LECTURE 6

Summary of the course so far

We have looked at sequential proof systems for refuting unsatisfiable CNF formulas

Proof is a sequence of lines presented on a blackboard

Derivation steps:

(1) Download of axiom (clause in F)

(2) Inference applying rules of proof system to lines currently on the board

(3) Erasure

What determines the properties of a concrete proof system?

- format of the lines
- type of inference rules

So far we have looked at resolution

- Lines are disjunctive clauses
- Only one rule resolution $\dfrac{x \vee y \quad \bar{y} \vee z}{x \vee z}$

Mainly interested in

- length — # derivation steps
- (clause) space — max # clauses needed on board simultaneously

But it turned out that width measure was very helpful in order to understand both length and space

In what follows, focus on k-CNFs, $k = O(1)$

LENGTH VS WIDTH

Strong upper bounds on length ⇒ (Fairly) strong upper bounds on width (and these bounds are right)

Strong upper bounds on width ⇒ Strong upper bounds on length (counting)

Trade-offs between length and width? (Q1)

SPACE VS WIDTH

Strong UB space ⇒ strong UB width

Strong UB width ⇒ ? (Q2)

Trade-offs? Yes, strong trade-offs

[started in class and finished in pebbling notes]

SPACE VS LENGTH

Strong UB space ⇒ Strong UB length

[using space ≥ width + counting]

Strong UB length ⇒ ? (Q3)

Trade-offs? (Q4)

Q1 is open (not any longer)

Q2–Q4 are not, but to keep suspense will have to wait til after Christmas break.

Now shift focus — for the rest of this term we will study polynomial calculus (PC) *

- Lines are polynomials or polynomial equations
- Inference rules addition and multiplication

calculations are done in some field $\mathbb{F}$
For applications $\mathbb{F}_2$ (a.k.a GF(2)) most relevant
In this course, will tend to think of $\mathbb{F}$ as finite.
In general, infinite fields (e.g. $\mathbb{Q}$, $\mathbb{R}$) are also fine.
Characteristic of field is important. Some results hold only for certain characteristics. (in literature)
E.g. for lower bounds often $\mathbb{F}_2$ seems hardest.
In this course, try to focus on results that hold regardless of characteristic.

Translate clauses to polynomial equations $p(x,y,\ldots)=0$
Clause C satisfied by α $\Leftrightarrow$ $p(\alpha) = 0$

So natural to think of $0 \equiv$ true
$1 \equiv$ false

$$x \vee \bar{y} \vee z \quad \rightsquigarrow \quad x(1-y)z = xz - xyz = 0$$

NB! Polynomials are always presented multiplied out as sums of monomials.

*) Introduced in [Clegg, Edmonds, and Impagliazzo '96]

Derivation rules

$$\frac{p = 0 \quad q = 0}{\alpha p + \beta q = 0} \quad \alpha, \beta \in \mathbb{F} \qquad \frac{p = 0}{xp = 0}$$

Because of this w.l.o.g all monomials multilinear

"Field axioms" / "Boolean axioms"

$$\frac{}{x^2 - x = 0}$$

forces 0/1 - solutions

Goal: Derive $1 = 0$ from clauses of F.

Easy direction: Soundness

If we can derive $1 = 0$ from F, then F is unsatisfiable.

Hard(er) direction: Completeness
If F is an unsatisfiable CNF formula, then there is a PC-derivation of $1 = 0$ from F

Can prove algebraically or via proof cplx simulations — for now let's just accept soundness and completeness

Proof cplx measures

Size — measure # monomials (off by linear factor but much cleaner)
Length — now different since lines can have exp size

Monomial space — analogue of clause space
Total space — total # symbols; again linearly related to (monomial) space

Main interest Size and (monomial) space

SAT solvers?

There are methods based on so-called Gröbner bases that search for PC-proofs, but state of the art is that they are (much) slower than CDCL solvers or DPLL solvers (based on resolution)

Another way of thinking:

Consider clauses in F as polynomials (not equations)

Look at <u>ideal</u> in $\mathbb{F}[x,y,z,\ldots]$ [polynomial ring] generated by these polynomials

<u>Ideal</u>: subring closed under multiplication by any element of whole ring

If the constant polynomial 1 is in — i.e. ideal is everything — the ideal generated by F, then F is unsat (and the other way round)

We will tend to use this perspective — just drop "$=0$" from all derivation rules

---

A technical issue:

Consider a clause $\overline{x}_1 \vee \overline{x}_2 \vee \cdots \vee \overline{x}_w$ (suppose axiom of F)

Encoded as

$$(1-x_1)(1-x_2)\cdots(1-x_w) = \sum_{S \subseteq [w]} (-1)^{|S|} \prod_{i \in S} x_i$$

Exponential # monomials measured in w!

So if F has clauses of linear width with "wrong sign", just downloading such axiom clause gives exponential LB on size & space.

Not very exciting...

What to do?

Two fixes:

(1) Focus on k-CNF formulas for $k = O(1)$
Essentially w.l.o.g.

This will always be our goal, but for various reasons good to be able to deal with formulas of unbounded width too (k-CNFs might be too hard, for instance)

(2) Change encoding by introducing ~~independent~~ new extra variables $\bar{x}, \bar{y}, \bar{z}$ corresponding to negated literals.

What is "the right fix" in practice? Should we even fix this in practice? What do GB algorithms do?

Not entirely clear. Some implementations suffer from this monomial blow-up. Others have more space-efficient encodings.

But when proving lower bounds, best to work in strong model where lower bounds will hold regardless of these issues.

Hence the following definition

POLYNOMIAL CALCULUS RESOLUTION (PCR)

Introduce extra variables $\bar{x}, \bar{y}, \bar{z}$
(new, independent variables, not related to anything, but we want them to encode negation)

Want $\bar{x} = 1 - x$

So add axiom $\dfrac{}{x + \bar{x} - 1}$ Complementarity

and also for simplicity $\dfrac{}{\bar{x}^2 - \bar{x}}$

for these new variables as well

Now $x \vee \bar{y} \vee z \rightsquigarrow x\bar{y}z$

All other derivation rules from PC are there as before

Size, length, (monomial) space defined as before

Upper bounds on $S_{PC}(F \vdash \bot)$, $Sp_{PC}(F \vdash \bot)$ will give upper bounds also for PCR

Lower bounds on $S_{PCR}(F \vdash \bot)$ and $Sp_{PCR}(F \vdash \bot)$ give lower bounds also for PC

PROPOSITION 1 PCR polynomially simulates resolution in essentially the same length, size, and space

Proof Exercise.

COROLLARY 2 For an unsat CNF F

1) $S_{PCR}(F \vdash \bot) = \exp(O(S(F)))$

2) $Sp_{PCR}(F \vdash \bot) = O(S(F))$

Where $S(F)$ = size of F, say, # literals counted with repetition.

Part 1) of Cor 2 holds also for PC by the proof of Prop 1.

PCR simulates resolution line by line and has one monomial corresponding to each clause. PC gets exponential blowup for this one monomial in worst case but there is only one monomial "per line".

It is not true that $Sp_{PC}(F \vdash 0) = O(S(F))$

Consider $F = (\bar{x}_1 \vee \cdots \vee \bar{x}_n) \wedge x_1 \wedge \cdots \wedge x_n$

The following results seems to be more or less folklore and we will not do it (probably)

THEOREM 3 PCR is exponentially stronger than resolution wrt proof size

OPEN QUESTION Is PCR asymptotically stronger than resolution wrt proof space.

Open AFAIK. Would expect the answer to be yes.

How much stronger is PCR than PC? Good question. The lower bound techniques we will introduce shortly automatically give us LB for both systems.

Also, for k-CNFs PC has "same worst-case behaviour" as we can show for resolution and PCR

LEMMA 4 [FLNTZ '12]

For any unsat k-CNF F there is a PC-refutation π with

$$S(\pi) \leq k^{O(S(F))}$$

$$S_p(\pi) = O(S(F))$$

Proof Left as a (not entirely trivial) exercise.

In PC/PCR, too, there is a very useful "auxiliary" measure: degree

largest degree of any monomial in a proof
= total # distinct variables (recall multilinearity)

Translating clauses ⟶ polynomials
width ⟶ degree

So analogue of width

Why useful: Not too hard to show bounded degree d GB-algorithms in size $n^{O(d)}$

Want to prove:

Large degree ⟹ large size

In fact, this result came before the width-length connection in resolution and [BW01] is essentially just a (smart and useful) translation of this theorem from polynomial calculus to resolution.

Formally: ~~[illegible] [IPS] [illegible]~~

THEOREM [IPS '99]

$$\mathrm{Deg}_{PCR}(F \vdash \Phi) \leq W(F) + O\left(\sqrt{n \ln S_{PCR}(F \vdash \Phi)}\right)$$

where n = # vars

COROLLARY For F k-CNF over n variables

$$S_{PCR}(F \vdash \Phi) = \exp\left(\Omega\left(\frac{\mathrm{Deg}_{PCR}(F \vdash \Phi)^2}{n}\right)\right)$$

Same theorems hold for PC
[In fact, originally proved for PC but we will use them for PCR]

QUESTION: If degree is the analogue of width, do we also have connections between degree and monomial space in PC/PCR?

Excellent question. No idea. Don't even know if anyone thought about this.

Before proving Thm 5 we take care of a couple of technicalities

PROPOSITION 7
Always $\mathrm{Deg}_{PC}(F \vdash \Phi) = \mathrm{Deg}_{PCR}(F \vdash \Phi)$
straightforward
Left as an exercise.

PROPOSITION 8
If $\pi : F \vdash \Phi$ is a PC- or PCR-refutation and ρ is a restriction, then $\pi|_\rho$ is a refutation of $F|_\rho$ in at most the same size, length, and mon space, and degree.

$x \mapsto$ true $\quad x = 0$, all polys for clauses C s.t. $x \in Lit(C)$ are satisfied & disappear

$x \mapsto$ false $\quad x = 1$, so literals x just disappear from clauses C s.t. $x \in Lit(C)$