Satisfiability Algorithms for Small Depth Circuits with Symmetric Gates

Suguru TAMAKI

Kyoto University & Simons Institute

Connections Between Algorithm Design and Complexity Theory, October 1, 2015, Simons Institute, Berkeley, CA

This Talk

is about Algorithm Design & Complexity Theory because

this workshop is about ``Connections Between Algorithm Design and Complexity Theory''

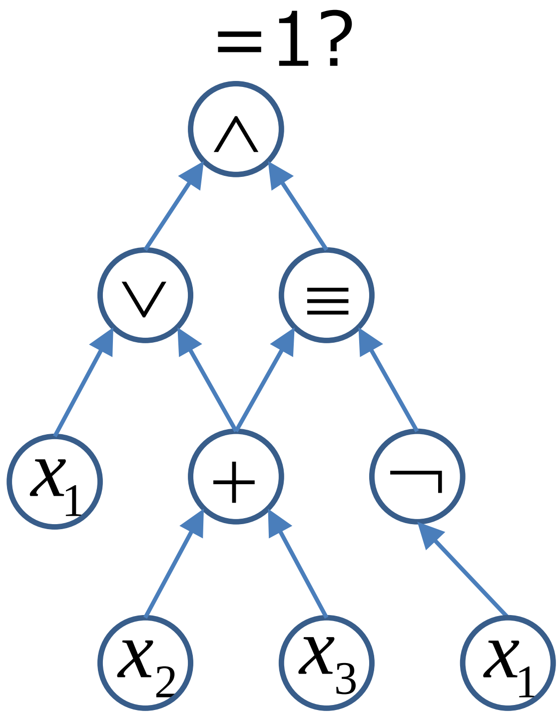
this program is about
``Fine-Grained Complexity and Algorithm Design''

Contribution (Algorithm Design)
Circuit SAT algorithms for ``interesting'' circuit classes

Implications (Complexity Theory)
Circuit size lower bounds (by known results)

Our Problem

Circuit Satisfiability (SAT)



Input:

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

Output:

$\exists x, C(x) = 1 \quad \Rightarrow \quad$ Yes

$\forall x, C(x) = 0 \quad \Rightarrow \quad$ No

- Canonical NPC problem
- Solved in time $\mathrm{poly}(|C|)2^n$ (n: #variables)
- $\mathcal{C}$-SAT: input only from circuit class $\mathcal{C}$

  e.g. $\mathcal{C} = (k\text{-})\mathrm{CNF}, \mathrm{AC}^0, \mathrm{AC}^0[p], \mathrm{ACC}^0, \mathrm{TC}^0, \mathrm{NC}^1$ (Formula),…

General Research Goals

1. Design non-trivial algorithms for a stronger circuit class

Non-trivial:

super-polynomially (exponentially) faster than 2^n

Stronger circuit class:

$(k\text{-})\mathrm{CNF} \subset \mathrm{AC}^0 \subset \mathrm{AC}^0[p] \subset \mathrm{ACC}^0 \subseteq \mathrm{TC}^0 \subseteq \mathrm{NC}^1 \subseteq \ldots \subseteq \mathrm{CKT}$

2. If non-trivial algorithms exist for $\mathcal{C}$-SAT, then

- improve the running time
- prove the difficulty of improvement

Why Study Circuit SAT?

1. Useful
- can encode many combinatorial problems efficiently
- (sometimes) inspires algorithms for other problems

  e.g. All-Pairs Shortest Paths (APSP) [Williams'14,...]

2. Connection to circuit lower bounds
- [black box] non-trivial $\mathcal{C}$-SAT algorithm ⇒ $\mathrm{NEXP} \not\subseteq \mathcal{C}$

  [Williams'10,11,...] (NEXP: nondeterministic exponential time)
- [white box] analysis of $\mathcal{C}$-SAT algorithm

  ⇒ average-case lower bounds

  [Santhanam'10,Seto-T'12,Chen-Kabanets-Kolokolova-Shaltiel-Zuckerman'14,...]

Circuit Classes

$(k\text{-})\mathrm{CNF} \subset \mathrm{AC}^0 \subset \mathrm{AC}^0[p] \subset \mathrm{ACC}^0 \subseteq \mathrm{TC}^0 \subseteq \mathrm{NC}^1 \subseteq \ldots \subseteq \mathrm{CKT}$

- $(k\text{-})\mathrm{CNF}$: conjunction of disjunctions of (at most k) literals
- AC^0: constant-depth, unbounded-fan-in, $\mathrm{AND}/\mathrm{OR}/\mathrm{NOT}$
- $\mathrm{AC}^0[p]$: $\mathrm{AC}^0 + \mathrm{mod}\ p$ gates (p: prime power)
- ACC^0: $\mathrm{AC}^0 + \mathrm{mod}\ m$ gates (m: integer ≥ 2)
- TC^0: constant-depth, unbounded-fan-in,
  linear threshold (THR) gates: $\mathrm{sgn}(\sum_{i=1}^{n} w_i x_i - \theta)$
- NC^1: fan-in 2, fan-out 1, $\mathrm{AND}/\mathrm{OR}/\mathrm{NOT}(/\mathrm{XOR})$
- CKT: fan-in 2, $\mathrm{AND}/\mathrm{OR}/\mathrm{NOT}(/\mathrm{XOR})$
- $\mathcal{C}_1 \circ \mathcal{C}_2$: composition of $\mathcal{C}_1 \circ \mathcal{C}_2$ e.g. $\mathrm{CNF} = \mathrm{AND} \circ \mathrm{OR}$

(Note: assume #gates $= \mathrm{poly}(n)$ unless otherwise specified)

Research Frontier

$\ldots \subset \mathrm{AC}^0[p] \subset \mathrm{ACC}^0 \subseteq \mathrm{ACC}^0 \circ \mathrm{THR} \subseteq \mathrm{TC}^0 \subseteq \mathrm{NC}^1 \subseteq \ldots \subseteq \mathrm{CKT}$

Non-trivial $\mathcal{C}$-SAT algorithm:

- $\mathrm{ACC}^0 \circ \mathrm{THR}$ with a super-poly #gates [Williams'14]
- $\mathrm{THR} \circ \mathrm{THR}$ with a linear #wires [Impagliazzo-Paturi-Schneider'13,...]

Lower Bounds for $\mathcal{C}$:

- majority, mod $q \notin \mathrm{AC}^0[p]$ [Razborov'87,Smolensky'87]
- $\mathrm{NEXP} \nsubseteq \mathrm{ACC}^0 \circ \mathrm{THR}$ [Williams'14]
- parity $\notin$ depth-d TC^0 with

  #wires $= n^{1+1/3^d}$ *or* #gates $= \ (n/2)^{1/(2d-1)}$

  [Impagliazzo-Paturi-Saks'93]

New $\mathcal{C}$-SAT algorithms

Attempts to handle TC^0 (probably $\mathcal{C} \not\subseteq \mathrm{ACC}^0 \circ \mathrm{THR}$)

- AC^0 with a limited #symmetric gates: $g(\sum_{i=1}^{n} x_i)$
- $\mathrm{THR} \circ \mathrm{THR}$ with a sub-quadratic #gates

Faster algorithms within $\mathrm{AC}^0[p]$ ($\mathcal{C} \subseteq \mathrm{ACC}^0 \circ \mathrm{THR}$)

- Systems of (degree-k) Polynomial Equations over $\mathrm{GF}(2)$
- $\mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR}$
- $\mathrm{AC}^0[p]$

AC^0 with a limited #symmetric gates

Motivation

Think about some interesting $\mathcal{C} \subset \mathrm{TC}^0 \setminus \mathrm{ACC}^0 \circ \mathrm{THR}$

i.e. $\mathcal{C}=$ ``AC^0 with $t(n)$ symmetric gates"

(Note: $\mathcal{C} \nsubseteq \mathrm{ACC}^0 \circ \mathrm{THR}$ is not known)

Definition:

■ $f: \{0,1\}^n \to \{0,1\}$ is symmetric (SYM)

if $\exists g: \mathbb{Z} \to \{0,1\}, f = g(\sum_{i=1}^n x_i)$

■ $f: \{0,1\}^n \to \{0,1\}$ is weighted symmetric

if $\exists g: \mathbb{Z} \to \{0,1\}, \exists w_i \in \mathbb{Z}, f = g(\sum_{i=1}^n w_i x_i)$

AND, OR, parity, mod m, majority are symmetric

THR ($\mathrm{sgn}(\sum_{i=1}^n w_i x_i - \theta)$) is weighted symmetric

Motivation

$\mathcal{C}$= ``AC^0 with $t(n)$ symmetric gates''

Interesting?

- contains Max SAT when depth-2, $t(n) = 1$ $(\mathrm{THR}\circ\mathrm{OR})$
  non-trivial algorithm for Max 3-SAT is open
  (cf. $2^{0.791n}$ time algorithm for Max 2-SAT [Williams'04])

- Lower bounds:
  generalized inner product (GIP) $\notin \mathrm{AC}^0$ with
  #symmetric gates $= n^{1-o(1)}$ *or* $\#\mathrm{THRs} = n^{1/2-o(1)}$
  [...,Lovett-Srinivasan'11]

New $\mathcal{C}$-SAT algorithms (1)

Theorem [Sakai-Seto-T-Teruyama]:

Let $\mathcal{C}$= ``AC^0 with $t(n)$ weighted symmetric gates"

where $t(n) = n^{o(1)}$, maximum weight $2^{n^{0.99}}$

There is a non-trivial deterministic algorithm for #$\mathcal{C}$-SAT

(Note: we assume evaluation of symmetric gate is easy)

Corollary:

Max SAT can be solved in deterministic time $2^{n-n^{1/O(k)}}$

when #clauses $= O(n^k)$

(Note: Max k-SAT $\Rightarrow$ #clauses $= O(n^k)$)

Implications

Corollary:

Let $\mathcal{C}$= ``AC^0 with $t(n)$ weighted symmetric gates"

where $t(n) = n^{o(1)}$, maximum weight $2^{n^{0.99}}$

Then $\mathrm{E}^{\mathrm{NP}} \not\subseteq \mathcal{C}$

Questions:

New? Interesting?

Core Technical Result

Lemma:

Let $\mathcal{C}$=(weighted SYM)$\circ$ AND

where $\#\mathrm{ANDs} = m$, maximum weight w

There is a deterministic algorithm for $\#\mathcal{C}$-SAT

that runs in time $\mathrm{poly}(n, m, \log w) 2^{n-\mu(n,m,w)}$

where $\mu(n, m, w) = (n/\log(mw))^{\Omega(\log n/\log m)}$

Based on ``Concentrated Shrinkage" & DP

(Note: Theorem follows from Lemma and transformation AC^0 with symmetric gates $\Rightarrow$ SYM$\circ$ AND using [Beigel-Reingold-Spielman'91,Beigel'92,Beame-Impagliazzo-Srinivasan'12])

$\mathrm{THR} \circ \mathrm{THR}$ with a sub-quadratic #gates

Motivation

Think about another interesting $\mathcal{C} \subset \mathrm{TC}^0 \setminus \mathrm{ACC}^0 \circ \mathrm{THR}$

i.e. $\mathcal{C} = \mathrm{THR} \circ \mathrm{THR}$

(Note: $\mathcal{C} \not\subseteq \mathrm{ACC}^0 \circ \mathrm{THR}$ is not known)

- $\mathcal{C}$-SAT can be solved in time $2^{n(1-\mu(c))}$
  where $\mu(c) = 1/c^{O(c)}$, #wires $= cn$
  [Impagliazzo-Paturi-Schneider'13, Chen-Santhanam'15]
  (non-trivial if $cn = o(n \log n / \mathrm{loglog}\, n)$)

- parity $\notin$ depth-d TC^0 with
  #wires $= n^{1+1/3^d}$ *or* #gates $= (n/2)^{1/(2d-1)}$
  [Impagliazzo-Paturi-Saks'93]

New $\mathcal{C}$-SAT algorithms (2)

Theorem [T]:

Let $\mathcal{C}$= $\mathrm{THR} \circ \mathrm{THR}$, where #gates $= m$

There is a randomized algorithm for $\mathcal{C}$-SAT

that runs in time $\mathrm{poly}(n, m)\ 2^{n-\mu(n,m)}$,

where $\mu(n, m) = \Omega\left(n/m^{1/2+o(1)}\right)^{c}$, $\exists c < 1/5$

(Note: #gates $\leq$ #wires)

Questions:

Derandomization?

(coNP algorithm is enough for $\mathrm{E}^{\mathrm{NP}} \not\subseteq \mathcal{C}$)

New lower bounds from analysis?

Proof Sketch

- based on the Polynomial Method in Circuit Complexity
- follow the framework for $\mathrm{ACC}^0 \circ \mathrm{THR}$ [Williams'14]
- use probabilistic polynomial for THR [Srinivasan'13]
- some transformation techniques due to
  [Maciel-Therien'98], [Beigel'92]
- use fast evaluation algorithm for $\mathrm{SYM} \circ \mathrm{SYM}$ [Williams'14]

Some details later

Faster algorithms within ACC^0

Motivation

$(k\text{-})\mathrm{CNF} \subset \mathrm{AC}^0 \subset \mathrm{AC}^0[p] \subset \mathrm{ACC}^0$

<u>$\mathcal{C}$: $\mathcal{C}$-SAT in time T, condition</u>

- k-CNF: $2^{n(1-\mu(k))}, \mu(k) = 1/O(k)$
  [Paturi-Pudlak-Zane'97,...]
- CNF: $2^{n(1-\mu(c))}, \mu(c) = 1/O(\log c)$, #clauses $= cn$
  [Schuler'05,Calabro-Impagliazzo-Paturi'06,...]
- AC^0: $2^{n(1-\mu(c,d))}$, $\mu(c,d) = 1/O(\log c + d \log d)^{\,d-1}$,
  depth-d, #gates $= cn$ [Impagliazzo-Matthews-Paturi'12]
- ACC^0: $2^{n-\mu(n,d)}, \mu(n,d) = n^{1/2^{O(d)}}$,
  depth-d, #gates $= 2^{n^{o(1)}}$ [Williams'11]

New $\mathcal{C}$-SAT algorithms (3)

Theorem [T-Williams]:

$\mathcal{C}$: $\mathcal{C}$-SAT in time T, condition

- Systems of degree-k Polynomial Equations over $\mathrm{GF}(2)$
  $(= \mathrm{AND} \circ \mathrm{XOR} \circ \mathrm{AND}_k \supset k\text{-CNF} = \mathrm{AND} \circ \mathrm{OR}_k)$:
  $2^{n(1-\mu(k))}, \mu(k) = 1/O(k)$
- $\mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR}$ $(\supset \mathrm{CNF} = \mathrm{AND} \circ \mathrm{OR})$:
  $2^{n(1-\mu(c))}, \mu(c) = 1/O(\log c)$, #ANDs $= cn$ at depth-4
- $\mathrm{AC}^0[p]$ $(\supset \mathrm{AC}^0)$:
  $2^{n(1-\mu(d,m))}$, $\mu(d,m) = 1/O(\log m)^{d-1}$,
  depth-d, #gates $= m$

Proof Sketch

- based on the Polynomial Method in Circuit Complexity
- use probabilistic polynomial for
  AND/OR [Razborov'87,Smolensky'87]
  $\mathrm{AC}^0[p]$ [Kopparty-Srinivasan'12]
- use fast evaluation algorithm for polynomial [Yates,…]
- first item is essentially due to [Lokshtanov-Paturi]
  (algorithm for k-CNF based on the polynomial method)
- second item is based on
  degree reduction for $\mathrm{AND}\circ\mathrm{XOR}\circ\mathrm{AND}\circ\mathrm{XOR}$
  extending Schuler's width reduction for CNF

Some details later

Algorithms via Polynomial Method

Polynomial Method

Example [Razborov'87,Smolensky'87]:

1. $\mathrm{AC}^0[p]$ can be well approximated
   by a low-degree $\mathrm{GF}(p)$ polynomial
2. majority, mod q cannot be well approximated
   by a low-degree $\mathrm{GF}(p)$ polynomial

1+2 ⇒ majority, mod $q \notin \mathrm{AC}^0[p]$

item 1 is useful in algorithm design
(``sparse" suffices instead of ``low-degree" in many cases)

Polynomial Method

Definition:

Let $f: \{0,1\}^n \to \{0,1\}$

A distribution P over polynomials is an ϵ-error probabilistic polynomial for f if $\forall x$, $\Pr_{p\sim P}[p(x) \neq f(x)] \leq \epsilon$

$\deg(P) \leq d$ if $\Pr_{p\sim P}[\deg(p) \leq d] = 1$

$\epsilon-$error probabilistic $\mathcal{C}$-circuit is defined analogously

Algorithm for $\mathcal{C}$-SAT

Input: $C\colon\{0,1\}^n \to \{0,1\}, C \in \mathcal{C}$

Step 1. Define $C'\colon\{0,1\}^{n-n'} \to \{0,1\}, C' \in \mathrm{OR}\circ\mathcal{C}$ as

$C'(y) \coloneqq \bigvee_{a\in\{0,1\}^{n'}} C(y,a)$ (Note: $|C'| \approx 2^{n'}|C|$)

Step 2. Construct (1/3)-error probabilistic polynomial p for C' in time $T(n,n',|C|)$

Step 3. Evaluate $p(y)$ for all $y \in \{0,1\}^{n-n'}$ in time $T'(n,n',|C|)$

Step 4. repeat 2-3 $O(n)$ times to reduce error probability

Output: truth table V of C' such that

$\forall y, \Pr[V(y) \neq C'(y)] \leq 2^{-2n}$

Running Time: $O(n(T(n,n',|C|) + T'(n,n',|C|)))$

Ingredients for $\mathrm{THR}\circ\mathrm{THR}$ (1/3)

Lemma [Razborov'87,Smolensky'87]:

There exists ϵ-error probabilistic polynomial for AND/OR of degree $\log(1/\epsilon)$ and it is efficiently samplable

Lemma [Srinivasan'13]:

There exists ϵ-error probabilistic polynomial for THR of degree $\sqrt{n}\ \mathrm{polylog}(n/\epsilon)$ and it is efficiently samplable

Ingredients for THR∘THR (2/3)

Lemma:

$\mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR} \circ \mathrm{AND} \subset \mathrm{XOR} \circ \mathrm{AND}$

Lemma [Maciel-Therien'98]:

$\mathrm{THR} \subset \mathrm{AC}^0[2] \circ \mathrm{SYM}$

Lemma [Beigel'92]:

$\mathrm{AND} \circ \mathrm{SYM} \subset$ weighted SYM

Ingredients for $\mathrm{THR} \circ \mathrm{THR}$ (3/3)

Lemma [Williams'14]:

Let C: $\{0,1\}^n \rightarrow \{0,1\}$, C ∈ SYM∘(weighted SYM) where

top gate has fan-in u

bottom gates have maximum weight w

such that $uw \leq 2^{0.1n}$

Then, truth table of C can be generated in time $\mathrm{poly}(n)2^n$

Algorithm for THR∘THR-SAT

Input: $C\colon \{0,1\}^n \to \{0,1\}, C \in \mathrm{THR}\circ\mathrm{THR}$

Step 1. Define $C'\colon \{0,1\}^{n'} \to \{0,1\}, C \in \mathrm{OR}\circ\mathcal{C}$ as

$C'(y) := \bigvee_{a\in\{0,1\}^{n'}} C(y,a)$ (Note: $|C'| \approx 2^{n'}|C|$)

Step 2. Construct (1/3)-error probabilistic circuit C'' for C'

$C'' \in$ (XOR∘AND)∘(XOR∘AND)∘ ⋯ ∘ (XOR ∘ AND) ∘SYM

Step 3. transform C'' into $C''' \in$ XOR∘(weighted SYM)

Step 4. Evaluate $C'''(y)$ for all $y \in \{0,1\}^{n-n'}$

Step 5. Repeat 2-4 $O(n)$ times to reduce error probability

Output: truth table of C'

Algorithm for $\mathrm{THR} \circ \mathrm{THR}$-SAT

Theorem [T]:

Let $\mathcal{C} = \mathrm{THR} \circ \mathrm{THR}$, where #gates $= m$

There is a randomized algorithm for $\mathcal{C}$-SAT

that runs in time $\mathrm{poly}(n, m)\ 2^{n-\mu(n,m)}$,

where $\mu(n,m) = \Omega\left(n/m^{1/2+o(1)}\right)^{c}, \exists c < 1/5$

Questions:

Derandomization?

(coNP algorithm is enough for $\mathrm{E}^{\mathrm{NP}} \not\subseteq \mathcal{C}$)

New lower bounds from analysis?

Other Techniques

Degree Reduction

AND∘XOR∘AND∘XOR with #ANDs at depth-3 $= m$

can be computed by a decision tree such that

1. internal node queries a linear equation $Ax = b$?
   where $\mathrm{rank}(A) \approx \log(m/n)$
2. leaf queries a system of degree-d polynomial equations
   where $d \approx \log(m/n)$
3. if leaf is reached after L times Yes,
   the system is defined over linear space of
   dimension $\approx n - \mathrm{L}\log(m/n)$
4. #leaves reached by L times Yes
   is at most $\binom{m+\mathrm{L}}{\mathrm{L}}$ (Note: $\mathrm{L} \leq n/\log(m/n)$)

Degree Reduction

degree reduction for $\mathrm{AND} \circ \mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR}$
is a generalization of Schuler's width reduction for CNF

Question:

degree reduction is useful in proving
average-case lower bounds for $\mathrm{AC}^0[p]$?

(as Schuler's width reduction is useful in proving
average-case lower bounds for AC^0)

Bottom Fan-In Reduction

(weighted SYM)∘AND with #ANDs $= m = O(n^k)$

can be computed by a decision tree such that

1. internal node queries a variable x_i
2. leaf queries a (weighted SYM)∘AND circuit
   whose bottom fan-in $= O(k)$
3. #leaves $\leq 2^{n-\sqrt{n}}$

Question:

useful in proving average-case lower bounds
for (weighted SYM)∘AND?

Conclusion

New $\mathcal{C}$-SAT algorithms

Attempts to handle TC^0 (probably $\mathcal{C} \not\subseteq \mathrm{ACC}^0 \circ \mathrm{THR}$)

- AC^0 with a limited #weighted symmetric gates
- $\mathrm{THR} \circ \mathrm{THR}$ with a sub-quadratic #gates

Faster algorithms within $\mathrm{AC}^0[p]$ ($\mathcal{C} \subseteq \mathrm{ACC}^0 \circ \mathrm{THR}$)

- Systems of (degree-k) Polynomial Equations over $\mathrm{GF}(2)$ ($\supset k$-CNF)
- $\mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR} \circ \mathrm{AND} \circ \mathrm{XOR}$ ($\supset \mathrm{CNF}$)
- $\mathrm{AC}^0[p]$ ($\supset \mathrm{AC}^0$)

Open Questions

- non-trivial algorithm for stronger $\mathcal{C}$:

  $\mathrm{ACC}^0 \circ \mathrm{THR} \subseteq \mathrm{TC}^0 \subseteq \mathrm{NC}^1 \subseteq \ldots \subseteq \mathrm{CKT}$

  e.g. AC^0 with more symmetric gates, $\mathrm{THR} \circ \mathrm{THR}$ with more gates, $\mathrm{THR} \circ \mathrm{THR} \circ \mathrm{THR}$,… or improve:

  NC^1 with n^3 (n^2) gates [Komargodski-Raz-Tal'13,…]

  CKT with $3n$ $(2.5n)$ gates [Chen-Kabanets'15]

- without polynomial method? (in polynomial space?)
- other lower bound techniques useful?

  e.g. communication complexity, proof complexity, mathematical programming,…

- lower bound for $\mathcal{C} \Leftrightarrow$ non-trivial $\mathcal{C}$-SAT algorithm?
- fine-grained reduction between $\mathcal{C}$-SAT and $\mathcal{C}$'-SAT?