Computing and Informatics, Vol. 27, 2008, 627–638

AVERAGE DEGREE IN THE INTERVAL GRAPH OF A RANDOM BOOLEAN FUNCTION

Eduard Toman, Daniel Olejár, Martin Stanek

Department of Computer Science
Faculty of Mathematics, Physics and Informatics
Comenius University
Mlynská dolina, 842 48 Bratislava, Slovak Republic
e-mail: {toman, olejar, stanek}@dcs.fmph.uniba.sk

Manuscript received 18 May 2006; revised 16 February 2007
Communicated by Vladimír Kvasnička

Abstract. We consider an n-ary random Boolean function f such that $\Pr[f(\tilde{\alpha}) = 1] = p$ for $\tilde{\alpha} \in \{0,1\}^n$ and study its geometric model, the so called interval graph. The interval graph of a Boolean function was introduced by Sapozhenko and has been used in construction of schemes realizing Boolean functions. Using this model, we estimate the number of maximal intervals intersecting a given maximal interval of a random Boolean function and prove that the asymptotic bound on the logarithm of the number is $(1 + \varphi(n)) \log_2 \log_{1/p} n \cdot \log_2 n$, where $\varphi(n) \to 0$ as $n \to \infty$.

Keywords: Random Boolean function, interval graph

1 PRELIMINARIES

Local algorithms form an important subclass of algorithms for construction of optimal schemes. The main idea of local algorithms is simple: they introduce a metric on the "space" of all elements (building blocks of schemes) and a "measure of quality" of elements. Then for every element of the scheme under construction they analyze its neighbouring elements, searching for better ones; if such elements exist, local algorithms chose the best of them and substitute the original element. The whole procedure is repeated until no replacement/improvement is possible.

Zhuravlev [12] studied the use of local algorithms in the minimization of disjunctive normal form (d.n.f.). He introduced the notion of a *conjunction neighbourhood*

and proved that the optimal d.n.f. cannot be constructed in general by means of local algorithms based on finite (local) conjunction neighbourhoods. Though it is impossible to find an optimal solution by means of local algorithms, we can construct a sub-optimal d.n.f. by analyzing the first order neighbourhoods in almost all d.n.f.'s of a given Boolean function.

We shall use the standard notation of Boolean function theory [12] and therefore we introduce only the notions and notation necessary for understanding the paper. Boolean variables and their negations are called *literals*. The literal of a variable x will be denoted by x^α, $(\alpha \in \{0,1\})$, where

$$x^\alpha = \begin{cases} x & \text{if } \alpha = 1, \\ \neg x & \text{if } \alpha = 0. \end{cases}$$

A conjunction $K = x_{i_1}^{\alpha_{i_1}} \dots x_{i_r}^{\alpha_{i_r}}$ of literals of different variables is called an *elementary conjunction*. The number of literals (r) in a conjunction K is called the *rank* of K. A special case is the conjunction of rank 0; it is called *empty* and its value is set to 1.

A formula $D = K_1 \vee \dots \vee K_m$, the disjunction of distinct elementary conjunctions is called a *disjunctive normal form*. The parameter m (the number of elementary conjunctions in D) is called the *length* of D. The d.n.f. with $m = 0$ is called *empty* and its value is 0. A d.n.f. D *represents* a Boolean function f if the truth tables of f and D coincide. Let us consider the class of all d.n.f.'s representing an n-ary Boolean function f; the d.n.f. with the minimal number of literals in this class is called a *minimal d.n.f. of* f, and the d.n.f. with the minimal length (in this class) is called a *shortest d.n.f. of* f.

We use a geometric representation of Boolean functions. The Boolean n-cube B^n is a graph B^n with 2^n vertices $\tilde{\alpha} = (\alpha_1, \dots, \alpha_n); \alpha_i \in \{0,1\}$ where edges are joining those pairs of vertices which differ in exactly one coordinate. For an n-ary Boolean function f let N_f denote the subset of vertices $\{(\alpha_1, \dots, \alpha_n);\ f(\alpha_1, \dots, \alpha_n) = 1\}$. As can be easily seen, there is a one-to-one correspondence between the sets N_f and Boolean functions f. The set of vertices $N_K \subseteq \{0,1\}^n$ corresponding to an elementary conjunction K od rank r is called *the interval of order* r. Obviously, to every elementary conjunction $K = x_{i_1}^{\alpha_{i_1}} \dots x_{i_r}^{\alpha_{i_r}}$ corresponds an interval of order r consisting of all vertices $(\beta_1, \dots, \beta_n)$ of B^n, such that $\beta_{i_j} = \alpha_{i_j}$ for $j = 1, \dots, r$ (the values of other vertex coordinates can be chosen arbitrarily). Consequently, every vertex of B^n represents an interval of order n and the vertex set of B^n itself corresponds to the interval of order 0. In the geometric model, every interval of order r represents an $(n-r)$-dimensional subcube of B^n . An interval N_K is called a maximal interval of a Boolean function f if $N_K \subseteq N_f$ and there is no such interval $N_{K'} \subseteq N_f$ such that $N_K \subseteq N_{K'}$. For every elementary conjunction K from the d.n.f. D the *neighbourhood* of K of the first order (with respect to the d.n.f. D) is defined as the set of all elementary conjunctions K_j from D, such that (in algebraic notation) $K \wedge K_j \not\equiv 0$ or (in our geometric model) $N_K \bigcap N_{K_j} \neq \emptyset$. (Since we study mainly the neighbourhoods of the first order in this paper, the notion "neighbourhood" will

denote the neighbourhood of the first order.) For an arbitrary Boolean function f and each of its d.n.f.'s $K_1 \vee K_2 \vee \cdots \vee K_m$ we have that

$$N_f = \bigcup_{j=1}^{m} N_{K_j}.$$

In other words, every d.n.f. of a Boolean function f corresponds to a covering of N_f by intervals $N_{K_1}, \ldots, N_{K_m}$ such that $N_{K_j} \subseteq N_f$. Conversely, every covering of N_f by intervals $N_{K_1}, \ldots, N_{K_m}$ contained in N_f corresponds to some d.n.f. of f. Using the geometric interpretation of d.n.f.'s, we can express the "irreducibility" of d.n.f.: the d.n.f. D of a Boolean function f cannot be simplified if and only if every interval N_K of the covering, corresponding to D contains at least one vertex belonging to just one interval of the covering. Let r_j denote the order of an interval N_{K_j}. Then the number of literals in a d.n.f. is $r = \sum_{j=1}^{m} r_j$ and the construction of the minimal d.n.f. can be formulated in the geometric model as a problem of constructing a covering of N_f by intervals $N_{K_j} \subseteq N_f$ with minimal r. On the other hand, the construction of the covering corresponding to a shortest d.n.f. requires to minimize the number of intervals in a covering of N_f. Various metrical parameters of "typical" Boolean functions have been studied in the context of Boolean functions minimization in the class of d.n.f.'s [5, 6, 7, 8, 9, 10]. We introduce a more general model of Boolean functions, a concept of *random Boolean function*, now. A random Boolean function is defined on vertices of the Boolean n-cube in the following way:

$$f(\alpha_1, \ldots, \alpha_n) = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } 1-p, \end{cases}$$

where the value $f(\tilde{\alpha})$ does not depend on the values which the Boolean function f attains on other vertices. Recall that the set N_f contains all vertices $\tilde{\alpha} \in \{0,1\}^n$ with $f(\tilde{\alpha}) = 1$. The subgraph of the Boolean n-cube induced by the set N_f is called *the graph of* f and will be denoted by $G(f)$. The probability that the graph $G(f)$ of random Boolean function f is realized by a subgraph G of the Boolean n-cube is

$$\Pr[G(f) = G] = p^m \cdot (1-p)^{2^n - m},$$

where m denotes the number of vertices in G. Let A be a certain property that a Boolean function may or may not have. If

$$\lim_{n\to\infty} \Pr[f \text{ has the property } A] = 1,$$

we say that a random Boolean function has the property A almost surely. Sapozhenko [5] studied some parameters of the graph $G(f)$ for $p = \frac{1}{2}$ associated with local minimization algorithms. Some of his results were generalized and improved by Toman in [8] for $0 < p \leq 1$. In [5] Sapozhenko introduced and studied the interval graph of a Boolean function f. The *interval graph* $\Gamma(f)$ is a graph associated with

a Boolean function f; its vertices correspond to maximal intervals of f and the vertices corresponding to intervals N_{K_i} and N_{K_j} are joined by an edge in $\Gamma(f)$ if and only if $K_i \wedge K_j \not\equiv 0$, or equivalently, $N_{K_i} \bigcap N_{K_j} \neq \emptyset$. We study the degree of a vertex in $\Gamma(f)$; namely we estimate the lower and upper bounds of this parameter. The main result of the paper is an asymptotic bound on the logarithm of the vertex degree of a interval graph $\Gamma(f)$. This bound enables us to estimate both the accuracy of the obtained results and the computational complexity of local algorithms for the minimization of d.n.f.'s.

Škoviera [6, 7], Glagoliev [2] and Weber [10] studied some properties of random Boolean functions. They used combinatorial-probabilistic methods, considering metric parameters of Boolean functions as random variables, estimated the expectations and variances of these variables and finally they estimated their values by means of Markov's and Chebyshev's inequalities. The same approach will be used in the present paper. Let X be a random variable and let the symbols $\mathrm{E}(X)$ and $\mathrm{Var}(X) = \mathrm{E}(X - \mathrm{E}(X))^2$ denote the expectation and variance of a random variable X, respectively. (We only use nonnegative random variables in the present paper.)

2 THE SIZE OF THE NIEGHBOUROOD OF A GIVEN MAXIMAL INTERVAL

Let $X_{n,k}^{\tilde{\alpha}}$ be a random variable denoting the number of $k-$dimensional intervals containing a fixed vertex $\tilde{\alpha}$.

Lemma 1.

$$\mathrm{E}(X_{n,k}^{\tilde{\alpha}}) = \binom{n}{k} \cdot p^{2^k}. \tag{1}$$

Proof. Let f be an n-ary random Boolean function. For every k-dimensional subcube (interval) N_K of the n-cube B^n we introduce a random variable $\eta_K(f)$ (called an indicator) defined as follows:

$$\eta_K(f) = \begin{cases} 1 & \text{if } N_K \subseteq N_f, \\ 0 & \text{otherwise.} \end{cases}$$

Obviously, the random variable $X_{n,k}^{\tilde{\alpha}}$ is the sum of all indicators:

$$X_{n,k}^{\tilde{\alpha}} = \sum_{N_K} \eta_K(f),$$

where the summation ranges over all k-dimensional subcubes of B^n. The Boolean function f attains the value 1 on all vertices of an k-dimensional interval N_K with probability

$$\Pr[N_K \subseteq N_f] = p^{2^k} = E(\eta_K(f)).$$

There are $\binom{n}{k}$ k-dimensional subcubes (intervals) in B^n containing a fixed vertex $\tilde{\alpha}$. Consequently,

$$\mathrm{E}(X_{n,k}^{\tilde{\alpha}}) = \binom{n}{k} \cdot p^{2^k}.$$

□

Now we estimate the variance of the random variable $X_{n,k}^{\tilde{\alpha}}$.

Lemma 2.

$$\mathrm{Var}\left(X_{n,k}^{\tilde{\alpha}}\right) \le \binom{n}{k}^2 p^{2^{k+1}} \left[\frac{k^3}{np^2} + \frac{k}{p^{2^k}\binom{n}{k}}\right].$$

Proof. We express the expectation of the random variable $(X_{n,k}^{\tilde{\alpha}})^2$ and then we compute its variance. Let N_K and $N_{K'}$ be two k-dimensional subcubes of B^n containing a vertex $\tilde{\alpha}$. To abbreviate the notation, we denote the probability that both subcubes N_K and $N_{K'}$ belong to N_f by the symbol $P_{n,k}(N_K, N_{K'})$. Since both N_K and $N_{K'}$ contain the vertex $\tilde{\alpha}$, they have a nonempty intersection, an interval of dimension j; $0 \le j \le k$. Therefore

$$\mathrm{E}((X_{n,k}^{\tilde{\alpha}})^2) = \sum_{N_K, N_{K'}} P_{n,k}(N_K, N_{K'}) = \sum_{j=0}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} p^{2^{k+1}-2^j},$$

and

$$\mathrm{Var}(X_{n,k}^{\tilde{\alpha}}) = \sum_{j=0}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} - \binom{n}{k}^2 p^{2^{k+1}}.$$

Now we can derive an upper bound on the variance.

$$\begin{aligned}
\mathrm{Var}\left(X_{n,k}^{\tilde{\alpha}}\right) &= \sum_{j=0}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} - \binom{n}{k}^2 p^{2^{k+1}} \\
&= \binom{n}{k}\binom{n-k}{k} p^{2^{k+1}-1} + \sum_{j=1}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} - \binom{n}{k}^2 p^{2^{k+1}} \\
&\le \binom{n}{k}^2 p^{2^{k+1}} - \binom{n}{k}^2 p^{2^{k+1}} + \sum_{j=1}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} \\
&= \sum_{j=1}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} = \binom{n}{k} p^{2^{k+1}} \sum_{j=1}^{k} \binom{k}{j}\binom{n-k}{k-j} p^{-2^j}.
\end{aligned}$$

We denote $\binom{k}{j}\binom{n-k}{k-j}p^{-2^j}$ by b_j and estimate the ratio

$$\frac{b_{j+1}}{b_j} = \frac{p^{-2^j}(k-j)^2}{(j+1)(n-2k+j+1)}.$$

Since

$$\frac{b_{j+1}}{b_j} : \begin{cases} < 1 & \text{if } j < \lfloor \log_2 \log_{1/p} n \rfloor, \\ > 1 & \text{otherwise,} \end{cases}$$

the maximal value of b_j is b_1 or b_k. Therefore

$$\sum_{j=1}^{k} b_j \leq k(b_1 + b_k) \leq k \left(k \binom{n-k}{k-1} p^{-2} + p^{-2^k} \right),$$

and

$$\mathrm{Var}(X_{n,k}^{\tilde{\alpha}}) \leq \binom{n}{k} p^{2^{k+1}} \left(k^2 \binom{n-k}{k-1} p^{-2} + k p^{-2^k} \right) \leq \binom{n}{k}^2 p^{2^{k+1}} \left[\frac{k^3}{np^2} + \frac{k}{p^{2^k} \binom{n}{k}} \right].$$

□

Škoviera [6] proved that dimension k of a maximal interval of a random Boolean function satisfies the following inequalities:

$$\lfloor \log_2 \log_{1/p} n \rfloor \leq k \leq \lfloor \log_2 \log_{1/p} n + \log_2 \log_2 \log_{1/p} n \rfloor + 1. \tag{2}$$

Corollary 1. Let k be an integer satisfying (2). Then

$$\mathrm{Var}(X_{n,k}^{\tilde{\alpha}}) \leq \mathrm{E}(X_{n,k}^{\tilde{\alpha}}) \frac{c_1 \log_{1/p} n}{n},$$

where c_1 is a positive constant.

Let $Y_{n,k}^{\tilde{\alpha}}$ be the random variable, expressing the number of k-dimensional maximal intervals containing a vertex $\tilde{\alpha}$ and let $E\left(Y_{n,k}^{\tilde{\alpha}}\right)$ be its expectation. Then we have

Lemma 3.

$$\mathrm{E}(Y_{n,k}^{\tilde{\alpha}}) = \binom{n}{k} p^{2^k} \left(1 - p^{2^k}\right)^{n-k}.$$

Proof. Let $P(N_K)$ denote the probability that a fixed maximal interval N_K containing a vertex $\tilde{\alpha}$ belongs to the set N_f of a random Boolean function f. Obviously, $P(N_K) = p^{2^k} \left(1 - p^{2^k}\right)^{n-k}$ and the number of such intervals is $\binom{n}{k}$. Combining these two facts completes the proof of our lemma. □

Now we shall estimate the variance $\mathrm{Var}(Y_{n,k}^{\tilde{\alpha}})$.

Lemma 4. Let k be an integer satisfying (2). Then

$$\mathrm{Var}(Y_{n,k}^{\tilde{\alpha}}) \leq c_2 \frac{\log_{1/p} n}{n} \mathrm{E}^2(X_{n,k}^{\tilde{\alpha}}),$$

where c_2 is a positive constant.

Proof. Using Lemma 3 we have

$$\mathrm{Var}(Y_{n,k}^{\tilde{\alpha}}) = \sum_{j=0}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} P'_j(N_{K_s}, N_{K_t}) - \left(\binom{n}{k} p^{2^k}(1-p^{2^k})^{n-k}\right)^2,$$

where $P'_j(N_{K_s}, N_{K_t})$ denotes the probability that a random Boolean function contains k-dimensional maximal intervals N_{K_s} and N_{K_t}, both containing a fixed vertex $\tilde{\alpha}$ and $N_{K_s} \bigcap N_{K_t}$ is a j-dimensional interval. The probability $P'_j(N_{K_s}, N_{K_t})$ can be estimated in the following way:

$$P'_j(N_{K_s}, N_{K_t}) \le P_j(N_{K_s}, N_{K_t}) = p^{2^{k+1}-2^j},$$

where $P_j(N_{K_s}, N_{K_t})$ denotes the probability that a random Boolean function contains k-dimensional intervals N_{K_s} and N_{K_t}, both containing the fixed vertex $\tilde{\alpha}$ and $N_{K_s} \bigcap N_{K_t}$ is a j-dimensional interval. We have

$$\begin{aligned}
&\mathrm{Var}(Y_{n,k}^{\tilde{\alpha}}) \le \sum_{j=0}^{k} \binom{n}{j}\binom{n-j}{k-j}\binom{n-k}{k-j} P_j(N_{K_s}, N_{K_t}) - \left(\binom{n}{k} p^{2^k}(1-p^{2^k})^{n-k}\right)^2 \\
&\le \binom{n}{k}\binom{n-k}{k} p^{2^{k+1}-1} + \binom{n}{k}\sum_{j=1}^{k}\binom{k}{j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} - \left(\binom{n}{k} p^{2^k}(1-p^{2^k})^{n-k}\right)^2 \\
&\le \left(\binom{n}{k} p^{2^k}\right)^2 - \left(\binom{n}{k} p^{2^k}(1-p^{2^k})^{n-k}\right)^2 + \binom{n}{k}\sum_{j=1}^{k}\binom{k}{j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} \\
&\le \left(\binom{n}{k} p^{2^k}\right)^2 \left(1-(1-p^{2^k})^{2(n-k)}\right) + \binom{n}{k}\sum_{j=1}^{k}\binom{k}{j}\binom{n-k}{k-j} p^{2^{k+1}-2^j} \\
&\le \mathrm{E}^2\left(X_{n,k}^{\tilde{\alpha}}\right)\left(1-(1-\frac{1}{n})^2\right) + c_1 \frac{\log_{1/p} n}{n}\mathrm{E}^2\left(X_{n,k}^{\tilde{\alpha}}\right) \\
&\le c_2 \frac{\log_{1/p} n}{n}\mathrm{E}^2\left(X_{n,k}^{\tilde{\alpha}}\right).
\end{aligned}$$

□

Definition 1. A vertex $\tilde{\alpha} \in N_f$, satisfying the condition

$$|(Y_{n,k}^{\tilde{\alpha}} - \mathrm{E}(Y_{n,k}^{\tilde{\alpha}})| \ge \frac{1}{\log_{1/p} n}\mathrm{E}(X^{\tilde{\alpha}_{n,k}})$$

will be called a *bad vertex* of random Boolean function f, otherwise, the vertex $\tilde{\alpha}$ will be called a *good vertex* of random Boolean function f.

The following lemma is a direct consequence of Chebyshev's inequality and Lemma 4.

Lemma 5. Let k be an integer satisfying (2), and let $P_n(\tilde{\alpha})$ be the probability that $\tilde{\alpha}$ is a bad vertex of an n-ary random Boolean function. Then

$$P_n(\tilde{\alpha}) \leq \frac{c_2 \log_{1/p}^3 n}{n}.$$

Now we estimate the number of bad vertices in a random Boolean function.

Lemma 6. Let $b_k(f)$ be a random variable expressing the number of bad vertices of a random Boolean function f. Let δ_n be the probability that

$$b_k(f) \leq 2^n \frac{\log_{1/p} n}{n}.$$

Then $\delta_n \geq 1 - c_3/\log_{1/p} n$, where c_3 is a positive constant.

Proof. A vertex $\tilde{\alpha}$ is a bad vertex of an n-ary random Boolean function f with probability $P_n(\tilde{\alpha})$. The expectation of the number of bad vertices in f is

$$\mathrm{E}(b_k(n)) = P_n(\tilde{\alpha}) \cdot 2^n \leq \frac{c_2 \log_{1/p}^3 n}{n} \cdot 2^n.$$

It follows from the Markov's inequality that the probability that

$$b_k(f) \leq \frac{\log_{1/p}^4 n}{n} \cdot 2^n$$

is at least $1 - c_3/\log_{1/p} n$. □

Let N_K denote a fixed maximal interval of a random Boolean function f. Let $b'_k(f)$ be a random variable expressing the number of bad vertices in N_K. Then

$$\mathrm{E}(b'_k(f)) \leq \frac{2^k \log_{1/p}^4 n}{n}$$

and consequently (by Markov's inequality)

$$b'_k(f) \leq \frac{2^k \log_{1/p}^4 n}{n} = o(2^k),$$

for $\lfloor \log_2 \log_{1/p} n \rfloor \leq k \leq \lfloor \log_2 \log_{1/p} n + \log_2 \log_2 \log_{1/p} n \rfloor + 1$ with probability tending to 1 as $n \to \infty$. Hence, an n-ary random Boolean function f, containing N_K as a maximal interval (and a k-dimensional subcube of n-cube B^n) contains at least one good vertex.

By the symbol $\Theta(N_K)$ we denote the neighbourhood of the first order of a maximal interval N_K, that is the set of all maximal intervals of a Boolean function f having a nonempty intersection with N_K.

Theorem 1. Let f be an n-ary random Boolean function and let N_K induce a maximal interval of f. Then the following inequalities hold with probability tending to 1 as $n \to \infty$:

$$n^{(1-\varepsilon_n)\log_2 \log_{1/p} n} \le |\Theta(N_K)| \le n^{(1+\varepsilon'_n)\log_2 \log_{1/p} n},$$

where $\varepsilon_n, \varepsilon'_n \to 0$ as $n \to \infty$.

Proof. If a fixed subcube of B^n induces a maximal interval of a random Boolean function f, then this interval contains at least one good vertex. The following inequalities follow from Lemma 5:

$$\mathrm{E}(Y^{\tilde{\alpha}}_{n,k}) - \frac{\mathrm{E}(X^{\tilde{\alpha}}_{n,k})}{\log_{1/p} n} \le Y^{\tilde{\alpha}}_{n,k} \le \mathrm{E}(Y^{\tilde{\alpha}}_{n,k}) + \frac{\mathrm{E}(X^{\tilde{\alpha}}_{n,k})}{\log_{1/p} n}.$$

Now we use the bounds (2) on the dimension of maximal intervals in N_f to finish the proof. To abbreviate the notation, we set $k_0 = \lfloor \log_2 \log_{1/p} n \rfloor$ and $k_1 = \lfloor \log_2 \log_{1/p} n + \log_2 \log_2 \log_{1/p} n \rfloor + 1$.

Upper bound.

$$\begin{aligned}
|\Theta(N_K)| &\le \sum_{\tilde{\alpha}} \sum_{k=k_0}^{k_1} \left[\mathrm{E}(Y^{\tilde{\alpha}}_{n,k}) + \frac{\mathrm{E}(X^{\tilde{\alpha}}_{n,k})}{\log_{1/p} n} \right] \le \sum_{k=k_0}^{k_1} \left[\mathrm{E}(Y^{\tilde{\alpha}}_{n,k}) + \frac{\mathrm{E}(X^{\tilde{\alpha}}_{n,k})}{\log_{1/p} n} \right] p^{2^k} \\
&\le \sum_{k=k_0}^{k_1} \left[\binom{n}{k} p^{2^k} (1-p^{2^k})^{n-k} + \binom{n}{k} \frac{p^{2^k}}{\log_{1/p} n} \right] \cdot p^{2^k} \\
&\le \sum_{k=k_0}^{k_1} \left[\binom{n}{k} p^{2^{k+1}} 2^k \left((1-p^{2^k})^{n-k} + \frac{1}{\log_{1/p} n} \right) \right] \\
&\le \sum_{k=k_0}^{k_1} \binom{n}{k} p^{2^{k+1}} \cdot 2^k \left((1-\frac{1}{n^2}) + \frac{1}{\log_{1/p} n} \right) \\
&\le (k_1 - k_0 + 1) \binom{n}{k_1} p^{2^{k_0+1}} 2^{k_1} \le (k_1 - k_0 + 1) p^{2^{k_0+1}} 2^{k_1} n^{k_1} \\
&\le n^{(1+\varepsilon'_n)\log_2 \log_{1/p} n},
\end{aligned}$$

where $\varepsilon'_n \to 0$ as $s \to \infty$.

Lower bound.

$$\begin{aligned}
|\Theta(N_K)| &\geq \sum_{\tilde{\alpha}} \sum_{k=k_0}^{k_1} \left[\mathrm{E}(Y_{n,k}^{\tilde{\alpha}}) - \frac{\mathrm{E}(X_{n,k}^{\tilde{\alpha}})}{\log_{1/p} n} \right] \geq \sum_{k=k_0}^{k_1} \left[\mathrm{E}(Y_{n,k}^{\tilde{\alpha}}) - \frac{\mathrm{E}(X_{n,k}^{\tilde{\alpha}})}{\log_{1/p} n} \right] p^{2^k} \\
&\geq \sum_{k=k_0}^{k_1} \binom{n}{k} p^{2^{k+1}} \left((1-p^{2^k})^{n-k} - \frac{1}{\log_{1/p} n} \right) \\
&\geq \sum_{k=k_0}^{k_1} \left(\frac{n}{k} \right)^k p^{2^{k+1}} \cdot 2^k \left((1-p^{2^k})^{n-k} - \frac{1}{\log_{1/p} n} \right) \\
&\geq (k_1 - k_0 + 1) \cdot \frac{n^{k_0}}{k_1^{k_1}} \cdot p^{2^{k_1+1}} \left(\left(1 - \frac{1}{n^2} \right)^{n-k_0} - \frac{1}{\log_{1/p} n} \right) \\
&\geq n^{(1-\varepsilon_n) \log_2 \log_{1/p} n},
\end{aligned}$$

where $\varepsilon_n \to 0$ as $n \to \infty$. □

3 CONCLUSIONS

Two important facts follow from the obtained results. In general, the application of the local analysis of the first order neighbourhood in the minimization algorithms increases their complexity at most by a multiplicative factor of $n^{(1+\varepsilon_n) \log_2 \log_{1/p} n}$, where $\varepsilon_n \to 0$ as $n \to \infty$. On the other hand, the local analysis improves the performance of approximative algorithms.

Let us consider the well known greedy algorithm. For every interval N_K under processing we create a list of the intervals from its neighbourhood of the first order and sort them with respect to their dimensions. Then we check the intervals from the list (starting with intervals of the smallest dimension); if all of the vertices of the interval being processed are covered by other intervals from the list, we do not need this interval and delete it from the list. Finally we find a vertex of N_f which is covered only by the tested interval (and maybe by some previously omitted intervals, too); the corresponding elementary conjunction must be included in the d.n.f. Having gradually checked all the intervals in the list, we obtain a d.n.f. which can not be simplified. The length of this d.n.f. is at most $c \cdot \log_2 \log_{1/p} n$-times larger (where c is a positive constant) than the length of the shortest d.n.f. of this Boolean function [7].

REFERENCES

[1] Feller, W.: An Introduction to Probability Theory and its Application. J. Willey and Sons, Vol. 1, 3rd Edition, New York, 1970.

[2] Glagolev, V. V.: Some Estimations for Disjunctive Normal Forms of Boolean Functions. Problemy kibernetiki, Vol. 19, Nauka, Moscow, pp. 75–94, 1967 (in Russian).

[3] Sapozhenko, A. A.: On the Complexity of d.n.f. Obtained by Means of Greedy Algorithm. Diskretnyj analiz, Vol. 21, 1972 (in Russian).

[4] Sapozhenko, A. A.: Disjunctive Normal Forms. Moscow University Press, Moscow, 1975 (in Russian).

[5] Sapozhenko, A. A.: Geometric Structure of Almost All Boolean Functions. Problemy kibernetiki, Vol. 30, Nauka, Moscow, pp. 227–261, 1975 (in Russian).

[6] Škoviera, M.: On the Minimization of Random Boolean Function I.: Computers and Artificial Intelligence, Vol. 5, 1986, No. 4, pp. 321–334.

[7] Škoviera, M.: On the Minimization of Random Boolean Function II.: Computers and Artificial Intelligence, Vol. 5, 1986, No. 6, pp. 493–509.

[8] Toman, E.: Geometric Structure of Random Boolean Functions. Problemy kibernetiki, Vol. 35, Nauka, Moscow, pp. 111–132, 1979 (in Russian).

[9] Toman, E.—Tomanová, J.: Some Estimates of the Complexity of Disjunctive Normal Forms of a Random Boolean Function: Computers and Artificial Intelligence, Vol. 10, 1991, No. 4, pp. 327–340.

[10] Weber, K.: Subcube Coverings of Random Graphs in the n-Cube. Annals of Discrete Math., Vol. 28, 1985, pp. 319–356.

[11] Yablonskii, S. V.—Lupanov, O. B.: Discrete Mathematics and Mathematical Problems of Cybernetics. Nauka, Moscow, 1974 (in Russian).

[12] Zhuravlev, J. I.: Set Theoretical Methods in the Algebra of Logic. Problemy kibernetiki, Vol. 8, Nauka, Moscow, pp. 5–44, 1962 (in Russian).

Eduard Toman received the Ph. D. degree in 1977 at Moscow University with the thesis Geometric properties of the random Boolean functions. His main topics of interest are complexity of Boolean functions, probabilistic methods in graph theory and combinatorics. At present he is associated professor at the Department of Computer Science of Comenius University.

Daniel Olejár graduated from Comenius University in cybernetics and computer science in 1980, currently he is professor of computer science at Comenius University. Scientific interests: combinatorics, coding theory, cryptology and information security.

Martin Stanek received his Ph. D. in computer science from Comenius University. At present he is a teaching assistant of the Department of Computer Science, Comenius University. His research interests include cryptography and information security.