DNF tautologies with a limited number of occurrences of every variable

P. Savický*   Jiří Sgall†

Abstract

It is known that every DNF tautology with all monomials of length k contains a variable with at least $\Omega(2^k/k)$ occurrences. It is not known, however, if this bound is tight, i.e. if there are tautologies with at most $O(2^k/k)$ occurrences of every variable. DNF tautologies with 2^k monomials of length k and with at most $2^k/k^\alpha$ occurrences of every variable, where $\alpha = \log_3 4 - 1 \geq 0.26$ are presented. This has the following consequence.

Let (k,s)-SAT be k-SAT restricted to instances with at most s occurrences of every variable. It is known that for every k, there is an s_k such that (k, s_k)-SAT is NP-complete and $(k, s_k - 1)$-SAT is trivial in the sense that every instance has positive answer. The above result implies that $s_k \leq 2^k/k^\alpha$. This improves the previously known bound $s_k \leq \frac{11}{32}2^k$.

1 Introduction

Let a (k,s)-formula be a formula in conjunctive normal form whose clauses consist of exactly k literals and such that every variable has at most s occurrences. Let (k,s)-SAT be the satisfiability problem restricted to (k,s)-formulas, see [3]. It is known that if there is an unsatisfiable (k,s)-formula for some k and s, then the (k,s)-SAT problem is NP-complete, see [2]. Let s_k be the minimum s for which an unsatisfiable (k,s)-formula exists. We prove that $s_k \leq 2^k/k^\alpha$, where $\alpha = \log_3 4 - 1 \geq 0.26$. This improves the previously known bound $s_k \leq \frac{11}{32}2^k$, see [1].

Using Lovasz local lemma, it is possible to prove (see [2]) that for any (k,s)-formula with $s \leq 2^k/(ek)$, a random assignment satisfies it with positive probability. It follows that $s_k \geq 2^k/(ek)$. It remains as an open problem to close the gap between this lower bound and the previously mentioned upper bound on s_k.

In order to prove the upper bound on s_k, we construct an unsatisfiable (k,s)-formula for appropriate s and k. For technical reasons, we formulate the construction in a dual form, i.e. we construct a formula in a disjunctive normal form that is a tautology, consists of monomials of length k and contains at most s occurrences of every variable. The negation of such a formula is the required unsatisfiable (k,s)-formula.

*Institute of Computer Science, Academy of Sciences of Czech Republic, Pod vodárenskou věží 2, 182 07 Praha 8, Czech Republic, e-mail: `savicky@uivt.cas.cz` . The research was supported by GA of the Czech Republic, grant No. 201/98/0717.

†Mathematical Inst., AS CR, Žitná 25, 115 67 Praha 1, Czech Republic and Dept. of Applied Mathematics, Faculty of Mathematics and Physics, Charles University, Praha, `sgall@math.cas.cz`, `http://www.math.cas.cz/~sgall/`. Partially supported by grant A1019602 of GA AV ČR, posdoctoral grant 201/97/P038 of GA ČR, and cooperative research grant INT-9600919/ME-103 from the NSF (USA) and the MŠMT (Czech republic).

2 DNF tautologies

Assume, a tautology over n variables consists of 2^k monomials of length k. Since every monomial is satisfied by 2^{n-k} assignments of the variables, it follows that every assignment of the variables satisfies exactly one monomial.

We define the *support* of a monomial to be the set of variables contained in it.

Lemma 1 *Let ϕ be a tautology with 2^k monomials of length k and let I be a set of variables of size k. Then, there is an even number of monomials with the support I. Moreover, among these monomials, the same number of monomials contain an odd and an even number of negated variables.*

Proof: Set all variables that do not belong to I to zero. After this setting, ϕ becomes a tautology, where the monomials with original support I remains without any change and all the others are either removed (set to zero) or their support is a proper subset of I. The settings of variables in I form a Boolean cube of size $2^{|I|}$. Every monomial with the support I covers exactly one point of the cube. Every other monomial covers a subcube of dimension at least one. It follows that such a monomial covers an even number of points and, moreover, it covers the same number of points with an odd and with an even parity. The lemma follows. $\square$

Similarly we can prove that for any support I, if there is no monomial with support intersecting I in more than i points then the number of monomials with support I must be divisible by $2^{|I|-i}$. This gives some information about the possible structure of the tautologies we study, however, we do not need this generalization for our result.

Definition 2 A DNF tautology is called a (k, ε)-tautology, if it consists of 2^k monomials of length k and every variable occurs in at most $\varepsilon 2^k$ monomials.

By an occurrence of a variable, we mean an occurrence of the variable or its negation. By a negative occurrence we mean an occurrence of the negation of the variable.

Lemma 3 *Assume that there is a (k, ε)-tautology and a (k', ε')-tautology. Then, there is also a $(kk', \varepsilon\varepsilon')$-tautology.*

Proof: Let ϕ be a (k, ε)-tautology over the variables $x_1, x_2, \ldots, x_n$ and let ϕ' be a (k', ε')-tautology over the variables $x'_1, x'_2, \ldots, x'_{n'}$.

By Lemma 1, the set of monomials of ϕ' may be split into two disjoint subsets of size $2^{k'-1}$ in such a way that every variable has at most $\varepsilon' 2^{k'-1}$ occurrences in each of the subsets. Consider such a splitting of monomials of ϕ' into two sets of size $2^{k'-1}$ and let ϕ'_0 and ϕ'_1 be the formulas formed by the disjunction of the first and the second set of monomials respectively. Moreover, let $\phi'_{a,j}$ for $a \in \{0, 1\}$ and $j = 1, 2, \ldots, n$ be a copy of ϕ'_a, where a new variable $x'_{i,j}$ is substituted for the variable x'_i for all $i = 1, 2, \ldots, n'$.

Consider a DNF formula ψ constructed from ϕ by replacing every negative occurrence of x_j by a new variable y_j for all $j = 1, 2, \ldots, n$. The formula ψ is such that it has the value 1, if $x_j \vee y_j = 1$ for all $j = 1, 2, \ldots, n$.

Finally, construct a formula ω over nn' variables $x'_{i,j}$, $i = 1, 2, \ldots, n'$, $j = 1, 2, \ldots, n$, from the formula ψ by substituting $\phi'_{1,j}$ for x_j and $\phi'_{0,j}$ for y_j for $j = 1, 2, \ldots, n$ and by expanding the conjunctions of disjunctions that appear after the substitution. The formula

ω is a tautology, since for every assignment of the variables of ω and every $j = 1, 2, \ldots, n$, the formulas substituted for x_j and y_j guarantee the condition $x_j \vee y_j = 1$ mentioned above.

In order to prove further properties of ω, let us partition the monomials of ω into 2^k groups. Each group consists of the monomials obtained by the expansion of the conjunction obtained from a single monomial of ψ by the above substitution. The formulas $\phi'_{a,j}$ substituted for the variables of ψ consist of $2^{k'-1}$ monomials of length k'. Every monomial of ψ has length k. It follows that every group consists of $2^{(k'-1)k}$ monomials of length kk'. Since there are 2^k groups, ω consists of $2^{kk'}$ monomials of length kk'.

Let us estimate the number of occurrences of single variables of ω. Select one of the pairs i, j, $i = 1, 2, \ldots, n'$, $j = 1, 2, \ldots, n$, and consider one of the groups of the monomials of ω described in the previous paragraph. The variable $x'_{i,j}$ has no occurrence in $\phi'_{a,l}$ for $l \neq j$. Hence, if the group corresponds to a monomial of ψ that has no occurrence of x_j and y_j, then the variable $x'_{i,j}$ has no occurrence in the considered group. Hence, at most $\varepsilon 2^k$ groups can contain an occurrence of $x'_{i,j}$. Consider a group corresponding to a monomial of ψ with an occurrence of x_j or y_j. There is only one such occurrence in the considered monomial. The formula $\phi'_{a,j}$ substituted for this occurrence contains at most $\varepsilon' 2^{k'-1}$ monomials containing $x'_{i,j}$. Hence, the group contains at most $\varepsilon' 2^{k'-1} \cdot 2^{(k'-1)(k-1)} = \varepsilon' 2^{(k'-1)k}$ occurrences of $x'_{i,j}$. Altogether, there is at most $\varepsilon 2^k \cdot \varepsilon' 2^{(k'-1)k} = \varepsilon\varepsilon' 2^{kk'}$ occurrences of $x'_{i,j}$ in ω. □

Theorem 4 *For every $l \geq 1$ and $k = 3^l$, there exists a DNF tautology with 2^k monomials of length k and at most $2^k/k^\alpha$ occurrences of every variable, where $\alpha = \log_3 4 - 1 \geq 0.26$.*

Proof: It is easy to see that for every $l \geq 1$ any $(3^l, (3/4)^l)$-tautology satisfies the conditions of the theorem. In order to prove the existence of such a tautology, we use Lemma 1. In view of this lemma, it is sufficient to demonstrate a $(3, 3/4)$-tautology. The required tautology is

$$x_1\bar{x}_2x_4 \vee x_2\bar{x}_3x_4 \vee x_3\bar{x}_1x_4 \vee x_1x_2x_3 \vee$$

$$x_1\bar{x}_2\bar{x}_4 \vee x_2\bar{x}_3\bar{x}_4 \vee x_3\bar{x}_1\bar{x}_4 \vee \bar{x}_1\bar{x}_2\bar{x}_3.$$

□

References

[1] O. Dubois, On the r,s–SAT satisfiability problem and a conjecture of Tovey, *Discr. Appl. Math.* 26 (1990), pp. 51-60.

[2] J. Kratochvíl, P. Savický, Zs. Tuza, One more occurrence of variables makes satisfiability jump from trivial to NP-complete, *SIAM J. Comput* Vol. 22, No. 1, pp. 203–210.

[3] C. A. Tovey, A simplified satisfiability problem, *Discrete Appl. Math.* 8 (1984), pp. 85–89.