Math-Net.Ru

Общероссийский математический портал

С. С. Мартынов, NP-полнота специальных задач редактирования слов, *Матем. вопр. криптогр.*, 2013, том 4, выпуск 4, 77–93

DOI: <https://doi.org/10.4213/mvk103>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 54.210.77.194

26 апреля 2023 г., 16:30:12

УДК 510.53

NP-полнота специальных задач редактирования слов

С. С. Мартынов

Лаборатория ТВП, Москва

Получено 20.IV.2012

В статье устанавливается NP-полнота задач редактирования слов относительно языка, заданного в терминах ограничений на множество подслов в его словах. Действие операций редактирования состоит в замене подслов редактируемого слова, принадлежащих некоторому блоковому коду, словами другого блокового кода.

Ключевые слова: задачи редактирования слов, NP-полные задачи, блоковый код

NP-completeness of special string editing problems

S. S. Martynov

TVP Laboratory, Moscow

Abstract. We establish the NP-completeness of the string editing problems with respect to a language defined by restrictions on a subwords of its words. The editing operations consists in a replacement of the substrings belonging to a specified block code, by the words of another block code.

Key words: string editing problems, NP-complete problems, block code

Citation: *Mathematical Aspects of Cryptography*, 2013, vol. 4, no. 4, pp. 77–93 (Russian).

© 2013 Мартынов С. С.

Введение

Пусть Σ — конечный алфавит, Σ^r — множество всех слов длины r в алфавите Σ , Σ^* — множество всех конечных слов в алфавите Σ , $L \subseteq \Sigma^*$ — некоторый язык [1]. Редактированием слова $S \in \Sigma^*$ относительно языка L называется процедура выбора минимальной последовательности операций из заданного набора операций Φ , переводящей S в какое-либо слово из L (см., например, [2, 3]). Ниже будем говорить о языке L как о базовом языке по отношению к соответствующей процедуре редактирования.

Рассмотрим язык $L = \Sigma^* \langle F \rangle$, заданный в терминах ограничения $\langle F \rangle$ на множество подслов, которые могут встречаться в его словах [3]: если $F(w_1, \dots, w_l)$ — предикат, определяемый отображением $F: \Sigma^l \rightarrow B$ (где $B = \{0, 1\}$), то язык L определяется как множество всех слов вида $S = s_1 \dots s_n$ ($n \geq l$), для которых $F(s_i, \dots, s_{i+l-1}) = 1$ ($i = 1, \dots, n - l + 1$). Будем использовать общее обозначение λ для набора $\lambda = (\lambda_1, \dots, \lambda_l)$ значений переменных $w_1, \dots, w_l$ и слова $\lambda = \lambda_1 \dots \lambda_l$. Таким образом, слово S содержится в языке $\Sigma^* \langle F \rangle$, если и только если все его подслова длины l содержатся в области истинности предиката F .

В предположении, что определены алфавит Σ , набор Φ операций редактирования и фиксированные натуральные числа a, T , ниже рассматриваются следующие задачи $\mathbf{K}_T$, связанные с задачами редактирования относительно заданного базового языка $\Sigma^* \langle F \rangle$, в котором предикат F существенно зависит не более, чем от T переменных (см. также [3]).

Условие. Заданы: слово $S \in \Sigma^*$ ($|S| \geq T$); натуральные числа l, k , удовлетворяющие условиям $T \leq l \leq |S|$, $k \leq a \cdot |S|$; предикат $F(w_1, \dots, w_l)$, существенно зависящий не более, чем от T переменных.

Вопрос. Существует ли последовательность ϕ операций редактирования из набора Φ , имеющая длину $|\phi| \leq k$ и переводящая слово S в некоторое слово из $\Sigma^* \langle F \rangle$?

Рассматриваемые задачи $\mathbf{K}_T$ относятся к задачам распознавания и могут иметь положительный или отрицательный ответ («да» или «нет»).

В работах [3, 4] установлена NP-полнота ряда задач $\mathbf{K}_T$, в которых набор операций редактирования Φ может содержать операции вставки знака, удаления знака, перестановки пары знаков, замены знака элементом алфавита Σ .

В ситуации (см. также [5, 6]), когда действие операций редактирования выражается в замене подслов редактируемого слова, являющихся словами некоторого блокового кода Ω [7], на слова другого блокового кода Θ , процесс редактирования напоминает процедуру подстрочного пере-

вода. В данной статье установлена NP-полнота соответствующих задач K_T , ниже называемых задачами о ПОДСТРОЧНОМ ПЕРЕВОДЕ.

Структура дальнейшего изложения включает три параграфа и заключение.

В § 1 вводятся необходимые термины и обозначения, а также приводится описание предложенной в работе [3] схемы кодирования условий NP-полной задачи ВЫПОЛНИМОСТЬ, используемой в данной статье при построении полиномиальных алгоритмов сведения указанной задачи к задачам редактирования слов.

В § 2 устанавливается NP-полнота задачи о ПОДСТРОЧНОМ ПЕРЕВОДЕ для $|\Sigma| \geq 3$ (при произвольной длине блоков в кодах Ω, Θ). Обсуждается применяемый метод построения полиномиальных алгоритмов сведения NP-полной задачи ВЫПОЛНИМОСТЬ к задачам редактирования слов относительно языков с ограничениями на множество подслов.

В § 3 в условиях, когда коды Ω, Θ являются подмножествами некоторого блокового кода без запятой Ξ [7], устанавливается NP-полнота задач о ПОДСТРОЧНОМ ПЕРЕВОДЕ, в том числе и для двоичного алфавита ($\Sigma = B$).

В заключении представлено беглое обсуждение возможных направлений практического использования полученных в статье результатов.

§ 1. Обозначения. Схема кодирования условий задачи ВЫПОЛНИМОСТЬ

Будем использовать следующие обозначения и термины.

Пусть $X, Y \in \Sigma^*$, $s \geq 2$ — натуральное число. Будем обозначать результат конкатенации слов X, Y через XY и использовать обозначения $XX = X^2$, $XX^{s-1} = X^s$.

Будем использовать запись $X \cdot \Phi$ для обозначения результата применения к слову X последовательности Φ операций редактирования.

Следуя [7], элементы множества Σ^n будем называть *блоками длины n* ; множество $\Xi \subseteq \Sigma^n$ — *блоковым словарем* или *кодом со словами длины n* . *Стыком* блоков $\alpha, \beta \in \Sigma^n$ ($\alpha = \alpha_1 \dots \alpha_n$, $\beta = \beta_1 \dots \beta_n$) называется слово $\sigma \in \Sigma^n$, представимое в виде конкатенации некоторых суффикса α' слова α и префикса β' слова β (т.е. $\sigma = \alpha_i \dots \alpha_n \beta_1 \dots \beta_{i-1}$ ($i = 2, \dots, n$)). Блоковый словарь Ξ называется *кодом без запятой* [7], если стык любых двух слов из Ξ не содержится в словаре Ξ .

Для $x, \alpha \in B$ будем использовать обозначение $x^\alpha = x \oplus \alpha \oplus 1$ (определения логических операций, применяемых ниже при проведении вычислений с элементами двоичного алфавита B , содержатся, например, в работе [8]).

Пусть $M = (m_{ij})_{i=1,\dots,h; j=1,\dots,w}$ — матрица размера $h \times w$, элементами которой являются блоки $m_{ij} \in \Sigma^s$ ($s \geq 1$). Обозначим через $W(M)$ слово длины $h \times w \times s$, являющееся результатом конкатенации блоков m_{ij} в порядке их следования в строках матрицы M :

$$W(M) = m_{1,1} m_{1,2} \dots m_{1,w} m_{2,1} \dots m_{2,w} \dots m_{h,1} \dots m_{h,w}.$$

Ниже в контексте исследования сложности задач редактирования слов рассматривается возможность построения алгоритмов сведения к указанным задачам NP-полной задачи ВЫПОЛНИМОСТЬ, которая формулируется следующим образом [1].

Условие. Заданы множество переменных U и набор C дизъюнкций над U .

Вопрос. Существует ли выполняющий набор значений истинности для C ?

Рассмотрим индивидуальную задачу $\tau \in$ ВЫПОЛНИМОСТЬ, в которой $U = \{x_1, \dots, x_n\}$, $C = \{c_1, \dots, c_m\}$. Для записи условий задачи τ будем использовать следующую схему кодирования, построенную в работе [3].

Построим матрицу $C(\tau)$ размера $(2m+1) \times (n+2)$, элементами которой являются двоичные блоки длины 2. Для нумерации строк используется пара индексов i, α ($i = 0, 1, \dots, m$, $\alpha = 0, 1$: если $i = 0$, то $\alpha = 1$; при $i = 1, \dots, m$ индекс α принимает значения 0, 1); для нумерации двоичных блоков длины 2 в строках матрицы $C(\tau)$ используются целые числа $j = 0, \dots, n+1$. В строке i, α на месте $j = 1, \dots, n$ находится блок $b_{i,\alpha j} = \beta_{i,\alpha j} \gamma_{i,\alpha j}$, в котором элемент $\beta_{i,\alpha j}$ принимает значение 1, если дизъюнкция c_i содержит литерал x_j^α ; в противном случае $\beta_{i,\alpha j} = 0$.

Набор значений элементов $\beta_{i,\alpha j}$ в блоках $b_{i,\alpha j}$ ($\alpha = 0, 1$; $j = 1, \dots, n$) содержит полную информацию о литералах дизъюнкции c_i ($i = 1, \dots, m$). В указанных блоках полагаем $\gamma_{i,\alpha j} = 0$. Остальные блоки $b_{i,\alpha j}$ имеют следующий вид:

$$b_{0,1j} = 01 \quad (j = 0, 1, \dots, n+1), \quad (1)$$

$$b_{i,0,0} = 11 \quad (i = 1, \dots, m), \quad (2)$$

$$b_{i,1,0} = b_{i,0,n+1} = 00 \quad (i = 1, \dots, m), \quad (3)$$

$$b_{i,1,n+1} = 01 \quad (i = 1, \dots, m). \quad (4)$$

Пусть $\epsilon = (\epsilon_1, \dots, \epsilon_n)$ — некоторый набор значений переменных $x_1, \dots, x_n$. Введем дополнительные переменные x_0, x_{n+1} и будем считать, что они принимают значения $\epsilon_0 = \epsilon_{n+1} = 0$. Построим расширенную матрицу $C(\tau, \epsilon)$, заменив блоки $b_{i,\alpha j}$ матрицы $C(\tau)$ двоичными блоками $b_{i,\alpha j}(\epsilon)$ длины 4

$$b_{i,\alpha j}(\epsilon) = \beta_{i,\alpha j} \gamma_{i,\alpha j} \epsilon_{i,\alpha j} \mu_{i,\alpha j}, \quad (5)$$

полагая

$$\varepsilon_{i,\alpha j} = \varepsilon_{\alpha j} = \varepsilon_j^\alpha \quad (j = 0, \dots, n+1). \quad (6)$$

Пусть $\beta\gamma\epsilon\mu, \beta\gamma\epsilon\mu'$ — пара подряд идущих блоков вида (5) в слове $W(C(\tau, \epsilon))$. Значения элементов $\mu_{i,\alpha j}$ в блоках $b_{i,\alpha j}(\epsilon)$ определяются следующими рекуррентными соотношениями:

$$\mu_{0,1,0} = 0, \quad (7)$$

$$\mu' = \mu(1 \oplus \beta\epsilon) \vee \gamma\beta. \quad (8)$$

Согласно Лемме 1 [3] набор значений переменных ϵ является выполняющим для задачи τ тогда и только тогда, когда для любой пары блоков $\beta\gamma\epsilon\mu, \beta\gamma\epsilon\mu'$ вида (5), подряд идущих в слове $W(C(\tau, \epsilon))$, выполняется равенство

$$\mu\gamma = 0. \quad (9)$$

Свойства описанной схемы кодирования задачи ВЫПОЛНИМОСТЬ используются ниже при обосновании возможности ее сведения к соответствующим задачам K_T редактирования слов. При рассмотрении возможности сведения задачи ВЫПОЛНИМОСТЬ к конкретным задачам K_T указываются правила, позволяющие, используя условия произвольной индивидуальной задачи $\tau \in$ ВЫПОЛНИМОСТЬ, строить условия равносильной ей индивидуальной задачи $\kappa(\tau) \in K_T$. В каждом рассматриваемом случае будет показано, что задача $\kappa(\tau)$ имеет положительное решение тогда и только тогда, когда такое же решение имеет задача τ . При этом временная сложность [1] алгоритма построения условий задачи $\kappa(\tau)$ ограничена полиномом от длины записи условия задачи τ (или, другими словами, построение условия задачи $\kappa(\tau)$ может быть выполнено за время, ограниченное полиномом от длины записи условия задачи τ).

§ 2. Сложность задач о подстрочном переводе при произвольной длине блоков в словарях Ω, Θ

Рассмотрим блоковые словари $\Omega, \Theta \subset \Sigma^*$, в которых длины блоков равны, соответственно, r_1, r_2 ; $|\Omega| = 4, |\Theta| = 16$. Для нумерации слов в словарях Ω, Θ будем использовать соответственно двоичные слова $\beta\gamma \in B^2, \beta\gamma\epsilon\mu \in B^4$. С учетом этого можно записать: $\Omega = \{\alpha(\beta\gamma): \beta\gamma \in B^2\}, \Theta = \{\alpha(\beta\gamma\epsilon\mu): \beta\gamma\epsilon\mu \in B^4\}$. Поскольку это не вызовет путаницы, для произвольных $\omega \in \Omega, \theta \in \Theta$ будем использовать записи $\beta(\omega), \gamma(\omega), \beta(\theta), \gamma(\theta), \epsilon(\theta), \mu(\theta)$.

Будем предполагать, что действие операций редактирования выражается в замене подслов редактируемого слова, являющихся словами кода Ω ,

на слова кода Θ . Как и выше, будем обозначать набор операций редактирования через Φ .

Ниже предполагается, что выполнены следующие условия:

- a1) алфавит $\Sigma = \{\#, \sigma_2, \dots, \sigma_{|\Sigma|}\}$ содержит $|\Sigma| \geq 3$ символов;
- a2) слова ω кода Ω имеют вид $\#\omega_2 \dots \omega_{r_1}$, где $\omega_i \neq \#$ ($i = 2, \dots, r_1$);
- a3) словарь Θ является кодом без запятой;
- a4) символ $\#$ не встречается в словах кода Θ ;
- a5) для произвольного $\beta\gamma\epsilon\mu \in B^4$ найдется такая операция редактирования $\varphi_{\beta\gamma\epsilon\mu} \in \Phi$, что $\varphi_{\beta\gamma\epsilon\mu}(\alpha(\beta\gamma)) = \theta(\beta\gamma\epsilon\mu)$;
- a6) для произвольных $\beta\gamma, \beta'\gamma' \in B^2$ ($\beta\gamma \neq \beta'\gamma'$)

$$\{\alpha(\beta\gamma) \cdot \varphi : \varphi \in \Phi\} \cap \{\alpha(\beta'\gamma') \cdot \varphi : \varphi \in \Phi\} = \emptyset.$$

Замечание 1. В силу условия a2 словарь Ω является кодом без запятой.

Теорема 1. Если выполнены условия a1 — a6, то задача $K_3 r_2$ NP-полна.

Доказательство. 1) Докажем, что $K_3 r_2 \in \text{NP}$. Для доказательства достаточно убедиться в том, что для произвольной индивидуальной задачи $k \in K_3 r_2$ недетерминированному алгоритму достаточно угадать последовательность Φ операций редактирования и что проверка соответствия последовательности Φ условиям задачи k может быть выполнена за время, ограниченное полиномом от длины записи условия задачи k (см. [1]). Процедура такой проверки состоит в вычислении значения слова $S \cdot \Phi$ с последующей проверкой справедливости включения $S \cdot \Phi \in \Sigma^*(F)$.

Учитывая неравенство $|\Phi| \leq k \leq a \cdot |S|$, а также вид применяемых операций редактирования, можно заключить, что слово $S \cdot \Phi$ может быть вычислено за время, ограниченное линейной функцией от длины записи слова S .

Замечая, что $|S \cdot \Phi| \leq |S| + k|r_2 - r_1| \leq |S|(1 + a \cdot |r_2 - r_1|)$, а также учитывая, что предикат $F(w_1, \dots, w_l)$, определяющий ограничение $\langle F \rangle$, существенно зависит от фиксированного числа $T \leq 3r_2$ переменных, убеждаемся, что проверка справедливости включения $S \cdot \Phi \in \Sigma^*(F)$ может быть выполнена за время, ограниченное линейной функцией от длины записи слова S .

2) Докажем сводимость задачи ВЫПОЛНИМОСТЬ к задаче $K_3 r_2$.

Для произвольной индивидуальной задачи $\tau \in \text{ВЫПОЛНИМОСТЬ}$ будет указано правило построения условий задачи $k(\tau) \in K_3 r_2$, которая имеет положительное решение, если и только если такое же решение имеет задача τ .

Для записи условий задачи $k(\tau)$ будем использовать матрицу $D(\tau)$, получаемую путем замены блоков b_{i, α_j} , являющихся элементами описанной

в § 1 матрицы $C(\tau)$, на слова $\omega_{i,\alpha j} = \omega(b_{i,\alpha j})$ ($i = 0, \dots, m$; $\alpha = 0, 1$; $j = 0, 1, \dots, n+1$). Соответственно, в качестве аналога матрицы $C(\tau, \epsilon)$ может рассматриваться некоторая матрица $D(\tau, \epsilon)$, получаемая путем замены блоков $b_{i,\alpha j}(\epsilon)$, являющихся элементами матрицы $C(\tau, \epsilon)$, на слова $\theta_{i,\alpha j}(\epsilon) = \theta(b_{i,\alpha j}(\epsilon))$. В рамках указанной конструкции слово $W(D(\tau, \epsilon))$ может быть получено из слова $W(D(\tau))$ путем применения последовательности операций редактирования $\varphi \in \Phi$, преобразующих слова кода Ω в слова кода Θ .

В условии задачи $\kappa(\tau)$ полагаем

$$S = QW(D(\tau)), \quad (10)$$

где $Q = \sigma^{r_2 \cdot (n+2)}$ (здесь $\sigma \in \Sigma$, $\sigma \neq \#$), $l = r_2(n+4)$,

$$k = (2m+1)(n+2). \quad (11)$$

Предикат $F = F(w_1, w_2, \dots, w_{r_2(n+4)})$ существенно зависит от $3r_2$ переменных $w_1, \dots, w_{r_2}, w_{r_2(n+2)+1}, \dots, w_{r_2(n+3)}, w_{r_2(n+3)+1}, \dots, w_{r_2(n+4)}$. Для удобства изложения будем использовать новые обозначения для части из указанных переменных:

$$(w_{r_2(n+2)+1}, \dots, w_{r_2(n+3)}) = (y_1, \dots, y_{r_2}),$$

$$(w_{r_2(n+3)+1}, \dots, w_{r_2(n+4)}) = (z_1, \dots, z_{r_2}),$$

а также будем использовать запись

$$F(w_1, w_2, \dots, w_{r_2(n+4)}) = g(w_1, \dots, w_{r_2}, y_1, \dots, y_{r_2}, z_1, \dots, z_{r_2})$$

(сохранив при этом обозначения $\Sigma^* \langle F \rangle$ и $\langle F \rangle$ для базового языка L и соответствующего ограничения).

Перечислим условия, которым должны удовлетворять наборы $\xi = \xi_1, \dots, \xi_{r_2}$, $\psi = \psi_1, \dots, \psi_{r_2}$, $\zeta = \zeta_1, \dots, \zeta_{r_2}$ значений переменных

$$w_1, \dots, w_{r_2}, y_1, \dots, y_{r_2}, z_1, \dots, z_{r_2}, \quad (12)$$

для которых выполняется равенство $g(\xi, \psi, \zeta) = 1$.

1°. Наборы значений ψ, ζ таковы, что или $\psi, \zeta \in \Theta$, или слово $\psi\zeta$ представимо в виде $\psi\zeta = \lambda_1\theta\lambda_2$, где $\theta \in \Theta$, а слова λ_1, λ_2 являются соответственно, суффиксом и префиксом некоторых слов из Θ . В последнем случае обозначим $|\lambda_1| = q$. Если же $\psi\zeta = \theta_1\theta_2$ ($\theta_1, \theta_2 \in \Theta$), то будем полагать $q = 0$.

Замечание 2. В силу условия $a3$, при выполнении условия 1° слово $\psi\zeta$ однозначно определяет значение q (поскольку согласно [7] блок-слово

варь без запятой Θ со словами длины r_2 является синхронизируемым с задержкой, не превышающей $2r_2 - 1$).

В случае $q = 0$ потребуем, чтобы наборы значений переменных (12) из описания предиката g удовлетворяли также следующим условиям.

2°. Если $q = 0$ и $\chi(\Psi) = 0$, то набор ξ, ψ, ζ значений переменных (12) должен удовлетворять условиям $\xi \in \Theta$, $\varepsilon(\xi) = \varepsilon(\psi) \oplus 1$.

3°. Если $q = 0$, то выполняется соотношение

$$\mu(\zeta) = \mu(\psi)(1 \oplus \beta(\zeta)\varepsilon(\zeta)) \vee \chi(\zeta)\beta(\zeta).$$

4°. Если $q = 0$, то выполняется соотношение $\mu(\psi)\chi(\zeta) = 0$.

Этим завершается построение условия задачи $\kappa(\tau)$, которое, очевидно, может быть выполнено за время, ограниченное полиномом от длины кода задачи τ .

3) Предположим, что существует выполняющий набор ε значений переменных $x_1, \dots, x_n$ для задачи τ . Пусть $d_{i,\alpha j}$ — некоторый блок, являющийся элементом матрицы $D(\tau)$. Рассмотрим соответствующий блок $c_{i,\alpha j}(\varepsilon)$ матрицы $C(\tau, \varepsilon)$, а также значения $\beta = \beta(d_{i,\alpha j})$, $\gamma = \gamma(d_{i,\alpha j})$ и $\varepsilon = \varepsilon(c_{i,\alpha j}(\varepsilon))$, $\mu = \mu(c_{i,\alpha j}(\varepsilon))$. Из условия $a5$ следует, что найдется такая операция редактирования $\varphi_{\beta\gamma\varepsilon\mu} \in \Phi$, что $d_{i,\alpha j} \cdot \varphi_{\beta\gamma\varepsilon\mu} = \alpha\beta\gamma\varepsilon\mu$. Будем рассматривать матрицу $D(\tau, \varepsilon)$, в которой на местах i, α, j находятся блоки $d_{i,\alpha j}(\varepsilon) = d_{i,\alpha j} \cdot \varphi_{\beta\gamma\varepsilon\mu}$. Покажем, что тогда

$$V = QW(D(\tau, \varepsilon)) \in \Sigma^*(F). \quad (13)$$

В самом деле, из определения матрицы $D(\tau, \varepsilon)$ следует выполнение условия 1° для всех наборов значений переменных $y_1, \dots, y_{r_2}, z_1, \dots, z_{r_2}$. Выполнение условия 2° обеспечивается формулой (6), условия 3° — соотношением (8). Выполнение условия 4° вытекает из соотношения (9) (по Лемме 1 работы [3]).

Поскольку для получения слова V из слова S (см. формулу (10)) достаточно $(2m+1)(n+2)$ операций замены блоков из словаря Ω на соответствующие блоки из словаря Θ , убеждаемся в том, что если исходная задача τ имеет положительный ответ, то такой же ответ имеет и задача $\kappa(\tau)$.

4) Предположим, что существует такая последовательность Φ операций редактирования из множества Φ , имеющая длину не более $(2m+1)(n+2)$, что слово V , являющееся результатом ее применения к слову $S = QW(D(\tau))$ ($Q = \sigma^{r_2 \cdot (n+2)}$), содержится в $\Sigma^*(F)$ (т. е. $V = (QW(D(\tau))) \cdot \Phi \in \Sigma^*(F)$). Покажем, что тогда существует и выполняющий набор значений переменных для задачи τ .

4.1. Из условия $a2$ и формулы (10) следует, что никакая часть слова $Q = \sigma^{r_2(n+2)}$ не может являться словом или префиксом слова $\omega \in \Omega$. Следовательно, префикс Q слова $S = QW(D(\tau))$ не может претерпеть изменений при применении к слову S произвольной последовательности Φ операций редактирования из набора Φ .

4.2. Из п. 4.1 и условия 1° следует, что слово V имеет вид $V = QH$, где $H = LMR$, L, R — суффикс и префикс некоторых слов из словаря Θ ; M — результат конкатенации некоторой последовательности слов из словаря Θ .

4.3. Из условий $a2$, $a4$ следует, что последовательность операций редактирования Φ должна содержать в точности $k = (2m + 1)(n + 2)$ операций, каждая из которых применяется к некоторому блоку $\alpha(b_{i,\alpha_j})$, являющемуся элементом матрицы $D(\tau)$.

4.4. Обозначим через θ_{i,α_j} блок из словаря Θ , получающийся в результате применения к блоку $\alpha(b_{i,\alpha_j})$ соответствующей операции редактирования из последовательности Φ . Тогда из условия $a6$ следует справедливость равенств

$$\beta(\theta_{i,\alpha_j}) = \beta(\alpha(b_{i,\alpha_j})), \quad \chi(\theta_{i,\alpha_j}) = \chi(\alpha(b_{i,\alpha_j})).$$

4.5. Из условия 2° следует, что $\varepsilon(\theta_{i,\alpha_j}) = (\varepsilon(\theta_{0,1,j}))^\alpha$.

4.6. Из условий 3° , 4° и Леммы 1 работы [3] следует, что набор

$$\varepsilon = (\varepsilon(\theta_{0,1,1}), \varepsilon(\theta_{0,1,2}), \dots, \varepsilon(\theta_{0,1,n}))$$

является выполняющим набором значений переменных для задачи τ .

Теорема 1 доказана.

В основе примененного при доказательстве теоремы 1 метода построения полиномиальных алгоритмов сведения NP-полной задачи ВЫПОЛНИМОСТЬ к соответствующим задачам редактирования K_T лежит использование схемы кодирования условий задачи $\tau \in \text{ВЫПОЛНИМОСТЬ}$, описанной в § 1 (см. также [3]), а также выбор, с учетом свойств применяемых операций редактирования, пары подходящих блоковых синхронизируемых кодов Ω , Θ ($|\Omega| = 4$, $|\Theta| = 16$), используемых, соответственно, для записи блоков, являющихся элементами матриц $D(\tau)$ и $D(\tau, \varepsilon)$. Словарь Ω используется при описании значения подлежащего редактированию слова S , входящего в условия задачи $K(\tau)$; словарь Θ в совокупности с набором чисел l, T, k ($l \geq T$) — при построении описания предиката $F = F(w_1, \dots, w_l)$ существенно зависящего не более, чем от T переменных и определяющего ограничение $\langle F \rangle$.

С точки зрения методологии доказательства NP-полноты, изложенной в работе [1], рассматриваемый метод может быть отнесен к категории методов построения компонент. Компонентой, «делающей выбор» набора значений переменных ε , служит строка 01 матрицы $D(\tau)$ в совокупности с пре-

фиксом Q слова $S = QW(D(\tau))$. Условие типа условия 2° из доказательства Теоремы 1 образует ограничитель, обеспечивающий присвоение переменным $(x_1, \dots, x_n)$ одинаковых для всех дизъюнкций значений из набора ϵ . Нулевой и $(n+1)$ -й столбцы блоков матрицы $D(\tau)$ в совокупности с условиями типа условий 3°, 4° из доказательства теоремы 1 выполняют функции компоненты проверки выполнения дизъюнкций из набора C при присваивании переменным $(x_1, \dots, x_n)$ значений из набора ϵ .

§ 3. Сложность задач о подстрочном переводе при $|\Sigma| \geq 2$

Использование в доказательстве теоремы 1 условий $a1$ (согласно которому $|\Sigma| \geq 3$), $a2$, $a4$ позволило, в частности, обеспечить обоснование справедливости п. 4.3, вызывавшего определенные трудности при произвольном выборе длин блоков в словарях Ω, Θ .

В данном параграфе при несколько более жестких условиях на выбор словарей Ω, Θ излагаются результаты о сложности задач о ПОДСТРОЧНОМ ПЕРЕВОДЕ, являющиеся справедливыми для алфавита Σ произвольной мощности ($|\Sigma| \geq 2$).

Будем предполагать, что выполняются следующие условия:

b1) При некотором r существует код без запятой Ξ ($\Xi \subset \Sigma^r$, $|\Xi| \geq 20$).

Выберем такие блоковые словари $\Omega_1, \Theta_1 \subset \Xi$, что $\Omega_1 \cap \Theta_1 = \emptyset$, $|\Omega_1| = 4$, $|\Theta_1| = 16$. Как и выше, для нумерации слов в кодах Ω_1, Θ_1 будем использовать двоичные слова $\beta\gamma \in B^2$, $\beta\gamma\epsilon\mu \in B^4$ соответственно: $\Omega_1 = \{\omega(\beta\gamma) : \beta\gamma \in B^2\}$, $\Theta_1 = \{\theta(\beta\gamma\epsilon\mu) : \beta\gamma\epsilon\mu \in B^4\}$.

b2) Набор Φ операций редактирования, используемых в рассматриваемых задачах о ПОДСТРОЧНОМ ПЕРЕВОДЕ K_T , имеет вид $\Phi = \{\phi_{\beta\gamma\epsilon\mu} : \beta\gamma\epsilon\mu \in B^4\}$, причем действие операции $\phi_{\beta\gamma\epsilon\mu}$ выражается в замене подслова редактируемого слова S , являющегося элементом кода Ω_1 , на слово кода Θ_1 в соответствии с правилом

$$\omega(\beta\gamma) \cdot \phi_{\beta\gamma\epsilon\mu} = \theta(\beta\gamma\epsilon\mu). \quad (14)$$

Теорема 2. Если выполнены условия b1, b2, то задача K_{3r} NP-полна.

Доказательство. В тексте доказательства будем использовать коды $\Omega = \Omega_1$, $\Theta = \Theta_1$.

1) $K_{3r} \in \text{NP}$. Доказательство аналогично п. 1 доказательства теоремы 1.

2) Рассмотрим индивидуальную задачу $\tau \in \text{ВЫПОЛНИМОСТЬ}$. Полагая $r_2 = r$, применим построение, аналогичное описанному в п. 2 доказательства теоремы 1 (за исключением ненужного в данной ситуации условия $\sigma \neq \#$ в формуле (10)). Это даст запись условия задачи $\kappa(\tau) \in K_{3r}$.

Покажем, что задача $\kappa(\tau)$ имеет положительное решение в том и только том случае, когда такое же решение имеет задача τ .

3) Если существует набор ϵ значений переменных $x_1, \dots, x_n$, являющийся выполняющим для задачи τ , то задача $\kappa(\tau)$ также имеет положительное решение.

Доказательство аналогично п. 3 доказательства теоремы 1.

4) Предположим, что существует такая последовательность Φ операций редактирования вида (14), имеющая длину $|\Phi| = v \leq k = (2m + 1)(n + 2)$, что справедливо включение $V = S \cdot \Phi \in \Sigma^*(F)$, где $S = QW(D(\tau))$ — слово вида (10). Покажем, что тогда существует и выполняющий набор значений переменных для задачи τ .

4.1. Из формул (10), (12) и условия 1° следует, что слово V имеет вид $V = Q'H$, где $|Q'| = |Q| = r(n + 2)$, $H = LMR$, L, R — суффикс и префикс некоторых слов из словаря Θ_1 , M — результат конкатенации некоторой последовательности слов из словаря Θ_1 . В силу способа выбора блоковых кодов Ω_1 , Θ_1 , слово H не может содержать подслов, являющихся словами кода Ω_1 .

4.2. Из п. 4.1 следует, что слово H получается путем применения к слову $W(D(\tau))$, являющемуся результатом конкатенации $k = (2m + 1)(n + 2)$ блоков из словаря Ω , в точности k операций редактирования вида (14), каждая из которых применяется к соответствующему блоку $\alpha(b_{i,\alpha_j})$, являющемуся элементом матрицы $D(\tau)$ ($i = 0, 1, \dots, m$, если $i = 0$, то $\alpha = 1$, при $i = 1, \dots, m$ индекс α принимает значения $0, 1$; $j = 0, 1, \dots, n + 1$).

4.3. В силу п. 4.2 слово V имеет вид $V = QM$, где слово M является конкатенацией $k = (2m + 1)(n + 2)$ блоков из словаря Θ_1 .

4.4. Обозначим через θ_{i,α_j} блок из словаря Θ_1 , получающийся в результате применения к блоку $\alpha(b_{i,\alpha_j})$ некоторой операции редактирования из последовательности Φ . Согласно формуле (14) справедливы равенства

$$\beta(\theta_{i,\alpha_j}) = \beta(\alpha(b_{i,\alpha_j})), \quad \chi(\theta_{i,\alpha_j}) = \chi(\alpha(b_{i,\alpha_j})).$$

4.5. Из условия 2° следует, что $\epsilon(\theta_{i,\alpha_j}) = (\epsilon(\theta_{0,1,j}))^\alpha$.

4.6. Из условий 3°, 4° и Леммы 1 [3] следует, что набор

$$\epsilon = (\epsilon(\theta_{0,1,1}), \epsilon(\theta_{0,1,2}), \dots, \epsilon(\theta_{0,1,n}))$$

является выполняющим набором значений переменных $x_1, \dots, x_n$ для задачи τ .

Теорема 2 доказана.

При конкретном выборе алфавита Σ представляет определенный интерес вопрос о том, можно ли уменьшить число T существенных переменных предиката F , задающего ограничение на множество подслов в словах базового языка, так, чтобы сохранилась возможность построения полиномиальных алгоритмов сведения задачи ВЫПОЛНИМОСТЬ к соответствующим задачам о ПОДСТРОЧНОМ ПЕРЕВОДЕ K_T , аналогичных алгоритму, примененному в доказательстве теорем 1 и 2.

В частности, при $|\Sigma| \geq 20$ справедливо следующее утверждение, которое может рассматриваться в качестве следствия из теоремы 2.

Следствие 1. Если $|\Sigma| \geq 20$, то задача $\mathbf{K}_3$ NP-полна.

Выбрав произвольное подмножество Σ' алфавита Σ ($|\Sigma'| \geq 20$), которое будем рассматривать в качестве кода Ξ , а также определив соответствующие операции редактирования, мы оказываемся в условиях теоремы 2 при значении параметра $r = 1$, что и доказывает справедливость Следствия 1.

Воспользовавшись верхней оценкой для числа слов заданной длины u из алфавита мощности v в синхронизируемом словаре (Теорема 11.14, [7]), можно показать, что блоковый двоичный код без запятой $\Xi \subset B^*$, содержащий 20 слов, нельзя построить если длина кодового слова меньше 8. Отсюда следует, что для двоичного алфавита ($\Sigma = B$) построение, изложенное в п. 2 доказательства теоремы 1, нельзя использовать для полиномиального сведения задачи ВЫПОЛНИМОСТЬ к задаче о ПОДСТРОЧНОМ ПЕРЕВОДЕ $\mathbf{K}_T$, если $T < 24$.

Покажем, что при $\Sigma = B$ и специальном выборе кода $\Xi \subset B^r$ за счет использования особенностей структуры слов кода Ξ можно построить полиномиальные алгоритмы сведения задачи ВЫПОЛНИМОСТЬ к задачам $\mathbf{K}_T$, в которых $T < 3r$. Для этого применим построение, аналогичное проведенному в работе [3] при доказательстве NP-полноты задач редактирования слов, в которых набор операций редактирования Φ содержит операции замены символа или перестановки пары символов в редактируемом слове (теоремы 5, 6 [3]).

Полагая $\Sigma = B$, выберем блоковые коды $\Xi = \Xi_2 \subset B^{10}$ ($|\Xi_2| = 20$), $\Omega = \Omega_2$, $\Theta = \Theta_2$ ($|\Omega_2| = 4$; $|\Theta_2| = 16$; $\Omega_2, \Theta_2 \subset \Xi_2$), имеющие вид

$$\begin{aligned} \Xi_2 = & \{1^2 0 \beta \gamma \theta^4 1 \mid \beta, \gamma \in B\} \cup \{1^2 0 \beta \gamma \theta^4 1 \mid \beta, \gamma \in B\} \cup \{1^2 0 \beta \gamma \theta^3 10 \mid \beta, \gamma \in B\} \cup \\ & \cup \{1^2 0 \beta \gamma \theta^2 10^2 \mid \beta, \gamma \in B\} \cup \{1^2 0 \beta \gamma \theta^3 10 \mid \beta, \gamma \in B\}, \end{aligned} \quad (15)$$

$$\Omega_2 = \{1^2 0 \beta \gamma \theta^4 1 \mid \beta, \gamma \in B\}, \quad (16)$$

$$\begin{aligned} \Theta_2 = & \{1^2 0 \beta \gamma \theta^4 1 \mid \beta, \gamma \in B\} \cup \{1^2 0 \beta \gamma \theta^3 10 \mid \beta, \gamma \in B\} \cup \\ & \cup \{1^2 0 \beta \gamma \theta^2 10^2 \mid \beta, \gamma \in B\} \cup \{1^2 0 \beta \gamma \theta^3 10 \mid \beta, \gamma \in B\}. \end{aligned} \quad (17)$$

Нетрудно проверить, что словарь Ξ_2 является кодом без запятой.

Для нумерации слов в словарях Ω_2 , Θ_2 будем использовать отображения $\alpha: B^2 \rightarrow \Omega_2$, $\theta: B^4 \rightarrow \Theta_2$ следующего вида:

$$\alpha(\beta\gamma) = 1^2 0 \beta \gamma \theta^4 1, \quad \theta(\beta\gamma\epsilon\mu) = 1^2 0 \beta \gamma \tilde{\theta}(\epsilon\mu), \quad (18)$$

где $\tilde{\theta}(\epsilon\mu)$ — результат применения к слову $\epsilon\mu \in B^2$ биективного отображения

$$\tilde{\theta}: B^2 \leftrightarrow \{10^4, 010^3, 0^210^2, 0^310\}.$$

Пусть $\tilde{\theta}(\epsilon\mu) = \eta_1\eta_2\eta_3\eta_4\eta_5$. Заметим, что значение слова $\epsilon\mu \in B^2$ однозначно определяется значениями любого из следующих двух подслов слова $\eta_1\eta_2\eta_3\eta_4\eta_5$:

$$\eta_1\eta_2\eta_3 = \tilde{\theta}'(\epsilon\mu), \quad \eta_2\eta_3\eta_4 = \tilde{\theta}''(\epsilon\mu).$$

Таким образом, при $\eta_1\eta_2\eta_3\eta_4\eta_5 \in \{10^4, 010^3, 0^210^2, 0^310\}$ можно рассматривать соответствующие обратные отображения

$$\varepsilon = \varepsilon(\tilde{\theta}) = \tilde{\varepsilon}'(\eta_1\eta_2\eta_3) = \tilde{\varepsilon}''(\eta_2\eta_3\eta_4),$$

$$\mu = \mu(\tilde{\theta}) = \tilde{\mu}'(\eta_1\eta_2\eta_3) = \tilde{\mu}''(\eta_2\eta_3\eta_4). \quad (19)$$

Будем, как и выше, рассматривать задачи о ПОДСТРОЧНОМ ПЕРЕВОДЕ $\mathbf{K}_T$, в которых действие операции редактирования $\varphi_{\beta\gamma\epsilon\mu}$ из набора $\Phi = \{\varphi_{\beta\gamma\epsilon\mu}: \beta\gamma\epsilon\mu \in B^4\}$ выражается в замене подслов редактируемого слова S , являющихся словами кода Ω_2 , на слова кода Θ_2 в соответствии с правилом (14):

$$\varphi_{\beta\gamma\epsilon\mu}(\omega(\beta\gamma)) = \theta(\beta\gamma\epsilon\mu).$$

Как следует из теоремы 2, соответствующая задача $\mathbf{K}_{30}$ является NP-полной. Используя особенности структуры слов кодов Ξ_2 , Ω_2 , Θ_2 , докажем справедливость следующего утверждения.

Теорема 3. Если коды $\Omega = \Omega_2$, $\Theta = \Theta_2$ соответственно имеют вид (16), (17), а действие операции редактирования $\varphi_{\beta\gamma\epsilon\mu}(\beta\gamma\epsilon\mu \in B^4)$ выражается в замене подслов редактируемого слова S , являющихся словами кода Ω_2 , на слова кода Θ_2 в соответствии с правилом (14), то задача $\mathbf{K}_{15}$ NP-полна.

Доказательство. Будем использовать коды $\Xi = \Xi_2$, $\Omega = \Omega_2$, $\Theta = \Theta_2$.

1) Покажем, что $\mathbf{K}_{15} \in \text{NP}$. Для произвольной индивидуальной задачи $\kappa \in \mathbf{K}_{15}$ после угадывания последовательности операций редактирования проверка ее соответствия условиям задачи κ может быть выполнена за время, ограниченное полиномом от длины записи задачи κ . Доказательство данного факта аналогично приведенному в п. 1 доказательства теоремы 1.

2) Построим алгоритм сведения задачи ВЫПОЛНИМОСТЬ к задаче $\mathbf{K}_{15}$.

Для записи условия задачи редактирования $\kappa = \kappa(\tau) \in \mathbf{K}_{15}$ будем использовать матрицу $D(\tau)$, получаемую посредством замены блоков b_{i,α_j} , яв-

ляющихся элементами матрицы $C(\tau)$, на соответствующие блоки $d_{i,\alpha j} = \alpha(b_{i,\alpha j}) \in \Omega_2$ ($i = 0, \dots, m$; $\alpha = 0, 1$; $j = 0, 1, \dots, n+1$).

При заданном наборе ϵ значений переменных $x_1, \dots, x_n$ в качестве аналога матрицы $C(\tau, \epsilon)$ будем рассматривать матрицу $D(\tau, \epsilon)$, получаемую посредством замены блоков $b_{i,\alpha j}(\epsilon)$, являющихся элементами матрицы $C(\tau, \epsilon)$, на слова $\theta_{i,\alpha j}(\epsilon) = \theta(b_{i,\alpha j}(\epsilon))$.

В условиях задачи $\kappa(\tau)$ полагаем

$$K = (2m+1)(n+2), \quad l = 10n+27, \quad S = QW(D(\tau)), \quad \text{где } Q = 1^{10n+15}. \quad (20)$$

Предикат $F = F(w_1, \dots, w_{10n+27})$ существенно зависит от $T = 15$ переменных

$$(w_1, w_2, w_3, w_{10n+16}, \dots, w_{10n+27}) = \\ = (w_1, w_2, w_3, y_1, \dots, y_{12}). \quad (21)$$

Описание наборов значений переменных (21), содержащихся в области истинности предиката F , представим в виде следующего набора условий.

c1°. Значения переменных $y_1, \dots, y_{12}$ таковы, что $(y_1, \dots, y_{12}) = (\eta_q, \dots, \eta_{q+11})$ при некотором $q \in \{1, \dots, 10\}$, а $\eta_1 \dots \eta_{10}$, $\eta_{11} \dots \eta_{20}$, $\eta_{21} \dots \eta_{30}$ — некоторые слова из Θ_2 .

Замечание 3. Проведя непосредственную проверку, можно убедиться в том, что если выполняется условие c1°, то набор значений переменных $y_1, \dots, y_{12}$ однозначно определяет значение q .

Поэтому в описание представленных ниже условий c2°–c4°, как и в доказательствах теорем 1, 2, включены ссылки на соответствующие значения параметра q .

c2°. Если $q = 1$ и $y_5 = 0$, то определено значение $\tilde{\epsilon}'_1(w_1, w_2, w_3)$ и выполняется равенство

$$\tilde{\epsilon}'_1(w_1, w_2, w_3) = \tilde{\epsilon}'_1(y_6, y_7, y_8).$$

c3°. Если $q = 7$, то

$$\tilde{\mu}'_1(y_{10}, y_{11}, y_{12}) = \tilde{\mu}''_1(y_1, y_2, y_3)(1 \oplus y_8 \tilde{\epsilon}'_2(y_{10}, y_{11}, y_{12})) \vee y_8 y_9.$$

c4°. Если $q = 7$, то

$$\tilde{\mu}''_1(y_1, y_2, y_3) y_9 = 0.$$

Построение условия задачи $\kappa(\tau)$ завершено.

3) Предположим, что существует набор ϵ значений переменных, являющийся выполняющим для задачи τ . Покажем, что тогда

$$V = QW(D(\tau, \epsilon)) = 1^{10n+15}W(D(\tau, \epsilon)) \in B^*(F). \quad (22)$$

Из описания применяемой схемы кодирования задачи ВЫПОЛНИМОСТЬ и леммы 1 [3] следует выполнение условий $c1-c4^\circ$, входящих в описание предиката F и соответствующего ограничения $\langle F \rangle$, что и доказывает справедливость включения в формуле (22).

Поскольку для получения слова V из S достаточно $(2m+1)(n+2)$ операций замены подслов редактируемого слова, являющихся словами кода Ω_2 , на слова кода Θ_2 , убеждаемся в том, что если исходная задача τ имеет положительное решение, то такое же решение имеет и задача $\kappa(\tau)$.

4) Пусть существует такая последовательность Φ операций редактирования, имеющая длину $|\Phi| \leq k = (2m+1)(n+2)$, что слово V' , являющееся результатом ее применения к слову $S = QW(D(\tau))$ вида (20), содержится в $B^*(F)$. Покажем, что тогда существует и выполняющий набор значений переменных для задачи τ .

4.1. Из формул (20), (21) и условия $c1^\circ$ следует, что слово V' имеет вид $V' = Q'H$, где $|Q'| = |Q| = 10n+15$, $H = LMR$, L, R — суффикс и префикс некоторых слов из словаря Θ_2 , M — результат конкатенации некоторой последовательности слов из словаря Θ_2 . В силу способа выбора блоковых кодов Ω_2 , Θ_2 слово H не может содержать подслов, являющихся словами кода Ω_2 .

4.2. Из п. 4.1 следует, что слово H получается путем применения к слову $W(D(\tau))$, являющемуся результатом конкатенации $k = (2m+1)(n+2)$ блоков из словаря Ω_2 , в точности k операций редактирования, каждая из которых применяется к соответствующему блоку $\alpha(b_{i,\alpha_j})$, являющемуся элементом матрицы $D(\tau)$.

4.3. В силу п. 4.2 слово V' имеет вид $V' = QM$, где слово M является конкатенацией $k = (2m+1)(n+2)$ блоков из словаря Θ_2 .

4.4. Обозначим через θ_{i,α_j} блок из словаря Θ_2 , получающийся в результате применения к блоку $\alpha(b_{i,\alpha_j})$ некоторой операции редактирования из последовательности Φ . Согласно формуле (14) справедливы равенства

$$\beta(\theta_{i,\alpha_j}) = \beta(\alpha(b_{i,\alpha_j})), \quad \chi(\theta_{i,\alpha_j}) = \chi(\alpha(b_{i,\alpha_j})).$$

4.5. Из условия $c2^\circ$ следует, что $\varepsilon(\theta_{i,\alpha_j}) = (\varepsilon(\theta_{0,1,j}))^\alpha$.

4.6. Из условий $c3^\circ$, $c4^\circ$ и Леммы 1 [3] следует, что набор

$$\epsilon = (\varepsilon(\theta_{0,1,1}), \varepsilon(\theta_{0,1,2}), \dots, \varepsilon(\theta_{0,1,n}))$$

является выполняющим набором значений переменных $x_1, \dots, x_n$ для задачи τ .

Теорема 3 доказана.

Заключение

Активные исследования алгоритмов редактирования слов стимулируются потребностями развития систем передачи, обработки и хранения данных, систем автоматизированной обработки текстов, а также приложений в некоторых отраслях биохимических и генетических исследований [9–11].

Слова языков $L = \Sigma^* \langle F \rangle$, заданных ограничениями $\langle F \rangle$ на множества подслов, которые могут встречаться в их словах [3], могут, в частности, служить естественной моделью рекуррентных последовательностей, образующихся при функционировании различных вычислительных устройств в процессе обработки данных.

Специфика процессов преобразования слов под действием операций редактирования, рассматриваемых при исследовании задач о ПОДСТРОЧНОМ ПЕРЕВОДЕ, имеет значительные аналогии также с процессами преобразования генетической информации в ходе функционирования живых организмов (транскрипция генетического кода при синтезе матричной РНК, трансляция генетического кода при синтезе белков [11]).

Полученные в данной статье результаты, устанавливающие NP-полноту задач о ПОДСТРОЧНОМ ПЕРЕВОДЕ, расширяют методическую базу для проведения исследований сложности задач редактирования слов, возникающих в конкретных приложениях, и обоснования трудоемкости их решения.

Автор благодарит профессора М. М. Глухова за полезные обсуждения полученных результатов.

Список литературы

1. Гэри, М., Джонсон, Д. Вычислительные машины и труднорешаемые задачи. — М.: Мир, 1982.
2. Wagner, R. A. Correcting regular and counter-recognizable languages // Proc. 8th Hawaii Intern. Conf. Syst. Sci. Honolulu, Hawaii — 1975. — P. 16–18.
3. Мартынов, С. С. О сложности некоторых задач редактирования слов // Дискретная математика. — 1989. — Т. 1. Вып. 4. — С. 104–112.
4. Мартынов, С. С. Сложность задачи о длине наибольшей подпоследовательности двоичного слова, удовлетворяющей заданному ограничению // Обозр. прикл. и пром. матем. — 2011. — Т. 18. Вып. 2. — С. 302.
5. Мартынов, С. С. О методе построения сводимостей задачи ВЫПОЛНИМОСТЬ к задачам редактирования слов // Обозр. прикл. и пром. матем. — 2004. — Т. 11. Вып. 2. — С. 372–373.
6. Мартынов, С. С. Сложность задач о подстрочном переводе // Обозр. прикл. и пром. матем. — 2012. — Т. 19. Вып. 3. — С. 455–456.

7. *Стиффлер, Дж. Дж.* Теория синхронной связи / Пер. с англ. Б. С. Цыбакова; под ред. Э. М. Габидуллина. — М.: Связь, 1975.
8. *Глухов, М. М., Шишков, А. Б.* Математическая логика. Дискретные функции. Теория алгоритмов. Учебное пособие. — СПб.: Лань, 2012, — 416 с.
9. Time warps, string edits and macromolecules: the theory and practice of sequence comparison. Ed. D. Sankoff, J.B. Kruskal. — London, 1983.
10. *Гасфилд, Дэн.* Строки, деревья и последовательности в алгоритмах. Информатика и вычислительная биология / Пер. с англ. И. В. Романовского. — СПб.: Невский диалект, БХВ-Петербург, 2003. — 654 с.
11. *Попов, В. В.* Геномика с молекулярно-генетическими основами. — М.: Книжный дом «Либроком», 2009. — 304 с.