GEORGIAN MATHEMATICAL JOURNAL: Vol. 3, No. 1, 1996, 97-100

COMPLEXITY OF THE DECIDABILITY OF ONE CLASS OF FORMULAS IN QUANTIFIER-FREE SET THEORY WITH A SET-UNION OPERATOR

M. TETRUASHVILI

ABSTRACT. We consider the quantifier-free set theory *MLSUn* containing the symbols U, $\setminus$, $=$, $\in$, *Un*. $Un(p)$ is interpreted as the union of all members of the set p. It is proved that there exists an algorithm which for any formula Q of the *MLSUn* theory containing at most one occurrence of the symbol *Un* decides whether Q is true or not using the space $cn^3\log_2 n$ (n is the length of Q).

Let *MLSUn* (*MLS*) be a quantifier-free set theory whose language contains the symbols U, $\setminus$, $=$, $\in$, *Un* (U, $\setminus$, $=$, $\in$), where *Un* is a unary functional symbol ($Un(p)$ is understood as a union of all members of the set p). Let $MLSUn^{(1)}$ be the class of formulas of the language *MLSUn* containing at most one occurrence of the symbol *Un*. The decidability problem for a class of formulas consists in finding an algorithm which decides whether an arbitrary formula of this class is true or not. For the class $MLSUn^{(1)}$ the problem of decidability can be easily reduced to testing the satisfiability of conjunctions of literals of the following types:

$$(=)\ \ x=y\cup z,\ \ x=y\setminus z,\ \ (\in)\ \ x\in y,\ \ \ (Un)\ \ u=Un(p),$$

where the conjunctions contain the (Un)-type literal not more than once.

Let Q be the formula of the $MLSUn^{(1)}$. If Q contains literals $u=Un(p)$ and $u=\varnothing$, then Q has a model if and only if Q_0 has a model, where Q_0 denotes the result of removing the clause $u=Un(p)$ from Q and adding either the literal $p=\varnothing$ or the clauses $\varnothing\in p$, $\varnothing=x\cap p\vee\varnothing=x\cap p\vee\varnothing=p\setminus x$ for every variable x of Q. The obtained Q_0 is a formula of the decidable theory *MLS* [1]. Therefore we can assume Q to contain the literal $u\neq\varnothing$.

Let $x,z_1,\dots,z_m,y$ be variables. By $P^+(x,z_1,\dots,z_m,y)$ we denote the disjunction $\bigvee\limits_{\langle,i_1,\dots,i_k\rangle\in\mathcal{I}_m^+}\big(x\in z_{i_1}\&z_{i_1}\in z_{i_2}\&\cdots\&z_{i_{k-1}}\in z_{i_k}\&z_{i_k}\in y\big)$ where $\mathcal{I}_m^+$ is the set of all non-empty ordered subsets of the set $\{1,\dots,m\}$.

1991 *Mathematics Subject Classification.* 03D15.
Key words and phrases. Complexity of decidability, set theory, set-union operator.

1072-947X/96/0100-0097$09.50/0

Let $x_1, \dots, x_m$ be all variables occurring in Q. It is proved in [2] that given a finite family $z_1, \dots, z_m$ of sets and the set p such that $Un(p) \neq \varnothing$, there exists a non-empty set $t \in p$ such that $P^+(t, z_1, \dots, z_m, Un(p))$ is false. Therefore we can add the statements $t \in p$, $t \neq \varnothing$ and $\neg P^+(t, x_1, \dots, x_m, u)$ to the formula Q not modifying its satisfiability (t is a new variable).

Finally, Q can be assumed to contain one full conjunct C of the disjunctive normal form (d.n.f.) of the statement $\mathop{\&}\limits_{i=1}^{m} \big(x_i \notin p \vee x_i \setminus u = \varnothing\big)$ which is true in any model of Q.

Let Q_- denote the result of removing the literal $u = Un(p)$ from Q. The obtained Q_- is a formula of MLS. It has been proved in [2] that Q has a model if and only if Q_- has a model. Therefore this theorem solves the problem of decidability for the class $MLSUn^{(1)}$. However, the formula $P^+(x, z_1, \dots, z_m, y)$ playing an important part in the investigation of satisfiability of Q is awkward: $2^m \le \mathrm{Card}(\mathcal{I}_m^+) \le m^m$, and the d.n.f. of this formula contains a large number of summands (exponential with respect to m).

Therefore the corresponding decidability procedure has an exponential computational complexity (by space).

Below we shall find an algorithm which solves the decidability problem for the class $MLSUn^{(1)}$ with a polynomial computational complexity (by space).

The formula $\neg P^+(x, z_1, \dots, z_m, y)$ is equivalent to the formula

$$\mathop{\&}\limits_{\langle, i_1, \dots, i_k\rangle \in \mathcal{I}_m^+} \big(x \notin z_{i_1} \vee z_{i_1} \notin z_{i_2} \vee \dots \vee z_{i_{k-1}} \notin z_{i_k} \vee z_{i_k} \notin y\big).$$

Its d.n.f. $\Delta^+(x, z_1, \dots, z_m, y) \equiv \Delta^+$ contains at most m^{m^2} summands.

Let $\mathcal{K}$ be an arbitrary summand of the above d.n.f. Δ^+. This $\mathcal{K}$ is a conjunction of literals of the following forms: $x \notin z_i$, $z_i \notin z_j$, $z_j \notin y$. The number of various ordered pairs composed of all variables of the formula Δ^+ is less than $(m+2)^2$. Therefore, although the number of factors in $\mathcal{K}$ is greater than 2^m, the number of different factors in it is less than $(m+2)^2$. Consequently, $\mathcal{K}$ contains the majority of recurring factors.

In the sequel, the disjunction will be assumed to be a well-ordered set of its summands. The set $\mathcal{I}_m^+$ is naturally well ordered with respect to the number k of elements in the system $\langle i_k, \dots, i_k\rangle$ and lexicographically for a given number k of elements in the system. Let M_ν denote the disjunction $x \notin z_1 \vee \dots \vee z_{i_k} \notin y$, where ν is a number of the system $\langle i_1, \dots, i_k\rangle$ in the set $\mathcal{I}_m^+$ and let m_ν be the last element of the set M_ν. The conjunction $\mathcal{K}$ is composed as follows: the number of all factors is equal to $\mathrm{Card}(\mathcal{I}_m^+)$ and the νth member of the conjunction is an element of the set M_ν.

It can be easily seen that the formula Q_- is equivalent to the disjunction

D of the formulas of the form

$$\mathcal{H}\&\Delta^+(x, z_1, \dots, z_m, y), \tag{1}$$

where $\mathcal{H}$ is $Q_0\&(t \in p)\&(t \neq \varnothing)\&C$, and Q_0 denotes the result of removing the literal $u = Un(p)$ from Q, length$(\mathcal{H}) \leq c_0 \cdot$ length(Q) with a constant c_0 and the number of summands in the disjunction is equal to 2^m. The formula Q is satisfiable if and only if one of the summands of D is satisfiable. In this case the passage from one summand to the next one is realized easily and mechanically. Therefore, having investigated one summand for satisfiability we can write out the next one $\mathcal{H}'$ and cancel the previous $\mathcal{H}$.

The satisfiability test for the given summand of D of the type (1) requires that for all summands of the d.n.f. of the formula (1). We test the satisfiability of each such conjunction with the linear space [3].

Obviously, an arbitrary summand of the d.n.f. of the formula of the type (1) has the form $\mathcal{H}\&\mathcal{K}$, where $\mathcal{K}$ is a conjunction of literals. Hence, to construct an algorithm testing the satisfiability of the formula Q_-, it suffices to construct an algorithm $\mathfrak{A}$ which for an arbitrary conjunction $\mathcal{K}'$ of the d.n.f. Δ^+ gives immediately the next conjunction $\mathcal{K}$. The formula Δ^+ can be regarded as a direct product of the well-known number of well-ordered finite sets M_ν and hence as a well-ordered set of conjunctions. Therefore, writing down the given conjunction $\mathcal{K}$ of the d.n.f. Δ^+ along with all its recurring factors (their number is greater than 2^m), we can easily write out the next conjunction immediately following it.

When constructing an algorithm $\mathfrak{A}$ it is natural to operate with reduced conjunctions $\mathcal{K}$ of the d.n.f. Δ^+. However, in order to reset all conjunctions of the d.n.f. Δ^+, it is necessary to keep some information about implicit recurring factors of the reduced conjunctions in the case where the conjunction $\mathcal{K}$ is reduced.

Let us remove from the conjunction $\mathcal{K}$ the recurring literals from the left to the right and keep the number of the removed literals. The obtained conjunction $\lambda_1\& \cdots \&\lambda_k$ of literals we write as follows:

$$\lambda_1\mathfrak{b}_1\lambda_2\mathfrak{b}_2\cdots\lambda_k\mathfrak{b}_k, \tag{2}$$

where $\mathfrak{b}_i$ is a binary notation of the number of those literals of $\mathcal{K}$ which iterate the literals $\lambda_1, \dots, \lambda_i$. Since $k \leq 2m^2$ and Card$(\mathcal{I}_m^+) \leq m^m$, the length of the whole writing (2) does not exceed $2m^3 \log_2 m$. Our aim can be achieved by proving the following

Lemma. *There exists an algorithm $\mathfrak{A}$ which for the variables $x, z_1, \dots,$ z_m, y and an arbitrary number μ (in binary notation) of the conjunction of the d.n.f. Δ^+ with the space $cm^3 \log_2 m$ gives this conjunction $\mathcal{K}_\mu$ in the form (2) and also determines in a non-last conjunction $\mathcal{K}_\mu$ the first from the right number ν_0 (in binary notation) of its literal which is different from m_{ν_0}.*

Proof. The binary notation length of the number μ of the conjunction of Δ^+ does not exceed $m^2 \log_2 m$. For $\mu = 1$ the validity of assertions of the lemma is obvious. Let for μ the conjunction $\mathcal{K}_\mu$ be already written out in the form (2) and let the first from the right number ν_0 of the literal λ from $\mathcal{K}_\mu$, which is different from m_{ν_0}, be also determined. We replace in $\mathcal{K}_\mu$ the literal λ with the number ν_0 by the next literal λ' of M_{ν_0}. Furthermore we replace all literals of $\mathcal{K}_\mu$ with the numbers $\nu > \nu_0$ by the first elements of M_ν and leave unchanged all literals with the numbers $\nu < \nu_0$. If $\lambda' \neq m_{\nu_0}$, then the required conjunction $\mathcal{K}_{\mu+1}$ and the number ν_0 are found for $\mu + 1$. Let $\lambda' = m_{\nu_0}$. Then we can determine the first, to the left of ν_0, number ν_1 of the literal of the complete conjunction $\mathcal{K}_{\mu+1}$ which is different from m_{ν_1}. To this end, starting with $\mu = 1$ we successively write out the conjunctions $\mathcal{K}_1, \dots, \mathcal{K}_\mu, \mathcal{K}_{\mu+1}$ and survey constantly the literal with the number $\nu_0 - 1$. This literal is included (temporarily) in the reduced writing of the conjunction, while when passing from $\mathcal{K}_i$ to $\mathcal{K}_{i+1}$ the previous writing relating to $\mathcal{K}_i$ is canceled. Thus, having reached $\mathcal{K}_{\mu+1}$, we can determine the literal $\widetilde{\lambda}$ of the conjunction $\mathcal{K}_{\mu+1}$ with the number $\nu_0 - 1$. If $\widetilde{\lambda} \neq m_{\nu_0-1}$, then our task is fulfilled for $\mu + 1$. If $\widetilde{\lambda} = m_{\nu_0-1}$, then the same can be repeated for the number $\nu_0 - 2$ of the literal, and in this case the previous writing for $\nu_0 - 1$ is canceled. If $\mu + 1$ is the number of the non-last conjunction, then there exists the unknown number ν_1 which can be found. □

It is easy to see that in computations with the use of the algorithm $\mathfrak{A}$ the length of the writing does not exceed the length of the writing of the conjunction from Δ^+ multiplied by a constant c, i.e., $cm^3 \log_2 m$. The main lemma and the reasonings preceding this lemma result in the validity of the following

Theorem. *There exists an algorithm which for any formula Q from the class MLSUn$^{(1)}$ determines the validity of Q and needs the space $cn^3 \log_2 n$, where n is the length of Q.*

References

1. A. Ferro et al., Decision procedures for elementary sublanguages of set theory. I. *Comm. Pure Appl. Math.* **33**(1980), 599–608.

2. M. Breban and A. Ferro, Decision procedures for elementary sublanguages of set theory. III. *Advances Appl. Math.* **5**(1984), 147–215.

3. M. R. Tetruashvili, The complexity of decision of the validity problem for an elementary sublanguage of quantifier-free set theory. (Russian) *Bull. Acad. Sci. Georgian SSR* **126**(1987), 245–247.

(Received 23.12.1993)

Author's address:
I. Vekua Institute of Applied Mathematics, Tbilisi State University
2, University St., Tbilisi 380043, Republic of Georgia