НЕКОТОРЫЕ ЗАМЕТКИ О СЛОЖНОСТЯХ ВЫВОДОВ В СИСТЕМАХ ФРЕГЕ

Чубарян А.А.
Факультет информатики и прикладной математики, Ереванский государственный университет,Российско-Армянский университет, Ереван, доктор физ.мат.наук,профессор
Петросян Г.В.
Факультет информатики и прикладной математики, Ереванский государственный университет, Ереван, магистрант

SOME NOTES ON PROOF COMPLEXITIES IN FREGE SYSTEMS

Chubaryan A.A.
Department of Informatics and Applied Mathematics, ,YerevanState University, Russian-Armenian University, Yerevan, Doctor of Sciences, Full Professor
Petrosyan G.W.
Department of Informatics and Applied Mathematics,Yerevan State University, Yerevan, Master student

АННОТАЦИЯ

В настоящей статье мы представляем некоторые результаты о сложностях выводов в системах Фреге. Сначала мы вводим понятие формулы *специального* вида и доказываем, что множество всех тавтологий длины *n* имеют полиномиально ограниченные сложности выводов в системах Фреге тогда и только тогда, когда таковыми являются выводы формул *специального* вида длины *n*. Далее мы доказываем, что все балансированные тавтологии в дизъюнктивной нормальной форме длины *n* также имеют полиномиально ограниченные сложности выводов в системах Фреге. В конце даются несколько заметок об соотношениях сложностей выводов тавтологий A_n , B_n и формул в форме A_n*B_n, где $*$ одна из цвязок $\wedge,\vee$ или $\supset$.

ABSTRACT

We present in this paper some results about Frege proof complexities. At first we introduce the notion of *specific* tautologies and show that Frege systems must have a polynomial size $p(n)$ proof for every tautology of size *n* iff the proofs of all specific tautologies of size *n* are polynomially bounded. Then we show, that all *balanced* tautologies in disjunctive normal form of size *n* also have Frege proofs with polynomially bounded sizes. Lastly we give some notes about relations between the proof complexities of tautologies A_n and B_n and proof complexities of the tautologies in a form A_n*B_n, where $*$ is $\wedge,\vee$ *or* $\supset$.

Ключевые слова: теория сложности, теория сложностей пропозициональных выводов, сложности выводов, системы Фреге, балансированные формулы.

Keywords: complexity theory, propositional proof theory, proof complexity, Frege systems, balanced formulas.

1. Introduction

One of the most fundamental problems of the proof complexity theory is to find an efficient proof system for classical propositional calculus. There is a wide spread understanding that polynomial time computability is the correct mathematical model of feasible computation. According to the opinion, a truly "effective" system must have a polynomial size $p(n)$ proof for every tautology of size n. In [1] Cook and Reckhow named such a system a *super system*. They showed that NP = *co*NP iff there exists a super system. Lately it is proved in [2] that NP=PSPACE by showing that arbitrary tautologies of Johansson's minimal propositional logic admit “small" polynomial-size dag-like natural deductions in Prawitz's system for minimal propositional logic. As corollary from this result follows that NP = *co*NP = PSPACE, hence it must be some propositional proof system, which is super system. It is well known that many systems are not super. This question about Frege system, the most natural calculi for propositional logic, is still open.

In this paper we present some results about Frege proof complexities. At first we introduce the notion of *specific* tautologies and show that Frege systems will be super iff the proofs of all specific tautologies of size *n* are polynomially bounded. Then we show, that all balanced tautologies in disjunctive normal form of size *n* also have proofs with polynomially bounded sizes. Lastly we give some results about relations between the proof complexities of tautologies A_n and B_n and proof complexities of the tautologies in a form A_n*B_n, where $*$ is $\wedge,\vee$ *or* $\supset$.

2. Preliminaries

We will use the current concepts of the unit Boolean cube (E^n), a literal, a propositional formula, a disjunctive normal form (DNF), a classical tautology, Frege proof systems for classical propositional logic, proof and proof complexity [1]. Let us recall some of them.

A Frege system $\mathcal{F}$ uses a denumerable set of propositional variables, a finite, complete set of propositional connectives; $\mathcal{F}$ has a finite set of inference rules

defined by a *figure* of the form $\frac{A_1A_2\dots A_m}{B}$ (the rules of inference with zero hypotheses are the axioms schemes); $\mathcal{F}$ must be sound and complete, i.e. for each rule of inference $\frac{A_1A_2\dots A_m}{B}$ every truth-value assignment, satisfying $A_1A_2\dots A_m$, also satisfies B, and $\mathcal{F}$ must prove every tautology.

The particular choice of a language for presented propositional formulas is immaterial in this consideration. However, because of some technical reasons we assume that the language contains the propositional variables p_i ($i \geq 1$) and (or) p_{ij} ($i \geq 1$; $j \geq 1$), logical connectives $\neg, \wedge, \vee, \supset$ and parentheses (,). Note that some parentheses can be omitted in generally accepted cases. Note that our convention for serial disjunction $A_1 \vee A_2 \vee \dots \vee A_k$ (conjunction $A_1 \wedge A_2 \wedge \dots \wedge A_k$) is associated from left to right.

By $|\varphi|$ we denote the size of a formula φ, defined as the number of all logical signs entries in it. It is obvious that the full size of a formula, which is understood to be the number of all symbols is bounded by some linear function in $|\varphi|$.

2.1. Proof complexity measures

In the theory of proof complexity four main characteristics of the proof are: *t*- complexity (length), defined as the number of proof steps, *l*-complexity (size), defined as sum of sizes for all formulas in proof (size), *s*-complexity (space), informal defined as maximum of minimal sum of sizes for formulas on blackboard needed to verify all steps in the proof (formal definitions are for example in [3]) and *w*- complexity (width), defined as the maximum of sizes of proof formulas.

Let ϕ be a proof system and φ be a tautology. We denote by $t_\varphi^\phi(l_\varphi^\phi, s_\varphi^\phi, w_\varphi^\phi)$ the minimal possible value of *t*-complexity (*l*-complexity, *s*-complexity, *w*-complexity) for all ϕ-proofs of tautology φ.

By analogy we can define the mentioned proof complexity characteristics for the proof of any formula A from premises Γ and denote them respectively by $t_{\Gamma\vdash A}^\Phi(l_{\Gamma\vdash A}^\Phi, w_{\Gamma\vdash A}^\Phi, s_{\Gamma\vdash A}^\Phi)$.

Let M be some set of tautologies.

Definition 2.1.1. We call the ϕ-proofs of tautologies from a set M *t*-polynomially (*l*- polynomially, *s*-polynomially, *w*- polynomially) bounded if there is a polynomial p() such that $t_\varphi^\phi \leq$ p(|φ|) ($l_\varphi^\phi \leq$ p(|φ|), $s_\varphi^\phi \leq$ p(|φ|), $w_\varphi^\phi \leq$ p(|φ|)) for all φ from M.

Definition 2.1.2. We call the ϕ-proofs of tautologies from a set M *t*-linearly (*l*- linearly, *s*- linearly, *w*-linearly) bounded if there is a linear function *f()* such that $t_\varphi^\phi \leq$ *f*(|φ|) ($l_\varphi^\phi \leq$ *f*(|φ|), $s_\varphi^\phi \leq$ *f*(|φ|), $w_\varphi^\phi \leq$ *f*(|φ|)) for all φ from M.

2.2. Essential subformulas of tautologies

For proving the main results we use also the notion of *essential subformulas*, introduced in [4].

Let *F* be some formula and $Sf(F)$ be the set of all non-elementary subformulas of formula F.

For every formula F, for every $\varphi \in Sf(F)$ and for every variable p by F_φ^p is denoted the result of the replacement of the subformulas φ everywhere in F by the variable p. If $\varphi \notin Sf(F)$, then F_φ^p is F.

We denote by $Var(F)$ the set of variables in F.

Definition 2.2.1. Let p be some variable that $p \notin Var(F)$ and $\varphi \in Sf(F)$ for some tautology F. We say that φ is an ***essential subformula*** in F iff F_φ^p is non-tautology.

We denote by $Essf(F)$ the set of essential subformulas in tautology F.

If F is minimal tautology, i.e. F is not a substitution of a shorter tautology, then $Essf(F) = Sf(F)$.

In [4] the following statement is proved.

Proposition 1. Let F be a minimal tautology and $\varphi \in Essf(F)$, then in every $\mathcal{F}$-proof of F subformula φ must be essential either at least in some axiom, used in proof or in formulae $A_1 \supset (A_2 \supset (\dots \supset A_m)\dots) \supset B$ for some used in proof inference rule $\frac{A_1A_2\dots A_m}{B}$.

Note (1) that for every Frege system the number of mentioned essential subformulas is bounded with some constant.

Definition 2.2.2 Let p and q be some variables that $p \notin Var(A)$ and $q \notin Var(A)$ for some tautology A, φ and ψ are subformulas of A such that neither φ nor ψ are subformula of each other. We say that φ and ψ are an ***essential pair of subformulas*** in A iff $A_{\varphi,\psi}^{p,q}$ is non-tautology.

By analogy to statement of Proposition 1. it is not difficult to prove the following statement.

Proposition 2. Let A be a minimal tautology and pair φ, ψ belong to the set of *essential pairs of subformulas* in A, then in every $\mathcal{F}$-proof of A pair φ, ψ must be essential either at least in some axiom, used in proof or in formulae $A_1 \supset (A_2 \supset (\dots \supset A_m)\dots) \supset B$ for some used in proof inference rule $\frac{A_1A_2\dots A_m}{B}$.

Definition 2.2.3 The set of *essential pairs of subformulas* in a tautology A is called ***canonic*** if every subformula of A has entry into no more than one pair of this set.

Note (2) that for every Frege system the number of pairs in the *canonic set of essential pairs of subformulas* both for every axiom and for formulae $A_1 \supset (A_2 \supset (\dots \supset A_m)\dots) \supset B$ of every proof inference rule $\frac{A_1A_2\dots A_m}{B}$ is bounded with some constant. Really by Note (1) the number of essential subformulas both for every axiom and for formulae $A_1 \supset (A_2 \supset$

$(\ldots\supset A_m)\ldots)\supset$ B of every proof inference rule $\frac{A_1A_2\ldots A_m}{B}$ is bounded with some constant and they can be only in one pair from *canonic set of essential pairs of subformulas*.

Some of definitions can be given further.

3. Main results.

Here we give the main results, mentioned in Introduction.

3.1.The role of specific formulas

Definition 3.1. Any propositional formula A is called *specific* if it is in the following form: $A = p \supset (A_1 \vee A_2 \vee \ldots \vee A_k)$ $(k \geq 1)$, where p is a literal (variable or negation of variable), neither $A_1 \vee A_2 \vee \ldots \vee A_k$ nor every $A_i (1 \leq i \leq k)$ are tautology or contradiction and $|A_i| \leq \frac{|A_1|}{2^{i-1}}$.

Theorem 1. Let M be the set of all specific tautologies. If $\mathcal{F}$-proofs of formulas from the set M are *t*-polynomially (*l*- polynomially, *s*- polynomially, *w*-polynomially) bounded, then $\mathcal{F}$-proofs of all tautologies are *t*-polynomially (*l*- polynomially, *s*- polynomially, *w*- polynomially) bounded.

For proving this theorem we at first give two auxiliary statements.

Lemma 3.1. If for a $\mathcal{F}$-proof of any formula B from premises Γ, A $t^{\mathcal{F}}_{\Gamma,A\vdash B}(l^{\mathcal{F}}_{\Gamma,A\vdash B}, w^{\mathcal{F}}_{\Gamma,A\vdash B}, s^{\mathcal{F}}_{\Gamma,A\vdash B}) \leq n$, then $t^{\mathcal{F}}_{\Gamma\vdash A\supset B}\left(l^{\mathcal{F}}_{\Gamma\vdash A\supset B}, w^{\mathcal{F}}_{\Gamma\vdash A\supset B}, s^{\mathcal{F}}_{\Gamma\vdash A\supset B}\right) \leq cn$ for some constant *c*.

Proof of this statement follows obviously from the proof of deduction theorem.

Lemma 3.2. The $\mathcal{F}$-proofs of formulas from the following set

1) $C \supset (B \supset C)$
2) $B \vee C \equiv \neg B \supset C$
3) $\neg B \supset (B \supset C)$
4) $C \supset (B \supset C \wedge B)$
5) $(\neg C \supset \neg B) \equiv (B \supset C)$
6) $\neg(B \vee C) \equiv \neg B \wedge \neg C$
7) $\neg(B \wedge C) \equiv \neg B \vee \neg C$
8) $\neg(B \supset C) \equiv B \wedge \neg C$

are *t*-linearly (*l*- linearly, *s*- linearly, *w*- linearly) bounded for every formulas *B* and *C*.

Proof of this statement is obvious.

Proof of Theorem 1. Given tautology *A* can be in one of the following form a) $A = B \wedge C$, b) $A = B \supset C$, c) $A = B \vee C$, d)$\neg A_1$. Use the tautologies 2),6),7),8) of Lemma 2. the formula A can be presented in the form a) or b), hence we can observe only these cases.

In the case a) the formulas B and C must be tautology, hence we can derive every of them and then use a proof of formula $C \supset (B \supset C \wedge B)$ derive the formula A. So in the case a) proof of A reduces to proofs of two smaller tautologies.

Let we have the case b). If the formula C is tautology, then we can derive the formula C, then use a proof of $C \supset (B \supset C)$ we can derive A. If the formula B is contradiction, then we can derive the formula $\neg B$, then use a proof of $\neg B \supset (B \supset C)$ we can derive A. In the other cases i) if $|B| \geq |C|$, then we can at first derive C from the premise B and by deduction theorem derive A with no more than linear increase of complexities characteristics (Lemma 3.1.), ii) if $|C| > |B|$, then we can at first derive $\neg B$ from the premise $\neg C$ and by deduction theorem derive $(\neg C \supset \neg B)$ also with no more than linear increase of complexities characteristics (Lemma 3.1.) and use the proof of formula $(\neg C \supset \neg B) \supset (B \supset C)$ derive A. So in the case b) proof of A reduces to proof C from the premise B or to proof $\neg B$ from the premise $\neg C$.

Later we must analyze as above the formulas B and C in the case a) and the formula C or the formula $\neg B$ in the case b). Note that in the last case we must already take into consideration the truth values of premises also.

We do all mentioned steps until we must obtain a proof of some literal p from the premises $A_1, A_2, \ldots, A_k$, every of each is either some subformula or negation of some subformula of formula . In the other words we must derive the formula $(A_1 \wedge A_2 \wedge \ldots \wedge A_k) \supset p$, which we can do using the proof of specific formula $\neg p \supset \neg A_1 \vee \neg A_2 \vee \ldots \vee \neg A_k$.

So the proof of tautology A reduces to proofs of no more than $|A|$ numbers specific tautologies. Every step of reducing gives to the size (the others complexities characteristics) of proof no more than linear increase with some constant *c* and number of steps is no more than $\log_2|A|$, hence general increasing is no more than $c^{\log_2|A|}$, which is no more than$|A|^{\frac{c}{2}}$. Note that every step of reducing can be also do inversely with the same property. □

3.2.Proof complexities of balanced DNF

Here we investigate the proof complexities characteristics for some set of tautologies, given in DNF (DNF-tautologies).

Definition 3.2.1. A propositional formula called *balanced* if every variable has only two occurrences in it: one positive and one with negative.

It is shown in [5] that problem on polynomially bounded sizes of proofs for all tautologies reduced to problem on polynomially bounded sizes of proofs for all balanced tautologies.

Definition 3.2.2. DNF-tautology is *correct* if DNF, obtained from it by removing of any conjunct, is no tautology already.

Lemma 3.2. The number of conjuncts in balanced correct DNF-tautology with *n* variables is *n+1*.

Proof is given by induction on number n of variables in a balanced DNF-tautology *A*. By n=1 we have $A = p \vee \neg p$. Suppose that statement is valid for number of variables $\leq$ n. If the number of variables is n+1, then correct DNF must have at least one conjunct with at least two literals $p_{i_1}^{\alpha_1}, p_{i_2}^{\alpha_2}$. After assigning $\boldsymbol{\alpha}_1$ to variable p_{i_1} everywhere in given DNF, both the number of variables and the number of conjuncts decrease with one, hence the number of conjuncts in primary DNF must be n+2.

Corollary. Every balanced correct DNF-tautology has at least one conjunct, consisting from one literal.

Theorem 2. The $\mathcal{F}$-proofs of all balanced correct DNF-tautologies are *t*-polynomially (*l*- polynomially, *s*- polynomially, *w*- polynomially) bounded.

Proof. Let the variables of given balanced correct DNF-tautology $A = K_1 \vee K_2 \vee \ldots \vee K_{n+1}$ are $p_1,p_2,\ldots,p_n$. To derive A we take $\neg A$ as premise and proof a contradiction. We have $\neg A = D_1 \wedge D_2 \wedge \ldots \wedge D_{n+1}$, where $D_i = \neg K_i$ in a form $D_i = p_{i_1}^{\alpha_1} \vee p_{i_2}^{\alpha_2} \vee \ldots \vee p_{i_r}^{\alpha_r}$ $(\alpha_i \in \{0,1\}; (1 \le i \le n+1)$. On the base of D_i we construct the formula E_i by adding instead of every variable p_j from $p_1,p_2,\ldots,p_n$, which has no occurrence in D_i, the formula $p_1 \wedge \neg p_1$ on the j-th place with disjunction. It is obvious that $D_i = E_i$ and this equivalence can derive with polynomially bounded characteristics of proof complexities. By these notation we have that formula $\neg A$ is equivalent to formula $A' = \wedge_{i=1}^{s} E_i$. Now we introduce the new propositional variables p_{ij} $(1 \le i \le s, 1 \le j \le n)$, where p_{ij} is true, if variable p_j has occurrence in D_i and p_{ij} is false for the opposite case, and construct on the base of E_i the new disjunctions D'_i by replacement both primary literals and formulas $p_1 \wedge \neg p_1$ by the corresponding variables p_{ij}.

Now we take with consideration that the well known formulas of Pigeon Hole Principle $PHP_n = \wedge_{i=0}^{n} \vee_{k=0}^{n-1} p_{ik} \supset \vee_{k=0}^{n-1} \vee_{0 \le i < j \le n} (p_{ik} \wedge p_{jk})$ have polynomially bounded size of $\mathcal{F}$-proofs [1]. Use this fact we can derive the formulas $\wedge_{i=1}^{n+1} D'_i \supset \vee_{j=1}^{n} \vee_{1 \le i < k \le n+1} (p_{ij} \wedge p_{kj})$ and then by *modus ponens* derive the formulas $C_n = \vee_{j=1}^{n} \vee_{1 \le i < k \le n+1} (p_{ij} \wedge p_{kj})$. Denote by H_n the formulas $\wedge_{j=1}^{n} \wedge_{1 \le i < k \le n+1} \neg(p_{ij} \wedge p_{kj})$. It is not difficult to derive formulas $H_n \supset \neg C_n$ with polynomially bounded size. If for every i, j we denote by $[p_{ij}]$ either the primary literal or formula $p_1 \wedge \neg p_1$ from formula E_i, then it is obvious, that

a) every formula $[H_n] = \wedge_{j=1}^{n} \wedge_{1 \le i < k \le n+1} \neg([p_{ij}] \wedge [p_{kj}])$ is tautology,

b) the formulas $[H_n]$, $[H_n] \supset \neg(\vee_{j=1}^{n} \vee_{1 \le i < k \le n+1} ([p_{ij}] \wedge [p_{kj}]))$ and

$[PHP_n] = \wedge_{i=0}^{n} \vee_{k=0}^{n-1} [p_{ik}] \supset \vee_{k=0}^{n-1} \vee_{0 \le i < j \le n} ([p_{ik}] \wedge [p_{jk}])$ have polynomially bounded size of $\mathcal{F}$-proofs [1] , hence we can derive contradiction. □

3.3.**Some properties of $\mathcal{F}$-proofs for the formulas in a form A∗B.**

Here we investigate the complexity characteristics of $\mathcal{F}$-proofs for formulas in the form $A_n = B_n * C_n$, where B_n and C_n are tautologies and $*$ is $\wedge, \vee$ *or* $\supset$.

Theorem 3.

1.a) There are tautologies B_n and C_n such that $t_{B_n}^{\mathcal{F}} = t_{C_n}^{\mathcal{F}} = \Theta(n)$, $l_{B_n}^{\mathcal{F}} = l_{C_n}^{\mathcal{F}} = \Theta(n^2)$, but for $A_n = B_n \vee C_n$ $t_{A_n}^{\mathcal{F}} = \Theta(1)$ and $l_{A_n}^{\mathcal{F}} = \Theta(n)$,

b) There are tautologies B_n and C_n such that $t_{B_n}^{\mathcal{F}} = t_{C_n}^{\mathcal{F}} = \Theta(n)$, $l_{B_n}^{\mathcal{F}} = l_{C_n}^{\mathcal{F}} = \Theta(n^2)$ and for $A_n = B_n \vee C_n$ $t_{A_n}^{\mathcal{F}} = \Theta(n)$ and $l_{A_n}^{\mathcal{F}} = \Theta(n^2)$ also.

2. For every tautologies B_n and C_n if $A_n = B_n \wedge C_n$, then $t_{A_n}^{\mathcal{F}} = \Theta(\max(t_{B_n}^{\mathcal{F}}, t_{C_n}^{\mathcal{F}}))$ and $l_{A_n}^{\mathcal{F}} = \Theta(\max(l_{B_n}^{\mathcal{F}}, l_{C_n}^{\mathcal{F}}))$.

3.a) For every tautologies B_n and C_n if $A_n = B_n \supset C_n$, then $t_{A_n}^{\mathcal{F}} = O(t_{C_n}^{\mathcal{F}})$ and $l_{A_n}^{\mathcal{F}} = O(l_{C_n}^{\mathcal{F}})$.

b) There are tautologies B_n and C_n such that for $A_n = B_n \supset C_n$ $t_{A_n}^{\mathcal{F}} = \Omega(t_{C_n}^{\mathcal{F}})$ and $l_{A_n}^{\mathcal{F}} = \Omega(l_{C_n}^{\mathcal{F}})$.

Proof. The main role for proof of some points of this theorem plays the tautologies $D_n(p) = \overbrace{\neg\neg \ldots \neg}^{2n}(\neg p \vee p)$, the number of essential subformulas of which is *n*.

For proving the point 1.a) we use the properties of essential subformulas for tautologies $B_n = D_n(p) \vee q$, $C_n = \neg q \vee D_n(p)$ and $A_n = B_n \vee C_n$, last of which can be derive from the formulas $D_n(p) \vee ((q \vee \neg q) \vee D_n(p))$ with constant steps and linear sizes.

For proving the point 1.b) we use the properties for the canonic set of essential pair of subformulas for the tautologies $A_n(p,q) = (D_n(p)) \vee (D_n(q))$, the canonic set of essential pair of subformulas for which are the pairs $\overbrace{\neg\neg \ldots \neg}^{i}(\neg p \vee p)$, $\overbrace{\neg\neg \ldots \neg}^{i}(\neg q \vee q)$ for every $i \in \{1,2,\ldots,2n\}$.

Proof of the point 2. is obvious. Really, the formula $A_n = B_n \wedge C_n$ can be derived by using the derivations of tautologies B_n , C_n and $B \supset (C \supset (B \wedge C)$. Every of the formulas B_n and C_n can be derived by using the derivations of tautologies $B_n \wedge C_n$, $B \wedge C \supset B$ and $B \wedge C \supset C$.

Proof of the point 3.a) is obvious.

For proving the point 3.b) we use the properties of essential subformulas for tautologies $A_n(p,q) = (D_n(p)) \supset (D_n(q))$. Really, every essential subformula of tautologies $D_n(q)$ must be esential for tautologies $A_n(p,q)$ also. □

Acknowledgments

This work arose in the context of propositional proof complexity research supported by the Russian-Armenian University from founds of MESRF

References

1. S.A.Cook, A.R.Reckhow: The relative efficiency of propositional proof systems, Journal of Symbolic logic, vol. 44, 1979, 36-50.

2. L. Gordeev, E. H. Haeusler, NP vs PSPACE, arXiv:1609.09562v1 [cs.CC] 30 Sep 2016

3. Nordstrom Jakob. Narrow proofs may be spacious: Separating space and width in resolution. SIAM Journal on Computing, 39(1):59–121, May 2009.

4. A.A. Chubaryan, On complexity of the proofs in Frege system, CSIT Conference, Yerevan, 2001, 129-132.

5. Lutz Straßburger: Extension without Cut,INRIA Saclay–ˆIle-de-France and Ecole Polytechnique, LIX, Rue de Saclay, 91128 Palaiseau Cedex, France