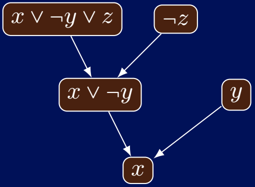
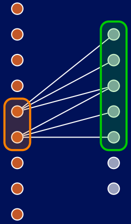
A Lower Bound for k-DNF Resolution on Random CNF Formulas via Expansion





Dmitry Sokolov
joint work with Anastasia Sofronova

Simons Institute
June 17, 2022

St Petersburg University

PDMI
RAS

Proof Systems

Definition[Cook, Reckhow 79]

Proof system for $L \Leftrightarrow$ poly-time algorithm Π: $\{0,1\}^* \times \{0,1\}^* \to \{0,1\}$:

- (completeness) $x \in L \Rightarrow \exists w\ \Pi(x,w) = 1$;
- (soundness) $\exists w\ \Pi(x,w) = 1 \Rightarrow x \in L$.

Resolution: proof of $\varphi \coloneqq \bigwedge_i C_i$ is a sequence of clauses $(D_1, D_2, D_3, \ldots, D_\ell)$:

Proof Systems

Definition[Cook, Reckhow 79]

Proof system for L $\Leftrightarrow$ poly-time algorithm Π: $\{0,1\}^* \times \{0,1\}^* \to \{0,1\}$:

- (completeness) $x \in L \Rightarrow \exists w\ \Pi(x,w) = 1$;
- (soundness) $\exists w\ \Pi(x,w) = 1 \Rightarrow x \in L$.

Resolution: proof of $\varphi \coloneqq \bigwedge_i C_i$ is a sequence of clauses $(D_1, D_2, D_3, \dots, D_\ell)$:

- $D_i \in \{C_i\}$;

Proof Systems

Definition[Cook, Reckhow 79]

Proof system for L $\Leftrightarrow$ poly-time algorithm Π: $\{0,1\}^* \times \{0,1\}^* \to \{0,1\}$:

- (completeness) $x \in L \Rightarrow \exists w\ \Pi(x,w) = 1$;
- (soundness) $\exists w\ \Pi(x,w) = 1 \Rightarrow x \in L$.

Resolution: proof of $\varphi \coloneqq \bigwedge_i C_i$ is a sequence of clauses $(D_1, D_2, D_3, \ldots, D_\ell)$:

- $D_i \in \{C_i\}$;
- $\frac{A \vee x \quad B \vee \bar{x}}{A \vee B}$, $D_i \coloneqq A \vee B$;

Proof Systems

Definition[Cook, Reckhow 79]

Proof system for $L \Leftrightarrow$ poly-time algorithm Π: $\{0,1\}^* \times \{0,1\}^* \to \{0,1\}$:

- (completeness) $x \in L \Rightarrow \exists w\ \Pi(x,w) = 1$;
- (soundness) $\exists w\ \Pi(x,w) = 1 \Rightarrow x \in L$.

Resolution: proof of $\varphi \coloneqq \bigwedge_i C_i$ is a sequence of clauses $(D_1, D_2, D_3, \ldots, D_\ell)$:

- $D_i \in \{C_i\}$;
- $\frac{A \vee x \quad B \vee \bar{x}}{A \vee B}$, $D_i \coloneqq A \vee B$;
- $D_\ell = \varnothing$.

Proof Systems

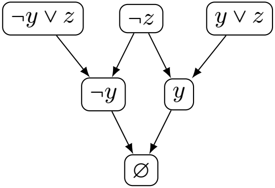
Definition[Cook, Reckhow 79]

Proof system for $L \Leftrightarrow$ poly-time algorithm Π: $\{0,1\}^* \times \{0,1\}^* \to \{0,1\}$:

- (completeness) $x \in L \Rightarrow \exists w\ \Pi(x,w) = 1$;
- (soundness) $\exists w\ \Pi(x,w) = 1 \Rightarrow x \in L$.

Resolution: proof of $\varphi \coloneqq \bigwedge_i C_i$ is a sequence of clauses $(D_1, D_2, D_3, \dots, D_\ell)$:

- $D_i \in \{C_i\}$;
- $\frac{A \vee x \quad B \vee \bar{x}}{A \vee B}$, $D_i \coloneqq A \vee B$;
- $D_\ell = \varnothing$.



Proof Systems

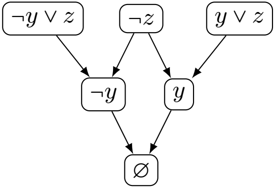
Definition[Cook, Reckhow 79]

Proof system for $L \Leftrightarrow$ poly-time algorithm $\Pi\colon \{0,1\}^* \times \{0,1\}^* \to \{0,1\}$:

- (completeness) $x \in L \Rightarrow \exists w\ \Pi(x,w) = 1$;
- (soundness) $\exists w\ \Pi(x,w) = 1 \Rightarrow x \in L$.

Resolution: proof of $\varphi \coloneqq \bigwedge_i C_i$ is a sequence of clauses $(D_1, D_2, D_3, \ldots, D_\ell)$:

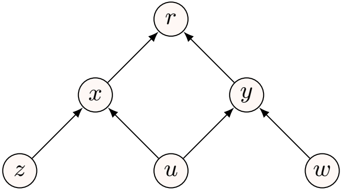
- $D_i \in \{C_i\}$;
- $\frac{A \vee x \quad B \vee \bar{x}}{A \vee B}$, $D_i \coloneqq A \vee B$;
- $D_\ell = \varnothing$.



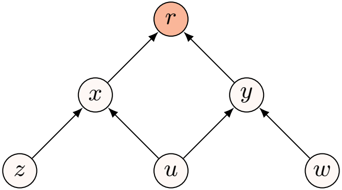
Cutting Planes: proof is a sequence of inequalities over $\mathbb{Z}$ $(p_1 \geq 0, p_2 \geq 0, p_3 \geq 0, \ldots, p_\ell \geq 0)$:

- p_i is an encoding of $C \in \varphi$, $x_k \geq 0$ or $-x_k + 1 \geq 0$;
- $\frac{p_i \quad p_j}{p_k}$, $(p_i \geq 0) \wedge (p_j \geq 0)$ imply $(p_k \geq 0)$ over $\mathbb{Z}^n$;
- $p_\ell = 1$.

Pebbling

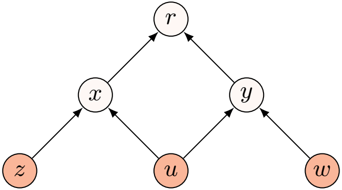


Pebbling



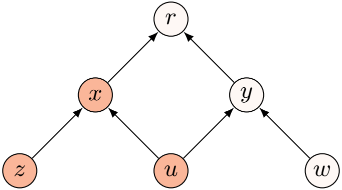
- $(\neg r)$;

Pebbling



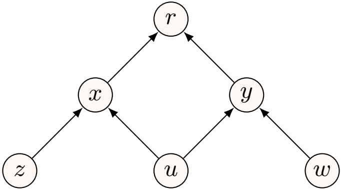
- $(\neg r)$;
- $(z), (u), (w)$;

Pebbling



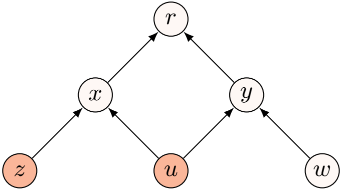
- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \vee \neg u \vee x)$, $(\neg u \vee \neg w \vee y)$, $(\neg x \vee \neg y \vee r)$.

Pebbling



- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \vee \neg u \vee x)$, $(\neg u \vee \neg w \vee y)$, $(\neg x \vee \neg y \vee r)$.

Pebbling



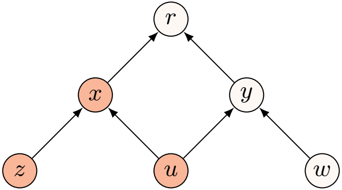
- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \vee \neg u \vee x)$, $(\neg u \vee \neg w \vee y)$, $(\neg x \vee \neg y \vee r)$.

u

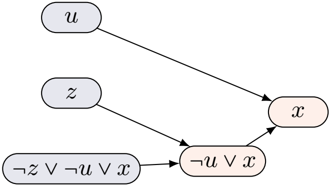
z

$\neg z \vee \neg u \vee x$

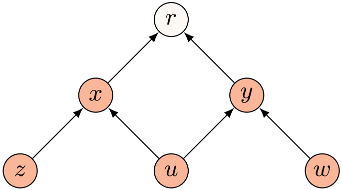
Pebbling



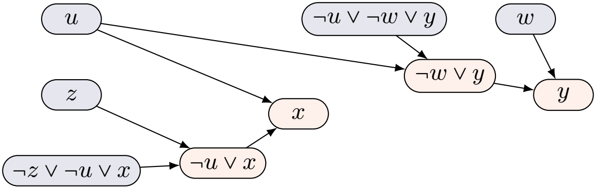
- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \vee \neg u \vee x),\ (\neg u \vee \neg w \vee y),\ (\neg x \vee \neg y \vee r)$.



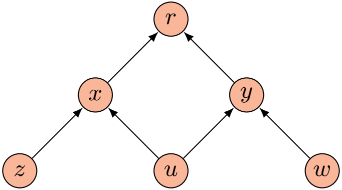
Pebbling



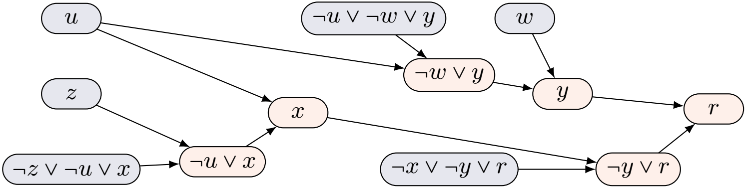
- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \lor \neg u \lor x)$, $(\neg u \lor \neg w \lor y)$, $(\neg x \lor \neg y \lor r)$.



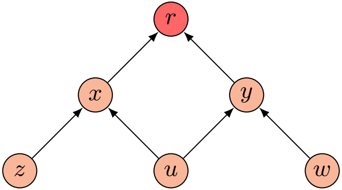
Pebbling



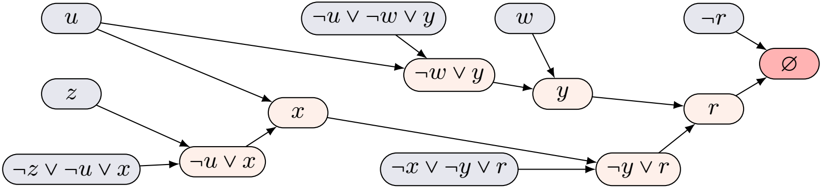
- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \lor \neg u \lor x)$, $(\neg u \lor \neg w \lor y)$, $(\neg x \lor \neg y \lor r)$.



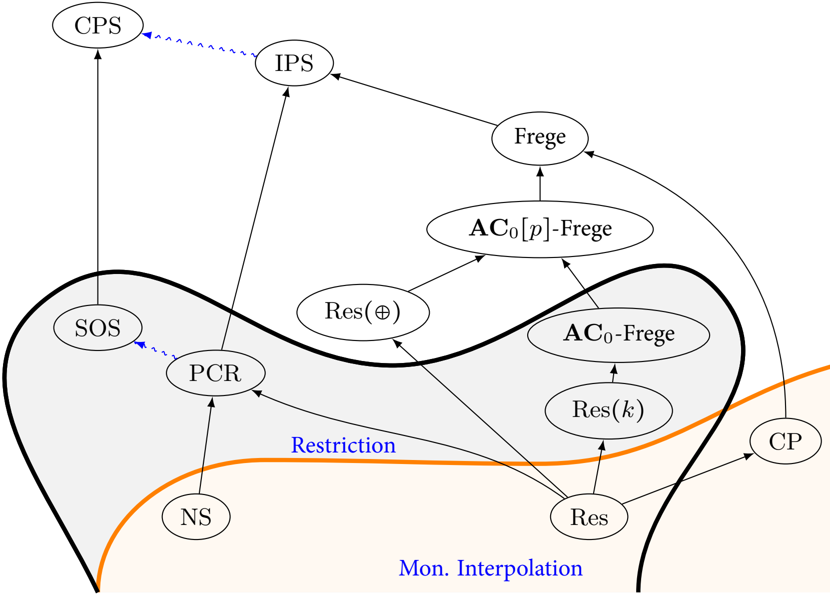
Pebbling



- $(\neg r)$;
- $(z), (u), (w)$;
- $(\neg z \lor \neg u \lor x)$, $(\neg u \lor \neg w \lor y)$, $(\neg x \lor \neg y \lor r)$.



Lower bounds in proof complexity



Hard formulas for all proof systems

- If φ is unsatisfiable then there is a “proof” of unsatisfiability.

Hard formulas for all proof systems

- If φ is unsatisfiable then there is a “proof” of unsatisfiability.
  - And we can realize it in some proof system...

Hard formulas for all proof systems

- If φ is unsatisfiable then there is a "proof" of unsatisfiability.
  - And we can realize it in some proof system...
- Distribution on formulas?

Hard formulas for all proof systems

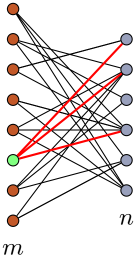
- If φ is unsatisfiable then there is a “proof” of unsatisfiability.
  - And we can realize it in some proof system...
- Distribution on formulas?
  - Fine. Counting argument do not work in proof complexity.

Hard formulas for all proof systems

- If φ is unsatisfiable then there is a "proof" of unsatisfiability.
  - And we can realize it in some proof system...
- Distribution on formulas?
  - Fine. Counting argument do not work in proof complexity.

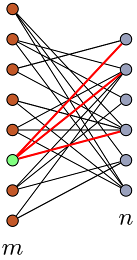
- Random Δ-CNF formulas
- Clique formulas
- Pseudorandom generator formulas

Random Δ-CNF



- m clauses;
- n variables;
- Δ neighbours: $\binom{n}{\Delta}$ possibilities;
- negations (uniformly at random);
- $\mathfrak{D} := \frac{m}{n}$ clause density.

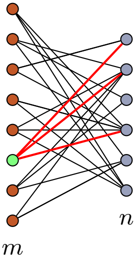
Random Δ-CNF



- m clauses;
- n variables;
- Δ neighbours: $\binom{n}{\Delta}$ possibilities;
- negations (uniformly at random);
- $\mathfrak{D} \coloneqq \frac{m}{n}$ clause density.

- $\mathfrak{D} > c_\Delta 2^\Delta \Rightarrow$ formula is unsat whp;

Random Δ-CNF



- m clauses;
- n variables;
- Δ neighbours: $\binom{n}{\Delta}$ possibilities;
- negations (uniformly at random);
- $\mathfrak{D} \coloneqq \frac{m}{n}$ clause density.

- $\mathfrak{D} > c_\Delta 2^\Delta \Rightarrow$ formula is unsat whp;
- Fiege's conjecture: $\mathfrak{D} = \mathcal{O}(1) \Rightarrow$ no poly-time algorithm may "prove" unsatisfiability of random $\mathcal{O}(1)$-CNF.
  - Non-approximability of many problems.

k-DNF Resolution

- Resolution with extension variables for conjunctions of k literals.

k-DNF Resolution

- Resolution with extension variables for conjunctions of k literals.
-
  - $\frac{F}{F \vee \ell}$;
  - $\frac{F \vee \ell_1, \ldots, F \vee \ell_w}{F \vee (\bigwedge_{i=0}^{w} \ell_i)}$;
  - $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i)}{F \vee \ell_i}$;
  - $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i) \quad G \vee (\bigvee_{i=0}^{w} \neg \ell_i)}{F \vee G}$.

k-DNF Resolution

- Resolution with extension variables for conjunctions of k literals.
-
  - $\frac{F}{F \vee \ell}$;
  - $\frac{F \vee \ell_1, \ldots, F \vee \ell_w}{F \vee (\bigwedge_{i=0}^{w} \ell_i)}$;
  - $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i)}{F \vee \ell_i}$;
  - $\frac{F \vee (\bigwedge_{i=0}^{w} \ell_i) \quad G \vee (\bigvee_{i=0}^{w} \neg \ell_i)}{F \vee G}$.
- Top-down (informal): decision "tree" with conjunctions of k literals.

Prior results

Proof system	Upper bound (poly)	Lower bound ($2^{n^{\varepsilon}}$)

Prior results

Proof system	**Upper bound** (poly)	**Lower bound** ($2^{n^{\varepsilon}}$)
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \leq n^{(\Delta-2)/4}, \Delta \geq 3$

Prior results

Proof system	Upper bound (poly)	Lower bound ($2^{n^{\varepsilon}}$)
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \leq n^{(\Delta-2)/4}, \Delta \geq 3$
PCR	$\cdot$	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$

Prior results

Proof system	Upper bound (poly)	Lower bound (2^{n^ε})
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \leq n^{(\Delta-2)/4}, \Delta \geq 3$
PCR	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$
SOS	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$

Prior results

Proof system	Upper bound (poly)	Lower bound ($2^{n^{\varepsilon}}$)
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \leq n^{(\Delta-2)/4}, \Delta \geq 3$
PCR	·	$\mathfrak{D} = \text{poly}(n), \Delta \geq 3$
SOS	·	$\mathfrak{D} = \text{poly}(n), \Delta \geq 3$
CP	·	$\mathfrak{D} = \text{poly}(n), \Delta = \Omega(\log n)$

Prior results

Proof system	**Upper bound** (poly)	**Lower bound** ($2^{n^{\varepsilon}}$)
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \leq n^{(\Delta-2)/4}, \Delta \geq 3$
PCR	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$
SOS	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$
CP	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta = \Omega(\log n)$
$\mathbf{TC}_0$-Frege	$\Delta = 3, \mathfrak{D} > n^{0.4}$	×

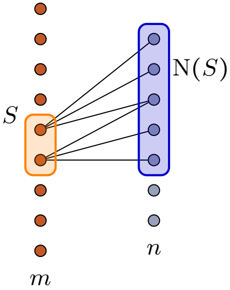
Prior results

Proof system	**Upper bound** (poly)	**Lower bound** $(2^{n^{\varepsilon}})$
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \le n^{(\Delta-2)/4}, \Delta \ge 3$
PCR	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \ge 3$
SOS	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \ge 3$
CP	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta = \Omega(\log n)$
$\mathbf{TC}_0$-Frege	$\Delta = 3, \mathfrak{D} > n^{0.4}$	×
$\mathrm{Res}(k)$	·	$\mathfrak{D} = \mathcal{O}(1), \Delta \ge 3, k = \mathcal{O}\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ $\mathfrak{D} = n^{1/6}, \Delta = \mathcal{O}(k^2), k = \mathcal{O}(1)$

Prior results

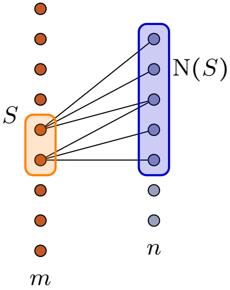
Proof system	Upper bound (poly)	Lower bound (2^{n^ε})
Resolution	$\mathfrak{D} > \frac{n^{\Delta-2}}{\log^{\Delta-2} n}$	$\mathfrak{D} \leq n^{(\Delta-2)/4}, \Delta \geq 3$
PCR	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$
SOS	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta \geq 3$
CP	·	$\mathfrak{D} = \mathsf{poly}(n), \Delta = \Omega(\log n)$
$\mathbf{TC}_0$-Frege	$\Delta = 3, \mathfrak{D} > n^{0.4}$	×
$\mathrm{Res}(k)$	·	$\mathfrak{D} = \mathcal{O}(1), \Delta \geq 3, k = \mathcal{O}\left(\sqrt{\frac{\log n}{\log \log n}}\right)$ $\mathfrak{D} = n^{1/6}, \Delta = \mathcal{O}(k^2), k = \mathcal{O}(1)$ $\mathfrak{D} = \mathsf{poly}(n), \Delta = \mathcal{O}(1), k = \mathcal{O}\left(\sqrt{\log n}\right)$

Expansion

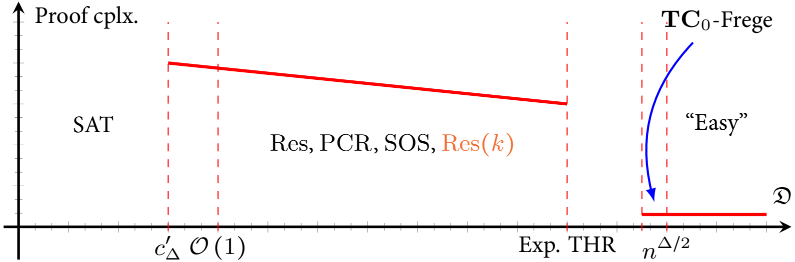


- (r, Δ, c)-expander;
- $\forall S \subseteq L, |S| \leq r \Rightarrow \mathrm{N}(S) \geq c|S|$.

Expansion



- (r, Δ, c)-expander;
- $\forall S \subseteq L, |S| \leq r \Rightarrow \mathrm{N}(S) \geq c|S|$.



Technical tools

Technical tools

- Induction on k.
- Restriction technique.
- “Independence” criteria.

Technical tools

- Induction on k.
- Restriction technique.
- "Independence" criteria.

Theorem

G_φ is an $(r, \Delta, 0.98\Delta)$-expander $\Rightarrow$ $\forall\delta > 0$ if:

$$n^\delta \left(\frac{n}{0.4r}\right)^{20k^2} = o(r/k)$$

then any $\mathrm{Res}(k)$ proof of φ has size at least 2^{n^δ}.

Open problems

Open problems

- Larger k?

Open problems

- Larger k?
- Weak pigeonhole principle in $\mathrm{Res}(2)$.

Open problems

- Larger k?
- Weak pigeonhole principle in $\mathrm{Res}(2)$.
- Other hard examples.