ОБОБЩЕННАЯ ВЫЧИСЛИМОСТЬ И ОПРЕДЕЛИМОСТЬ
(Вычислительные системы)
1998 год **Выпуск 161**

УДК 510.532:519.766.23

РАЗНОСТНЫЕ ИЕРАРХИИ РЕГУЛЯРНЫХ ЯЗЫКОВ[1]

А.Г. Шукин

Введение

Существует несколько способов описания и классификации регулярных языков (связанных, например, с логикой, регулярными выражениями или конечными автоматами) [10,7]. В подклассе "беззвездных" языков (соответствующих расширенным регулярным выражениям без символов "*") разными способами построено несколько иерархий. Д.Вжозовский и Р.Кнаст [2] доказали нетривиальность так называемой dot–depth–иерархии (основанной на подсчете числа чередований операций конкатенации с булевыми операциями в регулярных "беззвездных" выражениях). В.Томас [8] нашел связь этой иерархии с кванторной иерархией, которая исчерпывает тот же класс "беззвездных" языков. В частности, нетривиальность одной из этих иерархий влечет нетривиальность другой. Позднее,используя простой и наглядный способ (основанный на играх Эренфойхта–Фрессе), В.Томас [9, 10] дал прямое доказательство нетривиальности кванторной иерархии и новое доказательство нетривиальности

[1]Работа выполнена при поддержке гранта РФФИ № 96-01-00257.

dot-depth-иерархии. В данной работе рассмотрены разностные иерархии [3,6] над каждым уровнем кванторной иерархии и доказана их нетривиальность, что усиливает результат о нетривиальности кванторной иерархии.

Мы будем использовать описание языков с помощью предложений формальной логики. Пусть A — конечный алфавит, содержащий не менее двух символов. Обозначим через A^+ множество всех конечных слов ненулевой длины в алфавите A. Языком будем называть любое подмножество множества A^+. Если слово рассматривать как конечную линейно упорядоченную структуру (в подходящей сигнатуре, включающей унарный предикат для каждого символа из алфавита A и бинарный предикат " $<$ "), то предложения формальной логики первого порядка можно использовать для описания "беззвездных" языков. (Чтобы охватить все регулярные языки, надо добавить предложения с монадическими переменными второго порядка.) Кроме основного алфавита A мы рассмотрим расширенный [5,7] алфавит $A \times 2^{\nu}$, где ν — конечное множество переменных. Будем называть ν-структурой слово $(a_1, U_1) \dots (a_r, U_r)$ (расширенного) алфавита $A \times 2^{\nu}$ такое, что множества $U_1, U_2, \dots, U_r$ являются разбиением множества ν. Формулы со свободными переменными в ν будут выражать свойства ν-структур. Иногда ν-структуры мы будем называть структурами. Унарные предикаты Q_a, $a \in A$, будем интерпретировать так: $w \models Q_a x$ тогда и только тогда, когда w содержит букву вида (a, U), где $x \in U$. Предикат " $<$ " будет иметь обычный смысл: $w \models x < y$ тогда и только тогда, когда $w = (a_1, U_1) \dots (a_i, U_i) \dots (a_j, U_j) \dots (a_r, U_r)$, где $x \in U_i$, $y \in U_j$, $i < j$. Пусть $w = (a_1, U_1) \dots (a_r, U_r)$. Тогда $w \models \exists x \phi$ тогда и только тогда, когда для некоторого i, $1 \leq i \leq r$, справедливо $(a_1, U_1) \dots (a_i, U_i \cup \{x\}) \dots (a_r, U_r) \models \phi$. Булевы операции интерпретируются естественным образом. Букву $(a, \emptyset)$ алфавита $A \times 2^{\nu}$ мы будем отождествлять с буквой a алфавита A.

Если ϕ — предложение, то ϕ можно интерпретировать в слове $w \in A^+$, поскольку это слово можно рассматривать как $\emptyset$–структуру. Каждое предложение описывает язык, состоящий из слов, которые удовлетворяют этому предложению. Каждому множеству формул соответствует класс языков, которые можно описать предложениями из этого множества. Мы будем использовать одни и те же обозначения для множества формул и для соответствующих классов языков.

Пусть класс S_k состоит из языков, которые можно описать предложениями вида $\exists x_1^1 \ldots \exists x_{n_1}^1 \forall x_1^2 \ldots \forall x_{n_2}^2 \ldots Q x_1^k \ldots \ldots Q x_{n_k}^k \phi$, где $Q \in \{\exists, \forall\}$, ϕ — формула без кванторов. Иерархию языков $(S_k)_{k \in N}$ будем называть кванторной. (Мы считаем множество натуральных чисел N равным $\{1, 2, \ldots\}$.) Очевидно, что $S_k \cup \check{S}_k \subset S_{k+1}$, где $k \in N$, $\check{S}_k = \{A^+ \setminus L | L \in S_k\}$ — двойственный класс для S_k. Обозначим через $D_{k,l}$ множество, состоящее из элементов $\bigcup_{i=1}^{m} (L_{2i-1} \setminus L_{2i})$, $m \in N$, $L_i \in S_k$, $1 \le i \le l$, и $L_i = \emptyset$, $i > l$. Иерархию языков $(D_{k,l})_{l \in N}$ будем называть разностной иерархией над k–м уровнем кванторной иерархии S_k. Очевидно, что $D_{k,l} \subset S_{k+1}$, $k, l \in N$. Известно [4,1], что разностные иерархии над любым классом исчерпывают замыкание этого класса относительно булевых операций и для этих иерархий справедливы соотношения: $D_{k,l+1} = \{(A^+ \setminus L_1) \cap L_2 | L_1 \in D_{k,l}, L_2 \in S_k\}$, $D_{k,l} \cup \check{D}_{k,l} \subset D_{k,l+1}$, $k, l \in N$.

Как уже отмечалось, кванторная иерархия языков нетривиальна, или другими словами, последовательность классов языков $(S_k)_{k \in N}$ не стабилизируется: $S_k \neq S_{k+1}$, $k \in N$. В данной работе мы докажем, что разностные иерархии $(D_{k,l})_{l \in N}$ над каждым уровнем S_k, $k \in N$, также нетривиальны. Сформулируем наш основной результат в виде теоремы, из которой следует, что $D_{k,l} \neq D_{k,l+1}$, $k, l \in N$.

ТЕОРЕМА 1. *При всех $k, l \in N$ класс $D_{k,l}$ не совпадает со своим двойственным классом $\check{D}_{k,l}$.*

Определим индукцией по k множество формул специального вида класса $S_{k,r}$, $k \ge 0$, $r \in N$. При $r \in N$ множество формул специального вида класса $S_{0,r}$ состоит из формул без кванторов, которые находятся в дизъюнктивной нормальной форме без повторений элементарных конъюнкций и без повторений конъюнктивных членов в элементарных конъюнкциях. При $k, r \in N$ формула специального вида класса $S_{k,r}$ — это дизъюнкция попарно различных формул вида $\exists z_1^k \exists z_2^k \ldots \exists z_p^k \neg\phi$, где $p \leq r$, ϕ — формула специального вида множества $S_{k-1,r}$. Множество $S_{k,r}$ состоит из формул, которые эквивалентны формулам специального вида множества $S_{k,r}$. (Эквивалентность понимается как равносильность на конечных структурах.) По построению связанные переменные формул специального вида множества $S_{k,r}$ находятся среди $z_1^i, z_2^i, \ldots, z_r^i$, $i \leq k$. Легко проверить, что число формул специального вида из $S_{k,r}$ с фиксированным набором свободных переменных конечно. Для каждого $k \in N$ классы языков $S_{k,r}$, $r \in N$, соответствующие множествам формул $S_{k,r}$, исчерпывают класс S_k.

1. Вспомогательные результаты

Пусть $k \geq 0$ и $r \geq 1$. Введем рефлексивное и транзитивное отношение $\leq_{k,r}$ на множестве структур.

ОПРЕДЕЛЕНИЕ. Пусть ν — множество переменных, u и v — ν-структуры. Если из $u \models \phi$ следует $v \models \phi$ для любой формулы $\phi \in S_{k,r}$ со свободными переменными из ν, то будем писать $u \leq_{k,r} v$.

Опишем теперь вариант игры Эренфойхта-Фрессе $C_{k,r}(u, v)$ с k раундами. В начале игры имеется две $\{y_1, \ldots, y_p\}$-структуры u и v. В игре участвуют два игрока. Цель первого игрока — показать, что эти две структуры различны, цель второго — показать, что они неразличимы. У каждого игрока есть kr фишек, помеченных переменными $z_1^j, \ldots, z_r^j$, $j = 1, \ldots, k$. В первом раунде первый игрок ставит свои фишки $z_1^k, \ldots, z_p^k$, $p \leq r$, на буквы структуры u. В результате $\{y_1, \ldots, y_p\}$-структура u превратится в $\{y_1, \ldots, y_p, z_1^k, \ldots, z_p^k\}$-структуру u'. Второй игрок в ответ на ход первого игрока ставит свои фишки $z_1^k, \ldots, z_p^k$ в другую структуру v. Получается еще одна $\{y_1, \ldots, y_p, z_1^k, \ldots, z_p^k\}$-структура v'. Продолжение этой игры ничем не отличается от игры $C_{k-1,r}(v', u')$. В конце игры получим две структуры u_0 и v_0. Считается, что второй игрок выиграл, если для каждой атомарной формулы α $u_0 \models \alpha$ тогда и только тогда, когда $v_0 \models \alpha$. Другими словами,второй игрок выигрывает, когда $u_0 \leq_{0,r} v_0$ и $v_0 \leq_{0,r} u_0$ для всех $r \in N$. В любой конкретной игре у одного из игроков есть выигрышная стратегия.

ТЕОРЕМА 2. *Пусть u, v — ν-структуры, $k \geq 0$, $r \geq 1$. Тогда $u \leq_{k,r} v$, если и только если у второго игрока есть выигрышная стратегия в игре $C_{k,r}(u, v)$.*

ДОКАЗАТЕЛЬСТВО. Пусть сначала $u \leq_{k,r} v$. Докажем индукцией по k, что у второго игрока есть выигрышная стратегия. Если $k = 0$, то второй игрок выигрывает, поскольку по предположению u и v удовлетворяют одним и тем же атомарным формулам. Пусть теперь $k > 0$. Предположим, что у второго игрока нет выигрышной стратегии. Тогда выигрышная стратегия есть у первого игрока. Пусть он поставил фишки $z_1^k, \dots, z_p^k$, $p \leq r$, в структуру u, следуя своей выигрышной стратегии. В результате получилась структура u' такая, что при любом ответном ходе второго игрока в структуру v (в результате которого получается структура v') первый игрок обладает выигрышной стратегией в игре $C_{k-1,r}(v', u')$. По индуктивному предположению $v' \not\leq_{k-1,r} u'$. Пусть ϕ — дизъюнкция всех формул специального вида из $S_{k-1,r}$, не удовлетворяющих u'. Тогда $v' \not\models \neg\phi$. Поскольку это справедливо для любой структуры v', которую можно получить из v добавлением фишек $z_1^1, \dots, z_p^1$, имеем

$$v \not\models \exists z_1^k, \dots, z_p^k \neg\phi,$$

$$u \models \exists z_1^k, \dots, z_p^k \neg\phi.$$

Пришли к противоречию. Таким образом, у второго игрока есть выигрышная стратегия.

Вторая часть доказательства проводится тоже индукцией по k. Если $k = 0$ и у второго игрока есть выигрышная стратегия, то u и v удовлетворяют одним и тем же атомарным формулам, а следовательно, и одним и тем же бескванторным формулам. Пусть теперь $k > 0$. Предположим, что у второго игрока есть выигрышная стратегия в игре $C_{k,r}(u, v)$. Если $u \not\leq_{k,r} v$, то найдется формула ψ из $S_{k,r}$ такая, что $u \models \psi$ и $v \not\models \psi$. Можно считать, что ψ имеет вид $\exists z_1^k, \ldots, z_r^k \neg\phi$, где $\phi \in S_{k-1,r}$. Поскольку $u \models \psi$, первый игрок может так поставить свои фишки $z_1^k, \ldots, z_r^k$ в u, что полученная структура u' будет удовлетворять формуле $\neg\phi$. Второй игрок делает ответный ход, следуя выигрышной стратегии. Полученная в результате структура v' не удовлетворяет формуле $\neg\phi$. Но теперь у второго игрока есть выигрышная стратегия в игре $C_{k-1,r}(v', u')$. По индуктивному предположению $v' \leq_{k-1,r} u'$ и $v' \not\models \phi$. Противоречие. Теорема доказана.

Докажем теперь с помощью этой теоремы некоторые свойства отношения $\leq_{k,r}$. Мы будем обозначать конкатенацию двух слов (структур) u и v через uv или $u \cdot v$. Если w — слово алфавита A, мы будем использовать обозначение $w^n = w^{n-1} \cdot w$, где $n \geq 2$, $w^1 = w$.

ЛЕММА 1. *Пусть $k \geq 0$, $r \geq 1$ и $u_1, u_2, v_1, v_2, u_1u_2, v_1v_2$ — структуры. Тогда если $u_1 \leq_{k,r} v_1$ и $u_2 \leq_{k,r} v_2$, то $u_1u_2 \leq_{k,r} v_1v_2$.*

ДОКАЗАТЕЛЬСТВО. В случае $k = 0$ утверждение очевидно. Предположим, что $k > 0$. По теореме 2 достаточно показать, что у второго игрока есть выигрышная стратегия в игре $C_{k,r}(u_1u_2, v_1v_2)$. Пусть в структуре u_1u_2 первый игрок поставил p фишек ($p \leq r$). Можно считать, что фишки $x_1, \ldots, x_{p'}$, $p' \leq p$, поставлены в структуру u_1, а фишки $x_{p'+1}, \ldots, x_p$ поставлены в структуру u_2 (u_1' и u_2' — полученные структуры). Из посылки по теореме 2 заключаем, что у второго игрока есть выигрышные стратегии в играх $C_{k,r}(u_1, v_1)$ и $C_{k,r}(u_2, v_2)$, т.е. он может так поставить фишки в структуры v_1 и v_2, что полученные структуры v_1' и v_2' будут удовлетворять отношениям $v_1' \leq_{k-1,r} u_1'$, $v_2' \leq_{k-1,r} u_2'$. По индуктивному предположению $v_1'v_2' \leq_{k_1,r} u_1'u_2'$. Следовательно, у второго игрока есть выигрышная стратегия в игре $C_{k-1,r}(v_1'v_2', u_1'u_2')$. Далее второй игрок должен следовать этой стратегии. Лемма доказана.

Введем обозначения: $c(0, r) = 1$ при всех $r \in N$, $c(k, r) = r + (r+1)c(k-1, r) + 1$ при $k, r \in N$.

ЛЕММА 2. *Пусть* $k \geq 0$, $r \geq 1$, $w \in A^+$, $N_1, N_2 \geq c(k,r)$. *Тогда* $w^{N_1} \leq_{k,r} w^{N_2}$.

ДОКАЗАТЕЛЬСТВО. Будем считать, что $N_1 \neq N_2$. Пусть сначала $|N_1 - N_2| = 1$. Проведем индукцию по k. В случае $k = 0$ утверждение очевидно. Предположим, что $k > 0$, и утверждение справедливо для $k-1$. Покажем, что у второго игрока есть выигрышная стратегия в игре $C_{k,r}(w^{N_1}, w^{N_1+z})$, где $z = N_2 - N_1 \in \{1, -1\}$. Пусть в первом раунде игрок поставил p фишек ($p \leq r$) в слово w^{N_1}. В силу выбора N_1 в полученном слове найдется подслово $w^{c(k-1,r)+1}$ без фишек. Таким образом, первое слово можно записать в виде $w_1 w^{c(k-1,r)+1} w_2$, где w_1 и w_2 — подслова, в которые в первом раунде поставлены все p фишек. После первого раунда первое слово имеет вид $w_1' w^{c(k-1,r)+1} w_2'$. Второе слово можно записать так: $w^{N_1+z} = w_1 w^{c(k-1,r)+1+z} w_2$. Второй игрок должен так добавить фишки во второе слово, чтобы получить $w_1' w^{c(k-1,r)+1+z} w_2'$. По индуктивному предположению $w^{c(k-1,r)+1+z} \leq_{k-1,r} w^{c(k-1,r)+1}$, следовательно, по лемме 1 $w_1' w^{c(k-1,r)+1+z} w_2' \leq_{k-1,r} w_1' w^{c(k-1,r)+1} w_2'$. Далее второй игрок может играть, следуя своей выигрышной стратегии в игре $C_{k-1,r}(w_1' w^{c(k-1,r)+1+z} w_2', w_1' w^{c(k-1,r)+1} w_2')$. По теореме 2 получаем $w^{N_1} \leq_{k,r} w^{N_2}$.

Докажем лемму индукцией по $|N_2 - N_1|$. Пусть $|N_2 - N_1| > 1$ и лемма справедлива при меньших значениях $|N_2 - N_1|$. Возьмем натуральное число N_3 между N_1 и N_2. По индуктивному предположению $w^{N_1} \leq_{k,r} w^{N_3} \leq_{k,r} w^{N_2}$. Лемма доказана.

Для фиксированных k и r из N возьмем $M = r + 2(r+1)c(k-1,r) + 1$ и $K = M + c(k-1,r)$. Введем операции F и G на множестве A^+ так: $F(u) = u^M$, $G(u,v) = v^K u v^K$.

ЛЕММА 3. *Для любых* $u, v \in A^+$ *из* $u \leq_{k-1,r} v$ *следует* $F(v) \leq_{k,r} G(u,v)$, *где* $k, r \in N$.

ДОКАЗАТЕЛЬСТВО. Воспользуемся тем же приемом, который применяли для доказательства лемм 1 и 2. Рассматриваем игру $C_{k,r}(F(v), G(u,v))$. Пусть в первом раунде первый игрок поставил p фишек ($p \le r$) в слово $w_1 = F(v) = v^M$. В силу выбора M в полученном слове найдется подслово $v^{2c(k-1,r)+1}$ без фишек. Таким образом, первое слово можно записать в виде $u_1 v^{2c(k-1,r)+1} u_2 = u_1 v^{c(k-1,r)} v v^{c(k-1,r)} u_2$, где u_1 и u_2 — подслова, в которые в первом раунде поставлены все p фишек. После первого раунда первое слово будет иметь вид $w_1' = u_1' v^{c(k-1,r)} v v^{c(k-1,r)} u_2'$. Второе слово $w_2 = G(u,v)$ можно записать так: $w_1 v^{K-M} u v^{K-M} w_1 = u_1 v^{m_1} u v^{m_2} u_2$, где $m_1, m_2 \ge c(k-1,r)$. Второй игрок должен так добавить фишки во второе слово, чтобы получить $w_2' = u_1' v^{m_1} u v^{m_2} u_2'$. В силу леммы 2, $v^{m_1} \le_{k-1,r} v^{c(k-1,r)}$ и $v^{m_2} \le_{k-1,r} v^{c(k-1,r)}$. Из посылки по лемме 1 получаем $w_2' \le_{k-1,r} w_1'$. Далее второй игрок должен следовать своей выигрышной стратегии в игре $C_{k-1,r}(w_2', w_1')$. Так же, как и ранее, по теореме 2 получаем $F(v) \le_{k,r} G(u,v)$. Лемма доказана.

2. Доказательство теоремы 1

Доказательство основано на построении цепочек слов, удовлетворяющих некоторым условиям. Подобные цепочки часто используются при рассмотрении разностных иерархий. Разобьем доказательство на несколько вспомогательных утверждений. Получим сначала некоторое достаточное условие того, что предложение не принадлежит классу $D_{k,l}$.

ЛЕММА 4. *Пусть даны натуральные числа k, l и предложение ϕ такое, что для любого натурального числа r существует цепочка слов $v_1, v_2, \ldots, v_{l+1}$, удовлетворяющая условиям:*

$$v_i \le_{k,r} v_{i+1}, \quad i = 1, 2, \ldots, l, \tag{i}$$
$$v_{2i-1} \models \phi, \quad i \in N, \quad 2i - 1 \le l + 1, \tag{ii}$$
$$v_{2i} \not\models \phi, \quad i \in N, \quad 2i \le l + 1. \tag{iii}$$

Тогда $\phi \notin D_{k,l}$.

ДОКАЗАТЕЛЬСТВО. Проведем доказательство по индукции. Пусть $l = 1$ и $\phi \in S_{k,r}$. Возьмем r такое, что $\phi \in S_{k,r}$. По условию $v_1 \leq_{k,r} v_2, v_1 \models \phi$, $v_2 \not\models \phi$, противоречие. Пусть теперь $l > 1$ и ϕ принадлежит множеству формул $D_{k,l}$. Поскольку для класса языков $D_{k,l}$ справедливо, что $D_{k,l} = \{(A^+ \setminus L_1) \cap L_2 | \ L_1 \in D_{k,l-1}, \ L_2 \in S_k\}$, предложение ϕ можно представить в виде $\phi = \neg\phi_1 \wedge \phi_2$, где $\phi_1 \in D_{k,l-1}$, $\phi_2 \in S_k$. Возьмем p такое, что $\phi_2 \in S_{k,p} \subset$ $\subset S_{k,r}$, $r \geq p$. В силу (ii), имеем $v_1 \models \phi_2$, следовательно, по определению, $v_i \models \phi_2$ для всех $r \geq p$, $i \leq l+1$. Таким образом, при $r \geq p$ соотношения (ii) и (iii) останутся верными, если в них заменить ϕ на $\neg\phi_1$. Покажем с помощью индуктивного предположения, что $\phi_1 \notin D_{k,l-1}$, и получим таким образом противоречие. Для $r \geq p$ цепочки слов $v_2, \ldots, v_{l+1}$ удовлетворяют условиям леммы. Остается заметить, что если цепочка слов с требуемыми свойствами существует для некоторого $r = p$, то она будет обладать этими свойствами и для всех номеров r, меньших этого p.

СЛЕДСТВИЕ. *Пусть даны $k \geq 1$, $l \geq 2$ и предложения $\phi_2, \ldots, \phi_l \in S_k$ такие, что для любого r существует цепочка слов $v_1, \ldots, v_l$, удовлетворяющая условиям:*

$$v_i \leq_{k,r} v_{i+1}, \ i = 1, 2, \ldots, l-1, \tag{iv}$$

$$v_i \models \phi_i, \ i = 2, \ldots, l, \tag{v}$$

$$v_{i-1} \not\models \phi_i, \ i = 2, \ldots, l. \tag{vi}$$

Тогда $D_{k,l-1} \neq \check{D}_{k,l-1}$.

ДОКАЗАТЕЛЬСТВО. Пусть $\psi = \bigvee\limits_{j=1}^{m} (\phi_{2j} \wedge \neg\phi_{2j+1})$, где $2m \geq l$ и $\phi_j = false$ при $j > l$. Здесь $false$ — предложение, которому не удовлетворяет ни одно слово. По построению $\psi \in D_{k,l-1}$. С помощью леммы 4 покажем, что $\neg\psi \notin D_{k,l-1}$. Имеем $\phi = \neg\psi \equiv \bigwedge\limits_{j=1}^{m} (\neg\phi_{2j} \vee \phi_{2j+1})$. Возьмем p такое, что $\phi_i \in S_{k,p}$ $i = 2, \ldots, l$. Зафиксируем произвольное $R \geq 1$ и проверим выполнение условий леммы 4 для $r = R$. Возьмем цепочку слов $v_1, \ldots, v_l$ для $r = \max(R, p)$ из условия. Выполнение условия (i) леммы 4 обеспечивается условием (iv) следствия. Проверим условие (ii) леммы 4. Пусть $i \in N$ и $2i - 1 \leq l$. Покажем, что $v_{2i-1} \models \neg\phi_{2j} \vee \phi_{2j+1}$

для всех $j \in \{1,\dots,m\}$ и, следовательно, $v_{2i-1} \models \phi$. Пусть $2i-1 \geq 2j+1$, тогда по условию (ii) $v_{2j+1} \models \phi_{2j+1}$. По определению 1 и условию (i) $v_{2i-1} \models \phi_{2j+1}$, следовательно, $v_{2i-1} \models \neg\phi_{2j} \vee \phi_{2j+1}$. Пусть теперь $2i-1 \leq 2j-1$ и $2j \leq l$, тогда по условию (iii) $v_{2j-1} \models \neg\phi_{2j}$ и по условию (i) $v_{2i-1} \models \neg\phi_{2j}$, следовательно, $v_{2i-1} \models \neg\phi_{2j} \vee \phi_{2j+1}$. В случае $2j > l$ предложению $\neg\phi_{2j} \vee \phi_{2j+1}$ удовлетворяет любое слово, в том числе и v_{2i-1}. Проверим условие (iii) леммы 4. Пусть $i \in N$ и $2i \leq l$, тогда, в силу (ii) и (iii), $v_{2i} \not\models \neg\phi_{2i}$, и $v_{2i} \not\models \phi_{2i+1}$, следовательно, $v_{2i} \not\models \phi$. Таким образом, $\psi \in D_{k,l-1} \setminus \check{D}_{k,l-1}$. Следствие доказано.

Используя это следствие, мы докажем теорему 1. Сначала для произвольного $i \geq 2$ построим предложения $\phi_2^k,\dots,\phi_l^k \in S_k$, $k \in N$. Для этого нам будет нужна

ЛЕММА 5. *Для каждого предложение $\phi \in S_k$ существует формула $\hat{\phi}(x,y) \in S_k$ с двумя свободными переменными x и y такая, что для любой структуры $w = u_1(a_1,T_1)u(a_2,T_2)u_2$, где $x \in T_1$, $y \in T_2$, u_1, u_2 — структуры и u — слово в алфавите A, $w \models \hat{\phi}(x,y)$ тогда и только тогда, когда $u \models \phi$.*

ДОКАЗАТЕЛЬСТВО. Можно считать, что в предложении ϕ нет переменных x и y. Заменим в формуле ϕ каждое вхождение $\exists z(\dots)$ на $\exists z(x < z < y \wedge \dots)$ и каждое вхождение $\forall z(\dots)$ на $\forall z(x < z < y \to \dots)$. Полученная в результате формула $\hat{\phi}(x,y)$ удовлетворяет требуемому свойству.

Пусть алфавит A состоит из букв $a, b, \dots$. Зафиксируем произвольное $l \geq 2$. Предложения $\phi_2^k,\dots,\phi_l^k \in S_k$ будем строить индукцией по k.

$$\phi_j^1 = \exists x_1^1,\dots,x_j^1\Big(\bigwedge_{i=1}^{j} Q_a x_i^1 \wedge \bigwedge_{i=1}^{j-1} x_i^1 < x_{i+1}^1\Big), \qquad j = 2,3,\dots,l,$$

$$\begin{aligned}\phi_j^k = \exists x_1^k,\dots,x_{2p}^k\Big(&\bigwedge_{i=1}^{2p} Q_b x_i^k \wedge \bigwedge_{i=1}^{2p-1} x_i^k < x_{i+1}^k \wedge \\ &\wedge\forall y^k\Big(\bigwedge_{i=1}^{p-1} \neg(x_i^k < y^k < x_{i+1}^k) \wedge \bigwedge_{i=p+1}^{2p-1} \neg(x_i^k < y^k < x_{i+1}^K)\Big)\wedge \\ &\wedge\neg\hat{\phi}_{l+2-j}^{k-1}(x_p^k, x_{p+1}^k)\Big),\end{aligned}$$

где $k > 1$, $p = 2(k-1)$, $j = 2,3,\dots,l$.

ЗАМЕЧАНИЕ. Слово $w \in A^+$ удовлетворяет предложению ϕ_j^k, $k > 1$, тогда и только тогда, когда w содержит подслово $b^{2(k-1)}ub^{2(k-1)}$ такое, что $u \models \neg\phi_{l+2-j}^{k-1}$.

ЛЕММА 6. *Пусть u, v, w — слова алфавита A, $j = 2, 3, \ldots, l$, $k \in N$. Тогда если $w \models \phi_j^k$, то $uwv \models \phi_j^k$.*

ДОКАЗАТЕЛЬСТВО. Если $w \models \phi_j^1$, то слово w содержит не менее j букв a. Но тем же свойством будет обладать и слово uwv, а это значит, что $uwv \models \phi_j^1$. Пусть теперь $w \models \phi_j^k$, $k > 1$. В силу замечания, слово $b^{2(k-1)}ub^{2(k-1)}$ является подсловом w. Но оно будет также подсловом слова uwv. Таким образом, uwv также будет удовлетворять предложению ϕ_j^k. Что и требовалось доказать.

Возьмем произвольное $r \geq 1$ и построим цепочки слов $v_1^k, v_2^k, \ldots, v_l^k$, удовлетворяющие условиям следствия для предложений $\phi_2^k, \ldots, \phi_l^k \in S_k$. Цепочки на уровне k, $k > 1$, будем строить на основе цепочек уровня $k-1$. В случае $k = 1$ возьмем $v_i^1 = a^i$, $i = 1, 2, \ldots, l$. Предположим теперь, что $k > 1$ и на $(k-1)$–м шаге мы построили цепочку слов $v_1^{k-1}, \ldots, v_l^{k-1}$.

Обозначим

$$u_i^l = bv_i^{k-1}b, \quad i = 1, 2, \ldots, l,$$

$$u_i^j = G(u_i^{j-1}, u_{l+2-j}^{j-1}),\ j = 2, 3, \ldots, l,\ i = 1, 2, \ldots, l+1-j,$$

$$v_i^k = F^{l-i}(u_{l+1-i}^i), \quad i = 1, \ldots, l,$$

где F и G — функции из леммы 3.

ЛЕММА 7. *Слова $v_1^k, v_2^k, \ldots, v_l^k$, $k \in N$ удовлетворяют условию* (i) *следствия.*

ДОКАЗАТЕЛЬСТВО. В случае $k = 1$ утверждение легко проверить, используя теорему 2. Предположим теперь, что $k > 1$ и утверждение справедливо для $k-1$. По предположению индукции $u_{l-i}^1 \leq_{k-1,r} u_{l+1-i}^1$, $i = 1, 2, \ldots, l-1$. В силу леммы 1, $u_{l-i}^i \leq_{k-1,r} u_{l+1-i}^i$, $i = 1, 2, \ldots, l-1$. В силу леммы 3, $F(u_{l+1-i}^i) \leq_{k,r} G(u_{l-i}^i, u_{l+1-i}^i) = u_{l-i}^{i+1}$. В силу леммы 1, $F^{l-i}(u_{l+1-i}^i) \leq_{k,r} F^{l-i-1}(u_{l-i}^{i+1})$, т.е. $v_i^k \leq_{k,r} v_{i+1}^k$, $i = 1, 2, \ldots, l-1$. Что и требовалось доказать.

Если w — слово в алфавите $A = \{a, b, \ldots\}$, то через $C(w)$ мы будем обозначать подслово слова w, которое получается удалением всех символов b из начала и конца слова w (если они там есть).Например, $bbabb = bbC(bbabb)bb$.

Непосредственно из построения следует, что слова v_i^k, $k \geq 2$, $i = 1, 2, \ldots, l$, удовлетворяют следующим свойствам:

1) $v_i^k = b^{k-1}C(v_i^k)b^{k-1}$,

2) v_i^k не содержит подслова $b^{2(k-1)+1}$,

3) $v_i^k = \bigodot_{j=1}^{t}(b^{k-1}u_j b^{k-1})$, $t \in N$,

$$\{u_j | j = 1, \ldots, t\} = \{C(v_{l+1-i}^{k-1}), C(v_{l+2-i}^{k-1}), \ldots, C(v_l^{k-1})\}.$$

Здесь используются обозначения: $\bigodot_{i=1}^{1} u_i = u_1$, $\bigodot_{i=1}^{t} u_i = \bigodot_{i=1}^{t-1} u_i \cdot u_t$, $t > 1$.

ЛЕММА 8. *Для того чтобы $v_i^k \models \phi_j^k$, необходимо и достаточно, чтобы $C(v_i^k) \models \phi_j^k$, $k \in N$, $i \in \{1, 2, \ldots, l\}$, $j \in \{2, 3, \ldots, l\}$.*

ДОКАЗАТЕЛЬСТВО. В случае $k = 1$ утверждение очевидно. Пусть $k > 1$. В силу п.1 имеем $v_i^k = b^{k-1}C(v_i^k)b^{k-1}$. В силу леммы 6, получаем доказательство достаточности. Заметим, что если $v = b^{2(k-1)}ub^{2(k-1)}$ является подсловом v_i^k, то v также является подсловом $C(v_i^k)$. Теперь для доказательства необходимости осталось воспользоваться замечанием.

Лемма доказана.

Осталось показать, что для построенных слов и предложений выполняются условия (v) и (vi) следствия.

ЛЕММА 9. *Для всех $i, j \in \{2, 3, \ldots, l\}$, $i \geq j$, справедливы следующие соотношения: $v_i^k \models \phi_j^k$ и $v_{i-1}^k \not\models \phi_i^k$.*

ДОКАЗАТЕЛЬСТВО. Проведем доказательство по индукции. Для случая $k = 1$ утверждение очевидно. Предположим теперь, что $k > 1$ и утверждение справедливо для $k - 1$.

Покажем сначала, что $v_i^k \models \phi_j^k$. По предположению индукции $v_{l+1-j}^{k-1} \models \neg\phi_{l+2-j}^{k-1}$. По лемме 8, $C(v_{l+1-j}^{k-1}) \models \neg\phi_{l+2-j}^{k-1}$.

В силу п.3 и замечания $v_i^k = \ldots b^{2(k-1)} C(v_{l+1-j}^{k-1}) b^{2(k-1)} \ldots \models$ $\models \phi_j^k$.

Покажем теперь, что $v_{i-1}^k \not\models \phi_i^k$. В силу п.3, $v_{i-1}^k = $ $= \bigodot_{j=1}^{i} (b^{k-1} u_j b^{k-1})$, $u_j \in \{C(v_{l+2-i}^{k-1}), C(v_{l+3-i}^{k-1}), \ldots, C(v_l^{k-1})\}$. Проведем доказательство от противного. Предположим, что $v_{i-1}^k \models \phi_i^k$, тогда найдется подслово

$$u_s b^{2(k-1)} u_{s+1} b^{2(k-1)} \ldots b^{2(k-1)} u_{s+p},$$

удовлетворяющее предположению $\neg\phi_{l+2-i}^{k-1}$. По лемме 6, $u_s \models \neg\phi_{l+2-i}^{k-1}$, $u_s = C(v_m^{k-1})$, $m \geq l+2-i$. В силу леммы 8, $v_m^{k-1} \models \neg\phi_{l+2-i}^{k-1}$, но по предположению индукции $v_m^{k-1} \models \phi_{l+2-i}^{k-1}$. Пришли к противоречию. Лемма доказана.

Теперь из следствия следует теорема 1.

3. Дополнительные замечания

1. Нетрудно проверить, что $\bigcup_{l=1}^{\infty} D_{k,l} \subset S_{k+1} \cap \check{S}_{k+1}$ для каждого $k \in N$. Можно показать, что для $k = 1$ включение будет собственное (это контрастирует с хорошо известным из логики фактом, что при любом k выполняется равенство $\bigcup_{l=1}^{\infty} D_{k,l} = S_{k+1} \cap \check{S}_{k+1}$, где классы $D_{k,l}, S_k$ определяются так же, только вместо эквивалентности на конечных структурах надо взять обычную эквивалентность формул). Действительно, рассмотрим язык L, состоящий из слов, которые начинаются с буквы $a \in A$. Этот язык можно описать предложением $\phi_1 \in S_2$ и предложением $\phi_2 \in \check{S}_2$, где $\phi_1 = \exists x \forall y (Q_a x \wedge \neg(y < x))$, $\phi_2 = $ $= \forall y \exists x (Q_a x \wedge \neg(y < x))$. Следовательно, $L \in S_2 \cap \check{S}_2$. Рассмотрим бесконечную цепочку слов $w_1, w_2, \ldots,$ где $w_1 = a$, $w_{2i} = b w_{2i-1}$, $w_{2i+1} = a w_{2i}$, $i \in N$. Поскольку каждое последующее слово в этой цепочке содержит предыдущее как подслово, будут выполнены соотношения $w_i \leq_{1,r} w_{i+1}, i, r \in N$. По лемме 4 получаем $L \notin D_{1,l}$, $l \in N$.

2. Аналог теоремы справедлив для случая ω-языков, состоящих из бесконечных слов. Для доказательства можно модифицировать наше доказательство теоремы 1. При этом цепочки слов $v_1^k, \ldots, v_l^k$ из нашего доказательства можно заменить на цепочки ω-слов $v_1^k b^\omega, \ldots, v_l^k b^\omega$, а предложения ϕ_i^k оставить без изменения.

Автор статьи выражает благодарность профессору В.Л.Селиванову за активную помощь в работе.

Литература

1. ADDISON J. The method of alternating chains //The Theory of Models (J.W.Addison, L.A.Henkin, and A.Tarski, eds.). — Amsterdam: North-Holland, 1965. — Proceedings of the 1963 International Symposium at Berkeley. — P.1–16.

2. BRZOZOWSKI J.A., KNAST R. The dot-depth hierarchy of star-free languages is infinite //J. Comput. System Sci. — 1978. — Vol. 16. — P.37–55.

3. ЕРШОВ Ю.Л. Об одной иерархии множеств //Алгебра и логика. — 1968. — Т. 7. — С. 15–46.

4. HAUSDORFF F. Grundzüge der Mengenlehre. — Leipzig, 1914.

5. PERRIN D., PIN J.E. First-order logis and star-free sets //J. Comput. System. Sci. — 1986. — Vol. 32. — P. 393–406.

6. СЕЛИВАНОВ В.Л. Тонкая иерархия формул //Алгебра и логика. — 1991. — Т. 30. — С. 568–582.

7. STRAUBING H. Finite automata, formal logic, and circuit complexity. — Birkhäuser, 1994.

8. THOMAS W. Classifying regular events in symbolic logic. //J. Comput. System Sci. — 1982. — Vol. 25. — P. 360–376.

9. THOMAS W. An application of the Ehrenfeucht—Fraïssé game in formal language theory. //Bull. Soc. Math. France, Mem. — 1984. — Vol. 16. — P. 11–21.

10. THOMAS W. Automata on infinite objects //Handbook of Theoretical Computer Science (J.van Leeuwen, ed.). — MIT Press, Cambridge, Massachusetts. — 1990.

Поступила в редакцию
4 марта 1998 года