2. Elimination of Division Gates

Friday, August 11, 2023 9:59 PM

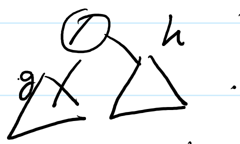
Ref: Shpilka-Yehudayoff '10, §2.5 (Strassen)

Thm 1. If $f \in \mathbb{F}[X_1, \dots, X_n]$ can be computed by an algebraic circuit of size s with divison gates, and $\deg(f) = d$, then f can be computed by an algebraic circuit of size $\mathrm{poly}(s, d, n)$ without division gates.

In particular, allowing divisions does not make VP more powerful.

Suppose C computes f with division gates.

Observation: It is easy to remove all division gates except one top gate:

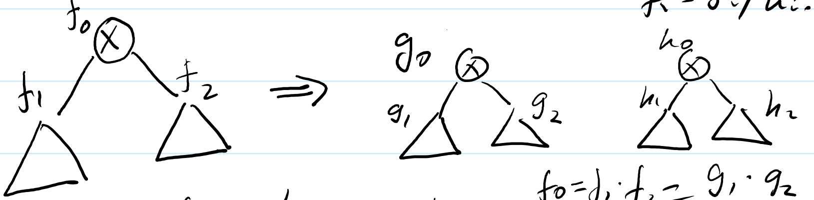


$f = g/h$.

Pf of the observation: In a bottom-up fashion, for each gate computing some f_i, turn it into g_i and h_i such that $f_i = g_i/h_i$.

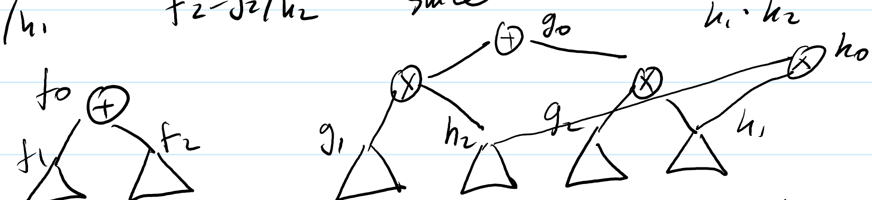
Transformation:

$\times$:



$f_1 = g_1/h_1$ $\quad f_2 = g_2/h_2$ $\quad$ since $f_0 = f_1 \cdot f_2 = \dfrac{g_1 \cdot g_2}{h_1 \cdot h_2}$

$+$:



since $f_0 = f_1 + f_2 = \dfrac{g_1 h_2 + h_1 g_2}{h_1 h_2}$ $\square$

Problems: (1) There is still one division gate at the top.

(2) $f = g/h$. Degrees of g and h can be very high.

Extracting homogeneous components.

Def: For $f \in \mathbb{F}[X_1, \dots, X_n]$ and an integer $i \geq 0$, define $\mathrm{Hom}_i(f)$ to be the homogeneous component of degree i of f.

Example: $f = x^3 + xy + x^2y + 3y^2 + 6$

Example: $f = X^3 + XY + X^2Y + 3Y^2 + 6$

Then $\mathrm{Hom}_2(f) = XY + 3Y^2$.

Lemma: Suppose f is computed by a circuit of size s and $\deg(f) = d$. Then there is a (multi-output) homogeneous circuit of size $O(d^2 \cdot s)$ that computes $\mathrm{Hom}_0(f), \cdots, \mathrm{Hom}_d(f)$.

Pf: Again, we do it gate by gate. For $f_0 = f_1 + f_2$, $\mathrm{Hom}_i(f_0) = \mathrm{Hom}_i(f_1) + \mathrm{Hom}_i(f_2)$.

And for $f_0 = f_1 \cdot f_2$, $\mathrm{Hom}_i(f_0) = \sum_{0 \le j \le i} \mathrm{Hom}_j(f_1) \cdot \mathrm{Hom}_{i-j}(f_2)$. $\square$

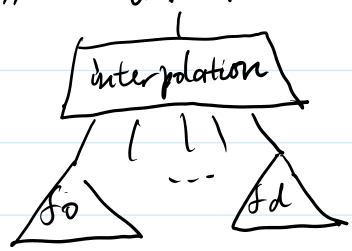
Remark: An alternative way (when $\mathbb{F}$ is large enough):

Pick distinct $a_0, \cdots, a_d \in \mathbb{F}$.

Compute $f_i := f(a_i X_1, a_i X_2, \cdots, a_i X_n)$ for $i = 0, 1, \cdots, d$.

Note $f_i = \sum_{j=0}^{d} a_i^j \mathrm{Hom}_j(f)$.

Extract $\mathrm{Hom}_0(f), \cdots, \mathrm{Hom}_d(f)$ from $f_0, \cdots, f_d$ via interpolation.



(Useful sometimes since the depth increase is additive.)

Lemma: Let $S \subseteq \mathbb{F}$ be a finite set of size k. Suppose $0 \ne f \in \mathbb{F}[X_1, \cdots, X_n]$ and $\deg_{X_i}(f) < k$ for $i = 1, 2, \cdots, n$. Then S^n contains a nonzero of f, i.e. $f(a_1, \cdots, a_n) \ne 0$ for some $(a_1, \cdots, a_n) \in S^n$.

Pf: For $i \in \{0, 1, \cdots, n\}$, we show $\exists\, a_1, \cdots, a_i \in S$ such that $f(a_1, \cdots, a_i, X_{i+1}, \cdots, X_n) \ne 0$.

Induct on i.

For $i = 0$, claim is obvious.

$i \to i+1$: We know $f(a_1, \cdots, a_i, X_{i+1}, \cdots, X_n) \ne 0$
$\in \mathbb{F}[X_{i+1}, \cdots, X_n]$

$$f(a_1, \cdots, a_i, X_{i+1}, \cdots, X_n) = \sum_{e = (e_{i+2}, \cdots, e_n)} X_{i+2}^{e_{i+2}} \cdots X_n^{e_n} \cdot f_e(X_{i+1})$$

$$f(a_1,\dots,a_i,X_{i+1},\dots,X_n) = \sum_{e=(e_{i+2},\dots,e_n)} X_{i+2}^{e_{i+2}} \cdots X_n^{e_n} \cdot f_e(X_{i+1})$$

where each f_e has degree $\le k$. and some $f_e \ne 0$. ← call it f_{e^*}

As f_{e^*} has at most k roots, we can choose some $a_{i+1} \in S$ such that $f_{e^*}(a_{i+1}) \ne 0 \Rightarrow f(a_1,\dots,a_{i+1},X_{i+2},\dots,X_n) \ne 0$.

Continue this process. $\square$

Remark: A related result called the Schwartz-Zippel lemma can also be used in place of Lemma 2.

Pf of Thm 1: Write $f = g/h$. Assume $\mathbb{F}$ is large enough. $(|\mathbb{F}| > \deg(h))$

Then $h(a) \ne 0$ for some $a = (a_1,\dots,a_n) \in \mathbb{F}^n$.

We may assume $a = \vec{0}$ by performing $X_i \mapsto X_i - a_i$ (and back)

So $h(\vec{0}) \ne 0$. By scaling, we may assume $h(\vec{0}) = 1$.

Write $h = 1 - t$ where $t := 1 - h \in \langle X_1,\dots,X_n \rangle$

That is, the constant term of t is zero. $(\Leftrightarrow t(0) = 0)$

$$f = \frac{g}{h} = \frac{g}{1-t} = g(1+t+t^2+\dots) \in \mathbb{F}[[X_1,\dots,X_n]]$$

← ring of formal power series over $\mathbb{F}$.

(Note: $h = 1-t$ is invertible in $\mathbb{F}[[X_1,\dots,X_n]]$ because $h(0) \ne 0$)

The expression $g(1+t+t^2+\dots)$ has infinitely many terms.

However, note $f = \mathrm{Hom}_{\le d}(g(1+t+t^2+\dots+t^d))$.

Pf: $f - \mathrm{Hom}_{\le d}(g(1+t+\dots+t^d)) = \mathrm{Hom}_{\le d}(f - g(1+t+\dots+t^d))$

$= \mathrm{Hom}_{\le d}(g(t^{d+1} + t^{d+2} + \dots))$

$= 0$ (why? b/c t is constant-free)

Compute f as $\mathrm{Hom}_{\le d}(g(1+t+\dots+t^d))$ and we are done.

Technical issue: need $|\mathbb{F}| > \deg(h)$.

How large can $\deg(h)$ be? At most $\exp(s)$. s = size of circuit.

If $\mathbb{F}$ is too small, choose an extension field $\mathbb{K}$ of $\mathbb{F}$

... view $\mathbb{K}$ as a vector space over ...

If $\mathbb{F}$ is too small, choose an extension field $\mathbb{K}$ of $\mathbb{F}$

simulate $\mathbb{K}$-operations over $\mathbb{F}$ by view $\mathbb{K}$ as a vector space over $\mathbb{F}$.

Just need $[\mathbb{K} : \mathbb{F}] \sim \log_{|\mathbb{F}|} \deg(h) \leq \log \deg(h) = O(s)$. $\square$

A similar result holds for algebraic formulas, but it requires a new technique called depth reduction. We will discuss this later.