# Testing Monotone Read-Once Functions

Dmitry V. Chistikov

Faculty of Computational Mathematics and Cybernetics
Moscow State University, Russia
dch@cs.msu.ru

**Abstract.** A checking test for a monotone read-once function $f$ depending essentially on all its $n$ variables is a set of vectors $M$ distinguishing $f$ from all other monotone read-once functions of the same variables. We describe an inductive procedure for obtaining individual lower and upper bounds on the minimal number of vectors $T(f)$ in a checking test for any function $f$. The task of deriving the exact value of $T(f)$ is reduced to a combinatorial optimization problem related to graph connectivity. We show that for almost all functions $f$ expressible by read-once conjunctive or disjunctive normal forms, $T(f) \sim n/\ln n$. For several classes of functions our results give the exact value of $T(f)$.

## 1 Introduction

A Boolean function of variables $X$ is called *monotone read-once* iff it can be expressed by a formula over $\{\wedge, \vee\}$ without repetitions of variables (such formulae are also called read-once). By definition, we say that 0 and 1 are monotone read-once functions too. One can see that $f$ depends essentially on a variable $x_i$ iff $x_i$ appears in a read-once formula for $f$.

Suppose that $f$ is a monotone read-once function depending essentially on all variables from $X$; then a set $M$ of input vectors is a *checking test* (or simply a *test*) for $f$ iff for each monotone read-once function $f' \not\equiv f$ of variables $X$ there exists a vector $\alpha \in M$ such that $f'(\alpha) \neq f(\alpha)$. In other words, $M$ is a checking test for $f$ iff values of $f$ on vectors from $M$ allow one to distinguish between $f$ and all other monotone read-once functions of variables $X$.

The *length* of a test is the number of vectors contained in it. For a read-once function $f$, the minimal length of a checking test for $f$ is denoted by $T(f)$. Any test for $f$ having this length is called *optimal* or, equivalently, *minimal*.

The problem of *checking* for read-once functions and study of minimal test length were suggested by A. A. Voronenko, whose paper [9] investigated this problem for the basis of all binary Boolean functions (in this setting the definition of a read-once function is appropriately generalized; most further results are available in English; see, e. g., [11]). It was proved that every $n$-variable read-once function over this basis has a checking test of length less or equal to $4\binom{n}{2}$. This universal bound was subsequently lowered to match a trivial individual lower bound of $\binom{n}{2} + n + 1$ for an $n$-ary disjunction $x_1 \vee \ldots \vee x_n$ [6]. It is known, though, that there exist individual functions allowing checking tests of length

$O(n)$ [8,12]. Generalizations of these results to the case of arbitrary Boolean bases are discussed in [11,13]. For the basis $\{\wedge, \vee, \neg\}$, a universal upper bound of $7n/2$ is proved in [10]. This bound has recently been improved to reach $2n+1$ [2], whereas the highest known individual lower bound is equal to $n+1$.

In this paper, we study minimal test length for individual monotone read-once functions (these functions are read-once over the basis $\{\wedge, \vee\}$). It is known [1] that for any such function $f$ depending essentially on $n$ variables it holds that $2\sqrt{n} \le T(f) \le n+1$. An example of a function requiring $n+1$ vectors in a checking test is an $n$-ary disjunction $x_1 \vee \ldots \vee x_n$, whereas the best known individual upper bound is $3\sqrt{n}-1$, for a special subsequence of *CNF-expressible* functions, i. e., those expressible by monotone read-once conjunctive normal forms (CNF). It is also known that all CNF-expressible functions $f$ have $T(f) \ge 2\sqrt{2}\sqrt{n}-1$.

We demonstrate that for any monotone read-once function $f$ sets of *zeros* and *ones* (false and true points) of $f$ in an optimal test can be chosen independently: for instance, one cannot reduce the number of zeros at the expense of adding several extra ones. If we denote by $T_0(f)$ and $T_1(f)$ the smallest possible number of zeros and ones, respectively, in a checking test for $f$, this result is stated as

$$T(f) = T_0(f) + T_1(f).$$

For CNF-expressible functions $f$, it turns out that the value of $T(f)$ is primarily determined by $T_0(f)$. More precisely, for almost all such functions $f$ depending essentially on $n$ variables (regarded as mappings from $\{0,1\}^n$ to $\{0,1\}$) it holds that

$$T(f) \sim T_0(f) = l(f) \sim \frac{n}{\ln n},$$

where $l(f)$ is the number of clauses in the CNF expressing $f$, and the equality in the center holds for *all* CNF-expressible $f$. Interestingly, the task of deriving the exact value of $T_1(f)$ reveals a curious combinatorial optimization problem related to graph connectivity. Our bounds on the optimal solution to this problem show that for a CNF with $r_1 \ge \ldots \ge r_l$ variables in its clauses it holds that

$$T_1(f) \ge \max\left\{ r_1 + r_2 - 1,\ \log_2\left(\sum_{i=1}^{l} 2^{r_i - 1} - l + 1\right) + 1 \right\},$$

$$T_1(f) \le \min\left\{ \sum_{i \ne 3k} (r_i - 1),\ 4\,(\max\{r_1, l\} - 1),\ \sum_{i=1}^{\lceil \log_2(l+1) \rceil} (r_i - 1) \right\} + 1.$$

For the general case of monotone read-once functions, we show that the task of deriving the exact value of $T(f)$ for an individual function $f$ can be performed by an inductive procedure traversing a read-once formula for $f$. Calculations on each step involve determining the optimal solution to the combinatorial problem mentioned above. For any function $f$ our results allow one to easily obtain individual lower and upper bounds on the value of $T(f)$. The known universal upper bound of $n+1$ can be deduced as a simple corollary. For one class of read-once functions, we present a simple way of computing exact values of $T(f)$ using read-once formulae for $f$.

Note that our definition of a checking test requires that all variables of $f$ be essential, while alternatives $f'$ may well have fictitious variables. This restriction is imposed so that the definition would be word-to-word identical to that for wider bases. Take, for instance, the basis $\{\wedge, \vee, \neg\}$. If the restriction on $f$ is not imposed, then any checking test for $f(x_1, \ldots, x_n) \equiv 0$ must distinguish it from all conjunctions of $n$ literals and, therefore, must contain all $2^n$ vectors. Such a setting should be considered degenerate.

For monotone read-once functions considered in this paper, however, such a restriction can be freely lifted. Indeed, if a monotone read-once function $f$ depends essentially on variables $x_1, \ldots, x_n$, and $g$ is obtained from $f$ by adding fictitious variables $y_1, \ldots, y_m$, then a checking test for $g$ can be obtained from a checking test $M$ for $f$ by extending all zeros of $f$ in $M$ with values $y_1 = \ldots = y_m = 1$, and all ones of $f$ in $M$ with values $y_1 = \ldots = y_m = 0$. Hence, the length of an optimal checking test for $g$ is equal to that for $f$.

## 2 Combinatorial Reduction

Let $f$ be a monotone read-once function of variables $X$. Denote by $R(f)$ a graph on vertices $X$ such that an edge $\{x_i, x_k\}$ is present in $R(f)$ iff $f$ has a projection equivalent to $x_i \wedge x_k$. A classic result of V. A. Gurvich [4] (see also [5]) states that all projections of $f$ that depend essentially on exactly two variables $x_i$ and $x_k$ are equal to each other, equivalent either to $x_i \wedge x_k$ or to $x_i \vee x_k$, and that at least one such projection exists. Every monotone read-once function is uniquely determined by its graph $R(f)$. We shall use a widely known fact that the complement of any nontrivial connected induced subgraph of $R(f)$ is disconnected ($R(f)$ is a *cograph* [3]). We say that a formula $\mathcal{F}$ is *of type* $\wedge$ ($\vee$) iff it is either a variable or a conjunction (a disjunction) of two or more subformulae. A monotone read-once function $f$ is *of type* $\circ \in \{\wedge, \vee\}$ iff it can be expressed by a read-once formula of type $\circ$ over $\{\wedge, \vee\}$. Clearly, $f$ is of type $\wedge$ iff $R(f)$ is connected.

We need the following notation. An *integer partition* is a way of representing an integer as a sum of several positive integers. When referring to integer partitions, we write $m = t_1 + \cdots + t_p$, where $m, p, t_1, \ldots, t_p \geq 1$. We also use one operation on equivalence relations. Suppose that $\epsilon'$ and $\epsilon''$ are equivalence relations on a set $S$; then by $\epsilon' \vee \epsilon''$ we denote the transitive closure of the union of $\epsilon'$ and $\epsilon''$. Thus, $\epsilon' \vee \epsilon''$ is itself an equivalence relation $\epsilon$ such that $a \sim_\epsilon b$ iff there exists a sequence $c_0, \ldots, c_k \in S$ such that $a = c_0$, $b = c_k$ and for each $i = 1, \ldots, k$ either $c_{i-1} \sim_{\epsilon'} c_i$ or $c_{i-1} \sim_{\epsilon''} c_i$. In other words, $\epsilon' \vee \epsilon''$ is the finest equivalence relation on $S$ that is coarser than both $\epsilon'$ and $\epsilon''$. By **true** we denote the binary all-true relation on a set. Finally, we use symbols $\mathbf{0}$ and $\mathbf{1}$ to denote vectors consisting only of zeros and ones, respectively.

Suppose that $l \geq 1$ and $r_1, \ldots, r_l \geq 1$. Take arbitrary positive integers $t_{i,j}$ for $1 \leq i \leq l$ and $1 \leq j \leq r_i$. Denote by $F$ the multiset of $l$ integer partitions $m_i = t_{i,1} + \cdots + t_{i,r_i}$, for $i = 1, \ldots, l$. Define $L(F)$ as the smallest number $t$ having the following property: there exist equivalence relations $\epsilon_1, \ldots, \epsilon_l$ on the set $\{1,\ldots,t\}$ such that each $\epsilon_i$ has $r_i$ equivalence classes of cardinality greater or equal to $t_{i,1},\ldots,t_{i,r_i}$, respectively, and $\epsilon_i \vee \epsilon_k = \mathbf{true}$ whenever $i \neq k$. If $t_{i,j} = 1$ for all possible $i,j$, the number $L(F)$ will also be referred to as $L(r_1,\ldots,r_l)$.

**Theorem 1.** *Let $l \geq 2$ and $r_1,\ldots,r_l \geq 1$. Suppose that $f_{i,j}$, $1 \leq j \leq r_i$, $1 \leq i \leq l$, are monotone read-once functions of type $\wedge$ depending on disjoint sets of variables, $f_i = f_{i,1} \vee \ldots \vee f_{i,r_i}$ for $i = 1,\ldots,l$, $f = f_1 \wedge \ldots \wedge f_l$, and $f_{i,1}$ is a single variable whenever $r_i = 1$. Then*

$$T(f) = T_0(f) + T_1(f),$$

$$T_0(f) = \sum_{i=1}^{l} T_0(f_i), \quad \text{and} \quad T_1(f) = L(F),$$

*where $F$ is the multiset of integer partitions $T_1(f_{i,1}) + \cdots + T_1(f_{i,r_i})$ for all $i = 1,\ldots,l$.*

*Proof.* First obtain the lower bounds. Suppose that a checking test $M$ for $f$ contains a vector $\alpha$ such that $f(\alpha) = 1$. Clearly, replacing $\alpha$ with any vector $\beta \leq \alpha$ such that $f(\beta) = 1$ yields a set of vectors that retains the property of being a checking test. Similarly, one can replace all vectors $\gamma$ such that $f(\gamma) = 0$ with vectors $\delta \geq \gamma$ such that $f(\delta) = 0$. The obtained set $M'$ will still constitute a checking test for $f$, and $|M'| \leq |M|$ (note that it may be the case that, e. g., different vectors $\alpha'$ and $\alpha''$ can be replaced with a single vector $\beta$). Further on, we assume that all possible replacements have been performed, i. e., $M'$ contains only *lower ones* and *upper zeros* of $f$.

Now take an arbitrary upper zero $\alpha$ of $f$. Let $\alpha$ be a concatenation of vectors $\alpha_1,\ldots,\alpha_l$ such that each $f_i$ depends essentially on the variables assigned by $\alpha_i$. Since $f$ is a read-once conjunction of all $f_i$, $1 \leq i \leq l$, this means that there exists a unique index $i$ such that $f_i(\alpha_i) = 0$ and $f_k(\alpha_k) = 1$ for all $k \neq i$. Moreover, it follows that $\alpha_i$ is an upper zero of $f_i$ and all $\alpha_k = \mathbf{1}$ for $k \neq i$. Denote by $z_i$ the number of all vectors $\alpha$ in $M'$ such that $f_i(\alpha_i) = 0$. Since $M'$ is a checking test for $f$, it follows that $z_i \geq T_0(f_i)$. (Indeed, if $z_i < T_0(f_i)$, then there exists a monotone read-once function $f_i' \not\equiv f_i$ which depends on the same variables as $f_i$ and agrees with it on all $z_i$ vectors $\alpha_i$. It then follows that $f$ cannot be distinguished from the function $f'$ obtained by substituting $f_i'$ for $f_i$ in the read-once formula for $f$.) Observe that no vector $\alpha$ can be counted more than once in $z_1,\ldots,z_l$, for all non-constant monotone read-once functions take the value of 1 at $\mathbf{1}$. Thus, $T_0(f) \geq \sum_{i=1}^{l} z_i \geq \sum_{i=1}^{l} T_0(f_i)$.

In order to prove the lower bound on $T_1(f)$, consider the set $\{\alpha^{(1)},\ldots,\alpha^{(t)}\}$ of all lower ones of $f$ contained in $M'$. One can see that an arbitrary lower one $\alpha$ of $f$ is a concatenation of $\alpha_1,\ldots,\alpha_l$ such that $f_i(\alpha_i) = 1$ for all $i$. Moreover, each $\alpha_i$ must be a lower one of $f_i$. Hence, each $\alpha_i$ is a concatenation of $\alpha_{i,1},\ldots,\alpha_{i,r_i}$, and there exists a unique index $j$ such that $f_{i,j}(\alpha_{i,j}) = 1$ and all $\alpha_{i,s} = \mathbf{0}$ for $s \neq j$. For each $\alpha^{(p)}$, these indices will be denoted by $j_1(p),\ldots,j_l(p)$.

For each $i = 1,\ldots,l$ consider an equivalence relation $\epsilon_i$ on $\{1,\ldots,t\}$ such that $p \sim_{\epsilon_i} q$ iff $j_i(p) = j_i(q)$. We claim that $\epsilon_i \vee \epsilon_k = \mathbf{true}$ if $i \neq k$. Assume the converse, then there exists a non-empty proper subset $S$ of $\{1,\dots,t\}$ such that for all $p\in S$ and $q\in\{1,\dots,t\}\setminus S$ it holds that $p \not\sim_{\epsilon_i} q$ and $p \not\sim_{\epsilon_k} q$. By definition, put $I=\{\, j_i(p) \mid p\in S\,\}$ and $K=\{\, j_k(p)\mid p\in S\,\}$. Construct a monotone read-once function $f'$ by replacing the conjunction $f_i\wedge f_k$ in the read-once formula for $f$ with a disjunction $(f_i'\wedge f_k')\vee(f_i''\wedge f_k'')$, where

$$f_i' = \bigvee_{j\in I} f_{i,j}, \qquad f_k' = \bigvee_{j\in K} f_{k,j},$$
$$f_i'' = \bigvee_{j\notin I} f_{i,j}, \qquad f_k'' = \bigvee_{j\notin K} f_{k,j}.$$

We see that $f$ is always greater or equal to $f'$ and disagrees with it only on vectors $\alpha$ such that either $f_i'(\alpha_i)\wedge f_k''(\alpha_k)=1$ or $f_i''(\alpha_i)\wedge f_k'(\alpha_k)=1$. Since $M'$ is a checking test, such a vector $\alpha$ must be present among $\alpha^{(1)},\dots,\alpha^{(t)}$. Assume without loss of generality that $f_i'\left(\alpha_i^{(p)}\right)\wedge f_k''\left(\alpha_k^{(p)}\right)=1$. By definition of $f_i'$ and $f_k''$, it holds that $j_i(p)\in I$ and $j_k(p)\notin K$, so $p\in S$ and $p\notin S$, which is a contradiction.

In order to prove that $T_1(f)\ge L(F)$, we show that $j_i(p)=j$ for at least $T_1(f_{i,j})$ numbers $p\in\{1,\dots,t\}$. Indeed, if this is not the case, then $M'$ contains fewer than $T_1(f_{i,j})$ vectors $\alpha$ such that $\alpha_{i,j}\ne\mathbf{0},\mathbf{1}$, $f_{i,j}(\alpha_{i,j})=1$ and $f(\alpha)=1$. By definition of $T_1(f_{i,j})$, this means that these $\alpha_{i,j}$ do not allow one to distinguish between $f_{i,j}$ and a certain monotone read-once function $f_{i,j}'$. (Note that $\mathbf{1}$ is included in a minimal checking test for $f_{i,j}$ iff $T_1(f_{i,j})=1$, otherwise it is obviously of no use.) Substituting $f_{i,j}'$ for $f_{i,j}$ in the read-once formula for $f$ yields a formula expressing a monotone read-once function $f'\not\equiv f$ such that $f$ agrees with $f'$ on all vectors from $M'$. This concludes the proof of the lower bounds.

We now prove the upper bounds and the equality $T(f)=T_0(f)+T_1(f)$. We use induction on the depth of the read-once formula for $f$. In the inductive step, we shall use only the fact that $T(f_i)=T_0(f_i)+T_1(f_i)$ as an inductive assumption. Therefore, we need to check this equality for single variables, conjunctions and disjunctions. It can easily be checked that for $n\ge1$,

$$T_0(x_1\vee\ldots\vee x_n)=1,\quad T_1(x_1\vee\ldots\vee x_n)=n,\quad T(x_1\vee\ldots\vee x_n)=n+1,$$
$$T_0(x_1\wedge\ldots\wedge x_n)=n,\quad T_1(x_1\wedge\ldots\wedge x_n)=1,\quad T(x_1\wedge\ldots\wedge x_n)=n+1,$$

so we proceed to the main part of the proof.

Let $M_1,\dots,M_l$ be optimal checking tests for $f_1,\dots,f_l$, respectively, all consisting of upper zeros and lower ones of $f_i$. Let $N$ consist of all vectors $\alpha=(\mathbf{1},\dots,\mathbf{1},\alpha_i,\mathbf{1},\dots,\mathbf{1})$ for all $\alpha_i\in M_i$ such that $f_i(\alpha_i)=0$ and $i=1,\dots,l$. Now suppose that $L(F)=t$ and equivalence relations $\epsilon_1,\dots,\epsilon_l$ on $\{1,\dots,t\}$ satisfy the conditions of $L(F)$ definition. Assume that equivalence classes of each $\epsilon_i$ are numbered 1 through $r_i$ so that $j$th class's cardinality is greater or equal to $T_1(f_{i,j})$. Now recall that every vector $\alpha_i\in M_i$ such that $f_i(\alpha_i)=1$ is a lower one of $f_i$ and thus has a special representation as a concatenation of $\alpha_{i,j}$, $1\le j\le r_i$. Put $U_{i,j}=\{\,\alpha_{i,j}\mid \alpha_i\in M_i,\ f_i(\alpha_i)=1,\ \alpha_{i,j}\ne\mathbf{0}\,\}$. For each $p=1,\dots,t$ put

$\alpha^{(p)} = \left(\alpha_1^{(p)}, \ldots, \alpha_l^{(p)}\right)$, where $\alpha_i^{(p)} = \left(\alpha_{i,1}^{(p)}, \ldots, \alpha_{i,r_i}^{(p)}\right)$ such that $\alpha_{i,j}^{(p)} \in U_{i,j}$ if the number $p$ belongs to the $j$th equivalence class of $\epsilon_i$ and $\alpha_{i,j}^{(p)} = \mathbf{0}$ otherwise. We also require that for each valid pair $i, j$ the set of all non-$\mathbf{0}$ vectors $\alpha_{i,j}^{(p)}$ be equal to $U_{i,j}$. This is possible because the number of elements in $j$th equivalence class of $\epsilon_i$ is greater or equal to $|U_{i,j}| = T_1(f_{i,j})$ (this equality follows from the inductive assumption and our choice of $M_i$, for it is easily checked that $\sum_{j=1}^{r_i} |U_{i,j}| = T_1(f_i)$). We claim that $M = N \cup \left\{\alpha^{(1)}, \ldots, \alpha^{(t)}\right\}$ is a checking test for $f$.

We show that if a monotone read-once function $f'$ coincides with $f$ on all vectors from $M$, then $f' \equiv f$. This is sufficient both for proving the claimed upper bound and, as a corollary, for establishing the equality $T(f) = T_0(f) + T_1(f)$. First, observe that $f'\left(\alpha^{(p)}\right) = 1$ for all $p = 1, \ldots, t$. Since $f'$ is monotone, it follows that $f'(\alpha) = 1$ for any $\alpha$ of type $\left(\mathbf{1}, \ldots, \mathbf{1}, \alpha_i^{(p)}, \mathbf{1}, \ldots, \mathbf{1}\right)$, where $p = 1, \ldots, t$. By our construction of $M$, all vectors from $M_i$ are present in $\left\{ \alpha_i^{(p)} \mid \alpha^{(p)} \in M \right\}$. Since $M$ contains all vectors from $N$, it follows that $f'(\alpha) = f(\alpha)$ for all $\alpha = (\mathbf{1}, \ldots, \mathbf{1}, \alpha_i, \mathbf{1}, \ldots, \mathbf{1})$, where $\alpha_i \in M_i$. One concludes, then, that for each $i = 1, \ldots, l$ the function $f'$ has a projection equivalent to $f_i$.

Now we are going to reconstruct the graph $R(f') = R(f)$ using the values of $f$ on vectors from $M$. Note that all subgraphs $R_i = R(f_i)$ are already known. Recall that all functions $f_{i,j}$ are of type $\wedge$, so all $R_{i,j}$ are connected. On the contrary, each $R_i$ is either disconnected or a single vertex. It suffices to show that each subgraph $R_i \cup R_k$ of $R(f')$ (a subgraph of $R(f')$ induced by vertices of $R_i$ and $R_k$) is connected if $i \neq k$, for then *all* edges between $R_i$ and $R_k$ must be present in it, otherwise any edge between $R_i$ and $R_k$ not present in $R_i \cup R_k$ would imply the connectivity of $R_i \cup R_k$'s complement, which contradicts the connectivity of $R_i \cup R_k$ (recall that $R(f')$ is a cograph). We now contract all vertices in each $R_{i,j}$ ($R_{k,j}$) to a single vertex $j$ ($j'$) and prove the connectivity of the obtained bipartite graph $R'$ on vertices $\{1, \ldots, r_i\} \cup \{1', \ldots, r'_k\}$.

We first claim that the equality $f'\left(\alpha^{(p)}\right) = 1$ implies that the edge $\{j_i(p), j_k(p)\}$ is present in $R'$. Indeed, since $f'\left(\alpha^{(p)}\right) = 1$, it holds that $f'(\beta) = 1$, where $\beta$ is obtained from $\alpha^{(p)}$ by changing all $\alpha_u^{(p)}$ to $\mathbf{1}$ for $u \neq i, k$. Suppose that $\gamma$ is obtained from $\beta$ by changing a 1 in $\alpha_i$ to 0. One now observes that replacing $\alpha_k$ with $\mathbf{1}$ in $\gamma$ yields a vector $\gamma'$ with a known value $f'(\gamma') = 0$, since every $\alpha_u$, where $u \neq i$, is now replaced by $\mathbf{1}$, and $\alpha_i$ is a lower one of the known projection $f_i$. Monotonicity of $f'$ implies $f'(\gamma) = 0$. Arguing as above, we see that $f'(\delta) = 0$, where $\delta$ is obtained from $\beta$ by changing a 1 in $\alpha_k$ to 0. The values of $f'$ on vectors $\beta$, $\gamma$, and $\delta$ are uniquely determined by its values on vectors from $M$ and imply that $f'$ has a projection of the type $x' \wedge x''$, where $f_{i,j_i(p)}$ and $f_{k,j_k(p)}$ depend on $x'$ and $x''$, respectively. Thus, $R'$ contains all edges $\{j_i(p), j_k(p)\}$ for $p = 1, \ldots, t$.

It remains to prove that these edges imply the connectivity of $R'$. Consider a graph $G$ on $2t$ vertices $\{1, \ldots, t\} \cup \{1', \ldots, t'\}$. Let $G$ contain all the edges $\{p, p'\}$ for $p = 1, \ldots, t$, edges $\{p, q\}$ whenever $p \sim_{\epsilon_i} q$, and $\{p', q'\}$ whenever $p' \sim_{\epsilon_k} q'$. Contracting all the edges within $\{1, \ldots, t\}$ and $\{1', \ldots, t'\}$ yields a graph $H$, which is known to be isomorphic to a subgraph of $R'$. Clearly, $H$ is connected if so is $G$. Contracting all the edges $\{p, p'\}$ in $G$ yields a graph $G'$ on vertices $\{1, \dots, t\}$ such that $\{p, q\}$ is present in $G'$ if and only if $p \sim_{\epsilon_i} q$ or $p \sim_{\epsilon_k} q$. Since $\epsilon_i \vee \epsilon_k = \mathbf{true}$, it follows that $G'$ is connected, and so are $G$, $H$ and $R'$. This concludes the proof.

**Corollary 2.** *For all non-constant monotone read-once functions $f$,*

$$T(f) = T_0(f) + T_1(f).$$

**Corollary 3.** *For a monotone read-once function*

$$f = (x_{1,1} \vee \ldots \vee x_{1,r_1}) \wedge (x_{2,1} \vee \ldots \vee x_{2,r_2}) \wedge \ldots \wedge (x_{l,1} \vee \ldots \vee x_{l,r_l}),$$

*where $l \geq 1$ and all $r_i \geq 1$, the following equality holds:*

$$T(f) = l + L(r_1, \dots, r_l).$$

**Remark 4.** The statements of Theorem 1 and Corollary 3 hold true if all symbols $\wedge$ and $\vee$ are exchanged and so are Boolean constants 0 and 1. Provided that an algorithm for determining $L(F)$ is known, one can use a simple inductive procedure to determine the value of $T(f)$ for any arbitrary monotone read-once function $f$. The induction basis is given by the values of $T_0(f)$ and $T_1(f)$ for disjunctions and conjunctions of $n \geq 1$ variables. Since $L(F)$ is evidently monotonically non-decreasing in all parameters in $F$, one can substitute lower and upper bounds for the unknown parameters to obtain lower and upper bounds on $L(F)$, respectively.

## 3 Obtaining Bounds on $L(F)$

**Proposition 5.** *$L(r_1, \dots, r_l) \leq L(F) \leq L(r_1, \dots, r_l) + \max\limits_{1 \leq i \leq l}(m_i - r_i) - d$, where $d = 1$ if there exist numbers $s \neq k$ such that $s$ is a maximum point of $m_i - r_i$ and $r_k \geq 2$, and $d = 0$ otherwise.*

*Proof.* The lower bound is obvious and the upper bound follows from the observation that in all non-trivial cases any equivalence relation $\epsilon_i$ from the definition of $t = L(r_1, \dots, r_l)$ has at least one equivalence class of size greater or equal to 2, so for each $i$ such that $m_i > r_i$ one needs less or equal to $m_i - r_i - 1$ new elements beyond $1, \dots, t$. The only exception is the case when $r_k = 1$ for all $k \neq i$.

When obtaining bounds on $L(F)$, we often use graph-theoretic terminology, as given by the following lemma. (Note that when we speak of graphs, we always mean undirected graphs without loops or multiple edges.)

**Lemma 6.** *The number $L(F)$ is the smallest number $t$ having the following property: there exist graphs $G_1, \dots G_l$ on vertices $\{1, \dots, t\}$ such that each $G_i$ has exactly $r_i$ connected components of size greater or equal to $t_{i,1}, \dots, t_{i,r_i}$, and all graphs $G_i \cup G_k$ are connected whenever $i \neq k$.*

The proof is trivial. Note that we can use arbitrary connected graphs as components of $G_i$. It is often convenient, though, to use only trees as these components; in this case all graphs $G_i$ are *required* to be forests.

**Theorem 7.** *Suppose that $F$ is a multiset of integer partitions $m_i = t_{i,1} + \cdots + t_{i,r_i}$ for $i = 1, \ldots, l$. Then the following inequality holds:*

$$L(F) \geq \max\left\{\max_i m_i,\ \max_{i\neq k}(r_i + r_k - 1),\ \log_2\left(\sum_{i=1}^{l} 2^{r_i-1} - l + 1\right) + 1\right\}.$$

*Proof.* Let $t$ be the value of $L(F)$. One can easily see that $t \geq m_i$ for all $i = 1, \ldots, l$. Indeed, since $G_i$ has to contain $r_i$ connected components of size at least $t_{i,1}, \ldots, t_{i,r_i}$, it follows that $t \geq \sum_{j=1}^{r_i} t_{i,j} = m_i$.

Now suppose that $i \neq k$ and assume without loss of generality that both $G_i$ and $G_k$ are forests. Then the number of edges in these two graphs is $t - r_i$ and $t-r_k$, respectively. Since $G_i \cup G_k$ is connected, we see that $(t-r_i)+(t-r_k) \geq t-1$ and $t \geq r_i + r_k - 1$.

Finally, consider partitions of $\{1, \ldots, t\}$ into two non-empty sets $S'$ and $S''$. The number of such partitions is $2^{t-1}-1$. For each graph $G_i$ having $r_i$ connected components there exist exactly $2^{r_i-1} - 1$ such partitions without edges between $S'$ and $S''$ in $G_i$. These sets of partitions must be disjoint for $i \neq k$, otherwise $G_i \cup G_k$ contains no edges between $S'$ and $S''$ and, therefore, is disconnected. Hence, $2^{t-1} - 1 \geq \sum_{i=1}^{l}(2^{r_i-1} - 1)$, which gives the desired.

In several cases the bounds of Theorem 7 turn out to be tight. Two next propositions show that all three expressions under max can give the exact value of $L(F)$ for some $F$.

**Proposition 8.** *If $l = 2$, then $L(F) = \max\{m_1,\ m_2,\ r_1 + r_2 - 1\}$.*

*Proof.* Let $F$ consist of partitions $m_1 = t_1 + \cdots + t_p$ and $m_2 = s_1 + \cdots + s_q$. By Theorem 7, $L(F) \geq m$, where $m = \max\{m_1, m_2, p+q-1\}$. Our goal is to prove an equal upper bound. Assume without loss of generality that $m_1 = m_2 \geq p+q-1$. (If this is not the case, increase some of the numbers $t_1, \ldots, t_p$ and $s_1, \ldots, s_q$ appropriately so that $m_1$ and $m_2$ would reach $m$ and observe that for the multiset $F'$ of the two obtained partitions it holds that $L(F) \leq L(F')$.) We claim that for any multiset $F$ satisfying this condition and any appropriate graph $G_1$ on $m = m_1$ vertices there exists a graph $G_2$ with the needed properties. The proof of this claim is by induction over $q \geq 1$. For $q = 1$, the desired is straightforward. Indeed, since $m = m_1$ and $m = m_2 = s_1$, one can take a complete graph on $m$ vertices for $G_2$. Now suppose that $q \geq 2$. Assume that $t_1 \leq t_2 \leq \ldots \leq t_p$ and $s = \max_i s_i$. If $s = 1$, then $p = (p+q-1) - (q-1) \leq m - q + 1 = q - q + 1 = 1$. So, $p = 1$ and $G_1$ is a complete graph on $m$ vertices, similarly to the case above. If $s \geq 2$, take connected components of size $t_1, \ldots, t_{s-1}$ and $t_p$ in $G_1$, choose one vertex from each of these components and form a clique of size $s$ on these vertices in $G_2$. If $p \leq s$, missing $s - p$ vertices are chosen arbitrarily and we have proved the desired without the inductive assumption. If $p > s$, the problem is reduced to a simpler one, with $q' = q-1$ (we use prime symbols for distinguishing a new instance of the problem from the old one), $p' = p-s+1$, $p'+q'-1 = (p+q-1)-s$ and $m' = m'_1 = m'_2 = m-s$. Indeed, components of size $t_1, \dots, t_{s-1}$ and $t_p$ in $G_1$ are connected with a clique in $G_2$, whose $s$ vertices are excluded from further consideration. The rest $t'_1 = (t_1-1)+\ldots+(t_{s-1}-1)+(t_p-1)$ vertices from these components (note that $t'_1 \geq 0+\ldots+0+1 = 1$) are considered to belong to the same component of $G'_1$ and so may be connected with, e. g., a clique in $G'_1$. Thus, $q$ is decreased by 1 and the property $m' = m'_1 = m'_2 \geq p'+q'-1$ still holds. This concludes the proof.

**Proposition 9.** $L(2, \dots, 2) = \lceil \log_2(l+1) \rceil + 1$.

*Proof.* The lower bound is given by Theorem 7. To prove the upper bound, take all possible graphs $G_i$ on vertices $\{1, \dots, t\}$ having exactly two connected components and all edges within each component (any $G_i$ is a union of two cliques). Clearly, if $G_i$ and $G_k$ are two such graphs, then $G_i \cup G_k$ is connected if and only if $G_i \neq G_k$. So, if $t$ is fixed, $l$ can be chosen as large as $2^{t-1}-1$. Hence, $t \leq \lceil \log_2(l+1) \rceil + 1$.

We now present an example of using Theorem 1 for deriving exact values of $T(f)$. This example is directly related to the result of Proposition 8. We say that a formula $\mathcal{F}$ over $\{\wedge, \vee\}$ is *strictly alternating* iff it is either a variable or a disjunction (respectively, a conjunction) of exactly two strictly alternating formulae that have type $\wedge$ (respectively, $\vee$). The structure of strictly alternating formulae can be represented by binary trees with alternating levels of internal vertices labeled with $\wedge$ and $\vee$. By $\mathcal{F}^*$ we denote a formula obtained from $\mathcal{F}$ by exchanging all symbols $\wedge$ and $\vee$.

**Proposition 10.** *Let $f$ be a monotone read-once function expressed by a strictly alternating read-once formula $\mathcal{F}$. By definition, put $x \star y = 3$ if $x = y = 2$ and $x \star y = \max\{x, y\}$ otherwise. Then $T(f) = E(\mathcal{F}) + E(\mathcal{F}^*)$, where $E(\mathcal{F})$ is the value of the formula obtained from $\mathcal{F}$ by changing all symbols $\vee$ to $+$, all symbols $\wedge$ to $\star$, and setting all variables' values to 1.*

*Proof.* Use the inductive procedure of Theorem 1. For conjunctions and disjunctions, there is nothing to prove. For all other cases, use Proposition 8. Observe that if $\max\{m_1, m_2\} < r_1 + r_2 - 1$, then $r_1 = r_2 = 2$, since each $r_i$ is less or equal to $\min\{m_i, 2\}$. It follows that $m_1 = m_2 = 2$, which gives the desired.

**Example 11.** Suppose that formulae $\mathcal{F}_n$ are represented by *perfect* binary trees with $n = 2^h$ leaves. For the corresponding read-once functions $f_n$, if $h \geq 3$, then $T(f_n) = 3\sqrt{n}$ if $h$ is even and $T(f_n) = 9\sqrt{2}\,/\,4 \cdot \sqrt{n}$ if $h$ is odd.

*Proof.* Without loss of generality, denote by $f_n$ a read-once function expressed by a formula $\mathcal{F}_n$ of type $\wedge$, and by $f_n^*$ a read-once function expressed by $\mathcal{F}_n^*$. Clearly, $T_0(f_n) = T_1(f_n^*)$ and $T_1(f_n) = T_0(f_n^*)$. Straightforward application of Proposition 10 shows that

$$T(f_8) = 9 = 9\sqrt{2}\,/\,4 \cdot \sqrt{8},$$
$$T(f_{16}) = 12 = 3\sqrt{16},$$

which proves the induction basis. For $n \geq 8$, Proposition 10 reveals that

$$T_0(f_{4n}) = 2\,T_0(f^*_{2n}) = 2\,T_0(f_n),$$
$$T_1(f_{4n}) = T_1(f^*_{2n}) = 2\,T_1(f_n),$$

and it follows that

$$T(f_{4n}) = 2\,T(f_n).$$

This completes the proof.

By $F_1 + F_2$ denote the sum of multisets $F_1$ and $F_2$, i. e., a multiset which contains each element of $F_1$ or $F_2$ as many times as $F_1$ and $F_2$ do together. Simple decomposition gives the following upper bound on $L(F_1 + F_2)$.

**Lemma 12.** $L(F_1 + F_2) \leq L(F_1) + L(F_2) - 1$.

*Proof.* Suppose that $t_1 = L(F_1)$ and $t_2 = L(F_2)$. Let $G^1_1, \ldots, G^1_{l_1}$ and $G^2_1, \ldots, G^2_{l_2}$ be graphs from the alternative definition of $L(F_1)$ and $L(F_2)$ given by Lemma 6. Assume without loss of generality that each $G^s_i$, where $s = 1, 2$, is a graph on vertices $\{(s, 1), \ldots, (s, t_s)\}$. Take each $G^s_i$ and extend it with a clique on vertices $\{(3 - s, 1), \ldots, (3 - s, t_{3-s})\}$, and then identify vertices $(1, 1)$ and $(2, 1)$. The obtained graph $H^s_i$ has $t_1 + t_2 - 1$ vertices and the same number of connected components as $G^s_i$. The number of vertices in each component is greater or equal to that in $G^s_i$. For all possible $i \neq k$, graphs $H^s_i \cup H^s_k$ are connected because so are $G^s_i \cup G^s_k$. Moreover, all $H^1_i \cup H^2_k$ for any possible $i, k$ are connected too, for they all contain cliques on vertices $\{(1, 1), \ldots, (1, t_1)\}$ and $\{(2, 1), \ldots, (2, t_2)\}$, where $(1, 1)$ and $(2, 1)$ are one vertex. This means that $L(F_1 + F_2) \leq t_1 + t_2 - 1$.

The results of Theorem 1 and Lemma 12 lead to an inductive proof of the following known result:

**Theorem 13.** *For all monotone read-once functions $f$ depending on $n$ variables,*

$$T(f) \leq n + 1.$$

*Proof.* Use induction on $n \geq 1$. For $n = 1$, the bound is clearly true. Suppose now that $n \geq 2$. Assume without loss of generality that $f = f_1 \wedge \ldots \wedge f_l$ and $f_i = f_{i,1} \vee \ldots \vee f_{i,r_i}$, where all $f_{i,j}$ are monotone read-once functions of type $\wedge$ depending on disjoint sets of variables, and $f_{i,1}$ is a single variable whenever $r_i = 1$. Let $n_i$ be the number of variables of $f_i$. By the inductive assumption, $T(f_i) \leq n_i + 1$. By Theorem 1, $T_1(f) \leq L(F)$, where $F$ is the multiset of integer partitions $T_1(f_{i,1}) + \cdots + T_1(f_{i,r_i})$ for $i = 1, \ldots, l$. By Lemma 12,

$$L(F) \leq \sum_{i=1}^{l} L(\{T_1(f_{i,1}) + \cdots + T_1(f_{i,r_i})\}) - l + 1 = \sum_{i=1}^{l}\sum_{j=1}^{r_i} T_1(f_{i,j}) - l + 1.$$

Applying Theorem 1 three more times yields

$$T_1(f) \le \sum_{i=1}^{l} T_1(f_i) - l + 1, \qquad T_0(f) = \sum_{i=1}^{l} T_0(f_i),$$

$$\text{and} \quad T(f) \le \sum_{i=1}^{l} (T(f_i) - 1) + 1 \le \sum_{i=1}^{l} n_i + 1 = n + 1,$$

which completes the proof.

To prove more accurate individual upper bounds, we need the following notation. If $m$ is a positive integer, then $\mathbf{Z}_2^m = \{ (x_1, \ldots, x_m) \mid x_1, \ldots, x_m \in \mathbf{Z}_2 \}$ is a vector space over $\mathbf{Z}_2 = \{0, 1\}$. For an arbitrary vector $\mathbf{x} \in \mathbf{Z}_2^m$, put $\operatorname{supp} \mathbf{x} = \{ i \mid 1 \le i \le m, \ x_i = 1 \}$. The following lemma gives an alternative definition of $L(r_1, \ldots, r_l)$.

**Lemma 14.** *The number $L(r_1, \ldots, r_l) - 1$ is the smallest integer $m$ having the following property: there exist linear subspaces $V_1, \ldots, V_l$ of $\mathbf{Z}_2^m$ such that $\dim V_i = r_i - 1$, $V_i \cap V_k = \{\mathbf{0}\}$ for $i \ne k$ and each $V_i$ has a basis $\mathbf{e}_{i,1}, \ldots, \mathbf{e}_{i,r_i - 1}$ with $\operatorname{supp} \mathbf{e}_{i,j'} \cap \operatorname{supp} \mathbf{e}_{i,j''} = \emptyset$ for $j' \ne j''$.*

The idea of the proof is that if $i = 1, \ldots, l$, then sets $\operatorname{supp} \mathbf{e}_{i,j}$ for $j = 1, \ldots, r_i - 1$ are equivalence classes of $\epsilon_i$ not containing $t = L(r_1, \ldots, r_l)$. Details are left to the reader and can be found in the Appendix of this preprint. Note that we can also reformulate the definition of $L(F)$ in a way similar to that of Lemma 14. Such a definition, though, is out of our scope now. We now aim to obtain an efficient upper bound on $L(r_1, \ldots, r_l)$. For convenience, by $\mathcal{L}(\mathbf{e}_1, \ldots, \mathbf{e}_s)$ we denote the linear subspace spanned by vectors $\mathbf{e}_1, \ldots, \mathbf{e}_s$.

**Proposition 15.** *If $r_1 \ge r_2 \ge r_3$, then $L(r_1, r_2, r_3) = r_1 + r_2 - 1$.*

*Proof.* The lower bound follows from Theorem 7. To prove the upper bound, put $m = r_1 + r_2 - 2$. By $\mathbf{e}_j$ denote a vector from $\mathbf{Z}_2^m$ with $m - 1$ zeros and an only one in $j$th position. Consider subspaces

$$V_1 = \mathcal{L}(\mathbf{e}_1, \ldots, \mathbf{e}_{r_1 - 1}), \qquad V_2 = \mathcal{L}(\mathbf{e}_{r_1}, \ldots, \mathbf{e}_m),$$
$$V_3 = \mathcal{L}(\mathbf{e}_1 + \mathbf{e}_{r_1}, \ \mathbf{e}_2 + \mathbf{e}_{r_1 + 1}, \ \ldots, \ \mathbf{e}_{r_3 - 1} + \mathbf{e}_{r_1 + r_3 - 2}).$$

Since $r_3$ is less or equal to both $r_1$ and $r_2$, the sets $V_i \setminus \{\mathbf{0}\}$ are pairwise disjoint. Each $V_i$ is spanned by $r_i - 1$ linearly independent vectors without common ones, so, by Lemma 14, we get the needed upper bound.

In the next proposition, the proof of the upper bound follows from a construction of [1] due to Voronenko, and the lower bound is given by Theorem 7.

**Proposition 16.** *If $r_1 \ge \ldots \ge r_l$, then $L(r_1, \ldots, r_l) \le 2\,\widehat{p}(\max\{r_1, l\}) - 1 \le 4 \max\{r_1, l\} - 3$, where $\widehat{p}(k)$ is the smallest prime greater or equal to $k$. In particular, if $r_1 = \ldots = r_l = l$ and $l$ is prime, then $L(l, \ldots, l) = 2l - 1$.*

Now we are ready to prove our main upper bounds.

**Theorem 17.** *If* $r_1 \ge r_2 \ge \ldots \ge r_l$, *then*

$$L(r_1,\ldots,r_l) \le \min\left\{ \sum_{i\ne 3k}(r_i-1),\ 4\,(\max\{r_1,l\}-1),\ \sum_{i=1}^{\lceil \log_2(l+1)\rceil}(r_i-1) \right\} + 1.$$

*Proof.* First combine the results of Propositions 8 and 15 using Lemma 12. Put $l = 3s+d$, where $d \in \{1,2,3\}$, and observe that the difference $L(r_1,\ldots,r_l) - L(r_{3s+1},\ldots,r_{3s+d})$ cannot be greater than

$$\sum_{i=0}^{s-1} L(r_{3i+1},r_{3i+2},r_{3i+3}) - s = \sum_{i=0}^{s-1}(r_{3i+1}+r_{3i+2}-1) - s = \sum_{\substack{1\le i\le 3s\\ i\ne 3k}}(r_i-1),$$

which gives the first needed inequality. The second inequality follows from Proposition 16. To prove the third inequality, use Lemma 14 directly. Choose $d = \lceil \log_2(l+1)\rceil$ as the number of digits in a binary representation of $l$. For $i = 1,\ldots,l$ by $\alpha_s(i)$ denote $s$th least significant bit in a $d$-bit binary representation $\alpha(i) = (\alpha_d(i)\ \ldots\ \alpha_1(i))$ of $i$. Let $\pi$ be any permutation on $\{1,\ldots,l\}$ such that $\pi(s) = 2^{s-1}$ for $s = 1,\ldots,d$. Put $u_{\pi(i)} = r_i - 1$ for $i = 1,\ldots,l$ and choose $m = \sum_{s=1}^{d} u_{\pi(s)} = \sum_{s=1}^{d}(r_s-1)$. For $s = 1,\ldots,d$ and $j = 1,\ldots,u_{\pi(s)}$ by $\mathbf{f}_{s,j}$ denote a vector from $\mathbf{Z}_2^m$ with $m-1$ zeros and an only one in $(u_{\pi(1)}+\ldots+u_{\pi(s-1)}+j)$th position. Take

$$\mathbf{e}_{i,j} = \sum_{s=1}^{d} \alpha_s(i)\,\mathbf{f}_{s,j}, \qquad j = 1,\ldots,u_i, \quad i = 1,\ldots,l,$$

and consider subspaces $U_i = \mathcal{L}(\mathbf{e}_{i,1},\ldots,\mathbf{e}_{i,u_i})$ for $i = 1,\ldots,l$. Clearly, $\dim U_i = u_i$, for $\sum_{j=1}^{u_i}\lambda_j\mathbf{e}_{i,j} = \mathbf{0}$ implies $\sum_{j=1}^{u_i}\sum_{s=1}^{d}\lambda_j\alpha_s(i)\,\mathbf{f}_{s,j} = \mathbf{0}$ and $\lambda_j = 0$ for all possible $j$, since vectors $\mathbf{f}_{s,j}$ are linearly independent. (Note the special case $i = 2^q$, where $j$ can assume values greater than $u_s$, but it turns out that all $\alpha_s(i)$ equal 0 except for one $s$ and we still see that all $\lambda_j$ must be equal to 0.) It is easily observed that for each $i$ sets $\operatorname{supp}\mathbf{e}_{i,j}$, where $j = 1,\ldots,u_i$, are disjoint. We claim that $U_i \cap U_k = \{\mathbf{0}\}$ whenever $i \ne k$, which is evidently sufficient for obtaining the desired bound.

Instead of using linear algebra in a straightforward way, we prove a special property of sets $U_i$. For each $\mathbf{x} \in \mathbf{Z}_2^m$, define $\operatorname{sig}\mathbf{x} = \{\, s \mid 1 \le s \le d,\ \exists j : \mathbf{x} \ge \mathbf{f}_{s,j}, 1 \le j \le u_s \,\}$. Take an arbitrary non-$\mathbf{0}$ vector $\mathbf{x} \in U_i$. Observe that

$$\mathbf{x} = \sum_{j=1}^{u_i}\lambda_j\mathbf{e}_{i,j} = \sum_{j=1}^{u_i}\lambda_j\sum_{s=1}^{d}\alpha_s(i)\,\mathbf{f}_{s,j} = \sum_{s=1}^{d}\alpha_s(i)\sum_{j=1}^{u_i}\lambda_j\mathbf{f}_{s,j}.$$

If $\alpha_s(i) = 1$, then $j$ assumes the values $1,\ldots,u_i \le u_{\pi(s)}$. Therefore, all vectors $\mathbf{f}_{s,j}$ have ones in different single components. It follows that $\operatorname{sig}\mathbf{x} = \operatorname{supp}\alpha(i)$, where $\alpha(i) = (\alpha_d(i)\ldots\alpha_1(i))$ is a $d$-bit binary representation of $i$. Hence, the sets $U_i \setminus \{\mathbf{0}\}$ are disjoint, which completes the proof.

In the next theorem, when we speak of *almost all* CNF- and DNF-expressible functions, these functions are regarded as mappings from $\{0,1\}^n$ to $\{0,1\}$. For each $n$, all *mappings* expressible by monotone read-once CNF or DNF are assigned equal non-zero probabilities. For instance, formulae $(x_1 \vee x_2) \wedge x_3$ and $(x_2 \vee x_1) \wedge x_3$ express the same function, different from one expressed by $(x_1 \vee x_3) \wedge x_2$.

**Theorem 18.** *For almost all monotone read-once CNF- and DNF-expressible functions $f$ depending essentially on $n$ variables,*

$$T(f) \sim \frac{n}{\ln n}.$$

*Proof.* Consider a monotone read-once CNF-expressible function $f$ which depends essentially on $n$ variables $x_1, \ldots, x_n$. One can easily see that there exists a one-to-one mapping $\phi$ between the set of all such functions and the set of all partitions of the set $\{1, \ldots, n\}$. For a function $f$ expressed by a CNF $\mathcal{F}$, indices $p$ and $q$ belong to the same block in $\phi(f)$ iff they belong to the same clause in $\mathcal{F}$. Denote by $l$ the number of clauses in $\mathcal{F}$. It is known that almost all partitions have asymptotically $n/\ln n$ blocks, all of which have size less or equal to $O(\ln n)$ (see, e. g., [7]). One observes then that the upper bound of Theorem 17 is almost always less or equal to

$$\left\lceil \log_2\left(\frac{n}{\ln n}\right) + o(1) \right\rceil \cdot O(\ln n) = O(\ln^2 n) = o\left(\frac{n}{\ln n}\right),$$

so the value of $T(f)$ given by Corollary 3 is asymptotically equivalent to $T_0(f) = l \sim n/\ln n$.

## 4 Conclusions

We reduced the problem of deriving the value of $T(f)$ to several instances of another combinatorial problem, that of determining the smallest number of vertices $L$ allowing the construction of a set of graphs with special properties. Our results give several explicit bounds on $L$ numbers and allow to deduce the implied bounds on $T(f)$ easily. For almost all CNF- and DNF-expressible functions, these bounds determine the asymptotic behaviour of $T(f)$. For arbitrary read-once functions $f$, one can use Theorem 1 repeatedly to obtain both lower and upper bounds on $T(f)$. For several classes of functions our results determine the exact value of $T(f)$. Finally, we remark that one can easily indicate a special class of read-once functions which shows that it is impossible to derive the exact values of $T(f)$ for *all* $f$ without computing $L$.

**Acknowledgements.** The author wishes to thank Prof. Bruce M. Kapron from the University of Victoria, Canada, who kindly agreed to present this paper at IWOCA 2011. This research was supported by Russian Presidential grant MD–757.2011.9.

## References


1. Bubnov, S. E., Voronenko, A. A., Chistikov, D. V.: Some test length bounds for nonrepeating functions in the $\{\&, \vee\}$ basis. Computational Mathematics and Modeling 21(2), 196–205 (2010)
2. Chistikov, D. V.: Testing read-once functions over the elementary basis. Moscow University Computational Mathematics and Cybernetics (to appear)
3. Corneil, D. G., Lerchs, H., Stewart Burlingham, L.: Complement reducible graphs. Discrete Applied Mathematics 3(3), 163–174 (1981)
4. Gurvich, V. A.: On repetition-free Boolean functions. Uspehi Matematicheskih Nauk 32(1), 183–184 (1977) (in Russian)
5. Karchmer, M., Linial, N., Newman, I., Saks, M., Widgerson, A.: Combinatorial characterization of read-once formulae. Discrete Mathematics 114 (1-3), 275–282 (1993)
6. Ryabets, L. V.: Checking test complexity for read-once Boolean functions. Ser. Diskretnaya Matematika i Informatika, vol. 18. Izdatel'stvo Irkutskogo gosudarstvennogo pedagogicheskogo universiteta (2007) (in Russian)
7. Sachkov, V. N.: Probabilistic methods in combinatorial analysis. Encyclopedia of Mathematics and its Applications, vol. 56. Cambridge University Press (1997)
8. Voronenko, A. A.: Estimating the length of a diagnostic test for some nonrepeating functions. Computational Mathematics and Modeling 15(4), 377–386 (2004)
9. Voronenko, A. A.: On checking tests for read-once functions. In: Matematicheskie Voprosy Kibernetiki, vol. 11, pp. 163–176. Fizmatlit, Moscow (2002) (in Russian)
10. Voronenko, A. A.: On the length of checking test for repetition-free functions in the basis $\{0, 1, \&, \vee, \neg\}$. Discrete Mathematics and Applications 15(3), 313–318 (2005)
11. Voronenko, A. A.: Recognizing the nonrepeating property in an arbitrary basis. Computational Mathematics and Modeling 18(1), 55–65 (2007)
12. Voronenko, A. A., Chistikov, D. V.: Learning read-once functions individually. Uchenye Zapiski Kazanskogo Universiteta, ser. Fiziko-Matematicheskie Nauki 151(2), 36–44 (2009) (in Russian)
13. Voronenko, A. A., Chistikov, D. V.: On testing read-once Boolean functions in the basis $B_5$. In: Proceedings of the XVII International Workshop "Synthesis and complexity of control systems", pp. 24–30. Izdatel'stvo Instituta matematiki, Novosibirsk (2008) (in Russian)

## Appendix

*Proof of Lemma 14.* Consider arbitrary positive integers $r_1, \ldots, r_l$. Suppose that $L(r_1, \ldots, r_l) = t$ and $\epsilon_1, \ldots, \epsilon_l$ are equivalence relations on the set $\{1, \ldots, t\}$ from the definition of $L(r_1, \ldots, r_l)$. Also suppose that equivalence classes of each $\epsilon_i$ are denoted by $S_{i,j}$, where $j = 1, \ldots, r_i$ and the class $S_{i,r_i}$ contains $t$. Take $m = t - 1$ and choose $\mathbf{e}_{i,j} \in \mathbf{Z}_2^m$ such that $\operatorname{supp} \mathbf{e}_{i,j} = S_{i,j}$, for $j = 1, \ldots, r_i - 1$. Clearly, $\operatorname{supp} \mathbf{e}_{i,j'}$ and $\operatorname{supp} \mathbf{e}_{i,j''}$ are disjoint for $j' \neq j''$. Vectors $\mathbf{e}_{i,1}, \ldots, \mathbf{e}_{i,r_i-1}$ are linearly independent and span an $(r_i - 1)$-dimensional subspace $V_i$ in $\mathbf{Z}_2^m$. Finally, if a non-zero vector $\mathbf{x} \in V_i \cap V_k$, then equivalence relations $\epsilon_i$ and $\epsilon_k$ both imply (are refinements of) another equivalence relation $\epsilon$ on $\{1, \ldots, t\}$ such that $p \sim_\epsilon q$ iff $p$ and $q$ are both elements of either $\operatorname{supp} \mathbf{x}$ or $\{1, \ldots, t\} \setminus \operatorname{supp} \mathbf{x}$, which is a contradiction.

Conversely, if $m$ is an integer satisfying the conditions of the lemma, take $t = m + 1$ and choose equivalence relations $\epsilon_i$ on $\{1, \ldots, t\}$ in the following way. Put $p \sim_{\epsilon_i} q$ iff $p$ and $q$ either belong to the same $\operatorname{supp} \mathbf{e}_{i,j}$ or belong to none of $\operatorname{supp} \mathbf{e}_{i,j}$ for $1 \leq j \leq r_i - 1$. Clearly, each $\epsilon_i$ splits $\{1, \ldots, t\}$ into $r_i$ equivalence classes. Now consider $\epsilon = \epsilon_i \vee \epsilon_k$. If $\epsilon \neq \mathbf{true}$, then any non-empty equivalence class in $\{1, \ldots, t\}$ (as partitioned by $\epsilon$) not containing $t$ can be represented both as a union of several $\operatorname{supp} \mathbf{e}_{i,j'}$ for some $j'$ and as a union of several $\operatorname{supp} \mathbf{e}_{k,j''}$ for some $j''$. In other words, there exists a non-zero vector $\mathbf{x}$ contained in both $V_i$ and $V_k$, which contradicts our initial assumptions.