Some results in Circuit Complexity

Johan Håstad

KTH Numerical Analysis
and Computer Science

CTW10, September 13, 2010

My interests

Circuit complexity, 1984-1990.

Cryptography, 1982-now, maybe less these days.

Complexity theory in general, 1982-now.

Approximation algorithms, 1993-now, main interest today.

One opinion

One should change topics every now and then, more often than I have done.

A sizeable investment, but usually pays off.

Life is long and it is fun to know different areas.

A related opinion

Learn in a broad area while young.

As one gets older, time gets scarcer and ones memory does not get better.

Maybe some help by better perspective but doubtful.

Today's topic

Circuit complexity.

Active area with lots of progress in late 1980'ies. Less now.

Today's topic

Circuit complexity.

Active area with lots of progress in late 1980'ies. Less now.

Maybe we solved all doable problems?

Today's topic

Circuit complexity.

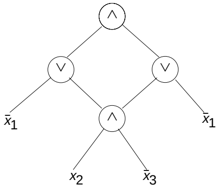
Active area with lots of progress in late 1980'ies. Less now.

Maybe we solved all doable problems?

Maybe we ran out of ambitious young researchers?

Basic definitions

A circuit is a directed acyclic graph from inputs to one output with n inputs.

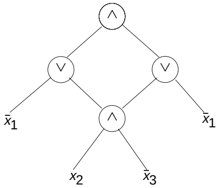


Size: Number of gates.

Depth: Longest path from input to output.

Basic definitions

A circuit is a directed acyclic graph from inputs to one output with n inputs.

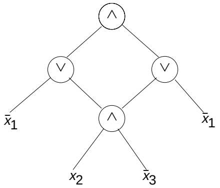


Size: Number of gates. 4

Depth: Longest path from input to output.

Basic definitions

A circuit is a directed acyclic graph from inputs to one output with n inputs.



Size: Number of gates. 4

Depth: Longest path from input to output. 3

Standard Gates

And-gates ($\wedge$), or-gates ($\vee$)

Usually negations (not counted in size). If not we have monotone circuits.

Fanin (number of inputs to a gate) can be bounded by two or unbounded.

Non-standard gates

Mod m gates (special case when m is a prime p).

Non-standard gates

Mod m gates (special case when m is a prime p).

Threshold gates $G(x) = \text{sign}(\sum_{i=1}^{t} w_i x_i - w)$.

Majority gates, all $w_i = 1$.

Non-standard gates

Mod m gates (special case when m is a prime p).

Threshold gates $G(x) = \text{sign}(\sum_{i=1}^{t} w_i x_i - w)$.

Majority gates, all $w_i = 1$.

Interesting in connection with very small depth circuits of great fanin.

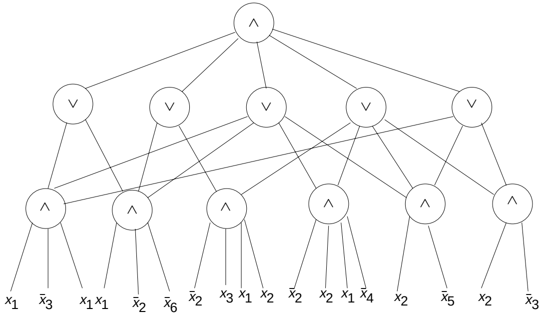
The first class

Unbounded fanin circuits with $\wedge$ and $\vee$-gates (and negations).

AC^0, alternating circuits of constant depth and polynomial size.

Naming due to $O((\log n)^0) = O(1)$.

The picture to have in mind



An old result

Theorem [S83, FSS84, Y85, H86] Computing parity of n inputs by a depth-d circuit requires size

$$2^{\Omega(n^{\frac{1}{d-1}})}.$$

Three proof approaches

Restrictions. Giving values to most variables, simplifying the circuit.

Approximations by polynomials. Output of circuit is close to a polynomial.

Top-down. Analysis starting with the output.

Restrictions

Idea by Sipser [S83]: Randomly give values to most of the variables.

Restrictions

Idea by Sipser [S83]: Randomly give values to most of the variables.

Formally: $\rho \in R_p$ for each variable x_i independently:

Keep it is a variable with probability p, otherwise fix to 0 and 1 with equal probability $(1-p)/2$.

What restrictions do

After a restriction parity turns into parity or negation of parity on the remaining variables.

A restriction simplifies the circuit by substituting the values for the variables that are fixed.

Simplifications of circuits

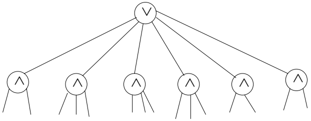
Restrictions greatly affect the bottom two layers of circuits with only $\wedge$-gates and $\vee$-gates.

- For an $\wedge$-gate of inputs of unbounded fanin. One input set to 0 makes it 0.
- If all inputs are set to one, it determines the value of the $\vee$-gate in the level above.

The switching lemma

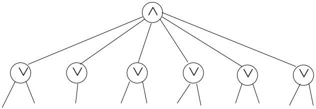
Lemma [Y85, H86] Any depth two circuit which is a $\vee$ of $\wedge$'s each of which is size $\leq t$ can, when hit with a random $\rho \in R_p$, with probability at least $1 - (5pt)^s$, be converted to a depth two circuit which is a $\wedge$ of $\vee$'s each of which is of size $\leq s$.

A picture



$\leq$ t inputs to each $\wedge$-gate

turns into



$\leq$ s inputs to each $\vee$-gate

with probability $1-(5pt)^s$ by $\rho \in R_p$.

Proof of switching lemma, idea

If the circuit is read-once it is a calculation as each $\wedge$ is independent.

Proof of switching lemma, idea

If the circuit is read-once it is a calculation as each $\wedge$ is independent.

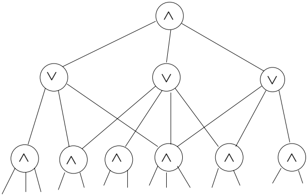
Correlation goes the right way. Do one $\wedge$ at the time and prove a more general, conditioned, statement by induction.

Switching gives parity lower bound

Induction with $p = n^{-1/(d-1)}$ and $s = t = \frac{1}{10} n^{1/(d-1)}$.

Each restriction wipes out one level.

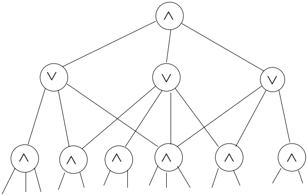
In pictures, I



bottom fanin $\leq t$

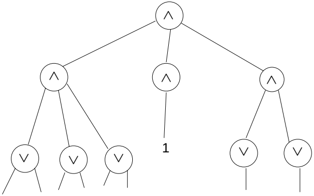
Apply $\rho \in R_p$ and use lemma on each depth 2 subcircuit.

In pictures, I



bottom fanin $\leq t$

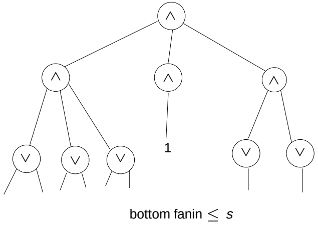
Apply $\rho \in R_p$ and use lemma on each depth 2 subcircuit.



bottom fanin $< s$

In pictures, II

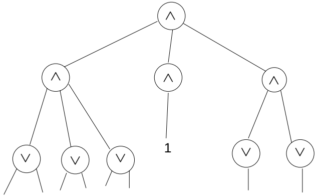
After switching we have



and we make shortcuts

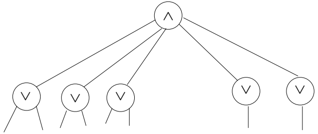
In pictures, II

After switching we have



bottom fanin $\leq s$

and we make shortcuts



bottom fanin $\leq s$

Punch line, switching

It is easy to see that a depth two circuit computing parity of m variables requires bottom fanin m and size 2^{m-1}.

We need to optimize p in R_p-restrictions to balance

- Making sure we can simplify circuit.
- Keeping many variables.

Polynomial approach

Large $\wedge$'s (and $\vee$'s) are not very useful. Let $\oplus$ be parity.

Compare

$$\vee_{j=1}^{m} F_j(x) \tag{1}$$

and

$$\oplus_{j=1}^{m} F_j(x) \tag{2}$$

Polynomial approach

Large $\wedge$'s (and $\vee$'s) are not very useful. Let $\oplus$ be parity.

Compare

$$\vee_{j=1}^{m} F_j(x) \tag{1}$$

and

$$\oplus_{j=1}^{m} F_j(x) \tag{2}$$

(1) is 0 then so is (2) and if (1) is 1 we have, heuristically speaking probability $1/2$ of getting a 1 also for (2).

Great idea

If instead of

$$\oplus_{j=1}^{n} F_j(x)$$

we take a random subset $S \subseteq [m]$ then

$$\oplus_{j \in S} F_j(x) = \vee_{j=1}^{n} F_j(x)$$

with probability $1/2$ (over S) for any fixed x.

The key lemma

Lemma: Let S_i, $1 \leq i \leq t$ be t independent subsets of $[m]$ then for any x

$$\vee_{i=1}^{t} \left(\oplus_{j \in S_i} F_j(x)\right) = \vee_{j=1}^{n} F_j(x).$$

with probability $1 - 2^{-t}$.

The key lemma

Lemma: Let S_i, $1 \leq i \leq t$ be t independent subsets of $[m]$ then for any x

$$\vee_{i=1}^{t} \left(\oplus_{j \in S_i} F_j(x)\right) = \vee_{j=1}^{n} F_j(x).$$

with probability $1 - 2^{-t}$.

The degree increases by only a factor t.

Consequence

Theorem (Razborov [R87]) If f is computed by a depth d circuit of size M then there exists a polynomial p mod 2 of degree S^d such that $f(x) = p(x)$ for all but a fraction $M2^{-S}$ of the inputs.

Consequence

Theorem (Razborov [R87]) If f is computed by a depth d circuit of size M then there exists a polynomial p mod 2 of degree S^d such that $f(x) = p(x)$ for all but a fraction $M2^{-S}$ of the inputs.

Remains true even if the circuit contains parity gates.

True, up to constants, if “mod 2” is replaced by “mod q” for any constant size prime q.

Punch line

Need to prove that some simple function is not approximated by a low degree polynomial.

Theorem by Razborov

Theorem (Razborov [R87]) Majority requires size

$$2^{\Omega(n^{\frac{1}{d+1}})}$$

to be computed by depth-d circuits containing $\wedge$, $\vee$ and $\oplus$-gates.

Theorem by Smolensky

Theorem (Smolensky [S87]) Mod m requires size

$$2^{\Omega(n^{\frac{1}{2d}})}$$

to be computed by depth-d circuits containing $\wedge$, $\vee$ and "mod p"-gates, as long as $m \neq p^r$.

Top-down approaches

The Karchmer-Wigderson communication game [KW90]. We are interested in computing f.

A gets an input x such that $f(x) = 0$ and B gets and input y such that $f(y) = 1$. By communicating they should find an i such that $x_i \neq y_i$.

A game for parity

Divide the input into s subsets. A computes the parity of each subset and sends to B.

B finds a subset where the parity of the two inputs differ.

A game for parity

Divide the input into s subsets. A computes the parity of each subset and sends to B.

B finds a subset where the parity of the two inputs differ.

Recurse.

Gives $n^{1/d}$ bits in each of d rounds.

The key theorem

Theorem: If f is computable by a depth-d circuit of size 2^s then the KW-game can be solved by a d move game where each player sends s bits in each round.

The proof

Induction. From output to an input find gates G_i in the circuit such that $G_i(x) = 0$ and $G_i(y) = 1$.

At $\vee$-gates B points to an input that is one and at $\wedge$-gates A points to an input that is 0.

Communication complexity

Even intuitively obvious facts are hard to prove and sometimes false.

Only proof with this method [HJP95] gives a $2^{\Omega(n^{1/2})}$ lower bound for depth three circuits.

Communication complexity

Even intuitively obvious facts are hard to prove and sometimes false.

Only proof with this method [HJP95] gives a $2^{\Omega(n^{1/2})}$ lower bound for depth three circuits.

Did anybody look at this for the last 15 years?

Communication complexity

Even intuitively obvious facts are hard to prove and sometimes false.

Only proof with this method [HJP95] gives a $2^{\Omega(n^{1/2})}$ lower bound for depth three circuits.

Did anybody look at this for the last 15 years?

Understanding in communication complexity has advanced, can it be useful?

Open problem

Get a better lower bound than $2^{n^{1/(d-1)}}$ for any function for depth d circuits.

Wide open even for $d = 3$.

Restrictions and approximation by polynomial will not do it.

Need more sophisticated properties of the function.

Some hope?

Rossman [R08] proved $\Omega(n^{k/4-o(1)})$ lower bound for constant depth circuits computing clique of size k for any constant depth circuit.

Exponent independent of depth and does use a more sophisticated property of the function.

However, far from the exponential bounds I want.

Other gates

The polynomial approach works with mod p gates for primes p.

One level of majority can be eliminated using correlation arguments.

Open problems

Lower bounds for depth 2-3 circuits with more general gates.

Open problems

Lower bounds for depth 2-3 circuits with more general gates.

Cannot rule out polynomial size circuits of depth 2 with mod 6 gates or threshold gates for any explicit function.

One Great Optimist

I had a bet with Andrew Yao around 1990 that someone would prove some explicit function not to be computable by polynomial size constant-depth circuits with threshold gates within two years.

One Great Optimist

I had a bet with Andrew Yao around 1990 that someone would prove some explicit function not to be computable by polynomial size constant-depth circuits with threshold gates within two years.

I won the bet but now I would have been happier to lose.

Monotone variants

Almost everything is simpler in the monotone variant but one problem I like.

For ordinary circuits, allowing weights in threshold circuits changes depth by at most additive one [GHR92].

What happens in the monotone case?

Changing gears

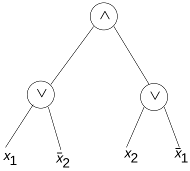
Formula size. Circuits where each gate has fanout 1.

A circuit that is a tree.

Depth the same as circuits, size much bigger.

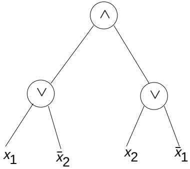
Parity formulas

Easy with two variables.



Parity formulas

Easy with two variables.



Recursive construction gives size n^2 when $n = 2^t$.

Classical counting

A random function on m inputs requires size $\Omega(2^m / \log m)$.

Seems hard to get good lower bounds for explicit functions.

Classical lower bound

Khrapchenko proved a general lower bound

$$\frac{|C|^2}{|A||B|}.$$

A is subset of $f(x) = 1$, B of $f(x) = 0$ and C is the set of $a \in A$ and $b \in B$ such that a and b only differ in one coordinate.

Bounds for parity

For parity we can have $|A| = |B| = 2^{n-1}$ and $|C| = n2^{n-1}$ giving n^2 lower bound.

Bounds for parity

For parity we can have $|A| = |B| = 2^{n-1}$ and $|C| = n2^{n-1}$ giving n^2 lower bound.

We know exactly the formula size of parity when n is a power of 2.

Beating n^2

Only known method invented by Subbotovskaya [S61] who designed a suitable function $S(x, y)$.

n bits x_i specifies a function f_x on $m = \log n$ bits.

n bits y_i to define $\log n$-bit input z to f_x.

$$z_j = \oplus_{i \in S_j} y_i.$$

Output: $f_x(z)$.

Proof idea

Fix the x to a function that requires formulas of size $2^m / \log m = \Theta(n/ \log \log n)$.

Use a restriction $\rho \in R_p$ on y simplifying the formula but keeping each z_j undetermined making the remaining as hard as f_x.

Proof idea

Fix the x to a function that requires formulas of size $2^m/\log m = \Theta(n/\log\log n)$.

Use a restriction $\rho \in R_p$ on y simplifying the formula but keeping each z_j undetermined making the remaining as hard as f_x.

Fixing x makes formula smaller but unclear how much.

ρ does shrink the formula.

First shrinking

Subbotovskaya proved that $\rho \in R_p$ shrinks a formula by a factor $p^{3/2}$.

This gives a lower bound of $n^{5/2-o(1)}$ for the formula size of $S(x,y)$.

Better shrinking

Subbotovskaya used local analysis.

More global analysis gives shrinking $p^{2-o(1)}$ [H98].

Up to $o(1)$ this is sharp (as seen from parity).

This gives lower bound $n^{3-o(1)}$ for the Subbotovskaya function.

Open problems

Get beyond n^3 for formula size.

Open problems

Get beyond n^3 for formula size.

Find another method that goes beyond n^2, getting bounds for nicer function.

A related question

Find an explicit function that cannot be computed by depth $O(\log n)$ and size $O(n)$ fanin-2 circuits.

The big question

Are circuits too complicated objects to understand?

The big question

Are circuits too complicated objects to understand?

Well if they are, we should prove this.

Natural proofs

Razborov and Rudich [RR97] considered natural proofs.

1. Works to give lower bounds for most function.
2. Needs a condition for hardness that is computationally easy to verify given truth-table of a function.

Natural proofs

Razborov and Rudich [RR97] considered natural proofs.

1. Works to give lower bounds for most function.
2. Needs a condition for hardness that is computationally easy to verify given truth-table of a function.

Cannot be used to prove lower bounds for any model which admits good pseudo-random generators.

Natural proofs

Razborov and Rudich [RR97] considered natural proofs.

1. Works to give lower bounds for most function.
2. Needs a condition for hardness that is computationally easy to verify given truth-table of a function.

Cannot be used to prove lower bounds for any model which admits good pseudo-random generators.

Does not rule out a proof but tells us where to look.

Final word

Circuit complexity has been mostly dormant for many years.

High risk of getting stuck.

But it is our duty to take another crack.

Final word

Circuit complexity has been mostly dormant for many years.

High risk of getting stuck.

But it is our duty to take another crack.

Need for young, optimistic researchers with new ideas.

My personal feelings

In the 1980'ies we thought we soon would know a lot more.

Now I would be extremely happy to see any major progress.

Thank you!