Adding Axioms to Bounded-Depth Frege Proof Systems:

Paul Beame
University of Washington
Seattle, WA 98195

Modular Counting versus PHP

$Count_q^n$ expresses fact that there is no perfect q-partition of domain of size $qn + 1$.

Thm: (Beame, Pitassi 1993; Riis 1993) Any depth d, Frege or $e\mathcal{F}$-proof of $Count_2$ requires size $\Omega(2^{n^{\epsilon_d}})$ even when augmented with instances of PHP_m^{m+1} as axiom schemas.*

Cor: $Count_2(R)$ is not provable in

$$S_2(R) + PHP.$$

Improves on superpolynomial size lower bounds from (Ajtai 1989).

* similar bounds apply for $Count_q$, $q > 2$.

$\underline{Count_q}$

Let $V = [1, qn + 1]$.
We have variables x_e for each $e \in \binom{V}{q}$, the set of q-subsets of V.
We say that $e \perp f$ if $e \neq f$ and $e \cap f \neq \emptyset$.

"There is no perfect q-partition of a set of size $qn + 1$."

$$Count_q^n = (\bigvee_{v \in V} \bigwedge_{v \in e} \neg x_e) \vee \bigvee_{e \perp f} (x_e \wedge x_f)$$

Restrictions for $Count_q^n$ are partial q-partitions or q-dimensional matchings on V.

q-Matching Decision trees

A *q-matching decision tree over V* is a rooted directed tree T whose:
Internal nodes are labelled by elements of V,
Leaves are labelled by 0 or 1 so that:

1.(a) If the root of T is labelled by $v \in V$ then for each $e \in \binom{V}{q}$ with $v \in e$ there is one out-edge from the root labelled e.

(b) There are no other out-edges from the root of T.

2. Let T^e be the tree whose root is the node connected to the root of T by the edge labelled e. T^e is a q-matching decision tree over $V \setminus e$.

In the case $q = 2$ these are general matchings over a set of size $2n + 1$ rather than bipartite matchings from $n + 1$ points to n points.

Define q-matching disjunctions as for bipartite matchings too.

A q-matching switching lemma

Let $M^{\ell}_{q,n}$ be the set of restrictions on x_e, $e \in \binom{[1,qn+1]}{q}$ consisting of all partial q-dimensional matchings of $[1, qn+q]$ with exactly $n-\ell$ edges.

There is a canonical conversion of an r-disjunction F to a q-matching decision tree $T_V(F)$.

Switching Lemma: (Beame 1994) Let $10 \leq \ell \leq (n/r)^{1/q^2}/3e$. Choose ρ at random from $M^{\ell}_{q,n}$. Given a q-matching disjunction F in variables x_e over vertex set V with terms of size at most r, the probability that $T_{V\lceil\rho}(F\lceil\rho)$ has height at least s is at most $(6e^q \ell^q (r/n)^{1/q})^s$.

k-evaluations using q-matchings

The definitions are essentially the same as for the bipartite matching restrictions except that the tree type has changed.

As in the PHP_n^{n+1} case, we can build a k-evaluation for $k = (2q)^{2d} \log_n S$ and can maintain that $k << n_d$ provided that $S \leq n^{n^{1/(5q)^{2d}}}$.

Note that if S is polynomial and d is constant then k is a constant.

Assume that we have k-evaluation consisting of q-matching decision trees for each subformula in the proof. As before, this k-evaluation preserves "true" lines in a Frege proof and the Frege axioms all k-evaluate to "true".

Furthermore, the tree associated with the restriction of $Count_q^n$ evaluates to "false".

The only parts of the proof whose k-evaluation we must try to understand are uses of the PHP_m^{m+1} axiom schema.

The interesting feature of the new argument is to show that these also evaluate to "true".

The overall argument

Suppose that F is an instance of the PHP_m^{m+1} axiom schema.

Let T_F be the q-matching decision tree in the k-evaluation and suppose there is some $\pi \in \text{br}_0(T_F)$.

Then the set of trees given by $T'_A = T_A\lceil_\pi$ is also a k-evaluation of the formulas in F and T'_F is "false".

(We'll be sloppy and write n for the size of the domain. This is not problem since $k << n_d$.)

Let F_{ij} for $i \in D = [1, m+1]$ and $j \in R = [1, m]$ be the formulas that substitute for P_{ij} in F.

Let $T_{ij} = T'_{F_{ij}}$.

Using the T_{ij}, we will see that if they 'locally' appear to define something that is not a total 1-1 function from D to R then we get that T'_F must be "true".

Then we show that it is impossible for the T_{ij} to describe something that locally does appear to be a total 1-1 function from D to R.

Lemma: (a) If $3k \leq n$ and there are compatible branches $\sigma_j \in \mathrm{br}_1(T_{ij})$ and $\sigma'_j \in \mathrm{br}_1(T_{ij'})$ then T'_F is "true".
(b) If $3k \leq n$ and there are compatible branches $\sigma_i \in \mathrm{br}_1(T_{ij})$ and $\sigma'_i \in \mathrm{br}_1(T_{i'j})$ then T'_F is "true".

Proof: (a) Let $\sigma = \sigma_j \cup \sigma'_j$ and apply it to all formulas in the k-evaluation T'. We see that it will make $T'_{\neg F_{ij} \vee \neg F_{ij'}} \lceil_\sigma$ "false". Thus $T'_{\neg(\neg F_{ij} \vee \neg F_{ij'})} \lceil_\sigma$ is "true" and thus $T'_F \lceil_\sigma = T'_F$ is "true".
(b) is similar. □

Lemma: If there is a restriction τ such that $|\tau| + k \leq n$ and τ is incompatible with all elements of $\bigcup_{j \in R} \mathrm{br}_1(T_{ij})$ then T'_F is "true".

Proof: Apply τ to all formulas in the k-evaluation T'. Then $T'_{\bigvee_{j \in R} F_{ij}} \lceil_\tau$ is "false" so $T'_{\neg \bigvee_{j \in R} F_{ij}} \lceil_\tau$ is "true" and thus $T'_F \lceil_\tau = T'_f$ is "true". □

Now suppose that none of these things happens.

We will build new q-matching decision trees T_i using the T_{ij} that will determine for each $i \in D$ how the image of point i depends on the input variables x_e.

A leaf in T_i will have label 'j' if, after the restriction coresponding to the branch is applied, the image of i is j and leaf label '$\emptyset$' if the image of i is undefined.

The basic idea for building these trees is to use the same kind of construction as the one which builds a Boolean decision tree of height cd for a formula that has CNF clause size $\leq c$ and DNF term size $\leq d$.

For $i \in D$ we construct the 'domain' tree T_i of height $\leq 2k^2$ as follows:

0. Let τ_0 be the empty restriction and $\ell = 0$.
1. Choose some $\pi \in \bigcup_{j \in R} \mathrm{br}_1(T_{ij})$ compatible with τ_ℓ. (One must exist by assumption.)
2. Query all endpoints of vertices in π and let σ be the restriction corresponding to the answers found.
3. If $\tau_\ell \sigma$ contains an element of $\mathrm{br}_1(T_{ij})$ for some $j \in R$ then output 'j'. (Note that there is at most one such j by assumption.)
4. Let $\tau_{\ell+1} = \tau_\ell \sigma$, $\ell \leftarrow \ell + 1$, and go to 1.

Observe that each T_i has all its leaves labelled by elements of $j \in R$ and that there are no compatible branches $\pi \neq \pi' \in \bigcup_{i \in D} \mathrm{br}(T_i)$ with the same leaf label.

We now build 'range' trees T_j^r of height $\leq 8k^4$ in a similar manner using the domain trees. Fix $j \in R$. Let $S_{ij} = \mathrm{br}_j(T_i)$ be the set of all branches of T_i with leaf label j.

0. Let τ_0 be the empty restriction and $\ell = 0$.
1. Choose some $\pi \in \bigcup_{i \in D} S_{ij}$ compatible with τ_ℓ. (If none exists then halt and output '$\emptyset$'.)
2. Query all endpoints of vertices in π and let σ be the restriction corresponding to the answers found.
3. If $\tau_\ell \sigma$ contains an element of S_{ij} for some $i \in D$ then make enough queries of other endpoints so that $8k^4$ edges have been found and then output 'i'. (Note that there is at most one such i by assumption.)
4. Let $\tau_{\ell+1} = \tau_\ell \sigma$, $\ell \leftarrow \ell + 1$, and go to 1.

Observe that
(i) every branch of T_j^r has length $2k^2$ and
(ii) for any branch $\pi \in T_i$ with leaf label j every branch $\pi' \in T_j^r$ compatible with π also contains π and has leaf label i. (Furthermore such branches exist.)

The contradiction

Now for each i and j, extend each branch π in T_i with leaf label j by the tree $T_j^r\lceil_\pi$ and call the resulting tree T_i^d. Note that T_i^d has all its branches of length $8k^4$ and that each branch in T_i^d also appears as a branch of exactly one T_j^r

Thus we get a 1-1 map from the set of branches in the T_i^d to the set of branches of the T_i^r.

Since every q-matching decision tree with all its branches of length $2k^2$ has the same number of branches and there are more domain trees than range trees we get a contradiction by the pigeonhole principle. $\square$

Modular Counting for Different Moduli

Thm: (Ajtai 1994; Beame, Impagliazzo, Krajíček, Pitassi, Pudlák 1994; Riis 1994) Any constant depth Frege proof of $Count_q^n$ requires superpolynomial size even when augmented with axiom schemas for $Count_p^m$ for $m \not\equiv 0$ (mod p) and p relatively prime to q.

Cor: For p relatively prime to q, $Count_q(R)$ is not provable in $I\Delta_0(R) + Count_p$

$Count_p$ **versus** $Count_q$

Let P be a bounded-depth Frege proof of $Count_q^n$ that uses axiom schemas for $Count_p^m$ for various m.

Again apply q-matching restrictions to P to obtain a k-evaluation of all the sub-formulas in P.

As before all Frege axioms and inference rules are converted to "true" provided that k-evaluations of all instances of $Count_p^m$ are converted to "true".

Notation: Use letters e and f for q-subsets of $[1, qn+1]$ and g and h for p-subsets of $[1, pm+1]$.

Fix one instance F of $Count_p^m$ and let F_g be the formulas that substitute for the variables x_g in F.

Suppose that there is a branch π of T_F labelled 0 and apply π to the entire k-evaluation as before to obtain a new k-evaluation.

Let T_g be the tree assigned by this new k-evaluation to the formula F_g.

As in previous argument:
(a) If $g \perp h$ and there are compatible branches $\sigma_g \in \mathsf{br}_1(T_g)$ and $\sigma_h \in \mathsf{br}_1(T_h)$ then T'_F is "true".
(b) For any $i \in [1, pm+1]$, if there is a restriction σ with $|\sigma| + k \leq n$ that is incompatible with all elements of $\bigcup_{i \in g} \mathsf{br}_1(T_g)$ then T'_F is "true".

Assume that neither of these happens.

We build trees T_i for each $i \in [1, pm+1]$ as we did in the PHP_m^{m+1} case.

By assumption, all of the leaves of T_i have labels that are p-subsets containing i. Also, the trees are consistent in that if branch σ of T_i has leaf label g then for all $j \in g$, every branch of $T_j\lceil_\sigma$ has leaf label g.

It is only one more step to build the following:

Def: A (p, m)-*generic system of height* ℓ *over* V is a collection of q-matching decision trees over V: T_i, $i \in [1, pm + 1]$, with leaf labels that are p-subsets of $[1, pm + 1]$ such that:
(1) each T_i has height at most ℓ;
(2) each branch in T_i with leaf label g has $i \in g$;
(3) for all $g \in [M]^p$, for all $i, j \in g$, $\mathrm{br}_g(T_i) = \mathrm{br}_g(T_j)$.

To get this: for each branch σ of T_i with leaf label g, extend σ in T_i by appending trees $T_j\lceil_\sigma$ for each $j \in g$ in turn and labelling all leaves of the resulting branches by g. This at most multiplies the height of the trees by p.

Generic systems $\Rightarrow$ Polynomials

To prove lower bound:
Show that no (p, m)-generic systems of height ℓ exist over V when $\ell << |V|$.

We first reduce the question to finding a degree lower bound for certain polynomials related to Hilbert's Nullstellensatz.

We show a non-constant degree lower bound on these polynomials
$\Rightarrow$ no generic systems with ℓ constant.
$\Rightarrow$ no polynomial-size bounded-depth Frege proof of $Count_q$ from $Count_p$ for p and q relatively prime.

More than a polylog degree lower bound would
$\Rightarrow$ better than a quasi-polynomial separation between $Count_p$ and $Count_q$ for bounded-depth Frege proofs.

Polynomials for $\neg Count_q^n$

Let $V = [1, qn + 1]$ and $E = \binom{V}{q}$.
We have variables x_e for each $e \in E$.

For each $v \in V$, let

$$Q_v(\vec{x}) = \sum_{v \in e \in E} x_e \ - 1.$$

For each $e, f \in E$ with $e \perp f$, let

$$Q_{e,f}(\vec{x}) = x_e \cdot x_f.$$

This family of polynomials (call it $\neg Count_q^n$) has no simultaneous 0 in any field by the mod-q counting principle.

A Translation Lemma

Lemma: If a (p, m)-generic system of height ℓ over V exists then there is a family of polynomials, P_v and $P_{e,f}$, of degree $\ell - 1$ over $GF(p)$ such that

$$\sum_{v \in V} P_v \cdot Q_v + \sum_{e \perp f} P_{e,f} Q_{e,f} \equiv 1.$$

over $GF(p)$.

Proof: The basic idea is to consider the polynomial whose monomials are the products of the variables associated with each branch of the trees in the generic system. That is, with each tree T_i we get a polynomial

$$R_i = \sum_{\pi \in \mathsf{br}(T_i)} \prod_{e \in \pi} x_e.$$

We will show that each R_i is $1 + L_i$ where L_i a linear combination of the Q polynomials of degree at most $\ell - 1$.

Assume this for the moment and consider $\sum_{i\in[1,pm+1]} R_i$ in $GF(p)$.

On the one hand it is

$$\sum_{i\in[1,pm+1]} (1 + L_i) = 1 + \sum_{i\in[1,pm+1]} L_i.$$

On the other hand, every branch in the generic system appears some multiple of p times; each time it has leaf label g there are p copies - one for each element of g. Therefore over $GF(p)$,

$$\sum_{i\in[1,pm+1]} R_i = 0.$$

We derive $1 + \sum_{i\in[1,pm+1]} L_i = 0$ - basically what we want. $\square$

Lemma: Let T be a q-matching decision tree. Then $R(T) = \sum_{\pi \in \mathsf{br}(T)} \prod_{e \in \pi} x_e$ is of the form $1 + L$ where L is a linear combination of the Q_v and $Q_{e,f}$ with coefficient polynomials of degree $\leq \ell - 1$.

Proof: Proof by induction on ℓ.

Base case: Height 1. Let v be the node queried at the root. Then the polynomial for T is just

$$\sum_{v \in e} x_e = 1 + Q_v$$

Induction step: (sketch) Let v be queried at the root of T. For each e with $v \in e$, group terms from T^e (the subtree reached by e) and factor out x_e. Apply the inductive hypothesis to each T^e over $V \setminus e$. Then observe that the missing terms would be cancelled by multiples of $Q_{e,f}$ to get the result for T. $\square$

Now show that any P_v, $P_{e,f}$ such that

$$\sum_v P_v Q_v + \sum_{e \perp f} P_{e,f} Q_{e,f} = 1$$

must have non-constant degree:

(Beame, Impagliazzo, Krajíček, Pitassi, Pudlák 1994) show by Ramsey theory that for any linear combination L of the $Q_v, Q_{e,f}$ with constant degree coefficients, there is some vector $\vec{x}$ of values so that $L(\vec{x}) = 0$ as a function.

Therefore no k-evaluation for constant k which gives the superpolynomial lower bound.

(Actual degree bound is $\Omega(\log^* n)$.)

Note: It is possible that for all $\vec{x}$, $L(\vec{x}) = 1$ but $L \not\equiv 1$ so this proves something that may be stronger than necessary.

Hilbert's Nullstellensatz

Given multivariate polynomials

$$Q_1(\vec{x}), \ldots, Q_m(\vec{x}) \in k[x_1, \ldots, x_n]$$

there is no solution to

$$\begin{aligned} Q_1(\vec{x}) &= 0 \\ \cdots &= \cdot \\ Q_m(\vec{x}) &= 0 \end{aligned}$$

over any extension field of k

$\Leftrightarrow \exists P_1(\vec{x}), \ldots, P_m(\vec{x}) \in k[x_1, \ldots, x_n]$
such that $\sum_{i=1}^{m} P_i(\vec{x}) \cdot Q_i(\vec{x}) \equiv 1$

$P_1, \ldots, P_m$ are *Nullstellensatz refutation* of (*).

The *degree* of the refutation
$=$ max degree of P_i

Note: Adding $x_1^p - x_1, \ldots, x_n^p - x_n$ to the Q_i's makes this work for solutions in $k = GF(p)$.

The original $Count_p$ versus $Count_q$ separation is due to (Ajtai 1994) who derives a 'symmetric' system of linear equations modulo p that are similar to the linear equations in the coefficients of the monomials in our polynomial equation.

He then rederives some of the basic results of the theory of representations of the symmetric group to show that they are constructive enough to forbid the solutions that these equations must have.

(His bound is also $\Omega(\log^* n)$ but his technique for deriving these equations does not apply to quasi-polynomial size trees.)

(Riis 1994) subsequently has applied the same Ramsey theory technique directly to the q-matching decision trees.

Open Problems and Future Work

Open Problem: Prove that poly-log height generic systems cannot exist.

Open Problem: Improve on the $\Omega(\log^* n)$ degree lower bound for the coefficient polynomials of the $\neg Count_q^n$ polynomials that produce 1.

Super-polylog degree lower bounds would show that $S_2(R)+Count_p$ cannot prove $Count_q$ where p and q are relatively prime.

Understand Nullstellensatz-based proof systems

Principles beyond $\bigcup_p Count_p$