Mathematical Problems of Computer Science 26, 2006, 117–122.

On Some Proporties of Frege Proofs

Sona R. Aleksanyan and Anahit A. Chubaryan

Department of Applied Mathematics, State Engineering University of Armenia,
Department of Informations and Applied Mathematics,
Yerevan State University,
e-mail: sonush@rambler.ru, achubaryan@ysu.am

Abstract

In [4] a measure s on propositional formula was defined such that for every tautology φ "high" value of $s(\varphi)$ requires the large size of proof in the "weak" propositional systems. In this paper it is shown, that there is a tautology φ, the measure $s(\varphi)$ of which has exponential dependence on the size of φ, but its proof complexity in Frege systems is polynomially bounded.

1 Introduction

It is well known that the investigations of the propositional proof complexity are very important due to their tight relation to the main problem of the complexity theory: $P \stackrel{?}{=} NP$. In particular, Cook and Reckhow proved, that $NP = coNP$ iff there is a polynomially bounded proof system for classical tautologies [6], therefore it is interesting to obtain "good" lower and upper bounds of proof complexities particularly in a Frege system – the most natural calculi for propositional logic.

In [4] a characteristic $s(\varphi)$ for every tautology φ was defined such that

1) for sufficiently large n and every $i \left(1 \le i \le \left[\frac{n}{2}\log_n 2\right]\right)$ there are sequences of tautologies φ_i^n of size n such that $s\left(\varphi_i^n\right) = \theta(n^i)$;

2) "high" measure of $s(\varphi)$ requires the large size of proof in the "weak" propositional systems (resolution, cut-free sequence, analytic tableaux) [2, 5];

3) the proof complexity of every tautology φ in the Frege system is no more than $c \cdot s(\varphi) \cdot |\varphi|$, where $|\varphi|$ is the size of φ and c is a constant.

From the third statement it follows that Frege proof complexity of a formula φ is polynomially bounded in the case if the measure $s(\varphi)$ has polynomial dependence on the size of φ.

It is interesting to investigate how long must be (can be) the size of Frege-proof, in the case if the measure of $s(\varphi)$ has exponential dependence on the size of φ for provable φ.

In this paper we prove that each of above mentioned formulae φ_i^n has polinomially bounded Frege-proof, i.e. the "high" measure of $s(\varphi)$ can be also "quickly arrived" in the Frege system.

2 Preliminary

In order to prove our main result, we recall some notions and notations. We shall use generally accepted concepts of unit Boolean cube (E^n), propositional formula, tautology, disjunctive normal form (DNF), propositional proof system, proof complexity.

The particular choice of language for representation of propositional formulae does not matter for our consideration, but because of technical reasons we assume that our language contains the propositional variables p_i ($i \geq 1$) and/or p_{ij} ($i \geq 1$, $j \geq 1$), logical symbols $\neg$, $\&$, $\vee$, $\supset$ and parenthesis (,). Note that some parentheses can be disregarded in generally accepted cases.

Following the usual terminology we call variables and negated variables *literals*; the conjunct $\mathcal{K}$ can be represented simply as a set of literals and is called a *clause* (no clause contains both a variable and the negation of that variable). A formula in DNF can be expressed as a set of clauses $\{\mathcal{K}_1, \mathcal{K}_2, \ldots, \mathcal{K}_\ell\}$.

The *elimination-rule* (ε-rule) infers $\mathcal{K}' \cup \mathcal{K}''$ from clauses $\mathcal{K}' \cup \{p\}$ and $\mathcal{K}'' \cup \{\bar{a}rp\}$, where $\mathcal{K}'$ and $\mathcal{K}''$ are clauses and p is a propositional variable.

We would like to say that the conjunct K is deduced from the DNF $\mathcal{D}$ if there is a finite sequence of such clauses, that every clause in the sequence is one of the clauses of $\mathcal{D}$ or inferred from earlier clauses in the sequence by ε-rule, and the last clause is $\mathcal{K}$.

If the empty conjunct (Λ) can be deduced from $\mathcal{D}$, then DNF $\mathcal{D}$ is called *full* (tautology).

The minimal number of the usages of ε-rule in the deduction of Λ from full DNF $\mathcal{D}$ is called *complexity* of $\mathcal{D}$ and denoted by $S(\mathcal{D})$.

Let φ be a propositional formula and the set of its distinct variables be $\{p_1, p_2, \ldots, p_n\}$. In [4] the following notions were introduced.

Definition 1. For some $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_m) \in E^m$ $(1 \leq m \leq n)$ the conjunct $K = \left\{p_{i_1}^{\sigma_1}, p_{i_2}^{\sigma_2}, \ldots, p_{i_m}^{\sigma_m}\right\}$ is called *φ-determinative* if the assignment of values σ_j to each p_{i_j} $(1 \leq m \leq n)$ induces the value of φ, without taking into consideration the values of the rest variables.

Let φ be a minimal tautology, i.e. φ is not an instance of a shorter tautology.

Definition 2. The full DNF $\mathcal{D}$ is called *φ-determinative* if every conjunct of $\mathcal{D}$ is φ-determinative.

Definition 3. Any φ-determinative DNF $\mathcal{D}$, having the minimal complexity, is called *minimal determinative* DNF for φ and is denoted by $\mathcal{D}_\varphi^{min}$.

Definition 4. The measure $S\left(\mathcal{D}_\varphi^{\min}\right)$ is called determinative complexity of φ for any minimal tautology φ and is denoted by $s(\varphi)$.

If φ is not minimal tautology, then $s(\varphi)$ is defined to be the minimal value $s(\varphi')$ for such minimal tautology φ', that φ is an instance of φ'.

Definition 5. If for a tautology φ there is a polynom $p()$ such that $s(\varphi) < p(|\varphi|)$, then φ is called *easy*, otherwise the tautology is *hard*.

The following statements about $\mathcal{D}_\varphi^{min}$ are proved in [4] and [5]:

1) If every φ-determinative conjunct contains at least m literals for any tautology φ, then $S\left(\mathcal{D}_\varphi^{min}\right) \geq 2^m$.

2) For every tautology φ of size n $\quad S\left(\mathcal{D}_\varphi^{min}\right) \leq 2^n$.

3) For sufficiently large n and every i $\left(1 \leq i \leq \left[\frac{n}{2}\log_n 2\right]\right)$ there are sequences of such tautologies φ_i^n, that $|\varphi_i^n| = \theta(n)$ and $s\left(\varphi_i^n\right) = \theta(n^i)$;

4) The proof complexity of formulae φ_i^n from mentioned sequence in some "weak" systems, like resolution system and cut-free sequence system, is $\theta(n^i)$.

The statements of the points 3) and 4) are proved, using the sequence of the following formulae:

$$\varphi_{n,m} = \bigvee_{(\sigma_1,\dots,\sigma_n)\in E^n} \mathop{\&}_{j=1}^{m} \bigvee_{i=1}^{n} p_{ij}^{\sigma_i} \qquad (n \geq 1,\ 1 \leq m \leq 2^n - 1).$$

It is not difficult to notice that for every fixed $n \geq 1$ and m $(1 \leq m \leq 2^n - 1)$ the formula $\varphi_{n,m}$ is tautology.

The above mentioned tautologies φ_i^n are accordingly φ_{n,n^i}. It is not difficult to notice also that $|\varphi_{n,m}| = \theta\left(2^n \cdot m \cdot n\right)$, the formulae $\varphi_{n,2^n-1}$ are minimal tautologies and every $\varphi_{n,2^n-1}$-determinative conjunct must contain $2^n - 1$ literals, therefore $s\left(\varphi_{n,2^n-1}\right) = 2^{\theta\left(\sqrt{|\varphi_{n,2^n-1}|}\right)}$, and hence the tautologies $\varphi_{n,2^n-1}$ are hard.

In this paper we investigate the relationship between the measure of $s(\varphi)$ and the proof complexity of φ in the Frege system.

We shall use the generally accepted concepts of Frege systems. Without loss of generality we assume, that the system $\mathcal{F}$ is a Frege system, the language of which contains the connectives $\neg$, &, $\vee$, $\supset$ perhaps together with the other connectives and *Modus ponens* is one of the deduction rules for it.

We use the well known notion of proof complexity.

We define for Frege proof ($\mathcal{F}$-proof) *t-complexity* to be its *length* (= the total number of lines) and *ℓ-complexity* to be its *size* (= the total number of symbols). The minimal n such that the formula φ has a proof of length $\leq n$ (of size $\leq n$) is called t-complexity (ℓ-complexity) of a formula φ in the system $\mathcal{F}$ and is denoted by $t_\varphi^{\mathcal{F}}$ $\left(\ell_\varphi^{\mathcal{F}}\right)$.

We would like to say that tautology φ has *t-polynomially bounded* (*ℓ-polynomially bounded*) $\mathcal{F}$-proof if there is a polynomial $p()$ such, that $t_\varphi^{\mathcal{F}} \leq p\left(|\varphi|\right)$ $\left(\ell_\varphi^{\mathcal{F}} \leq p\left(|\varphi|\right)\right)$.

It is obvious that the proof of Λ from $\mathcal{D}_\varphi^{min}$ for every minimal tautology φ can be transformed into a Frege proof using Kalmar's proof of deducibility of tautologies in classical propositional calculus [7].

In [4] it was shown, that $t_\varphi^{\mathcal{F}} \leq c \cdot s(\varphi)|\varphi|$ and $\ell_\varphi^{\mathcal{F}} \leq c \cdot s(\varphi)|\varphi|^2$ for every tautology φ, where $s(\varphi)$ is the above defined determinative complexity of φ, $|\varphi|$ is the size of φ and c is a constant, therefore every *easy* tautology has t-polynomially (ℓ-polynomially) bounded $\mathcal{F}$-proof.

3 Main results

Here it will be proved that the measure $s(\varphi)$ is insufficient to prove super-polynomial lower bounds for the Frege proof.

Theorem 1. For every $n \geq 1$ and $1 \leq m \leq 2^n - 1$ the tautology $\varphi_{n,m} = \bigvee_{(\sigma_1,\dots,\sigma_n)\in E^n} \mathop{\&}_{j=1}^{m} \bigvee_{i=1}^{n} p_{ij}^{\sigma_i}$ has t-polynomially (ℓ-polynomially) bounded $\mathcal{F}$-proof.

To prove this statement, it will be shown, that the $\mathcal{F}$-proof of $\varphi_{n,m}$ can be "polynomially reduced" to $\mathcal{F}$-proof of PHP_m – the well-known "Pigeonhole principle", which, in its turn, has ℓ-polynomially (hence t-polynomially) bounded $\mathcal{F}$-proof (see [1]).

We use also the following statement

Lemma 1. For every formulae α, β, γ, α_i, β_i, α_{ij}, β_{ij} each of the following formulae

1. $\alpha \vee \bar{a}r\alpha$

2. $\alpha \supset \alpha \vee \beta$

3. $(\alpha \supset \beta) \supset ((\beta \supset \gamma) \supset (\alpha \supset \gamma))$

4. $\left(\bar{\beta} \supset \alpha\right) \supset (\bar{a}r\alpha \supset \beta)$

5. $\alpha_1 \supset (\alpha_2 \supset (\ldots \supset (\alpha_k \supset \alpha_1 \& \alpha_2 \& \ldots \& \alpha_k) \ldots)) \; (k \geq 2)$

6. $\alpha \vee \bar{a}r\alpha \supset \beta_1 \vee \ldots \vee \beta_k \vee \alpha \vee \beta_{k+1} \vee \ldots \vee \beta_{k+r} \vee \bar{a}r\alpha \vee \beta_{k+r+1} \vee \ldots \vee \beta_{k+r+t}$
$(k \geq 1, r \geq 1, t \geq 1)$

7. $\neg\left(\bigvee\limits_{i=1}^{k} \mathop{\&}\limits_{j=1}^{m} \alpha_{ij}\right) \supset \mathop{\&}\limits_{i=1}^{k} \bigvee\limits_{j=1}^{m} \bar{a}r\alpha_{ij} \quad (k \geq 1, \; m \geq 1)$

8. $\mathop{\&}\limits_{i=1}^{k} (\beta_{1i} \vee \beta_{2i}) \supset \neg\left(\bigvee\limits_{i=1}^{k} \left(\bar{\beta}_{1i} \& \bar{\beta}_{2i}\right)\right) \qquad (k \geq 1)$

has t-polynomially (ℓ-polynomially) bounded $\mathcal{F}$-proof.

Proof. Proof of this Lemma is obvious.

For each fixed n and $1 \leq m \leq 2^n - 1$ the expanded form of the formula $\varphi_{n,m}$ is

$$\begin{aligned}
\varphi_{n,m} = {} & (p_{11} \vee p_{21} \vee \ldots \vee p_{n1}) \& (p_{12} \vee p_{22} \vee \ldots \vee p_{n2}) \& \ldots \& (p_{1m} \vee p_{2m} \vee \ldots \vee p_{nm}) \vee \\
& \vee (p_{11} \vee p_{21} \vee \ldots \vee \bar{p}_{n1}) \& (p_{12} \vee p_{22} \vee \ldots \vee \bar{p}_{n2}) \& \ldots \& (p_{1m} \vee p_{2m} \vee \ldots \vee \bar{p}_{nm}) \vee \\
& \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \; \ldots \\
& \vee (\bar{p}_{11} \vee \bar{p}_{21} \vee \ldots \vee \bar{p}_{n1}) \& (\bar{p}_{12} \vee \bar{p}_{22} \vee \ldots \vee \bar{p}_{n2}) \& \ldots \& (\bar{p}_{1m} \vee \bar{p}_{2m} \vee \ldots \vee \bar{p}_{nm})
\end{aligned}$$

Let $q_{i,j} = p_{1j}^{\sigma_{i1}} \vee p_{2j}^{\sigma_{i2}} \vee \ldots \vee p_{nj}^{\sigma_{in}}$ for $1 \leq i \leq 2^n$ and $1 \leq j \leq m$, where $\sigma_{i1}, \sigma_{i2}, \ldots, \sigma_{in}$ is the binary notation of the integer $2^n - i$.

Using the notation $q_{i,j}$ we obtain

$$\begin{aligned}
\varphi_{n,m} = {} & q_{1,1} \& q_{1,2} \& \ldots \& q_{1,m} \vee \\
& \vee q_{2,1} \& q_{2,2} \& \ldots \& q_{2,m} \vee \\
& \ldots \; \ldots \; \ldots \; \ldots \\
& \vee q_{2^n,1} \& q_{2^n,2} \& \ldots \& q_{2^n,m} \,.
\end{aligned}$$

Let $\psi_{n,m}$ be the following subformula of $\varphi_{n,m}$

$$\begin{aligned}
\psi_{n,m} = {} & q_{1,1} \& q_{1,2} \& \ldots \& q_{1,m} \vee \\
& \vee q_{2,1} \& q_{2,2} \& \ldots \& q_{2,m} \vee \\
& \ldots \; \ldots \; \ldots \; \ldots \\
& \vee q_{m+1,1} \& q_{m+1,2} \& \ldots \& q_{m+1,m} \,.
\end{aligned}$$

Using point 7. of Lemma 1, we obtain that the formula $\bar{\alpha}r\psi_{n,m} \supset \mathop{\&}\limits_{i=1}^{m+1} \bigvee\limits_{j=1}^{m} \bar{\alpha}rq_{i,j}$ has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof **(1)**.

The formula, which presents the well-known "Pigeonhole Principle", is the following

$$PHP_n = \mathop{\&}\limits_{i=1}^{n+1} \bigvee\limits_{j=1}^{n} p_{ij} \supset \bigvee\limits_{1\le i<k\le n+1} \bigvee\limits_{1\le j\le n} (p_{ij} \,\&\, p_{kj}) \,.$$

In [1] it is proved, that PHP_n has ℓ-polynomially bounded (hence also t-polynomially bounded) $\mathcal{F}$-proof.

Let $PHP'_m = \mathop{\&}\limits_{i=1}^{k} \bigvee\limits_{1\le j\le m} \bar{\alpha}rq_{i,j} \supset \bigvee\limits_{1\le i<k\le m+1} \bigvee\limits_{1\le j\le m} (\bar{\alpha}rq_{i,j} \,\&\, \bar{\alpha}rq_{k,j})$. It is obvious, that the formula PHP'_m, which is corresponding instance of the PHP_m, also has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof **(2)**.

Let $A = \bigvee\limits_{1\le i\le k\le m+1} \bigvee\limits_{1\le j\le n} (\bar{\alpha}rq_{i,j} \,\&\, \bar{\alpha}rq_{k,j})$, then using (1), (2) and point 3. of Lemma 1, we obtain that the formula $\bar{\alpha}r\psi_{n,m} \supset A$ has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof **(3)**.

Now we shall show, that the formula $\bar{\alpha}rA$ also has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof.

Lemma 2. For every fixed j $(1 \le j \le m)$ and i, k $(1 \le i < k \le 2^n)$ the formula $q_{i,j} \vee q_{k,j}$ has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof.

Proof. Proof follows from the fact of the existence of such t $(1 \le t \le n)$, that $q_{i,j}$ contains p_{tj} and $q_{k,j}$ contains $\bar{\alpha}rp_{tj}$, and from points 1. and 6. of Lemma 1.

From Lemma 2 and point 5. of Lemma 1 we obtain, that the formula

$$B = \mathop{\&}\limits_{m\le i<k\le m+1} \mathop{\&}\limits_{1\le j\le m} (q_{i,j} \vee q_{k,j})$$

has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof, and from point 8. of Lemma 1 it follows that the formula $\neg A$ has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof **(4)**. From (3), (4) and point 4. of Lemma1 we obtain that $\psi_{n,m}$ has t- (ℓ-) polynomially bounded $\mathcal{F}$-proof and, finally, from point 2. of Lemma 1 follows the proof of Theorem.

Conclusion. There is a hard tautology, which has polynomially bounded $\mathcal{F}$-proof.

References

[1] S. R. Buss, Polynomial size proofs of the propositional pigeonhole principle, *Journal of Symbolic Logic*, 52, 1987, 916–927.

[2] A. A. Chubaryan, On the proof complexity in some system of classical propositional logic, *Izvestija NAN Armenii, Mathematika*, Vol. 37, N 5, 1999, 16–26.

[3] A. A. Chubaryan, On the complexity of proofs in Frege systems, *CSIT Conference, Yerevan*, 2001, 129–132.

[4] A. A. Chubaryan, Relative efficiency of a proof system in classical propositional logic, *Izvestija NAN Armenii, Mathematika*, Vol. 37, N 5, 2002, 71–84.

[5] S. A. Cook, A. R. Reckhow, The relative efficiency of propositional proof systems, *Journal of Symbolic Logic*, 1979, 44, 36–50.

[6] E. Mendelson, Introduction to Mathematical Logic, *D. Van Nostrand company, INC.*

Ֆրեգեյի արտածումների մի հատկության մասին

Ս. Ռ. Ալեքսանյան և Անահիտ Ա. Չուբարյան

Ամփոփում

Յուրաքանչյուր φ նույնաբանության համար [4]-ում ներմուծվել է այնպիսի $s(\varphi)$ մեծություն, որ $s(\varphi)$-ի "բարձր" արժեքի դեպքում φ-ի արտածման երկարությունը եբույլե համակարգերում նույնպես մեծ է: Սույն հոդվածում ցույց է տրվում, որ գոյություն ունի այնպիսի φ նույնաբանություն, որի համար $s(\varphi)$-ն ունի ցուցչային կախվածություն φ-ի երկարությունից, սակայն նրա արտածումը Ֆրեգեյի համակարգերում ունի բազմանդամային բարդություն: