Satisfiable Tseitin formulas are hard for nondeterministic read-once branching programs

Ludmila Glinskih, Dmitry Itsykson

St.Petersburg Department of V.A.Steklov Institute of Mathematics

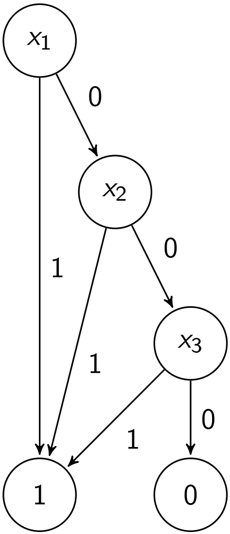
Estonian-Latvian Theory Days 2017
Tartu
November 24, 2017

Outline

- Branching programs
- Tseitin formulas
- Lower bounds

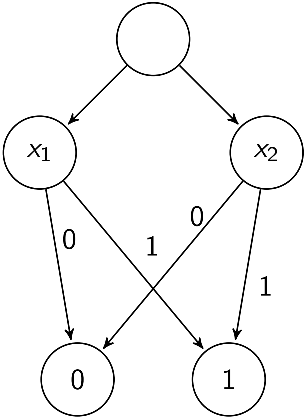
Branching program

- BP is the way to represent Boolean function
  - directed graph without cycles;
  - two sinks: labeled with 0 and 1, one root;
  - every internal vertex labeled with a variable and has two outgoing edges: labeled with 0 and 1;
  - the value of function equals label of the sink in the end of corresponding path.
- BP size is between circuit and formula sizes
  - $C(f) \leq 3BP(f) \leq O(L^{1+\epsilon}(f))$,
  - $C(f)$ circuit complexity of f, $L(f)$ size of minimal formula, BP size of minimal branching program.



Non-deterministic Branching program

- NBP additionally has non-deterministic nodes:
  - non-deterministic nodes are unlabeled
  - the value equals 1 iff there exists a path to 1-sink
- BP corresponds to L/poly
- NBP corresponds to NL/poly
- k-(N)BP if for every path every variable occurs no more than k times
- OBDD is 1-BP with fixed order for variables in every path



Known lower bounds for k-BPs

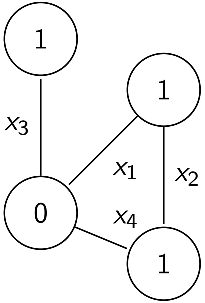
- (Borodin, Razborov, Smolensky, 1993)
  $CLIQUE\_ONLY_n(G) : \{0,1\}^{\frac{n\cdot(n-1)}{2}} \to 0, 1$
  - $CLIQUE\_ONLY_n(G) = 1$ iff graph G on n vertices is exactly $\frac{n}{2}$-clique
  - 1-NBP$(CLIQUE_ONLY_n) = 2^{\Omega(n)}$
  - 2-BP$(CLIQUE_ONLY_n) = poly(n)$
- (Borodin, Razborov, Smolensky, 1993) first strongly exponential lower bound
  - f artificially constructed function
  - 1-NBP$(f) = 2^{\Omega(n)}$
- (Thathachar, 1998) an explicit functions f_k for every k:
  - k-NBP$(f_k) = 2^{\Omega(n^{1/k})}$
  - (k+1)-BP$(f_k) = O(n)$

Known lower bounds for k-BPs (2)

- (Duris, Hromkovic, Jukna, Sauerhoff, Schnitger, 2004)
  - $\oplus$ parity of the number of triangles in a graph, $\overline{\triangle}$ is 1 iff graph has no triangles
  - 1-NBP($\oplus$) $= 2^{\Omega(n)}$
  - 1-NBP($\overline{\triangle}$) $= 2^{\Omega(n)}$
- (Juhna, 1995)
  - f_k characteristic function of error-correcting codes
  - k-NBP(f_k) $= 2^{\Omega(\sqrt{n})}$

Tseitin formulas

- Tseitin $TS_{G,c}$ formula defined for a graph
  - every edge is labeled with a variable
  - every vertex has a 0-1 label: $c : V \to 0, 1$
  - $TS_{G,c}(x) = 1 \iff$
    $$\bigwedge_{v \in V} \left(\sum_{e \text{ incident } v} x_e = c(v) \bmod 2 \right)$$
- A Tseitin formula is satisfiable iff for every connected component the sum of labels is even
- Unsatisfiable Tseitin formulas are classical hard examples for proof systems:
  - Resolution [Tseitin, 1968], [Urquhart, 1987]
  - Bounded depth Frege [Ben-Sasson, 2002], [Pitassi, Rossman, Servedio, Tan, 2016]
  - Polynomial Calculus over field with char != 2 [Alekhnovich, Razborov, 2001]
  - Tree-like Lovasz-Schrijver [Itsykson, Kojevnikov, 2006]



$$\begin{cases} x_1 + x_2 = 1 \\ x_1 + x_3 + x_4 = 0 \\ x_2 + x_4 = 1 \\ x_3 = 1 \end{cases}$$

Tseitin formulas (2)

- (Itsykson, Knop, Romashchenko, Sokolov, 2017) Exponential lower bound for OBDD(join) proof systems for unsatisfiable Tseitin formulas
  - key step: satisfiable Tseitin formulas are hard for OBDDs.
  - $OBDD(TS_{G,c}) = 2^{\Omega(n)}$
- OBDD is a partial case of 1-BP
  - (Bryant, 1991) (Sieling, Wegener, 1995) $f : OBDD(f) = 2^{\Omega(n)}$, 1-BP(f)=$poly(n)$

The goal: what is the size of 1-BP for satisfiable Tseitin formulas?

Main Theorem

Lower bound for 1-NBP:
1-NBP for satisfiable Tseitin formula for (n, d, α)-expander is $2^{\Omega(n)}$, where

- $\alpha < \frac{1}{3}$
- (n, d, α)-expander is a d-regular graph on n vertices with absolute value of the second largest eigenvalue $\leq d \cdot \alpha$

Corollary: The size 1-NBP for satisfiable Tseitin formula for complete graph is $2^{\Omega(n)}$
Upper bound: Tseitin formula for a graph with n vertices and m edges can be represented by OBDD of size $m \cdot 2^{O(n)}$

The comparison with other results

Satisfiable Tseitin formulas $TS_{G,w}$ on (n, d, α)-expander can be represented by:

- CNF($TS_{G,c}$)$=O(n)$
- 1-NBP($\rceil TS_{G,c}$)$=O(n)$
- 2-BP($TS_{G,c}$)$=O(n)$

- 1-NBP($TS_{G,c}$)$=2^{\Omega(n)}$

Best previously known gaps:

CNF(f)$=O(n^3)$	1-NBP(f)$=2^{\Omega(n)}$ (Duris et al, 2004)
1-NBP($\rceil f$)$=O(n)$	1-NBP(f)$=2^{\Omega(\sqrt{n})}$ (Juhna, 2009) explicit construction
1-NBP($\rceil f$)$=O(n)$	1-NBP(f)$=2^{\Omega(n)}$ (Duris et al, 2004) probabilistic construction
2-BP(f)$=O(n)$	1-NBP(f)$=2^{\Omega(\sqrt{n})}$ (Thathachar, 1998)

Plan of the proof

Theorem: 1-NBP for satisfiable Tseitin formula for (n, d, α)-expander is $2^{\Omega(n)}$, where $\alpha < \frac{1}{3}$
Note: proof for deterministic case
Plan of the proof:
Choose level l of the 1-BP

- show that there are at least $2^{C_1 n}$ non-zero paths to level l
- show that for every node v on the l-th lever only $2^{C_2 n}$ non-zero paths go from source to the node v

Get that at least $2^{(C_1 - C_2)n}$ different nodes are on the l-th level, so is in the 1-BP

Technical lemma

Lemma: Let $G(V, E)$ be a graph with k connected components. If the Tseitin formula $TS_{G,c}$ is satisfiable, then the number of its satisfying assignments equals $2^{|E|-|V|+k}$.
Proof:

- fix a spanning forest T of G: it has $|V| - k$ edges
- choose arbitrary assignment for all variables on edges of $G \setminus T$: $|E| - |V| + k$ edges
- the values of T can be determined unambiguously to satisfy formula

Lower bounds for number of all paths

Lemma: Let k be the maximum number of connected components that can be obtained after deleting of l edges from G, then there are at least $2^{l-(k-1)}$ non-zero paths of length l

- by the technical lemma we have at least $2^{|E|-|V|+1}$ satisfying assignments
- if we assign l values corresponding to non-zero path we get satisfiable Tseitin formula on G' with:
  - at most k connected components
  - at most $2^{|E|-|V|-l+k}$ satisfying assignments
- hence we need at least $\frac{2^{|E|-|V|+1}}{2^{|E|-|V|-l+k}} = 2^{l-(k-1)}$ different non-zero paths to reach level l of 1-BP.

Lemma: In (n, d, α)-expander parameter $k \leq \frac{2l}{d(1-\alpha)} + 1$, if $l \leq \frac{1}{4}n$

Upper bound for the number of paths to the fixed vertex

Lemma: For every node v there are at most 2^{l-t} different non-zero paths of length l to this node, where t is the number of edges in the spanning forest of graph that obtained from G by deleting all edges except l

- if two paths go to one node, then they have similar set S of tested variables
- assignment, that corresponds to a path, is a satisfying assignment of a Tseitin formula that is defined on a graph on edges from S
- this Tseitin formula has no more than 2^{l-t} satisfying assignments

Lemma: In (n, d, α)-expander parameter t for level l can be bounded with $\frac{l}{d(\alpha+\beta)}$, where $l \leq \beta n - 1$

Result

- we showed that there are at least $2^{l\cdot(1-\frac{2}{d(1-\alpha)})}$ paths to the level l
- no more than $2^{l\cdot(1-\frac{1}{d(\alpha+\beta)})}$ paths meet in one node
- from these we get that there are at least $2^{\frac{l}{d}\cdot(\frac{1}{\alpha+\beta}-\frac{2}{1-\alpha})}$ nodes on level l
- if $\beta < \frac{1-3\alpha}{2}$ we get $2^{\Omega(n)}$ nodes in 1-BP

Open question

Prove non-trivial lower bound for semantic 1-NBP

- only consistent paths should contain every variable no more than once
- that means that all paths that consist repeated variable should not be reachable by any assignment

No exponential lower bounds for semantic 1-NBP for a boolean function

All known results with exponential lower bounds are for functions with non-boolean domain [Cook, Edmonds, Medabalimi, Pitassi, 2016], [Jukna, 2009]