Mathematical Problems of Computer Science 57, 47–55, 2022.

doi: 10.51408/1963-0086

UDC 510.64

On Sizes of Linear and Tree-Like Proofs for any Formulae Families in Some Systems of Propositional Calculus

Levon A. Apinyan¹ and Anahit A. Chubaryan²

¹Russian-Armenian University, Yerevan, Armenia ²Yerevan State University, Yerevan, Armenia e-mail: apinlev00@gmail.com, achubaryan@ysu.am

Abstract

The sizes of *linear and tree-like proofs* for any formulae families are investigated in some systems of propositional calculus: in different sequent systems (with quantifier rules, with the substitution rule, with the cut rule, without the cut rule, monotone) and in the generalization splitting system. The comparison of results obtained here with the bounds obtained formerly for the steps of proofs for the same formulas in the mentioned systems shows the importance of the *size of proof* among the other characteristics of proof complexities.

Keywords: The varieties of propositional sequent systems, The generalization splitting system, The proof size and number of proof steps, Exponential speed-up.

Article info: Received 15 March 2022; accepted 12 May 2022.

1. Introduction

The existence of a propositional proof system, which has polynomial-size proofs for all tautologies, is equivalent to saying that N P = co -N P [1]. This simple observation has drawn attention in recent years to the formalisms of propositional logic for the study of questions of computational complexity A hierarchy of propositional proof systems has been defined in terms of two main complexity characteristics (*size* and *lines*), and the relations between these systems are currently being analyzed. New systems are discovered and, as a consequence, the computational power of the old ones is better understood. It was shown in [2] that the addition of quantifier rules to the propositional sequent calculus induces, for some sequences of formulas, an exponential speed-up by lines over Substitution Frege systems when proofs are considered as *trees*. It was shown in [3] that the lines for *linear* proofs of the same formulae families both in quantifier systems and in the systems with substitution systems are the same by order. In this paper, we

investigate the sizes of linear and tree-like proofs for the mentioned sequence of formulas in some sequent systems (QPK – the system with quantifier rules, SPK – the system with substitution rule, PK – the system with cut-rule, PK – the system without cut-rule, Pmon- the monotone system) and in the system GS, based on the generalized splitting method. The comparative analysis of our results shows that the size of proofs is a more important complexity characteristic of proofs and the linear proofs are preferable to the tree-like proofs.

2. Preliminaries

We will use the current concepts of a propositional formula, *quantified* propositional formula, a free variable in a quantified formula, sequent, different sequent systems and proof complexities. The language of the considered systems contains the propositional variables, logical connectives \neg , &, \lor , \triangleright and parentheses (,). Note that some parentheses can be omitted in generally accepted cases. In some systems, we can use the symbols T for «true» and \bot for «false».

2.1 Definition of Considered Sequent Systems

The sequent system uses the denotation of sequent $\Gamma \to \Delta$, where Γ (antecedent) and Δ (succedent) are finite (may be empty) sequences of propositional formulas.

For every propositional variable p, the sequents $p \to p$, $\to T$ are axioms of PK. For every formulas A, B, for any sequence of formulas Γ and sequence Δ , the logic rules are as follows:

The system **PK**— is obtained from the system **PK** by removing the cut rule. The system **SPK** is obtained from the system **PK** by adding a substitution rule:

$$S_p^{\mathrm{B}} \quad \frac{\mathrm{C}(\mathrm{p}), \Gamma \to \Delta, \mathrm{A}(\mathrm{p})}{\mathrm{C}(\mathrm{B}), \Gamma \to \Delta, \mathrm{A}(\mathrm{B})},$$

where the variable p has no occurrences either in Γ or in Δ , B is the formula, which is substituted everywhere for the variable p.

The system **QPK** is obtained from the system **PK** by adding the following rules :

$$\frac{A(q), \Gamma \to \Delta}{(\exists p) A(p), \Gamma \to \Delta} (\exists \to) \qquad \frac{\Gamma \to \Delta, A(B)}{\Gamma \to \Delta, (\exists p) A(p)} (\to \exists)$$

$$\frac{A(B) \Gamma \to \Delta}{(\forall p) A(p), \Gamma \to \Delta} (\forall \to) \qquad \qquad \frac{\Gamma \to \Delta, A(q)}{\Gamma \to \Delta, (\forall p) A(p)} (\to \forall),$$

where B is any quantified propositional formula. The application of the rules $\exists \to \text{ and } \to \forall$ is restricted to the following requirements: the eigenvariable q does not occur free in the lower sequent of the rule, and all occurrences of q in A(q) are substituted by p. The rules $\to \exists$ and $\forall \to$ require B not to contain variables, which are under the scope of some quantifier.

All formulas in the antecedents and succedents of the system **Pmon** use only monotone logical functions, therefore the rules for implication and negation are not used here.

2.2 Definition of the System GS

Following the usual terminology, we call the variables and negated variables literals.

The following notions were introduced in [4]. Each of the under-mentioned trivial identities for a propositional formula ψ is called *a replacement rule*:

Application of the replacement rule to some words consists in the replacing of some of its subwords having the form of the left-hand side of one of the above identities by the corresponding right-hand side.

The proof system GS. Let φ be some formula and p be some of its variables. Results of the splitting method of formula φ by the variable p (splinted variable) are the formulas $\varphi[p^{\delta}]$ for every δ from the set $\{0,1\}$, which are obtained from φ by assigning δ to each occurrence of p and successively using replacement rules. The generalization of the splitting method allows every formula φ to associate some tree with a root, the nodes of which are labeled by formulas and edges, labeled by literals. The root is labeled by the formula φ itself. If some node is labeled by the formula v and α is some of its variable, then both edges outgoing from this node, are labeled by one of the literals α^{δ} for every δ from the set $\{0,1\}$, and each of 2 "sons" of this node is labeled by the corresponding formula $v[\alpha^{\delta}]$. Each of the tree's leaves is labeled with some constant from the set $\{0,1\}$. The tree, which is constructed for the formula φ by the described method, we will call *a splitting tree* (s.t.) of φ . It is obvious, that by changing the order of splinted variables in the given formula φ , we can obtain different splitting trees of φ .

The GS proof system can be defined as follows: for every formula φ must be constructed some s.t. and if all the tree's leaves are labeled by the value 1, then the formula φ is a tautology, and therefore we can consider the pointed constant 1 as an axiom, and for every formula v, which is a label of some s.t. node, and p is its splinted variable, then the following figure $v[p^0]$, $v[p^1] \vdash v$

can be considered as some inference rule, hence, every above-described s.t. can be considered as some proof of φ in the system GS.

2.3 Proof Complexities

By $|\varphi|$ we denote the size of a formula φ , defined as the number of all logical signs 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 $|\varphi|$.

In the theory of proof complexity, the two main characteristics of the proof are: t-complexity (lines) defined as the number of proof steps, l-complexity (size) defined as the sum of sizes for all formulas (sequents) in the proof [1, 2].

Let ϕ be some proof system, ϕ be some tautology. By $t^{\phi}(\phi)(l^{\phi}(\phi))$ is denoted the minimal possible value of *t*-complexity (*l*-complexity) for all Φ -proofs of ϕ (sequent $\to \phi$).

If for some sequence of sequents $\rightarrow \phi_n$ in two systems ϕ_1 and ϕ_2 for sufficiently large n is valid $t^{\phi_1}(\varphi_n) = \Omega(2^{t^{\phi_2}(\varphi_n)})$, then we say that the system ϕ_2 has exponential sped-up by lines over the system ϕ_1 .

2.4. Results of the Papers [2,3]

Some family of tautologies is introduced in [2]. For propositional variable p, the formula p^m is defined inductively as $p^0 \equiv p \bowtie p^{i+1} \equiv (p^i \& p^i)$ for $i \geq 0$. It is easy to verify that the formula p^m has $2^m - 1$ logical signs and m distinct subformulas.

To simplify further notes, we introduce the following denotations. Let Φ be some sequent system, *t- complexity* (*l-complexity*) for tree-like proofs of the sequent $p \to p^m$ is denoted by $Tt(m)(Tl^{\varphi}(m))$, and for linear proofs, accordingly, by $Lt^{\varphi}(m)(Ll^{\varphi}(m))$.

Theorem 1: ([2]). For sufficiently large n and sequence of sequents $p \to p^{2^n}$ the following holds:

$$T\mathbf{t}^{\mathbf{QPK}}(2^n) = O(n); \ T\mathbf{t}^{\mathbf{SPK}}(2^n) = \Omega(2^n); \ T\mathbf{t}^{\mathbf{PK}}(2^n) = \Omega(2^n); \ T\mathbf{t}^{\mathbf{PK}-}(2^n) = \Omega(2^n)$$

For the lines of linear proofs of the same sequence, the following was proved in [3]..

Theorem 2: ([3]). For sufficiently large n and sequence of sequents $p \to p^{2^n}$ the following holds:

$$\boldsymbol{Lt^{\text{QPK}}}(2^n) = O(n); \ L\boldsymbol{t^{\text{SPK}}}(2^n) = O(n); \ L\boldsymbol{t^{\text{PK}}}(2^n) = \theta(2^n); \ L\boldsymbol{t^{\text{PK}-}}(2^n) = \theta(2^n).$$

The comparative analysis results of both above theorems shows that the system **QPK** has no preference by lines of proof over the system SPK, and the latter system has a well-known speed-up by lines over PK. Analogous sped-up was first fixed in [5].

3. The Main Results

3.1. The *l-complexities* of **linear** proofs for the same family of sequents $p \to p^{2^n}$ in abovementioned sequent systems are investigated here.

Theorem 3: For sufficiently large n and sequence of sequents $p \to p^{2^n}$ the following holds:

$$Ll^{\mathbf{QPK}}(2^n) = \theta(2^{2^n}); \quad Ll^{\mathbf{SPK}}(2^n) = \theta(2^{2^n}); \quad Ll^{\mathbf{PK}}(2^n) = \theta(2^{2^n}); \quad Ll^{\mathbf{PK}-}(2^n) = \theta(2^{2^n})$$
 and $Ll^{\mathbf{Pmon}}(2^n) = \theta(2^{2^n}).$

To prove the mentioned results, we should evaluate the *sizes* of proofs for the sequents $p \to p^{2^n}$ in all the mentioned systems. Note that $|p^{2^n}| = 2^{2^n} - 1$ and as the derivable sequent itself must be in every proof, then the lower bounds $\Omega(2^{2^n})$ are obvious for all systems. To prove the upper bounds, we investigate the "good" linear proofs in the mentioned systems.

Linear proof in QPK

We use the tree-like proofs of $p \to p^{2^n}$ in the system **QPK** with O(n) lines [2]. At first, we consider the provable sequent $\forall q (q \supset q^k) \to \forall q (q \supset q^{2k})$, where k is an arbitrary integer and $q^{2k} = (q^k)^k$. The proof of this sequent will not depend on k and can be obtained in a constant number of lines as follows (note, that their sizes are written to the right of every sequent):

$p \to p p^k \to p^k$	$2^{k+2}+2$
$\overline{p\supset p^k,p\to p^k}$	$2^{k+2}+2$
$\forall q(q \supset q^k), p \to p^k p^{2k} \to p^{2k}$	$2^{k+2}+2^{2k+2}+5$
$\forall q(q\supset q^k), p^k\supset p^{2k}, p\to p^{2k}$	$2^{k+2}+2^{2k+2}+5$
$\forall q(q\supset q^k), p^k\supset p^{2k}\to p\supset p^{2k}$	$2^{k+2}+2^{2k+2}+6$
$\forall q(q \supset q^k), \forall q(q \supset q^k) \rightarrow p \supset p^{2k}$	$2^{k+2}+2^{2k+1}+8$
$\forall q(q\supset q^k)\to p\supset p^{2k}$	$2^{k+1}+2^{2k+1}+5$
$\forall q(q\supset q^k) \to \forall q(q\supset q^{2k})$	$2^{k+1}+2^{2k+1}+7$

Note, that this proof is also *linear*. By combining the above sequents *n* times, one obtains

$$\forall q(q\supset q^2)\to \forall q(q\supset q^{2^n}),$$

and since $\forall q (q \supset q^2)$ is provable in constant lines, one infers $\forall q (q \supset q^{2^n})$, and therefore $\rightarrow p \supset p^{2^n}$ in O(n) lines. The number of all logical signs in the pointed part of the proof is

 $7 \cdot 2^{k+2} + 9 \cdot 2^{2k+1} + 40$, and as such steps are repeated n times with $k = 2^i$, for i = 0, 1, 2, ..., n, then the size of all proofs must be $\sum_{i=0}^{n} (7 \cdot 2^{2^i+2} + 9 \cdot 2^{2^{i+1}+1} + 40)$. The bound of the major addendum is $7 \sum_{i=0}^{n} 2^{2^i+2} \le 7 \sum_{i=0}^{2^{n+2}} 2^i \le 7 \cdot (2^{2^n+3} - 1)$, and hence the upper bound is $O(2^{2^n})$. So, $Ll^{QPK}(2^n) = \theta(2^{2^n})$.

Linear proof in SPK

2n + 1
$$p^{2^{n-1}} \rightarrow p^{2^n}$$
 subst. $2^{2^{n-1}} - 1 + 2^{2^n} - 1$
2n + 2 $p^0 \rightarrow p^{2^n}$ cut $2^{2^n} - 1$

2n + 2 $p^{0} \rightarrow p^{2}$ cut 2 - 1It is not difficult to see that the size cannot be more, than $3 \sum_{i=0}^{n} 2^{2^{i}} \le 3 \sum_{i=0}^{2^{n}} 2^{i} \le 2^{2^{n}+3}$, hence the upper bound is $O(2^{2^{n}})$. So, $Ll^{SPK}(2^{n}) = \theta(2^{2^{n}})$.

Linear proof in PK

The size of such linear proof must be no more, than $\sum_{i=0}^{2^n} 2^i \le 2^{2^{n+1}}$, hence the upper bound is $O(2^{2^n})$. So, $Ll^{PK}(2^n) = \theta(2^{2^n})$.

As in this proof we do not use the cut rule, but only the rule $(\rightarrow \&)$, then the bounds both in **PK**⁻ and in **Pmon** are analogous.

Theorem 1 is proved.

3.2. The *l-complexities* of **tree-like** proofs for the same family of sequents $p \to p^{2^n}$ in the abovementioned sequent systems are investigated here.

Theorem 4: For sufficiently large n and sequence of sequents $p \to p^{2^n}$ the following holds:

$$Tl^{\mathbf{QPK}}(2^n) = \theta(2^{2^n}); \quad Tl^{\mathbf{SPK}}(2^n) = \theta(2^{2^n});$$
 $\log_2(Tl^{\mathbf{PK}}(2^n)) = \theta(2^n); \quad \log_2(Tl^{\mathbf{PK}-}(2^n)) = \theta(2^n) \text{ и } \log_2(Tl^{\mathbf{Pmon}}(2^n)) = \theta(2^n).$

To **prove** these results, we transform every linear proof above into a tree-like proof in the same system.

The size of tree-like proof in QPK: As we noted above, the proof in QPK is linear and tree-like simultaneously, hence the bound is the same.

The size of tree-like proof in SPK: We should transform the above proof (1) into tree-like. It is enough to change every part « $p^0 \rightarrow p^{2^i}$, $p^{2^i} \rightarrow p^{2^{i+1}}$ (*substitution*), $p^0 \rightarrow p^{2^{i+1}}$ (cut)» for $0 \le i \le n-1$ of linear proof with the part «tree-like proof of $p^0 \rightarrow p^{2^i}$, $p^{2^i} \rightarrow p^{2^{i+1}}$ (*substitution*), tree-like proof of $p^0 \rightarrow p^{2^i}$, $p^0 \rightarrow p^{2^{i+1}}$ (cut)». After such transformation we have

$$Tl^{\text{SPK}}\!\!\left(2^i\right) \leq \!\! 2Tl^{\text{SPK}}\!\!\left(2^{i-1}\right) + \left|p^{2^{i-1}} \rightarrow p^{2^i}\right| + \left|p^0 \rightarrow p^{2^i}\right| \; for \; 1 \leq i \leq n,$$

hence

$$Tl^{\text{SPK}}(2^n) \le 2^n l_{\mathcal{A}}^{\text{SPK}}(2^0) + \sum_{i=1}^n 2^{n-i} (2^{2^i+2}) \le 2^n + 2^{n+2} 2^{2^n+2}.$$

So, the upper bound is $O(2^{2^n})$, hence $Tl^{SPK}(2^n) = \theta(2^{2^n})$.

The size of tree-like proof in PK: Here we should transform the above proof (2) into tree-like. It is enough to change every part « $p^0 \rightarrow p^i$, $p^0 \rightarrow p^{i+1}$ (\rightarrow &)» for $0 \le i \le n-1$ of linear proof with the part « tree-like proof of $p^0 \rightarrow p^i$, tree-like proof of $p^0 \rightarrow p^i$, $p^0 \rightarrow p^{i+1}$ (\rightarrow &)», then it is obvious that

$$Tl^{\mathbf{PK}}(\mathbf{i}) \leq 2Tl^{\mathbf{PK}}(\mathbf{i} - \mathbf{1}) + |p^0 \rightarrow p^i| \text{ for } 1 \leq i \leq 2^n$$

hence we have

$$Tl^{PK}(2^n) \le 2^{2^n} Tl^{PK}(2^0) + \sum_{i=1}^{2^n} 2^i (2^{n-i}) \le 2^{2^n} + 2^n 2^{2^n (2^n+1)/2}.$$

So, the upper bound for $\log_2(Tl^{PK}(2^n))$ is $O(2^n)$, and hence $\log_2(Tl^{PK}(2^n)) = \theta(2^n)$.

As above, in this proof we do not use the cut rule, but only the rule $(\rightarrow \&)$, then the bounds both in **PK**⁻ and in **Pmon** are analogous.

Theorem 4 is proved.

Note that we do not have any exponential speed-up here (it may only be quadratic).

The size of linear and tree-like proofs in GS:

Theorem 5: For sufficiently large n and sequence of formulas $p \supset p^{2^n}$ the following holds: $Lt^{GS}(2^n) = \theta(1)$; $Tt^{GS}(2^n) = \theta(1)$;

$$Ll^{GS}(2^n) = \theta(2^{2^n}); Tl^{GS}(2^n) = \theta(2^{2^n}).$$

Note that every formula $p \supset p^{2^n}$ has only one variable for split, hence the **proof** of Theorem 5 is obvious.

4. Conclusion

The analysis of all the results shows that in the theory of proof complexity, the investigations of *l*-complexity in linear proofs are important.

References

- [1] S. A. Cook and A. R. Reckhow, "The relative efficiency of propositional proof systems", *Symbolic Logic*, vol. 44, pp. 36-50, 1979.
- [2] A. Carbone, "Quantified propositional logic and the number of lines of tree-like proofs", *Studia Logica*, vol. 64, pp. 315-321, 2000.

- [3] А. А. Тамазян и А. А. Чубарян, "Об отношениях сложностей выводов в ряде систем исчисления высказываний", *Математические вопросы кибернетики и вычислетьной техники*, vol. 54, pp. 138-146, 2020.
- [4] Ан. А. Чубарян и Арм. А. Чубарян, "Оценки некоторых сложностных характеристик выводов в системе обобщенных расщеплений", НАУ, Отечественная наука в эпоху изменений: постулаты прошлого и теории нового времени, часть 10, 2(7), стр.11-14, 2015.
- [5] Г. Цейтин и Ан. Чубарян, "Некоторые оценки длин логических выводов в классическом исчислении высказываний", ДАН Арм. ССР, том 55, но.1, стр. 10-12, 1972.

Ասույթային հաշվի մի շարք համակարգերում բանաձևերի որոշ ընտանիքների գծային և ծառատիպ արտածումների երկարությունների մասին

Լևոն Ա. Ափինյան¹ և Անահիտ Ա. Չուբարյան²

¹Ռուս-հայկական համալսարան, Երևան, Հայաստան ²Երևանի պետական համալսարան, Երևան, Հայաստան e-mail: apinlev00@gmail.com, achubaryan@ysu.am

Ամփոփում

Բանաձների մի քանի ընտանիքների համար ուսումնասիրված են *զծային և ծառատիպ արտածումների* երկարությունները ասույթային հաշվի մի քանի համակարգերում՝ սեկվենցիալ համակարգերի տարատեսակներում /ծավալիչներով, տեղադրման կանոնով, հատույթի կանոնով, առանց հատույթի կանոնի, մոնոտոն/, ինչպես նաև ընդհանրացված տրոհումների համակարգում։ Սույն աշխատանքում ստացված արդյունքների համեմատումը նախկինում ստացված նույն բանաձների նշված համակարգերում արտածումների քայլերի համար ստացված արդյունքների հետ փաստում են *արտածումների երկարության* որպես բարդության բնութագրիչի արժեքավորումը։

Բանալի բառեր՝ ասույթային հաշվի սեկվենցիալ համակարգերի տարատեսակներ, ընդհանրացված տրոհումների համակարգ, արտածման քայլերի քանակ և երկարություն, էքսպոնենցիալ արագացում։

О длинах линейных и древовидных выводов некоторых семейств формул в ряде систем исчисления высказываний

Левон А. Апинян¹ и Анаит А. Чубарян²

¹Российско-Армянский университет, Ереван, Армения ²Ереванский государственный университет, Ереван, Армения e-mail: apinlev00@gmail.com, achubaryan@ysu.am

Аннотация

Для некоторых семейств формул исследованы *длины* линейных и древовидных выводов в ряде систем исчисления высказываний: в разновидностях секвенциальных систем (с кванторами, с правилом подстановки, с правилом сечения, без правила сечения, монотонных), а также в системе обобщенных расщеплений. Сравнение полученных результатов с ранее полученными оценками для *шагов* тех же разновидностей выводов тех же формул и в тех же системах указывают на определенную значимость именно длины вывода как основной сложностной характеристики выводов.

Ключевые слова: разновидности секвенциальных систем исчисления высказываний; система обобщенных расщеплений; количество шагов и длина вывода; экспоненциальное ускорение.