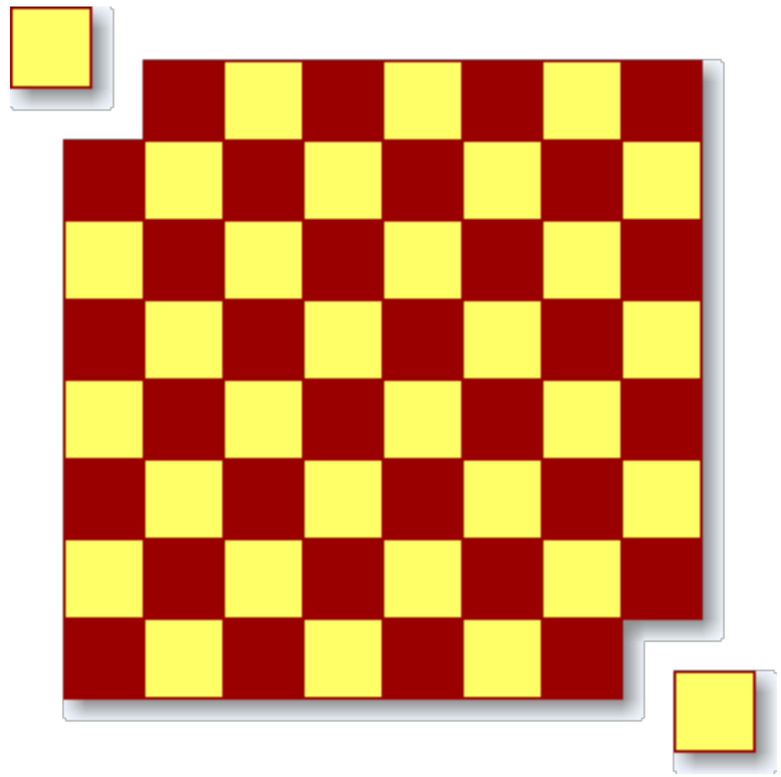
Russell Impagliazzo, Sasank Mouli, Toniann Pitassi:

Lower bounds for Polynomial Calculus with extension variables over finite fields.

You land on a hostile world filled with dangerous androids

The androids' only weakness is logic puzzles— they will be helpless while solving them



Can you cover a chessboard with two diagonal squares removed
with non-overlapping dominoes?

Problem proposed by McCarthy as a challenge to automated reasoning

How long will this puzzle stun androids?

ChatGPT: not long

- However:
- ``How about a seven by seven chessboard?''
- ChatGPT: A 7x7 chessboard is a two-dimensional board with 49 squares. Therefore, there are 49 unit squares or cubes, each with a volume of 1 cubic unit, in a three-dimensional 7x7x1 cuboid.
- If we remove two opposite corner cubes from this 7x7x1 cuboid, we are left with 47 cubes. Each 2x1x1 prism will cover one white and one black cube, so we need 23 prisms to cover all the remaining cubes. It's easy to see that we can arrange these 23 prisms in a way that covers all the remaining cubes without overlap or leaving any cube uncovered.
- So, it is possible to cover the remaining cubes of a 7x7x1 cuboid with 2x1x1 prisms, whereas it is impossible to do so for an 8x8x8 cube or cuboid as discussed earlier.

Circuit complexity vs. proof complexity

- Intuition: Proofs in formal systems are limited to reasoning that can be expressed with allowable lines in the systems
- Proving limitations on the types of circuits allowed as lines should help prove lower bounds in proof complexity, and hence lower bounds on SAT algorithms using reasoning expressible within the system.

To put this into practice, we need some circuit lower bounds

- CNF lower bounds (folklore) -→ Resolution lower bounds (Haken, ……)
- AC_0- lower bounds (Ajtai, FurstSaxeSipser…Hastad) → $AC_0 - Frege$ lower bounds (BCIKPW)
- $AC_0[p]$-lower bounds (Razborov, Smolensky) → $AC_0[p]$-Frege lower bounds (oops, we're still trying).

Algebraic proof systems

- Since the RS lower bound method involved approximating circuits with polynomials, it seemed natural to try to introduce proof systems using polynomials as lines as a stepping stone towards $AC_0[p]$-Frege lower bounds.
- Nullstellensatz (BeameIKraijicekPitassiPudlak) and Polynomial Calculus (CleggEdmondsI) were partially motivated as attempts to do this.

Polynomial calculus (PCR)

- Represent constraints as polynomials over a field F, p$(x_1, \overline{x}_1 .. x_n, \overline{x}_n)(= 0)$ in variables representing literals.
- Add $x + \overline{x} = 1, x\overline{x} = 0$
- Rules: linear combinations of previous lines, multiply previous line by variable
- Measure: maximum degree, size= total non-zero monomials.

Good news

- Strong lower bounds
- Degree lower bounds follow from ``pseudo-ideals''
- Strong enough degree lower bounds → exponential size lower bounds.
- Lower bound shows limits of Groebner basis algorithm and other ``algebraic reasoning'' mod p .

Bad news

- Doesn't seem to help with $AC_0[p]$-Frege lower bounds, which require some form of ``approximate'' or ``randomized'' polynomials, or extension variables.
- Size lower bounds very brittle, even small changes of variables can make exponential proofs polynomial size

Tseitin graph tautologies

- ``No edge induced subgraph of G has exactly one odd degree vertex''
- Hard when G is a sparse expander, linear degree, when F= mod p, p odd prime.
- Intuition: need to look at parities of edges in cuts to prove it.

Change of variables

- Introduce $y_e = 1 - 2x_e$ for each edge e
- Parities of edges in a cut are products of

$$y_e$$

Refutation is still linear degree, but now polynomial size.

Previous size-degree connections strongly require variables possibly be zero, because then high degree monomials can be easily removed.

Challenge

- Come up with size lower bounds for non-Boolean variables mod p
- Come up with size lower bounds that allow changes of variables or introduction of extension variables.
- Each round of extension variables can code another depth of circuit, and even small constant depth lower bounds have implications for very strong proof systems (GrigorievHirsh05, RazTzameret08, IMP20)

Sokolov’s breakthrough

- Sokolov (20) answered the first challenge by coming up with a new technique for proving size lower bound for 1, -1 valued variables.
- Intuitively, it proved an exponential size lower bound whenever a degree lower bound held for any restriction of the original formula.
- However, the lower bound seemed particular to 1, -1 variables, and didn’t allow for example , both the original Boolean variable and the 1,-1 version.

Extensions over characteristic zero

- Forbes, Shpilka , Tzameret, Wigderson , Andrews and Forbes, and Alekseev proved lower bounds over IPS, a stronger algebraic proof system, over the rationals. However, the formulas they prove lower bounds for aren’t derived from CNFs or other translations of Boolean tautologies, and do not have modular equivalents.

K-local extensions

- Here , we consider proofs augmented with one round of extension variables.
- Given a tautology in variables $x_1 .. x_n$, the proof can define $z_1, .. z_m$ with each z_i defined as a function of at most k of the inputs x . These definitions can be used as axioms in deriving a PCR contradiction .

Our results

- Theorem 1 (high-end). There is a family of CNF tautologies ψ with poly(N) clauses of width O(logN) so that for any prime p, any PC refutation of ψ with any O(NlogN) O(1)-local extensions over Fp requires size $\exp(\frac{N}{polylog(N)})$
- Theorem 2 (low-end). For the same family of tautologies above, for any prime p, there are 0 < α, β, γ < 1, with γ < 1 − α − β so that any PC refutation of Φ together with any $N^{1+\alpha}\ \beta logn$ -local extensions over Fp requires size $\exp(N^{\gamma})$

Main ideas

- We follow Sokolov's lead strongly, but figuring out how to generalize his approach and what order to apply things was complicated.
- In proofs involving 1,-1 variables Z, we have
- $Z^2 = 1$ and can mod out by this to make polynomials multi-linear.
- For general extension variables that cannot take on value 0 (non-singular), we can mod out by some $Z^k = c$ for some smallest k. k=p-1, c=1 always works. This makes the degree of each variable constant in all polynomials.

Quadratic degree → Factored degree

- Sokolov looked at the following complexity measure (roughly) : the maximum degree of the square of a line of a proof.
- We use an idea that generalizes this.
- The factored degree of a polynomial P is the min degree of Q so that we can write P= M Q

where M is a monomial. Note that if M is in 1,-1 variables, P=MQ → $P^2 = Q^2$, so small factored degree generalizes small quadratic degree.

Small factored degree refutations imply small degree refutations

- Claim: If there is a refutation where each line has factored degree d, then there is a refutation of degree O(d).

Proof: Write monomials as vectors of their exponents; hamming distance of these vectors is the number of variables with different powers.

- For each line P=MQ , we will derive Q. At the end, P=1 implies Q=1. If P=MQ then the Hamming distance of each term in P to M is small. If we add P=MQ and P'=M'Q' and get P''=M''Q'', we must have M, M', M'' close in hamming distance, M'= Mt', M''t''=M, for t' and t'' monomials of degree O(d). P+P'= MQ +M'Q'= M(Q+t'Q')= M'' (t'' Q + t''t'Q') =M'' Q'', so we can derive Q'' as t''Q+t''t'Q' in degree O(d).

Pairs of violating monomials

- We say that M_1 and M_2 are a factored degree d violating pair if they have hamming distance > d and appear in the same line of a proof.
- We say the violation number for a proof is the number of such pairs (no matter how many lines they appear in, we count the pair once).
- Each such pair has d different degree variables
- We look for a Z that is a different degree variable for many violating pairs, and degrees I and J that are the exponents for Z of many of these.

Split

- If Z is a 1,-1 variable that doesn't appear in any unsatisfied axioms, except $Z^2 = 1$, we can write each line as QZ+R , where Q and R do not contain Z. Sokolov: There is a proof containing each Q and R as separate lines.
- Axioms: Just R's, since Q=0. Sums: sum Q's, sum R's. Multiply by non-Z variable: multiply Q, and R separately. Multiply by Z: switch Q and R.
- This ``split'' removes all violating pairs where Z has a different exponent in the pair.

Problem

- If Z can take on say values 1, 2,

We have $Z^2 - 3Z + 2 = 0$ as the axiom.

If we multiply QZ+R by Z and mod out, we get Q(3Z-2)+ RZ = (3Q+R) Z – 2Q . Except for no longer disagreeing on Z, splitting doesn’t reduce the number of violating pairs.

New form of split

- Say Z= (a-b)x+b, where x is Boolean, and there are no other axioms about x or Z.
- We write every polynomial as :
- P=$P_{\{k-1\}}Z^{k-1}$+$P_{k-2}Z^{k-2} + \cdots P_0$
- Since Z can only be a or b, we can also

write this as $P = \ Q_i Z^i + Q_j Z^j$ if Z = a or Z=b. Q_i, Q_j will be linear in the P_l. P_i will only appear in the expression for Q_i and P_j in Q_j

Splitting

- First, remove x (or set x to 0 if it appears frequently).
- Then derive each Q_i, Q_j line by line. Axioms A

Are of the form cA, dA, so don't change proof.

Adding previous lines or multiplying by a non-Z variable is done component-wise.

Multiplying by Z is a linear function, so we have to take two linear combinations.

Splitting removes violating pairs

- If we have terms MZ^i and M'Z^j as a violating pair, $M\ and\ M'$ might appear in the proof,

but never in the same line (unless there were multiple violating pairs involving M and M').

So we reduce the number of violating pairs by the number of such pairs where Z has exponent I and J respectively. For an average Z, this is at least a $\frac{d}{Mp^2}$ fraction of all violating pairs.

Robustly high degree tautologies

- What kind of formula do we get a lower bound for? Once we find the Z we want to split on, we need to simplify its definition to depend on one Boolean variable by setting other variables in the definition. We need to do this o(N/polylogN) times, and still have degree close to $\Omega(\frac{N}{polylogN})$.

Tautology

- $\Phi(x_1 .. x_n)$: CNF where any Cn clauses of m=Dn are unsatisfiable, each x_i appears in O(1) clauses $\Psi_{\mathrm{j}}\left(x_{\{i_1\}}, .. x_{\{i_l\}}\right), l = O(1)$

- Variables $y_1 \ldots y_{\{cn\}}$ taking on values in 1…m, expressed as log m length binary vectors.
- Axioms: All $y_i \neq y_j, \Psi_{\{y_i\}}$ (i.e., $y_i = j \rightarrow \Psi_j$)
- The $y_i's\ pick\ Cn\ true\ clauses, but\ there\ aren't$

enough true clauses for them to pick.

Open problems

- Go beyond $M = N^2$ barrier in all our size degree connections.
- Prove lower bound for much larger locality
- Prove lower bound for multiple rounds of extension variables.

Thanks!