropy
  contradicts extractor property QED

- Theorem: AC^0
  min-entropy C
  ➔ C(X) close
  with min-e

Rules out Statistical-Distance 0, but not 0.1

Possible:
Statistical-Distance(C(X), (Y,f(Y)) ≤ 0.1,
but min-entropy C(X) = O(1)

Example next

- Bit-block sour
  Example: (0,

- Corollary: f bit

- Proof: C(X) = (Y, f(Y)) ➔ min-entropy C(X) $\geq |Y| = n$
  ➔ convex combination high min-entropy bit-block sources
  can fix "f(Y)" bit leaving high min-entropy
  contradicts extractor property QED

Example

- Circuit C: “On input x:

  If first 4 bits are 0 output the all-zero string

  Otherwise sample (Y, f(Y)) exactly”

- Statistical-Distance(C(X) , (Y, f(Y)) ≤ 0.1,

  but min-entropy C(X) = O(1)

- Observation: If you fix first 4 bits,

  min-entropy polarizes: either zero or very large

  We show this happens for every AC^0 circuit

Polarizing min-entropy

- Theorem: For every AC^0 circuit $C : \{0,1\}^L \to \{0,1\}^n$
  $\exists$ set S of $\exp(n - n^{0.9})$ restrictions such that:

  (1) preserve output distribution

  $C|_r (X) \approx C(X)$ for uniform $r \in S$

  (2) polarize min-entropy

  $\forall\, r \in S$, $C|_r$ has min-entropy 0 or $n^{0.8}$

- Note: $|S| = \exp(n)$ useless and trivial:
  S := one input for each of $\leq 2^n$ outputs, entropy always 0

Proof steps

- AC^0
- Decision trees
- Polarized decision trees

Small set of restrictions that
(1) preserve output distribution
hypercontractivity +
specific concentration of measure

(2) collapse AC^0 to decision trees
switching lemma

Further restrict
tree either fixed or has high min entropy

Conclusion

- Open problem: Statistical distance 1/2 - $\exp(-n^{0.1})$
  Neither in reduction to bit-block nor entropy polarization

- Much more to chart...