On The Method of Approximations

A. A. Razborov
Steklov Mathematical Institute,
117966, Moscow, GSP-1, Vavilova, 42, USSR

1 Introduction

Recently, there has been progress in proving lower bounds for certain restricted models of Boolean computations. In particular, by the so-called method of approximations, exponential lower bounds have been obtained for monotone complexity [Raz85b, Raz85a, AE85, AB87, Tar89, And87] and for bounded-depth circuits over several bases [Raz87, Smo87, Pat86] (an informal account of the method can be found in [Raz86, BS88]). The most interesting question about approximations is, doubtless, how useful can they be for arbitrary circuits. This question can be made quite precise because the method of approximations is easy to formalize (actually this was done already in [Raz85b]).

In the present paper we prove some results clarifying the situation. It turns out that the answer to the question we are interested in depends very much on whether we allow auxiliary variables or not.

First we prove that lower bounds which could be obtained by the method of approximations never exceed $O(n_0 n)$ where n is the total number of variables of the Boolean function whose complexity we want to estimate, and n_0 is the number of essential variables of this function (theorem 3.1). Moreover, if one were to use probabilistic arguments for estimating the distance $\rho(f, \mathcal{M})$ then this restriction can be strengthened to $O(n_0)$ (theorem 2.4).

We see that these results do not exclude a possibility of proving good lower bounds for the circuit size of a Boolean function by introducing a large number of auxiliary variables (note, however, that theorem 2.4shows the uselessness of probabilistic methods for this purpose). We prove that this does work in the strongest sense. Namely, for any Boolean function we define some "universal" legitimate model $\mathcal{M}_{max}$ such that the effective bound $\rho(f, \mathcal{M}_{max})$ for the circuit size of f *coincides* with this size up to a polynomial (theorem 2.6).

The total number of variables involved in the model $\mathcal{M}_{max}$ is huge (double exponential in $n_0!$). It can be decreased to a single exponent without affecting universal properties of the model (theorem 2.7). But, unlike $\mathcal{M}_{max}$, the construction of this model is not canonical and makes use of what could be called "the axiom of choice."

The paper is organized as follows. In section 2 we give precise definitions to formalize the method of approximations and formulations of all our theorems. In section 3 we prove results illustrating the weakness of approximations, i.

Permission to copy without fee all or part of this material is granted provided that the copies are not made or distributed for direct commercial advantage, the ACM copyright notice and the title of the publication and its date appear, and notice is given that copying is by permission of the Association for Computing Machinery. To copy otherwise, or to republish, requires a fee and/or specific permission.

© 1989 ACM 0-89791-307-8/89/0005/0167 $1.50

e. theorems 2.1, 2.4. In section 4 we prove theorems 2.6, 2.7. The paper is completed by a short discussion in section 5.

2 The method of approximations. Main results.

Throughout the paper B^n denotes an n-dimensional boolean cube and F_n the set of all boolean functions in n variables. For $u \in B^n$, $u^i (1 \leq i \leq n)$ means the ith bit in u. Let $X_i^\varepsilon = \{u \in B^n | u^i = \varepsilon\}$ for $1 \leq i \leq n$, $\varepsilon \in \{0,1\}$. Given a variable x_i set $x_i^1 = x_i$; $x_i^0 = (\neg x_i)$. It is convenient for our purposes to regard *circuits* as having gates $\{\&, \vee\}$ and inputs $0, 1, x_i^\varepsilon (1 \leq i \leq n, \varepsilon \in \{0,1\})$ i.e. negations are allowed only on variables. By $|C|$ we denote the function computed by a circuit C; by $L(f)$ - the circuit size of f; by $L_{mon}(f)$ - the monotone circuit size of a monotone f.

An $\mathcal{M} \subseteq F_n$ supplied with two binary operations $\bar{\&}$ and $\bar{\vee}$ is called a *legitimate model* if

$$\{0, 1, x_i^\varepsilon (1 \leq i \leq n, \varepsilon \in \{0,1\})\} \subseteq \mathcal{M}. \quad (1)$$

Given a circuit C we denote by $\bar{C}$ the circuit obtained from C after replacing $\{\&, \vee\}$ by $\{\bar{\&}, \bar{\vee}\}$. We keep notations $\bar{g}, \bar{h}$ and so on for Boolean functions from $\mathcal{M}$. Let

$$\begin{aligned} \delta^+_{\&}(\bar{g}, \bar{h}) &= (\bar{g} \& \bar{h}) \backslash (\bar{g} \bar{\&} \bar{h}); \\ \delta^+_{\vee}(\bar{g}, \bar{h}) &= (\bar{g} \vee \bar{h}) \backslash (\bar{g} \bar{\vee} \bar{h}); \quad (2) \\ \delta^-_{\&}(\bar{g}, \bar{h}) &= (\bar{g} \bar{\&} \bar{h}) \backslash (\bar{g} \& \bar{h}); \\ \delta^-_{\vee}(\bar{g}, \bar{h}) &= (\bar{g} \bar{\vee} \bar{h}) \backslash (\bar{g} \vee \bar{h}). \quad (3) \end{aligned}$$

Set

$$\Delta^+ = \{\delta^+_*(\bar{g}, \bar{h}) | \bar{g}, \bar{h} \in \mathcal{M}; * \in \{\&, \vee\}\}; \quad (4)$$

$$\Delta^- = \{\delta^-_*(\bar{g}, \bar{h}) | \bar{g}, \bar{h} \in \mathcal{M}; * \in \{\&, \vee\}\} \quad (5)$$

Further, given $f \in F_n, \bar{f} \in \mathcal{M}$ let $\rho(f, \bar{f})$ be the minimal t for which there exist t triples $< *_i, \bar{g}_i, \bar{h}_i > (*_i \in \{\&, \vee\}; \bar{g}_i, \bar{h}_i \in \mathcal{M})$ such that

$$f \leq \bar{f} \vee \bigvee_{i=1}^{t} \delta^+_{*_i}(\bar{g}_i, \bar{h}_i), \quad (6)$$

$$\bar{f} \leq f \vee \bigvee_{i=1}^{t} \delta^-_{*_i}(\bar{g}_i, \bar{h}_i). \quad (7)$$

The main property of approximations is that $\rho(|C|, |\bar{C}|) \leq size(C)$.

In order to show this, let $*_i$ be the operation computed by the ith gate of the circuit C, $\bar{g}_i, \bar{h}_i$ be inputs of this gate in the circuit $\bar{C}$. Say that a gate of C has *+- error* on an n-bit string v if $g(v) = 1, \bar{g}(v) = 0$ and has *- - error* on an n-bit string u if $g(u) = 0, \bar{g}(u) = 1$ where g and $\bar{g}$ are outputs of this gate computed by circuits C and $\bar{C}$ respectively. For proving (6) we need to consider a string v such that $f(v) = 1, \bar{f}(v) = 0$ (i.e. the output gate of C has +- error on v) and prove that $\delta^+_{*_i}(\bar{g}_i, \bar{h}_i)(v) = 1$ for some i. Since inputs of C do not have + -errors, there exists a gate which has + - error on v whereas both its predecessors do not have. It is easy to check that $\delta^+_*(\bar{g}, \bar{h})(v) = 1$ for the corresponding $< *, \bar{g}, \bar{h} >$. (7) is proved similarly. Therefore, by letting $\rho(f, \mathcal{M}) = min_{\bar{f} \in \mathcal{M}} \rho(f, \bar{f})$ we have

$$\rho(f, \mathcal{M}) \leq L(f). \quad (8)$$

Similarly, for any $\bar{g}_1, ..., \bar{g}_l \in \mathcal{M}$ and $f \in F_l$

$$\rho(f(\bar{g}_1, ..., \bar{g}_l), \mathcal{M}) \leq L_{mon}(f) \quad (9)$$

Theorem 2.1 For any legitimate model $\mathcal{M}$ and boolean function $f \in F_n$ actually depending on $n_0 (n_0 \leq n)$ variables the inequality $\rho(f, \mathcal{M}) \leq \mathcal{O}(n_0 n)$ holds.

Remark 2.2 Let us note for comparison that if (1) is weakened to $\{0, 1, x_i (1 \leq i \leq n)\} \subseteq \mathcal{M}$ then $\rho(f, \mathcal{M})$ for monotone f can be exponential in n. This is the circumstance that gave rise

to the successful use of approximations in the monotone case.

Remark 2.3 In the *ACC*-papers [Raz87], [Smo87], [Pat86] a "symmetrized" version of the method was used. It is obtained by coupling(2) and(3);(4) and(5);(6)and(7).

$$\delta_{\vee}(\bar{g},\bar{h}) = (\bar{g}\vee\bar{h})\oplus(\bar{g}\bar{\vee}\bar{h}), \quad (10)$$

$$\delta_{\&}(\bar{g},\bar{h}) = (\bar{g}\&\bar{h})\oplus(\bar{g}\bar{\&}\bar{h}); \quad (11)$$

$$\Delta = \{\delta_*(\bar{g},\bar{h})|\bar{g},\bar{h}\in\mathcal{M}; \quad (12)$$

$$* \in \{\&,\vee\}\}; \quad (13)$$

$$f\oplus\bar{f} \leq \vee_{i=1}^{t}\delta_{*_i}(\bar{g}_i,\bar{h}_i). \quad (14)$$

Denoting the corresponding distance by ρ_{sym} we clearly have $\rho_{sym}(f,\mathcal{M})\leq\wp(f,\mathcal{M})$ i.e. theorems 2.1 and 2.4 also hold for this symmetrized model.

In all the previous works lower bounds for $\rho(f,\mathcal{M})$ were extracted from the following source. Assume that we are given a pair of random n-bit strings $< v,u >$ where $v\in f^{-1}(1)[u\in f^{-1}(0)$ respectively] with probability 1. Set

$$d^+ = max_{\delta^+\in\Delta^+}P[\delta^+(v)=1]; \quad (15)$$

$$d^- = max_{\delta^-\in\Delta^-}P[\delta^-(u)=1]; \quad (16)$$

and

$$\wp(f,\mathcal{M},v,u) = min_{f\in\mathcal{M}}max\left(\frac{P[\bar{f}(v)=0]}{d^+},\frac{P[\bar{f}(u)=1]}{d^-}\right) \quad (17)$$

Let us check that for any (v,u) we have $\wp(f,\mathcal{M},v,u)\leq\wp(f,\mathcal{M})$. Indeed, it is sufficient to prove that for any $\bar{f}\in\mathcal{M}$ $max[\frac{P[\bar{f}(v)=0]}{d^+},\frac{P[\bar{f}(u)=1]}{d^-}]\leq\wp(f,\bar{f})$. So, we have to prove that $\frac{P[\bar{f}(v)=0]}{d^+}\leq\wp(f,\bar{f})$ and $\frac{P[\bar{f}(u)=1]}{d^-}\leq\wp(f,\bar{f})$. But from (6) we obtain

$$\begin{aligned}1 &= P[f(v)=1]\leq P[\bar{f}(v)=1]\\ &+ \wp(f,\bar{f})\cdot max_{1\leq i\leq t}P[\delta^+_{*_i}(\bar{g}_i,\bar{h}_i)(v)=1]\\ &\leq P[\bar{f}(v)=1]+\wp(f,\bar{f})\cdot d^+ \quad (18)\end{aligned}$$

which yields $\frac{P[\bar{f}(v)=0]}{d^+}\leq\wp(f,\bar{f})$. The second statement $\frac{P[\bar{f}(u)=1]}{d^-}\leq\wp(f,\bar{f})$ is proved similarly.

Example: Look from this point at lower bounds for the function $CLIQUE(m,s)$ (which tests whether an m-vertex graph contains an s-vertex complete subgraph) in the monotone case [Raz85b, AB87]. We obtain the random input v by picking at random an s-vertex complete subgraph and the input u by picking at random an $(s-1)$-partite complete subgraph. In order to estimate $\delta(f,\mathcal{M},v,u)$ from below we consider two cases in (17). If $\bar{f}=0$ then $P[\bar{f}(v)=0]=1$ and the first term under max in (17) is large. If $\bar{f}\neq 0$ then, using information about $\mathcal{M}$ and the fact $\bar{f}\in\mathcal{M}$, we can prove that $P[\bar{f}(u)=1]$ is large comparative to d^- therefore the second term in (17) is large. Surely, this is only the general idea and in order to put it into effect one needs sufficiently complicated combinatorial arguments (see [Raz85b, AB87]).

Theorem 2.4 For any legitimate model $\mathcal{M}$, boolean function $f\in F_n$ actually depending on $n_0(n_0\leq n)$ variables and distributions v,u such as those described above the inequality $\rho(f,\mathcal{M},v,u)\leq\wp(n_0)$ holds.

We see that theorem 2.1 does not exclude the possibility of proving good lower bounds for a Boolean function $f(x_1,x_2,...,x_{n_0})$ by introducing a large number of auxiliary variables $x_{n_0+1},x_{n_0+2},...,x_n$ followed by estimating the distance $\rho(f,\mathcal{M})$ for some suitable model $\mathcal{M}\subseteq F_n$. Let us describe some special kind of models which turn out to be good for this purpose. From now on assume that a function f whose complexity we want to estimate

is given. Set $U = f^{-1}(0), V = f^{-1}(1)$. Given $v \in V$, denote by $\Im_v$ the set of all monotone $(0-1)$-valued functionals F defined on subsets of U and satisfying

$$F(\emptyset) = 0, \quad F(U) = 1,$$
$$F(U \cap X_i^\varepsilon) = \quad v^i \oplus \varepsilon \oplus 1 \qquad (19)$$

Let $\Im = \bigcup_{v \in V} \Im_v$. The meaning of all these restrictions to a functional F is, as we shall see a little later, that they force (1) in the model $\mathcal{M}$ which we are constructing.

Given $\Im_0 \subseteq \Im$, construct a legitimate model $\mathcal{M}(\Im_0)$ as follows. First set $n = n_0 + \log_2 |\Im_0|$ and fix arbitrary surjective enumeration $\mu : B^{n-n_0} \to \Im_0$. We shall drop μ throughout and identify a Boolean string $y \in B^{n-n_0}$ with the functional $\mu(y) \in \Im$ which it encodes. By definition, $< \mathcal{M}(\Im_0), \bar{\vee}, \bar{\&} >$ is isomorphic (as algebraic system) to $< F_{n_0}, \vee, \& >$. $g \in F_{n_0}$ corresponds via this isomorphism to the function $\bar{g} \in F_n$ given by

$$\bar{g}(x, F) = \begin{cases} g(x), & x \neq v(F) \\ F(U_g), & x = v(F) \end{cases} \qquad (20)$$

where $|x| = n_0, v(F)$ is such that $F \in \Im_{v(F)}$ and $U_g = \{u \in U | g(u) = 1\}$. Then (19) just means $\bar{0} = 0, \bar{1} = 1, \bar{x}_i^\varepsilon = x_i^\varepsilon$, i.e. $\mathcal{M}(\Im_0)$ actually is a legitimate model. By the *maximal model* $\mathcal{M}_{max}$ we mean $\mathcal{M}(\Im)$. Let us emphasize that the construction of $\mathcal{M}_{max}$ *depends* on f.

It turns out that general concepts become especially simple in these specific models $\mathcal{M}(\Im_0)$. Namely, the only function from $\mathcal{M}$ which one can successfully use for approximating f is the $\bar{f}$ given by (20); three of four kinds of δ-functions (namely, $\delta_\vee^+, \delta_\vee^-$ and $\delta_\&^-$) vanish and we have to care only on $\delta_\&^+$; the only Boolean inputs we should consider are of the form $(v(F), F)$ and therefore there is the obvious one-to-one correspondence between these inputs and $\Im$ and so on (see the proof of the claim in Section 4). All these simplifications lead to the following combinatorial interpretation of $\mathcal{G}(f, \mathcal{M}(\Im_0))$. Say that a pair (A, B) of subsets of U *covers* some $F \in \Im$ if

$$F(A) = 1, F(B) = 1, F(A \cap B) = 0. \qquad (21)$$

Claim 2.5 $\rho(f, \mathcal{M}(\Im_0))$ equals to the minimal possible number of pairs (A, B) covering all members of $\Im_0$.

This Claim (as well as theorems 2.6, 2.7 below) will be proved in 4.

Theorem 2.6 For any boolean function $f_0 \in F_{n_0}$ such that $L(f) \geq \omega(n_0^3)$ and corresponding maximal model $\mathcal{M}_{max}$, we have $\rho(f, \mathcal{M}_{max}) \geq \Omega(L(f)^{\frac{1}{3}})$.

The number of auxiliary variables involved in $\mathcal{M}_{max}$ is $\exp(\Omega(|U|))$ where U is the set of inputs on which f outputs 1. Note that $L(f) = \mathcal{O}(n_0|U|)$ therefore $\mathcal{M}_{max}$ has exponentially more auxiliary variables than could be expected in view of theorem 2.1. Somehow this situation can be improved by the following result:

Theorem 2.7 For any boolean function $f \in F_{n_0}$ such that $L(f) \geq \omega(n_0^3)$ and $\Im$ defined as above there exists $\Im_0 \subseteq \Im$ such that

$$\rho(f, \mathcal{M}(\Im_0)) \geq \Omega(L(f)^{\frac{1}{3}}) \qquad (22)$$

and

$$|\Im_0| \leq exp(\mathcal{O}(L(f)|U|)). \qquad (23)$$

So, $\mathcal{M}(\Im_0)$ has $\mathcal{O}(L(f_0)|U|)$ variables.

But we can exhibit no explicit construction of $\Im_0$ satisfying these conditions.

3 Weakness of approximations.

In this section we prove theorems 2.1, 2.4. The main component of the proofs is the following lemma:

Lemma 3.1 Assume that we are given a boolean function $f \in F_n$ actually depending on $n_0 (n_0 \leq n)$ variables and a legitimate model $\mathcal{M}$. Then there exist random functions $\bar{f} \in \mathcal{M}$, $\delta^+ \in \Delta^+$ and $\delta^- \in \Delta^-$ such that for any $v \in f^{-1}(1)$

$$P[\bar{f}(v) = 0] = \mathcal{O}(n_0 \cdot P[\delta^+(v) = 1]) \qquad (24)$$

and, similarly, for any $u \in f^{-1}(0)$

$$P[\bar{f}(u) = 1] = \mathcal{O}(n_0 \cdot P[\delta^-(u) = 1]) \qquad (25)$$

First we deduce theorems 2.1 and 2.4 from the lemma; its own proof is postponed.

Proof of theorem 2.1. First note that $\rho(f, \mathcal{M}) \leq L(f) \leq \mathcal{O}(2^{n_0})$ therefore the theorem trivially holds it $n \geq 2^{n_0}$. So, we can assume

$$\log n \leq n_0. \qquad (26)$$

Let $f, \mathcal{M}, \bar{f}, \delta^+, \delta^-$ be as in lemma 3.1. Call an $v \in f^{-1}(1)$ $[u \in f^{-1}(0)]$ *bad* if $P[\bar{f}(v) = 0] \geq \frac{1}{3}$ $[P[\bar{f}(u) = 1] \geq \frac{1}{3}$ respectively] and *good* otherwise. Consider the (random) function $MAJ(\bar{f}_1, \bar{f}_2, ..., \bar{f}_r)$ where $\bar{f}_i$ are independent copies of $\bar{f}$. If an $v \in f^{-1}(1)$ is good then $P[MAJ(\bar{f}_1, \bar{f}_2, ..., \bar{f}_r)(v) = 0] = \exp(-\Omega(r))$ and similarly for a $u \in f^{-1}(0)$. Hence, taking $r = \mathcal{O}(n)$ (the constant being large enough) we get that the function $MAJ(\bar{f}_1, \bar{f}_2, ..., \bar{f}_r)$ coincides with f on all good inputs with a non-zero probability. Fix some $\bar{f}_1, \bar{f}_2, ..., \bar{f}_r$ with this property and set $f' = MAJ(\bar{f}_1, \bar{f}_2, ..., \bar{f}_r)$. By (9) and (26)

$$\mathcal{Y}(f', \mathcal{M}) = \mathcal{O}(n \log n) \leq \mathcal{O}(n_0 n) \qquad (27)$$

since $L_{mon}(MAJ(\bar{x}_1, \bar{x}_2, ..., \bar{x}_r)) = \mathcal{O}(r \log r)$ [AKS83]. It means that there exists and $\bar{f} \in \mathcal{M}$ such that those good inputs $v \in f^{-1}(1)$ for which $\bar{f}(v) = 0$ can be covered by $\mathcal{O}(n_0 n)$ functions from Δ^+ and similarly for $u \in f^{-1}(0)$.

Turning to bad inputs $v \in f^{-1}(1)$ we see that by (24) for each of them

$$P[\delta^+(v) = 1] = \Omega(1/n_0) \qquad (28)$$

Therefore, taking independent copies δ_1^+, δ_2^+,..., δ_s^+ of δ^+ we obtain $P[\bigvee_{i=1}^s \delta_i^+(v) = 0] = \exp(-\Omega(s/n_0))$. So, if $s = \mathcal{O}(n_0 n)$ (the constant being large enough) we conclude that (with non-zero probability) all bad inputs $v \in f^{-1}(1)$ are covered and similarly for $u \in f^{-1}(0)$. Hence $\rho(f, \bar{f}) \leq \mathcal{O}(n_0 n)$ and the proof of theorem 2.1 is completed.

Proof of theorem 2.4. Now we are additionally given two distributions v, u. Averaging (24) on v, we obtain $P[\bar{f}(v) = 0] = \mathcal{O}(n_0 \cdot P[\delta^+(v) = 1]) \leq \mathcal{O}(n_0 \cdot d^+)$. This implies $E_{(\bar{f})}[\frac{P_{(v)}[\bar{f}(v)=0]}{d^+}] = \mathcal{O}(n_0)$ (subscripts to E and P indicate under what distribution they are taken). Similarly from (25) we get $E_{(\bar{f})}[\frac{P_{(u)}[\bar{f}(u)=1]}{d^-}] = \mathcal{O}(n_0)$. By adding these two inequalities,

$$E_{(\bar{f})}\left[\frac{P_{(v)}[\bar{f}(v)=0]}{d^+} + \frac{P_{(u)}[\bar{f}(u)=1]}{d^-}\right] = \mathcal{O}(n_0). \qquad (29)$$

Taking as $\bar{f}$ in (17) such $\bar{f}$ that the value under $E_{(\bar{f})}$ in (29) is $\mathcal{O}(n_0)$, we obtain $\rho(f, \mathcal{M}, v, u) \leq \mathcal{O}(n_0)$. The proof is complete.

Proof of lemma 3.1. Assume for simplicity that $x_1, x_2, ..., x_{n_0}$ is the complete list of essential variables of f. Given a function $g(x_1, x_2, ..., x_d) (d \leq n_0)$ we define by induction on d a circuit C_g of exponential size such that $|C_g| = g$. If $d = 0$ then g is a constant and we let C_g consist of a single gate.

Assume $g(x_1, x_2, ..., x_d)(d > 0)$ is given. Set $g^\varepsilon = g(x_1, x_2, ..., x_{d-1}, \varepsilon)$. Put, by definition,

$$C_g = (C_{(g^0)} \& x_d^0) \vee (C_{(g^1)} \& x_d^1). \tag{30}$$

In other words, C_g is the random function corresponding to the universal decision tree for C_g. Denote by g_d the uniform distribution on F_d. Consider the following (random) circuit C which computes **EXCLUSIVE-OR** of g_{n_0} and $g_{n_0} \oplus f$ in a monotone manner:

$$\begin{aligned} C &= (C_{g_{n_0}} \& C_{g_{n_0} \oplus f \oplus 1}) \\ &\vee (C_{g_{n_0} \oplus 1} \& C_{g_{n_0} \oplus f}). \end{aligned} \tag{31}$$

Clearly, $|C| \equiv f$. Take $|\bar{C}|$ as $\bar{f}$. To construct the distributions δ^+ and δ^- first pick at random

$$\begin{aligned} d &\in \{1, 2, ..., n_0, \oplus\} \text{and} \\ t &\in \{0, 1, \vee\}. \end{aligned} \tag{32}$$

Then set:

$$\delta^o = \delta^o_{\&}(|\widetilde{C_{g_d}}|, x_d^t) \tag{33}$$

if

$$d \in \{1, 2, ..., n_0\}, \tag{34}$$
$$t \in \{0, 1\}, \tag{35}$$
$$o \in \{+, -\}, \tag{36}$$
$$\delta^o = \delta^o_{\vee}(|\widetilde{(C_{g_d^0} \& x_d^0)}|, |\widetilde{(C_{g_d^1} \& x_d^1)}|) \tag{37}$$

if

$$d \in \{1, 2, ..., n_0\}, \tag{38}$$
$$t = \vee(g_d^0 \text{ and } g_d^1 \text{ are independent copies of } g_d), \tag{39}$$
$$\delta^o = \delta^o_{\&}(|\widetilde{C_{g_{n_0} \oplus t}}|, |\widetilde{C_{g_{n_0} \oplus t \oplus f \oplus 1}}|) \tag{40}$$

if

$$d = \oplus, \tag{42}$$
$$t \in \{0, 1\}, \tag{43}$$
$$\delta^o = \delta^o_{\vee}(|\widetilde{(C_{g_{n_0}} \& C_{g_{n_0} \oplus f \oplus 1})}|, |\widetilde{(C_{g_{n_0} \oplus 1} \& C_{g_{n_0} \oplus f})}|) \tag{44}$$

if $d = \oplus, t = \vee$. In short, we have taken δ-functions corresponding to the computations (30),(31).

Fix $v \in f^{-1}(1)$. Recall that a gate of the circuit C has +- error on v if $g(v) = 1, \bar{g}(v) = 0$ where g and $\bar{g}$ are functions computed by this gate in circuits C and $\bar{C}$ respectively. Surely, if $f(v) = 0$ then there is at least one gate which has +- error on v whereas both its predecessors do not and this implies that $\delta^+(v) = 1$ for the function δ^+ corresponding to this gate. At first sight this observation is quite useless because the size of C is exponential in n_0. But it turns out (and this is the crucial point of the whole proof that does not have any analogies in the monotone case!) that it is possible to distinguish a small set of gates (with size only $O(n_0)$) such that one of these distinguished gates satisfies this +- error property.

To be more precise let

$$g_d^v = g(x_1, x_2, ..., x_d, v^{d+1}, ..., v^{n_0}). \tag{45}$$

All gates $C_{(g_d^v)}$ of the form (30) $(1 \le d \le n_0)$ as well as all other intermediate gates in (30) for such g_d^v will be called *v-gates* of $C_g(g \in F_n)$. *v-gates* of the circuit (31) are the v-gates of the four subcircuits it consists of as well as the results of the three intermediate computations in (31). Clearly the whole number of v-gates in C is $O(n_0)$.

Statement 3.2 If $|\bar{C}|(v) = 0$ then there exists at least one v-gate such that the function $\delta^+_*(\bar{g}_1, \bar{g}_2)$ corresponding to this gate outputs 1 on v. The dual statement for - -errors also holds.

Proof of Statement. Assume that for some $g_{n_0} \in F_{n_0}$ the function $|\bar{C}|$ where C is given by (31) outputs 0 on v. If all four subcircuits in (31) do not have +- errors on v then at least one of three v-gates in (31) has +- error on v whereas both its predecessors do not. The statement (just for this v-gate) follows automatically.

So, we can assume that one of the four subcircuits in (31) (say, $C_{g_{n_0}}$) has +- error on v. Following through the v-path in the underlying tree, we find some d such that $C_{(g^v_d)}$has +- error on v whereas $C_{(g^v_{d-1})}$ does not. Let $g'_{d-1} = g^v_d(x_1, x_2, ..., x_{d-1}, x_d \oplus 1)$. The point is that the $\&$-gate

$$C_{(g'_{d-1})} \& x_d^{v_d \oplus 1(21)} \tag{46}$$

... *never* has +- error on v because it outputs 0 on v. This allows us to find a gate in (30) which has +- error on v whereas both its predecessors do not and to finish the proof of Statement for +- errors. Let us remark that for +- errors we did not use δ^+-functions corresponding to the gates (46) at all. They appear in the proof of the statement for - - errors because in that case the gates (46) (with replacing v by u) *can* have - - errors on u. But if the gate (46) has - - error on u then the corresponding δ^--function outputs 1 on u independently of whatever $\bar{C}_{g'_{d-1}}$ is. All other arguments are carried over the case of - - errors without any changes. The statement is proved.

The statement directly implies (24) and (25) because all the v-gates of the random circuit C can be, in a natural way, enumerated and supplied with a type (d,t) such that the distribution $\delta^+_*(\bar{g}_1, \bar{g}_2)$ corresponding to a gate with number i and type (d,t) coincides with the conditional distribution, the condition being just "(d,t) is the type obtained in drawing $(d,t) \in \{1,2,...,n_0\} \times \{0,1,\vee\}$". The probability of this condition is $\Omega(1/n_0)$.

The proof of lemma 3.1 is complete.

4 Power of approximations.

This section is devoted to proofs of the theorems 2.6-2.7. Let us remember that in section 2 we assigned to a Boolean function f (assumed to be fixed throughout) some set of functionals $\Im$ and to any $\Im_0 \subseteq \Im$ assigned some legitimate model $\mathcal{M}(\Im_0)$. Here we start with proving the claim from section 2 giving a combinatorial interpretation of $\rho(f, \mathcal{M}(\Im_0))$.

Claim 4.1 $\rho(f, \mathcal{M}(\Im_0))$ equals to the minimal possible number of pairs $(A,B)(A,B \subseteq \mathcal{V} = f^{-1}(\sigma))$ covering all members of $\Im_0$.

Proof.

a) Assume that $(A_i, B_i)(1 \le i \le t)$ cover all functionals from $\Im_0$. We have to construct $\bar{f}, \delta^o_1, ..., \delta^o_t \in \Delta^o (o \in \{+,-\})$ such that (6), and (7) hold for f. Take $\bar{f}$ accordingly to (20). Note that $f(x)$ can differ from $\bar{f}(x,F)$ only if $x \in V$ i.e. $f(x) = 1$ therefore $f \ge \bar{f}$ and (7) holds automatically. To get (6), set

$$\delta^+_i = \delta^+_{\&}(\bar{g}_i, \bar{h}_i) \tag{47}$$

where $\bar{g}_i, \bar{h}_i \in \mathcal{M}(\Im_0)$ correspond via (20) to functions $g_i, h_i \in F_{n_0} (1 \le i \le t)$ which output 1 just on the sets of inputs A_i, B_i respectively. If $f(v) = 0$ whereas $\bar{f}(x,F) = 1$ then $x \in V$ and $F \in \Im_x$ (see (20)). Choose i such that (A_i, B_i) covers F. It is easy to check that $\delta^+_i(x,F) = 1$. So, (6) also holds and we are done in one direction.

b) Assume now $f' \in F_{n_0}$ is taken in such a way that $\rho(f, \mathcal{M}(\Im_0)) = \rho(f, \bar{f}') = t$. By (20), $\bar{g}(u,F) = g(u)$ for any $g \in F_{n_0}$ and $u \in U$. Therefore any element of Δ^- always outputs 0 on U and (7) implies that f' outputs 0 on U i.e. $f' \le f$. Moreover,

the mapping $g \to \bar{g}$ is monotone, hence, in view of satisfying (6), we can put $f' = f$. The same observation about the monotonicity implies $\delta^+_\vee(\bar{g}, \bar{h}) = 0$ for all $\bar{g}, \bar{h}$. So, we have t functions of the form (47) satisfying (6). Set $A_i = U \cap g_i^{-1}(1), B_i = U \cap h_i^{-1}(1)$. Arguments reversed to those used above show that $(A_i, B_i)(1 \le i \le t)$ cover all functionals from $\Im_0$. The Claim is proved.

Now we turn directly to proofs of the theorems 2.6-2.7.

Proof of theorem 2.6. By the Claim, we are given a set $(A_i, B_i)(1 \le i \le t, t = \rho(f, \mathcal{M}_{max}))$ covering all functionals from $\Im$ and have to design a circuit of size $\mathcal{O}(t^3)$ computing f. Note that it is sufficient to do with size $\mathcal{O}((t+n_0)^3)$ because then the assumption $L(f) = \omega(n_0^3)$ would give us $t = \omega(n_0)$.

Extending if necessary the list (A_i, B_i) we may suppose that all the pairs $(U \cap X_i^0, U \cap X_i^1)$ occur among (A_i, B_i). Set $S = \{A_i\} \cup \{B_i\} \cup \{A_i \cap B_i\}$. Observe that $|S| = \mathcal{O}(t+n_0)$. Consider the following rules of inference on S:

$$A_i, \quad B_i \vdash A_i \cap B_i, \tag{48}$$

$$C \vdash D, \qquad \text{if } C \subseteq D \tag{49}$$

(cf. the rule used in [Raz 85b, Raz 85a, AE 85, AL 85, Tar 89, And 87] when $r = 2$). These rules define in the usual way notions of the *closure* of a $S_0 \subseteq S$ (denoted by $cl(S_0)$) and a *closed* subset of S.

Statement 4.2 Given $w \in B^{n_0}$, $f_0(w) =$ 1 iff

$$\emptyset \in cl(\{U \cap X_i^{(w^i)} | 1 \le i \le n_0\}). \tag{50}$$

Proof. ($\Rightarrow$) Assume $w \in V$. Denote $cl(\{U \cap X_i^{(w^i)} | 1 \le i \le n_0\})$ by S_0. Suppose that $\emptyset \notin S_0$. Define a monotone functional F on subsets of U by $F(D) = 1 \equiv \exists C \subseteq D (C \in S_0)$. All properties in (19) for this F are clear except perhaps $F(U \cap X_i^{(w^i \oplus 1)}) = 0$. But if there existed some $C \subseteq U \cap X_i^{(w^i \oplus 1)}$ such that $C \in S_0$, then $C \vdash U \cap X_i^{(w^i \oplus 1)}$ would imply $U \cap X_i^{(w^i \oplus 1)} \in S_0$, and $U \cap X_i^{(w^i \oplus 1)}, U \cap X_i^{(w^i)} \vdash \emptyset$ would imply $\emptyset \in S_0$, in contradiction with our assumption. Therefore (19) holds and $F \in \Im_w$. Clearly, F is covered by none of the pairs (A_i, B_i). The contradiction proves that $\emptyset \in S_0$.

($\Leftarrow$) Assume $w \in U$. Then $\{C \in S | w \in C\}$ is closed, contains all the sets $U \cap X_i^{(w^i)}$, but does not contain $\emptyset$. Hence $\emptyset \notin S_0$.

Now it is quite obvious how to prove theorem 2.6; we only have to recognize the property (50) by a small circuit. It is done directly: given an input w let $f_{D,k}(D \in S, 1 \le k \le |S|)$ mean that D can be deduced from $\{(U \cap X_i^{(w^i)}) | 1 \le i \le n_0\}$ within k steps. Then

$$f_{D,0} = \begin{cases} x_i^\epsilon & \text{if } D = U \cap X_i^\epsilon \\ 0 & \text{otherwise} \end{cases} \tag{51}$$

$$f_{D,k+1} = \begin{cases} (f_{A_i,k} \& f_{B_i,k}) \vee \bigvee_{C \subseteq D} (f_{C,k}) \\ \text{if } D = A_i \cap B_i \\ \bigvee_{C \subseteq D} (f_{C,k}) \text{ otherwise} \end{cases} \tag{52}$$

This circuit has size $\mathcal{O}(|S|^3) = \mathcal{O}((t+n_0)^3)$ and $f = f_{\emptyset,|S|}$ by the Statement. The proof of theorem 2.6 is completed.

Proof of theorem 2.7. We only have to note that, in order to make the proof of theorem 2.6 work, it is sufficient to consider only those functionals which appear in the proof of the Statement (part ($\Rightarrow$)) instead of the whole $\Im$. The number of these functionals does not exceed the number of possible choices of the system $(A_i, B_i)(1 \le i \le t, t = \rho(f, \mathcal{M}_{max}))$ multiplied by $|V|$. This is $exp(\mathcal{O}(t|U| + n_0)) \le exp(\mathcal{O}(L(f)|U|))$ because $L(f) \ge max(t, n_0)$. The theorem is proved.

Remark 4.3 If we consider the measure $L^\&(f)$ (the minimal possible number of &-gates over all circuits computing f; $\vee$-gates are free) rather

than $L(f)$ then we can slightly improve theorem 2.6. Namely,

$$\Omega(L^{\&}(f)^{\frac{1}{2}}) \leq \rho(f, \mathcal{M}_{max}) \leq L^{\&}(f). \quad (53)$$

The right-hand inequality follows from the Claim and its proof ($\delta_\vee$-functions do not contribute to the list (A_i, B_i)) whereas the left is obtained from the estimate $O(t^2)$ of the number of &-gates in the circuit above.

Surely, the same observation can be also applied to theorem 2.7.

5 Concluding Remarks.

We have seen above that if all variables involved in the approximating model are essential variables of the function whose complexity we are estimating then the method of approximations is practically useless for handling arbitrary circuits (theorems 2.1 and 2.4). On the other hand, theorem 2.6 states that the complexity $L(f)$, up to polynomial, can be regarded as the output of an explicitly given instance of the **MINIMAL COVERING**, namely: *"How many pairs (A,B) does one need to cover all functionals from* $\Im$*?* This instance is extremely singular, in the sense that there exists a distribution on covering sets (provided by lemma 3.1) under which any point is covered with probability $\Omega(\frac{1}{n_0})$. In particular, direct probabilistic methods appear to be useless in dealing with this instance. It is this circumstance which probably will be a big obstacle to resolving this instance or, more modestly, to proving nonlinear lower bounds.

6 Acknowledgements

I am thankful to Faith Fich for many helpful suggestions and remarks. Also, my thanks are due to M. S. Paterson for many remarks concerning the English of this paper and to M. Sipser and C. Brownlie for assistance in preparing the manuscript.

References

[AB87] N. Alon and R. Boppana. The monotone circuit complexity of boolean functions. *Combinatorica*, 7(1):1–22, 1987.

[AE85] Andreev A. E. On a method for obtaining lower bounds for the complexity of individual monotone functions. *Doklady Akademii Nauk*, 282(5):1033–1037, 1985. English translation in: *Soviet Mathematics Doklady* 31:3, 530-534.

[AKS83] M. Ajtai, J. Komlos, and E. Szemeredi. An $0(n \log n)$ sorting network. In *Proceedings of the* 15^{th} *Annual ACM Symposium on Theory of Computing, Boston, Massachusetts*, pages 1–9, ACM SIGACT, ACM, 1983.

[And87] A. E. Andreev. On one method of obtaining constructive lower bounds for the monotone circuit size. *Algebra and Logics*, 26(1):3–26, 1987.

[BS88] R. Boppana and M. Sipser. Complexity of finite functions. Cambridge, USA, 1988. preprint.

[Pat86] M. Paterson. Bounded depth circuits over $\{\&, \oplus\}$. Warwick, Britain, 1986. preprint.

[Raz85a] A. A. Razborov. A lower bound on the monotone network complexity of the logical permanent. *Mathematicheskie Zametki*, 37(6):887–900, 1985. English translation in: *Mathematical Notes of the Academy of Sciences of the USSR* 37:6, 485-493.

[Raz85b] A. A. Razborov. Lower bounds on the monotone complexity of some boolean functions. *Doklady Akademii Nauk*, 281(4):798–801, 1985. English translation in: *Soviet Mathematics Doklady* 31, 354-357.

[Raz86] A. A. Razborov. Lower bounds for the monotone complexity of boolean functions. In *Proceedings of the International Congress of Mathematicians*, pages 1478–1487, Berkeley, California, 1986.

[Raz87] A. A. Razborov. Lower bounds on the size of bounded depth networks over a complete basis with logical addition. *Mathematicheskie Zametki*, 41(4):598–607, 1987. English translation in: *Mathematical Notes of the Academy of Sciences of the USSR* 41:4, 333-338.

[Smo87] R. Smolensky. Algebraic methods in the theory of lower bounds for boolean circuit complexity. In *Proceedings of the* 19^{th} *Annual ACM Symposium on Theory of Computing, New York City*, pages 77–82, ACM SIGACT, ACM, 1987.

[Tar89] Eva Tardos. The gap between monotone and non-monotone circuit complexity is exponenial. 1989. to appear in *Combinatorica*.