Entropy of Operators
or
Nechiporuk for Depth-2 Circuits

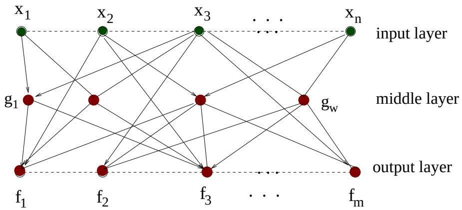
Stasys Jukna

Institute of Mathematics, Vilnius, Lithuania
University of Frankfurt, Germany

Dagstuhl 2008

Model = Circuits with Arbitrary Gates

- Operator $F : \{0,1\}^n \to \{0,1\}^n$
  - $\Rightarrow$ set of boolean functions $F = \{f_1, \ldots, f_n\}$
- Want to simultaneously compute all functions in F
- $Wires_2(F)$ = Number of wires in depth-2 circuit
- Arbitrary boolean functions as gates !
  - $\Rightarrow$ $Wires_2(F) \leq n^2$, even in depth-1
  - $\Rightarrow$ Problem not computation but information transfer



- Depth-2 already non-trivial !
- $f_i(\vec{x}) = \varphi(\vec{x}, g_1(\vec{x}), \ldots, g_k(\vec{x}))$, $\varphi, g_1, \ldots, g_k$ arbitrary boolean functions

Why Depth-2 Interesting?

Valiant 1977 (Reduction to Depth-2)

Let $r = O(n/\ln\ln n)$ gates on the middle layer are given for free. If $Wires_2(F) = \Omega(n^{1+\epsilon})$ then F cannot be computed in log-depth and linear size.

- Best known: $\Omega((n^2/r)\ln(n/r))$ [Pudlák/Rödl/Sgal 1997]
- Weakening: $Wires_2(F)$ = count all wires (nothing "for free")
- Graph-theoretical approach based on *superkoncentrators*

  Long history: Pippenger, Valiant, Wigderson, Pudlák, Raz, Shpilka, ...
- Best: $\Theta(n\ln^2 n/\ln\ln n)$ [Radhakrishnan/Ta-Shma 2003]
- Too weak since too strong: Establish *graph structure* of circuits

Direkt Approach

- Idea: count the wires directly, ignore the graph-structure
- [Cherukhin 2005]: $Wires_2(Conv_n) = \Omega(n^{3/2}) \quad \Rightarrow$ a breakthough
- Question 1: What about other operators? Say, matrix product $F = X \cdot Y$?
- Question 2: Why some operators are difficult to compute?

Our result

For all $F : \{0,1\}^n \to \{0,1\}^m$ we have $Wires_2(F) \geq Entropy(F)$

- Conection to classics:
  - Proof = reminiscent of Nechiporuk's argument for formula size
- Yields new and/or higher lower bounds:
  - Tight bound for matrix product and a lot of other operators
  - Entropy + Raz&Shpilka $\Rightarrow$ higher lower bounds for any constant depth [Cherukhin 2008]

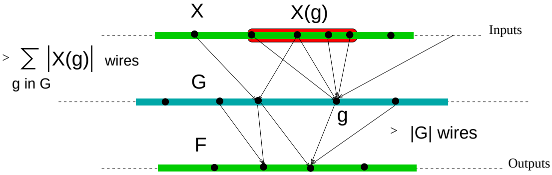
Nechiporuk's Argument for Formula Size

- Boolean function $f(X)$
- $L(f)$ = formula size universal basis (all fanin ≤ 2 functions)
- Partite variables $X = X_1 \cup \cdots \cup X_p$
- $\mathrm{Subf}(f|X_i) := \Big\{ f(X_i, \vec{a}) | \text{ assignments } \vec{a} : X \setminus X_i \to \{0,1\} \Big\}$

Nechiporuk 1966

$$L(f) \geq \sum_{i=1}^{p} \log \#\mathrm{Subf}(f|X_i)$$

Proof Idea

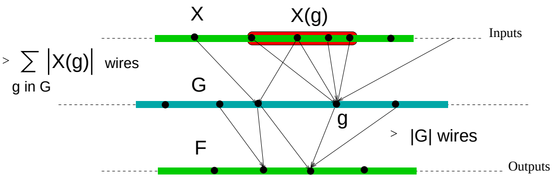


- Subfunctions $\mathrm{Subf}(g) := \{g(\vec{0}, Y), g(\vec{e}_1, Y), \ldots, g(\vec{e}_n, Y)\}$
- If no wire from x_i to g then $g(\vec{e}_i, Y) = g(\vec{0}, Y)$

  $\Rightarrow$ one and the same subfunction for all non-wires !
- $\Rightarrow$ number of different subfunctions $|\mathrm{Subf}(g)| \leq \deg(g)$
- $\Rightarrow$ $\#$ Wires $:= \sum_{g \in G} \deg(g) \geq \sum_{g \in G} |\mathrm{Subf}(g)| \geq |G^*|$

  where $G^* = \bigcup_{g \in G} \mathrm{Subf}(g)$ set of all subfunctions

Proof (end)



- Know: $Wires_2(F) \geq$ number $|G^*|$ of subfunctions on the middle layer
- But ... don't know what G, and hence, what G^* is ...
- Still know: F^* must be computable from G^*

  $\Rightarrow$ If $S \subseteq \{0,1\}^Y$ is separated by F^*

  $$\text{i.e. } \forall \vec{a} \neq \vec{b} \in S \; \exists h \in F^* : \quad h(\vec{a}) \neq h(\vec{b})$$

  then S must be also separated by G^*
- Define $Entropy(G) := \max\{\log_2 |S| : S \text{ is separated by } G^*\}$

  $\Rightarrow$ $Entropy(F) \leq Entropy(G) \leq |G^*| \leq Wires_2(F)$ □

Why Entropy, Not just Number of Subfunctions?

- We need: If F computable from G then $|G| \geq \mu(F)$
- Why not take $\mu(F) =$ number $|F|$ of different subfunctions, as Nechiporuk?
- Answer: Because gates are arbitrary boolean functions
- If $G = \{g_1, \ldots, g_k\}$ then

$$F = \{\varphi_1(g_1, \ldots, g_k), \ldots, \varphi_N(g_1, \ldots, g_k)\}$$

  is computable from G but

$$|F| \geq 2^{2^k} \gg k = |G| \quad \Rightarrow \quad |G| \ll \mu(F)$$

- Nechiporuk could do this because $k = 2$ for formulas

How to Apply?

- Boolean function $f(X)$
- Partite variables $X = X_1 \cup \cdots \cup X_p$
- Define $\mathit{Entropy}(f) := \sum_{i=1}^{p} \log \#\mathrm{Subf}(f|X_i)$
- Nechiporuk: $L(f) \geq \mathit{Entropy}(f)$

- Operator $F(X, Y)$ = set of boolean functions $f(X, Y)$
- Partite variables $X = X_1 \cup \cdots \cup X_p$ and operator $F = F_1 \cup \cdots \cup F_p$
- Know: For all $i = 1, \ldots, p$

$$\#(\text{Wires from } X_i) + \#(\text{Wires to } F_i) \geq \mathit{Entropy}(F_i|X_i)$$

$$\Rightarrow \quad \mathit{Wires}_2(F) \geq \sum_{i=1}^{p} \mathit{Entropy}(F_i|X_i)$$

What Operators have Large Entropy?

- Entropy of $F(X, Y) = \log_2 \#$ Range of suboperator
- $F(X, Y)$ isolates a variable $y \in Y$ if $\exists i$: $y \in F(\vec{e}_i, Y)$
- F isolates all variables in $Y \quad \Rightarrow \quad Entropy(F) \geq |Y|$
- Scalar product $f(X, Y) = \sum_{i=1}^{m} x_i y_i$ isolates all variables $y \in Y$

$$f(\vec{e}_j, Y) = 0 \cdot y_1 + \cdots + 0 \cdot y_{j-1} + 1 \cdot y_j + +0 \cdot y_{j+1} + \cdots + +0 \cdot y_m = y_j$$

$\Rightarrow$ operators based on quadratic forms have large entropy

- Natural candidate: $m \times m$ Matrix product = set of $n = m^2$ scalar products

Entropy of Matrix Product

- $Mult_n$:=Product of two $m \times m$ matrices; input size $2n$ with $n = m^2$
- Trivial upper bound $Wires_2(Mult_n) \leq nm = n^{1.5}$ (even in depth-1)
- Any depth: $O(n^{1.2})$ wires enough (fanin-2) [V. Strassen 1973]
- $2.5n$ gates necessary (fanin-2) [N. Bshouthy 1982]
- Best for depth-2: $Wires_2(Mult_n) = \Omega(n \log n)$ [Raz-Spilka 2003]

Entropy of Matrix Product

$Entropy(Mult_n) = \Omega(n^{1.5}) \quad \Rightarrow \quad Wires_2(Mult_n) = \Theta(n^{1.5})$

Proof.

$F = X \cdot Y \quad \Rightarrow$ partite row-wise $X = X_1 \cup \cdots \cup X_m$, $F = F_1 \cup \cdots \cup F_m$
Scalar products $\quad \Rightarrow \quad Entropy(F_i|X_i) \geq |Y| = n$ for all $i = 1, \ldots, m = \sqrt{n}$
$\Rightarrow \quad Entropy(F) \geq m \cdot n = n^{1.5}$ □

Problems

Problem 1 (probably YES, try counting)

What is $\max_F Wires_2(F)$ over all $F : \{0,1\}^n \to \{0,1\}^n$? Is it $\Omega(n^2)$?

- The same for linear operators $f_A(\vec{x}) = A\vec{x}$
- Best explicit: $Wires_2(f_A) = \Omega(n \ln n)$ [Pudlák (*Combinatorica* 1994)]

Problem 2 (surprising if YES, much harder but tractable)

What is $\max_A Wires_2(f_A)$? Is it $\Omega(n^2 / \ln n)$?

Problem 3 (THE challenge, hard)

Extend the entropic approach to Valiant's depth-2 circuits, where $O(n/\ln\ln n)$ gates on the middle are given for free.

- Would yield super-linear lower bound for NC^1-circuits

END