Lecture 17: On Balancing Circuits

Anup Rao

March 9, 2023

In this lecture, we finally prove something that I mentioned in my very first lecture: it is possible to balance every arithmetic circuit.

Homogenization

First, we need the concept of a *homogenous* polynomial/circuit. A polynomial is homogenous if all of its monomials have the same degree. An arithmetic circuit is homogenous if every gate computes a homogenous polynomial. Given a polynomial f of degree d, we write f_i to denote its i'th homogenous part. So, $f = f_0 + \ldots + f_d$.

A useful fact is that every circuit can be made homogenous in the following sense:

Theorem 1. *If f is a degree d polynomial that can be computed by a circuit of size s, then $f_0, \ldots, f_d$ can all be computed by a homogenous arithmetic circuit of size $O(sd^2)$.*

Proof The idea of the proof is to compute $g_0, \ldots, g_d$ for every gate g in the circuit of size s. If $g = u + v$, then $g_i = u_i + v_i$, so the homogenous parts of g can be computed from the homogenous parts of u, v. If $g = u \cdot v$, then $g_i = u_0 \cdot v_i + u_1 \cdot v_{i-1} + \ldots + u_i \cdot v_i$, so once again the homogenous parts of g can be computed. All of these operations may increase the size of the circuit by a factor of $O(d^2)$. ■

The key claim

The key claim we shall make is the following:

Theorem 2. *Suppose $f(X_1, \ldots, X_n)$ is a degree d homogenous polynomial computed by a homogenous arithmetic circuit of size s. Then we can express*

$$f = \sum_{i=1}^{s} u_i v_i,$$

where for every i, u_i and v_i both have degree at least $d/3$ and at most $2d/3$, u_i occurs as a gate in the original circuit, and v_i can be computed by the same circuit after replacing some of the gates with the constants 0 or 1.

Balancing

THEOREM 2 IS extremely powerful. In particular, it implies that one can compute f using a circuit of depth at most $O((\log s)(\log d))$. To see this, generate a circuit of depth $O(\log s)$ that computes f from inputs u_i, v_i as above. Then, since each of u_i, v_i can be computed by circuits of size s, we can recursively apply the Theorem to these polynomials and continue. In each step, the degree of the polynomials we are working with drops by a constant factor, so there can be at most $O(\log d)$ steps.

Even if f is not homogenous, we can use Theorem 1 to make a homogenous circuit computing the homogenous parts of f in size $O(sd^2)$. Then, applying Theorem 2, we obtain a circuit of depth $O((\log sd^2) + \log d) \leq O((\log s + \log d) \log d)$ computing the homogenous parts of f. We can then sum up these parts adding another $O(\log d)$ to the depth to recover f. As a consequence, we obtain:

Theorem 3. *If f is a polynomial of degree d that can be computed using an arithmetic circuit of size s, then f can be computed by an arithmetic circuit of depth $O((\log s + \log d) \log d)$.*

Proving the theorem

FINALLY, LET US TURN to proving the theorem. The given circuit is assumed to be homogenous. In fact, it is no loss of generality to assume that every gate of the circuit computes a polynomial of degree at most d. This is because if the circuit contains a $+$ gate that computes the polynomial 0, then we can eliminate that gate. Once all such gates have been eliminated, we see that every gate computes a polynomial whose degree is larger than the degrees of its inputs. Thus, any gate computing a polynomial of degree larger than d cannot be connected to the output gate, and it can be dropped.

Next we run a process similar to what we have seen when found a way to balance Boolean formulas. Let $a_1, a_2, \ldots$ be a sequence of gates, where a_1 the output gate, and given a_i, a_{i+1} is the gate that feeds into a_i of larger degree (breaking ties arbitrarily). Since the product of two gates adds the degrees, the degree of the polynomial computed by a_{i+1} must be at least $1/2$ of the degree of a_i. Let a_{i+1} be the first gate in this sequence with

$$d/3 \leq \mathsf{deg}(a_{i+1}) \leq 2d/3.$$

By construction, we must have $a_i = a_{i+1} \cdot b$, and the degree of a_i must be greater than $2d/3$. Now, imagine replacing the gate a_i with a new variable Y. Let $g(X_1, \ldots, X_n, Y)$ denote the output of the circuit after making this change, so $f(X_1, \ldots, X_n) = g(X_1, \ldots, X_n, a_i)$, where here a_i denotes the polynomial computed by the gate a_i.

We claim:

Claim 4. *If a gate r in the circuit computing g computes a polynomial containing the monomial $Y \cdot h$, then the degree of r in the circuit for f must be $\mathsf{deg}(a_i) + \mathsf{deg}(h)$.*

The claim holds by induction. It is true for the gate a_i, and given that the claim holds for the inputs of r, it must hold for r, since we have eliminated all gates of the circuit for f that compute the 0 polynomial.

Next, we claim that the degree of Y in g is at most 1. Indeed, if the circuit ever multiplies a polynomial containing Y with another polynomial containing Y, then the degree of this gate in the original circuit has to be at least $4d/3$, but there are no such gates, since we got rid of them in the first step of the proof. Thus, we must have

$$g = h \cdot Y + q,$$

for some polynomials $h(X_1, \ldots, X_n), q(X_1, \ldots, X_n)$.

Now, set $u_1 = a_{i+1}$, $v_1 = h \cdot b$. Then we have

$$f = u_1 \cdot v_1 + q.$$

v_1 can be computed by considering the path from b to the output gate, replacing the gate a_i by 1, and replacing every polynomial that is added to this path by 0.

Moreover, q can be computed by substituting $Y = 0$ in the circuit computing g. Thus, q must be homogenous and have the same degree as f (or be 0). Since q can be computed by a circuit of size at most $s - 1$, the proof is completed by induction.