11. Width-3 ABP, Nisan's Characterization

Tuesday, September 26, 2023 11:05 PM

Today: <u>Thm</u> [Ben-Or, Cleve '88]: Algebraic formulas and width-3 ABPs are equivalent up to poly-size. (or width-c, $c \geq 3$ is a constant).

<u>Pf</u>: Suppose F is a poly(n)-size formula in $X_1, \dots, X_n$. We turn it into a poly(n)-size width-3 ABP.

By depth reduction, we may assume the depth of F is $O(\log n)$.

Recursively transform F into a product M of 3×3 matrices that equals $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ F & 0 & 1 \end{pmatrix}$ whose entries are degree-≤ 1 polynomials

If F is a leaf $\Rightarrow$ $M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ F & 0 & 1 \end{pmatrix}$

If $F = F_1 + F_2$ $\Rightarrow$ $M = M_1 M_2$, where $M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ F_1 & 0 & 1 \end{pmatrix}$, $M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ F_2 & 0 & 1 \end{pmatrix}$

If $F = F_1 \cdot F_2$ $\Rightarrow$

$$M = \underbrace{\overset{=A}{\begin{pmatrix} 1 & 0 & 0 \\ -F_2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \cdot \overset{=B}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & F_1 & 1 \end{pmatrix}}} \cdot \underbrace{\overset{C}{\begin{pmatrix} 1 & 0 & 0 \\ F_2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}} \cdot \overset{=D}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -F_1 & 1 \end{pmatrix}}}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ -F_2 & 1 & 0 \\ 0 & F_1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ F_2 & 1 & 0 \\ 0 & -F_1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ F_1 F_2 & 0 & 1 \end{pmatrix}$$

A, B, C, D can be computed from M_1 or M_2 via elementary row and column operations.

The final ABP is $(0, 0, 1)\, M \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$. size = poly($n$).

Conversely, suppose B is an ABP of size poly(n).

$B = u \cdot M_1 \cdots M_\ell \cdot v$. Each entry of $u M_1 \cdots M_{\ell/2}$ and $M_{\ell/2+1} \cdots M_\ell \cdot v$ (only 6 entries in total.) is computed by an ABP with ℓ replaced by $\ell/2$.

Transform B into a formula recursively. size $S(\ell) = 6S(\ell/2) + O(1)$

(when $\ell = 0$, just compute $B = u \cdot v$ as a formula.) $\Rightarrow S(\ell) = O(\ell^{\log_2 6}) \leq \text{poly}(n)$.

$\square$.

The proof also shows that $(\mathrm{IMM}_{3,n})$ is VF-complete. (3: width, n: length)

The proof also shows that $\{\mathrm{IMM}_{3,n}\}$ is VP-complete,

$\underset{\text{width}}{\overset{\nearrow}{3}}$, $\underset{\text{length}}{n}$

Related work:

Thm (Allender-Wang '11) For $k \geq 8$, $\sum_{i=1}^{k} X_{2i-1} \cdot X_{2i} + \underbrace{\ell(X_1, \ldots, X_n)}_{\text{arbitrary linear form.}}$ is not computable by width-2 ABPs of <u>any</u> length.

suppose $\mathrm{char}(\mathbb{F}) \neq 2$,

Thm (Bringmann-Ikenmeyer-Zuiddam '17) width-2 ABPs of poly(n) size can <u>approximately compute</u> any $f \in VF$, in the sense that

there exists a width-2 ABP of poly(n) size over $\mathbb{F}(\varepsilon)$ (ε a new symbol.) computing $f + \varepsilon g$, where $g \in (\mathbb{F}[\varepsilon])[X_1, \ldots, X_n]$ $\left(\text{intuitively, } \lim_{\varepsilon \to 0} (f + \varepsilon g) = f\right)$.

Nisan's characterization of (non-commutative) ABP complexity.

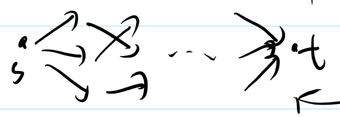
The noncommutative polynomial ring $\mathbb{F}\{X_1, \ldots, X_n\}$ is the <u>free algebra</u> generated by $X_1, \ldots, X_n$ over $\mathbb{F}$. (In contrast, $\mathbb{F}[X_1, \ldots, X_n]$ is the free commutative algebra generated by $X_1, \ldots, X_n$ over $\mathbb{F}$, i.e. we have the relations $X_i X_j = X_j X_i$.)

An ABP in the non-commutative setting computes some $f \in \mathbb{F}\{X_1, \ldots, X_n\}$.

Let $f \in \mathbb{F}\{X_1, \ldots, X_n\}$ be homogeneous of degree d.

Consider the computation of f in the homogeneous model:

Layers 0 1 2 ... d-1 d



← #vertices at layers are allowed to vary.

← weights are homogeneous linear forms in $X_1, \ldots, X_n$.

For $0 \leq k \leq d$, let M_k^f be a $n^k \times n^{d-k}$ matrix whose rows are indexed by $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$, and columns are indexed by $(j_1, \ldots, j_{d-k}) \in \{1, \ldots, n\}^{d-k}$,

and $M_k^f\big((i_1, \ldots, i_k), (j_1, \ldots, j_{d-k})\big) =$ Coefficient of $X_{i_1} \cdots X_{i_k} \cdot X_{j_1} \cdots X_{j_{d-k}}$ in f.

Thm (Nisan '91) Let $0 \neq f \in \mathbb{F}\{x_1, \dots, x_n\}$.

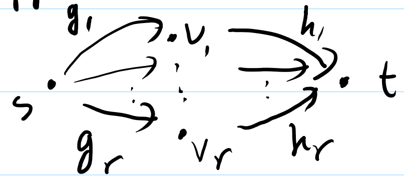
(1) Let B be a homogeneous ABP computing f. Then

$$\#\text{ vertices at layer } k \text{ of } B \geq \operatorname{rank}(M_k^f) \quad \text{for } k = 0, \dots, d.$$

(2) There exists a homogeneous ABP B computing f with

$$\#\text{ vertices at layer } k \text{ of } B = \operatorname{rank}(M_k^f) \quad \text{for } k = 0, \dots, d.$$

Pf : (1) Suppose B has $v_1, \dots, v_r$ at layer k.



Then $f = \sum_{i=1}^{r} g_i \cdot h_i$.

$$\text{Then } M_k^f = M_k^{\sum_{i=1}^r g_i h_i} = \sum_{i=1}^{r} M_k^{g_i h_i}.$$

By definition, $M_k^{g_i h_i} = \underbrace{\left(\operatorname{coeff}_{x_{\mathbf{i}}} g_i\right)_{\mathbf{i}}^{T}}_{\text{column vector}} \cdot \underbrace{\left(\operatorname{Coeff}_{x_{\mathbf{j}}} h_i\right)_{\mathbf{j}}}_{\text{row vector}}$ ← rank-1 matrix.

$\mathbf{i} = (i_1, \dots, i_k) \in \{1, \dots, n\}^k \quad \mathbf{j} = (j_1, \dots, j_{n-k}) \in \{1, \dots, n\}^{n-k}$.

So $\operatorname{rank}(M_k^f) \leq r$.

(2) We iteratively build the k-th layer of B for $k = 0, 1, \dots, d$
with vertices $v_{k,1}, \dots, v_{k,r_k}$ s.t.
$f_{v_{k,1}}, \dots, f_{v_{k,r_k}}$ form a basis of the column space of M_k^f (*)
where $f_{v_{k,i}}$ denotes the polynomial computed at $v_{k,i}$,
and the columns of M_k^f are viewed as deg-k polynomials
via the correspondence $(c_{(i_1, \dots, i_k)}) \mapsto \sum_{(i_1, \dots, i_k)} c_{(i_1, \dots, i_k)} x_{i_1} x_{i_2} \cdots x_{i_k}$.

Base case: $k = 0$. $v_{k,1} = s$. $f_{v_{k,1}} = 1$. $f \neq 0$.

M_k^f is a $1 \times n^d$ matrix. So $f_{v_{k,1}}$ is a basis of the column space of M_k^f.

Now suppose layer $k-1$ is built with vertices $v_{k-1,1}, \dots, v_{k-1,r_{k-1}}$.

Let $f_1, \dots, f_r \in \mathbb{F}\{x_1, \dots, x_n\}$ form a basis of the column space of M_k^f

Let $f_1, \dots, f_{r_k} \in \mathbb{F}\{x_1, \dots, x_n\}$ form a basis of the column space of M_k^f.

Add $v_{k,1}, \dots v_{k,r_k}$ to the k-th layer.

We need to add wires between $(k-1)$ th and k-th layer such that $v_{k,i}$ computes f_i. Then $(*)$ would hold.

For this, we just need to prove each $f_i \in \operatorname{span}\{f_{v_{k-1,j}} \cdot x_{j'}\}_{1 \le j \le r_{k-1},\ 1 \le j' \le n}$.

Write $f = \sum_{(i_k, \dots, i_d) \in \{1,\dots,n\}^{d-k+1}} c_{(i_k,\dots,i_d)} x_{i_k} \cdots x_{i_d}$ where $c_{(i_k,\dots,i_d)}$ is a deg-$(k-1)$ polynomial in $\mathbb{F}\{x_1, \dots, x_n\}$.

Then both $\{f_{v_{k-1,1}}, \dots, f_{v_{k-1,r_{k-1}}}\}$ and $\{c_{(i_k,\dots,i_d)}\}_{(i_k,\dots,i_d)}$ span the column space of M_{k-1}^f.

Moreover, both $\{f_1, \dots, f_{r_k}\}$ and $\{c_{(i_k,\dots,i_d)} \cdot x_{i_k}\}_{(i_k,\dots,i_d)}$ span the column space of M_k^f.

So it suffices to show each $c_{(i_k,\dots,i_d)} \cdot x_{i_k}$ is in $\operatorname{span}\{c_{(i_k,\dots,i_d)} \cdot x_{j'}\}$ which is obviously true by letting $j' = i_k$.

<u>Lemma</u>: Let $f = PERM_n = \sum_{\sigma \in S_n} \prod_{i=1}^n x_{i\sigma(i)}$ or $f = DET_n = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^n x_{i\sigma(i)}$.

Then $\operatorname{rank}(M_k^f) = \binom{n}{k}$ for $k = 0, \dots, n$.

<u>Pf</u>: For $i_1, \dots, i_k \in \{1, \dots, n\}$ and $j_1, \dots, j_{n-k} \in \{1, \dots, n\}$.

The $\big((1,i_1), \dots, (k,i_k)\big), \big((1,j_1), \dots, (n-k, j_{n-k})\big)$-the entry of $PERM_n$ is 1 if $\{i_1, \dots, i_k\}$ and $\{j_1, \dots, j_{n-k}\}$ form a partition of $\{1, \dots, n\}$, and 0 otherwise.

The other entries of $PERM_n$ are zero.

So/ <u>we</u> may choose $\binom{n}{k}$ k-sets as rows and $\binom{n}{n-k} = \binom{n}{k}$ $(n-k)$-sets as columns after removing duplicate rows / columns,

[illegible] identity submatrix and this is the largest nonsingular submatrix.

after removing duplicate rows / columns,

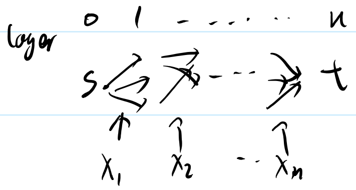
to form an identity submatrix, and this is the largest nonsingular submatrix.

$\Rightarrow \text{rank} = \binom{n}{k}$.

For DET_n it is the same except that the duplicate rows / columns may be multiplied by -1 due to $\mathrm{sgn}(\sigma)$. □

Cor: In the above ABP model, size of ABPs computing $PERM_n$ or $DET_n \geq \exp(n)$.

A related model: read-once oblivious branching programs (ROABPs)



weights of edges from $(k-1)$th to kth layer are univariate polynomials $f_e \in \mathbb{F}[X_k]$ with $\deg(f_e) \leq d$. $d \in \mathbb{N}$.

Polynomials are in the commutative polynomial ring $\mathbb{F}[X_1, \ldots X_n]$. In fact, it does not matter since X_i always gets multiplied before X_j if $i < j$.

Nisan's characterization also holds in this model, with $X_1, \ldots, X_n$ at the k-th position replaced by $X_k^0, X_k^1, \ldots, X_k^d$. The proof is the same and left as an exercise.