Напомним, что конечным недетерминированным преобразователем $\mathfrak{A}$ называется пятерка $\langle A, Q, Q_0, F, \delta\rangle$, где A — конечный алфавит, Q — конечное множество состояний, $Q_0 \subseteq Q$ — множество начальных состояний, F — множество заключительных состояний, δ — множество переходов. Каждый переход это четверка $\langle q_1, a, v, q_2\rangle$. Здесь q_1 — состояние до перехода, q_2 — состояние после перехода, $a \in A \bigcup \{\Lambda\}$ (Λ — пустое слово) — вход перехода, $v \in A^*$ — выход перехода. Будем представлять $\mathfrak{A}$ ориентированным графом, вершины которого являются состояниями, а ребра — переходами. Каждое ребро помечено входом и выходом соответствующего перехода. Путями в этом графе будем называть ориентированные пути. Слово, полученное конкатенацией входов вдоль некоторого пути l, назовем входом пути l (обозначение: $вх(l)$), а конкатенацией выходов — выходом пути l (обозначение: $вых(l)$). Путь, начинающийся в начальном состоянии и заканчивающийся в заключительном будем называть допускающим. Графиком $\Gamma(\mathfrak{A})$ преобразователя $\mathfrak{A}$ назовем множество пар $\langle u, v\rangle$ таких, что $u = вх(l)$, $v = вых(l)$ для некоторого допускающего пути l. Будем говорить, что преобразователь $\mathfrak{A}_1$ вложен в преобразователь $\mathfrak{A}_2$, если $\Gamma(\mathfrak{A}_1) \subseteq \Gamma(\mathfrak{A}_2)$. Преобразователи $\mathfrak{A}_1$ и $\mathfrak{A}_2$ назовем эквивалентными, если $\Gamma(\mathfrak{A}_1) = \Gamma(\mathfrak{A}_2)$. Размером $\mathfrak{A}$ (обозначение: $|\mathfrak{A}|$) будем называть сумму числа его состояний, переходов и длин их выходов. Очевидно, что по преобразователю $\mathfrak{A}$ можно за полиномиальное от $|\mathfrak{A}|$ время построить эквивалентный ему преобразователь, в котором через любое состояние проходит хотя бы один допускающий путь. Поэтому, будем считать, что все рассматриваемые далее преобразователи обладают этим свойством. Основным для дальнейшего изложения будет следующее определение.

ОПРЕДЕЛЕНИЕ. Преобразователь $\mathfrak{A}$ называется *конечнозначным*, если существует константа c такая, что для любого слова u существует не более c различных слов v таких, что $\langle u, v\rangle \in \Gamma(\mathfrak{A})$. Минимальное такое c называется значностью $\mathfrak{A}$. Если значность равна 1, то $\mathfrak{A}$ называется однозначным.

Назовем преобразователь $\mathfrak{A}$ редуцированным, если у него ровно одно начальное и ровно одно конечное состояние, и все переходы с пустым входом идут из начального состояния в конечное. А.Вебер доказал в [1] следующее утверждение.

ЛЕММА 1. *По преобразователю* $\mathfrak{A}$ *можно за полиномиальное время либо построить эквивалентный ему редуцированный преобразователь, либо констатировать бесконечнозначность* $\mathfrak{A}$.

ДОКАЗАТЕЛЬСТВО. Добавив не более двух состояний и не более $2|\mathfrak{A}|$ переходов, очевидным образом добъемся чтобы в $\mathfrak{A}$ было ровно одно начальное состояние, из которого переходы только выходят, и ровно

одно заключительное, куда переходы только входят. Далее, зададим на состояниях $\mathfrak{A}$ частичный порядок: $q_1 > q_2$ если существует путь из q_1 в q_2 с пустым входом. Предположим, что q_1 и q_2 эквивалентны относительно этого порядка. Легко видеть, что если $\mathfrak{A}$ конечнозначный, то любой путь из q_1 в q_2 с пустым входом должен иметь пустой выход. Очевидно, что проверка последнего условия для всех пар $\langle q_1, q_2\rangle$ выполняется за полиномиальное время. Пусть это условие выполнено. Каждый класс эквивалентных состояний естественным образом склеим в одно состояние. Получим преобразователь $\mathfrak{A}'$, который, очевидно, эквивалентен $\mathfrak{A}$. Пусть в $\mathfrak{A}'$ есть не начальное и не заключительное состояние q такое, что существует переход с пустым входом, который входит в q или выходит из q. Для каждой пары переходов $\langle q_1, a_1, v_1, q\rangle$,$\langle q, a_2, v_2, q_2\rangle$ такой, что $a_1 = \Lambda$ или $a_2 = \Lambda$, добавим переход $\langle q_1, a_1a_2, v_1v_2, q_2\rangle$. После этого, удалим все переходы с пустым входом, входящие в q и выходящие из q. Так как в $\mathfrak{A}'$ нет переходов вида $\langle q, \Lambda, v, q\rangle$, то описанная процедура не изменяет график преобразователя. Кроме того, если до нее у некоторого состояния q' не было инцидентных ему переходов с пустым входом, то и после нее их не будет. Поэтому, не более чем за $|\mathfrak{A}'|$ таких операций мы получим требуемый редуцированный преобразователь. Лемма 1 доказана.

Скажем, что конечнозначный преобразователь $\mathfrak{A}$ *размагается в декомпозицию* однозначных преобразователей $\mathfrak{A}_1, \mathfrak{A}_2, \ldots, \mathfrak{A}_k$, если $\Gamma(\mathfrak{A}) = \bigcup\limits_{i=1}^{k} \Gamma(\mathfrak{A}_i)$. А.Вебер доказал в [2], что любой конечнозначный преобразователь $\mathfrak{A}$ размера n можно эффективно разложить в декомпозицию не более $\exp(\mathrm{poly}(n))$ однозначных преобразователей за время $\exp(\exp(\mathrm{poly}(n)))$. При этом размер каждого однозначного преобразователя не более $\exp(\exp(\mathrm{poly}(n)))$. Мы предлагаем в следующей теореме более простой путь декомпозиции с лучшей оценкой на время и размер компонент.

ТЕОРЕМА 1. *Существует алгоритм, который по любому конечнозначному преобразователю $\mathfrak{A}$ размера n находит его декомпозицию из не более $k \le \exp(\mathrm{poly}(n))$ однозначных преобразователей $\mathfrak{A}_1, \ldots, \mathfrak{A}_k$ за время не более $\exp(\mathrm{poly}(n))$. Размер каждого $\mathfrak{A}_i$ не более $\exp(\mathrm{poly}(n))$.*

ДОКАЗАТЕЛЬСТВО. Из утверждения леммы 1 следует, что достаточно рассмотреть случай, когда в $\mathfrak{A}$ нет переходов с пустым входом. Если l — путь в $\mathfrak{A}$ со входом u и $u' \subseteq u$ (т.е. u' — начало u), то через $l(u')$ будем обозначать начало пути l такое, что $вх(l(u')) = u'$. Если $u = v_1v_2v_3$, то через $l(v_2)$ будем обозначать путь такой, что конкатенация $l(v_1)l(v_2)$ есть $l(v_1v_2)$. Если l_1 и l_2 — два пути такие, что $вх(l_1) = вх(l_2) = u$ и $u' \subseteq u$, то будем обозначать через $d(l_1, l_2, u')$ число,

равное $|\,вых(l_1(u'))| - |\,вых(l_2(u'))|$. Сформулируем критерий конечнозначности для преобразователя без переходов с пустым входом.

ТЕОРЕМА 2. *Преобразователь* $\mathfrak{A}$ *размера* n *без пустых входов конечнозначен тогда и только тогда, когда для любых двух его состояний* s_1, s_2 *(не обязательно различных) и любых трех путей* p_1, p_2, p_3 *с одним и тем же входом* u *и таких, что* p_1 *начинается в* s_1 *и кончается в* s_1, p_2 *начинается в* s_1, *кончается в* s_2, p_3 *начинается в* s_2 *и кончается в* s_2, *выполняются следующие два условия:*
(1) *Если* $вых(p_1) \neq \Lambda$, *то* $вых(p_2)$ *является началом бесконечного слова* $(вых(p_1))^\infty$.
(2) *Для любого* $u' \subseteq u$ *выполняется:* $|d(p_1, p_2, u')| \leq n^3$.

Сходный критерий сформулировал и доказал в [1] А.Вебер. Докажем необходимость сформулированного критерия. Предположим, что условие (1) не выполняется. Пусть $k > c$, где c — значность $\mathfrak{A}$. Рассмотрим множество $\{p_1^{i-1}p_2p_3^{k-i} \mid i = 1, \ldots, k\}$ путей из s_1 в s_2. Легко видеть, что все эти пути имеют один и тот же вход u^k и попарно различные выходы. Это противоречит тому, что $k > c$. Предположим, что условие (2) не выполняется. Зафиксируем $m = n^3 + 1$ начал слова u: $u_1, u_2, \ldots, u_m$ так, чтобы все $d(p_1, p_2, u_i)$ были попарно различны. Каждому u_i соответствует тройка состояний, в которых кончаются пути $p_1(u_i)$, $p_2(u_i)$, $p_3(u_i)$. Возьмем такие $u_{i'}$ и $u_{i''}$ $(u_{i'} \subset u_{i''})$, которым соответствует одна и та же тройка. Таким образом, $u = v_1v_2v_3$, где $v_1 = u_{i'}$, $v_1v_2 = u_{i''}$. По построению $|\,вых(p_1(v_2))| \neq |\,вых(p_2(v_2))|$. Если $|\,вых(p_1)| \neq |\,вых(p_3)|$, то длины выходов путей $p_1^{i-1}p_2p_3^{k-i}$, очевидно, попарно различны, что противоречит выбору k. Если $|\,вых(p_1(v_1))| + |\,вых(p_1(v_3))| \neq |\,вых(p_3(v_1))| + |\,вых(p_3(v_3))|$, то такое же рассуждение верно для путей $[p_1(v_1)p_1(v_3)]^{i-1}p_2(v_1)p_2(v_3)[p_3(v_1)p_3(v_3)]^{k-i}$. Иначе, рассмотрим различные пути из q_1 в q_2 со входом $v_1v_2v_3v_1v_2v_2v_3 \ldots v_1v_2^kv_3$, проходящие по p_1, p_2, p_3. Легко видеть, что длина выхода этих путей монотонно зависит от того, какой блок $v_1v_2^iv_3$ приходится на путь p_2, что противоречит выбору k. Необходимость критерия доказана.

При доказательстве теоремы 1 будем предполагать вместо конечнозначности $\mathfrak{A}$ выполненность данного критерия. Поскольку из возможности декомпозиции преобразователя следует его конечнозначность, то будет доказана достаточность критерия.

Продолжим доказательство теоремы 1. Пусть q_1 и q_2 — состояния $\mathfrak{A}$. Скажем, что $q_1 \geq q_2$, если существует путь из q_1 в q_2. Это отношение частичного порядка разбивает множество состояний Q на классы эквивалентных состояний. Переход $\langle q_1, a, v, q_2\rangle$ будем называть *переходом в состояниях*, если q_1 не эквивалентно q_2. Пусть Q_1 и Q_2 — подмножества

Q. Скажем, что Q_2 достигается из Q_1 на слове w, если существует множество L путей из Q_1 в Q_2 со входом w такое, что для любого $q_1 \in Q_1$ существует путь из L, начинающийся в q_1, а для любого $q_2 \in Q_2$ существует путь из L, кончающийся в q_2. Скажем, что $Q_1 \geq Q_2$, если существует слово, на котором Q_2 достигается из Q_1. Очевидно, что введенное отношение является отношением частичного порядка на подмножествах Q, разбивающее их на классы эквивалентных подмножеств. Пусть есть слово u и $u' \subseteq u$. Множеством двусторонней достижимости $M_u(u')$ назовем множество таких состояний q, для которых существует допускающий путь l такой, что $вх(l) = u$ и путь $l(u')$ кончается в q. Легко видеть, что если $u_1 \subseteq u_2 \subseteq u$, то $M_u(u_1) \geq M_u(u_2)$. Пусть на допускающем пути l есть переход p и l_1 — начало l, кончающееся непосредственно до p, а l_2 — начало l, кончающееся сразу после p. Будем называть p *переходом во множествах* на l, если $M_u(u_1)$ не эквивалентно $M_u(u_2)$, где $u_1 = вх(l_1)$, $u_2 = вх(l_2)$. Докажем две леммы.

ЛЕММА 2. *Пусть в* $\mathfrak{A}$ *два состояния* q_1 *и* q_2 *эквивалентны. Тогда для любых двух путей* l_1 *и* l_2 *из* q_1 *в* q_2 *таких, что* $вх(l_1) = вх(l_2)$ *выполняется:* $вых(l_1) = вых(l_2)$. *Для любого начала* u' *слова* $u = вх(l_1)$ *выполняется:* $|d(l_1, l_2, u')| \leq n^3$.

ДОКАЗАТЕЛЬСТВО. Обозначим: $w_1 = вых(l_1)$, $w_2 = вых(l_2)$. В силу эквивалентности q_1 и q_2 существует путь l из q_2 в q_1. Если $|w_1| = |w_2|$, применим условие (1), где $s_1 = s_2 = q_1$, $p_1 = l_1 l$, $p_2 = p_3 = l_2 l$. Получим, что $вых(l_1 l) = вых(l_2 l)$, откуда следует $w_1 = w_2$. Пусть $|w_1| \neq |w_2|$. Будем считать, что $|w_1| > 0$. Получаем противоречие с условием (2), положив в нем $s_1 = s_2 = q_1$, $p_1 = (l_1 l)^k$, $p_2 = p_3 = (l_2 l)^k$ для $k > c$. Последнее утверждение леммы следует из условия (2). Лемма 2 доказана.

Скажем, что два слова согласованы, если одно из них является началом другого.

ЛЕММА 3. *Пусть* l *— допускающий путь в* $\mathfrak{A}$, u_1, u_2 *— два начала слова* $u = вх(l)$, $u_1 \subseteq u_2$, $M_u(u_1)$ *эквивалентно* $M_u(u_2)$. *Обозначим через* l' *отрезок пути* l, *дополняющий* $l(u_1)$ *до* $l(u_2)$, *а через* q_1 *и* q_2 *— соответственно начальное и конечное состояния* l'. *Пусть* $|вых(l')| > n^3$. *Тогда для любых двух путей* l_1, l_2 *из* q_1 *в* q_2 *со входом* $v = вх(l')$ *слова* $вых(l_1)$ *и* $вых(l_2)$ *согласованы и для любого* $v' \subseteq v$ *выполняется:* $|d(l_1, l_2, v')| \leq 2n^3$.

ДОКАЗАТЕЛЬСТВО. Обозначим: $M_1 = M_u(u_1)$, $M_2 = M_u(u_2)$. Будем считать, что q_1 не эквивалентно q_2, иначе утверждение леммы сразу следует из леммы 2. Из эквивалентности M_1 и M_2 следует, что M_1 достигается из M_2 на некотором слове v_1. Тогда M_1 достигается из самого себя на слове vv_1. Поэтому существует бесконечная последовательность состояний из M_1: $q_0, q_{-1}, q_{-2} \ldots$ такая, что для любого $i \leq 0$ существует

путь γ_i из q_i в q_{i+1} со входом vv_1. Зафиксируем состояние $b_1 = q_i = q_j$ для некоторых $i < j \leq 1$. Получим замкнутый путь из b_1 в b_1, который обозначим через p. Аналогично, существует бесконечная последовательность состояний из M_1: $q_3, q_4, \ldots$ такая, что существует путь γ_2 из q_2 в q_3 со входом v_1, и для любого $i \geq 3$ есть путь γ_i из q_i в q_{i+1} со входом vv_1. Здесь $q_i \neq q_j$ для всех $i \leq 1$, $j \geq 2$ в силу неэквивалентности q_1 и q_2. Зафиксировав $b_2 = q_i = q_j$ для некоторых $3 \leq i < j$, получим замкнутый путь из b_2 в b_2, который обозначим через p'. Путь из b_1 в b_2, состоящий из трех отрезков: из b_1 в q_1 по путям γ_i $(i \leq 0)$, из q_1 в q_2 по l' и из q_2 в b_2 по путям γ_i $(i \geq 2)$, обозначим через γ. Очевидно: $|\,\textit{вх}(p)| = kn_1$, $|\,\textit{вх}(\gamma)| = kn_2$, $|\,\textit{вх}(p')| = kn_3$, где $k = |vv_1|$, а n_1, n_2, n_3 — натуральные числа. Возьмем на p состояние s такое, что путь, начинающийся в s, идущий все время по циклу p и имеющий длину $k(n_1n_2n_3 - n_2)$, оканчивается в b_1 (обозначим этот путь через γ'). Рассмотрим три пути: путь p_1 из s в s, совершающий n_2n_3 оборотов по циклу p, путь p_2 из s в b_2, равный $\gamma'\gamma$ и путь p_3 из b_2 в b_2, совершающий n_1n_2 оборотов по циклу p'. Очевидно, что $\textit{вх}(p_1) = \textit{вх}(p_2) = \textit{вх}(p_3)$. Так как $|\,\textit{вых}(p_2)| \geq |\,\textit{вых}(l')| > n^3$, то по условию (2) $|\,\textit{вых}(p_1)| > 0$. Путь p_2 можно провести по l_1 или l_2 вместо l'. По условию (1) выходы любого из этих трех возможных вариантов пути p_2 являются началами слова $(\textit{вых}(p_1))^\infty$. Слова $\textit{вых}(l_1)$ и $\textit{вых}(l_2)$ продолжают одно и то же начало этого слова, следовательно, они согласованы друг с другом. Из условия (2) легко видеть, что $|d(l_1, l_2, v')| \leq 2n^3$. Лемма 3 доказана.

Пусть l — допускающий путь в $\mathfrak{A}$. Каждому началу l' пути l соответствует пара (q, M), где q — состояние, в котором оканчивается l', а $M = M_u(u')$, где $u = \textit{вх}(l)$, $u' = \textit{вх}(l')$. Будем отмечать пары, соответствующие некоторым началам пути l. Отметим пары, соответствующие пустому началу и всему пути. Для каждого перехода в состояниях на l, который не является переходом во множествах на l, возьмем ближайшие слева и справа на l (мы представляем l направленным слева направо) переходы во множествах (если они есть). Для каждого из этих двух переходов отметим по две пары, соответствующие началам пути l, кончающимся непосредственно до перехода и сразу после него. Для каждого перехода в состояниях, который является и переходом во множествах на l, отметим две указанные пары, соответствующие этому переходу. Обозначим через D последовательность всех отмеченных пар, расположенных в порядке возрастания соответствующих им начал. Легко видеть, что для любых двух соседних в D пар (q_1, M_1), (q_2, M_2) возможны следующие три случая:

1) q_1 эквивалентно q_2.

2) Начала l, соответствующие этим парам, различаются на один пе-

реход пути l, который является переходом и в состояниях и во множествах.

3) M_1 эквивалентно M_2, q_1 не эквивалентно q_2.

Если имеет место i-тый случай, будем говорить, что отрезок $[(q_1, M_1), (q_2, M_2)]$ имеет i-тый тип. В случае 2 припишем этому отрезку слово, являющееся выходом соответствующего перехода. Случай 3 разбивается на два подслучая. Пусть l' — участок пути l между положениями (q_1, M_1) и (q_2, M_2), $v = вх(l')$, $w = вых(l')$. Подслучай 1: $|w| \leq n^3$. В этом подслучае приписываем данному отрезку слово w. Подслучай 2: $|w| > n^3$. В этом подслучае припишем данному отрезку число k, равное разности между $|w|$ и минимальной длиной выхода среди выходов всех путей из q_1 в q_2 со входом v. Из утверждения леммы 3 следует, что $k \leq 2n^3$.

Последовательность D с приписанной отрезкам указанной информацией будем называть диаграммой пути l и обозначать через $D(l)$. Информация включает также номер типа каждого отрезка, а в случае третьего типа — номер подтипа.

ЛЕММА 4. *Пусть l_1 и l_2 — допускающие пути в $\mathfrak{A}$ такие, что $вх(l_1) = вх(l_2)$ и $D(l_1)$ совпадает с $D(l_2)$. Тогда $вых(l_1) = вых(l_2)$.*

ДОКАЗАТЕЛЬСТВО. Очевидно, что любая не крайняя пара из диаграммы соответствует положению пути непосредственно до или после перехода во множествах, причем из диаграммы видно, какой из этих случаев имеет место. Любое множество двусторонней достижимости M, о котором известно, что оно соответствует положению пути непосредственно слева (справа) от перехода во множествах однозначно определяет начало u' слова $u = вх(l_1)$ такое, что $M = M_u(u')$. Поэтому, пары из диаграммы однозначно разбивают вход u на соответствующие участки и достаточно доказать равенство выходов l_1 и l_2 на каждом участке входа. Если отрезок диаграммы имеет первый тип, то совпадение выходов на соответствующем участке входа следует из леммы 2. При отрезке второго типа и первого подтипа третьего типа выход указан непосредственно в диаграмме. На отрезке второго подтипа третьего типа выходы l_1 и l_2 согласованы по лемме 3, поэтому одинаковая разность с одним и тем же числом обеспечивает их совпадение. Лемма 4 доказана.

Так как все диаграммы имеют полиномиальную длину, то всего возможных диаграмм не более $\exp(\mathrm{poly}(n))$. Поэтому, для доказательства теоремы 1 достаточно для каждой возможной диаграммы D построить за экспоненциальное время преобразователь $\mathfrak{A}(D)$, график которого состоит из тех и только тех элементов $\Gamma(\mathfrak{A})$, которые реализуются путями с диаграммой D. По лемме 4 $\mathfrak{A}(D)$ будет однозначным.

Множеством левосторонней достижимости для слова u назовем множество таких состояний q, что существует путь из начального состояния в q со входом u. Если дана диаграмма D, то множеством локальной двусторонней достижимости на отрезке $[(q_1, M_1), (q_2, M_2)]$ диаграммы D для входа u и его начала u' назовем множество таких состояний q, что существует путь l из q_1 в q_2 такой, что $вх(l) = u$ и путь $l(u')$ кончается в q. Множеством локальной левосторонней достижимости на этом отрезке для входа u назовем множество таких состояний q, что существует путь из q_1 в q со входом u. Будем считать, что в первой паре (q, M) диаграммы D множество M является подмножеством множества начальных состояний, а в последней паре — подмножеством множества заключительных состояний, поскольку иначе, очевидно, не существует пути с диаграммой D.

Опишем $\mathfrak{A}(D)$. Состояниями $\mathfrak{A}(D)$ будут следующие совокупности: $\langle Q_1, Q_2, q, L\rangle$. Здесь в Q_1 будет вычисляться множество левосторонней достижимости для прочтенного начала входа, в Q_2 — множество двусторонней достижимости для всего входа, в $q \in Q_2$ — текущее состояние угадываемого пути с читаемым входом и диаграммой D. Очевидно, что множество Q_2 однозначно определяет текущий отрезок $[(q_1, M_1), (q_2, M_2)]$ диаграммы D первого или третьего типа такой, что $M_1 \geq Q_2 \geq M_2$, если он существует. L непусто тогда и только тогда, когда этот отрезок существует и имеет третий тип. Для первого подтипа L — число от 0 до n^3, в котором будет вычисляться длина выхода на текущем отрезке. Для второго подтипа L — совокупность $\langle Q_1', Q_2', i, P\rangle$. Здесь в Q_1' будет вычисляться множество локальной левосторонней достижимости на текущем отрезке, в Q_2' — множество локальной двусторонней достижимости, $q \in Q_2' \subseteq Q_1'$, i — число от 0 до $n^3 + 1$, в котором будет проверяться, что длина выхода на отрезке больше n^3. P это множество пар $\langle q', m\rangle$ по одной для каждого состояния $q' \in Q_2'$, где $0 \leq m \leq 2n^3$, и еще одна текущая пара с q' равным текущему состоянию q. Обозначим через u_t отрезок входа, прочтенный к текущему моменту в $\mathfrak{A}(D)$ на участке, где $M_1 \geq Q_2 \geq M_2$. В текущей паре $\langle q, m\rangle$ в m будет вычисляться разность между длиной выхода угадываемого в $\mathfrak{A}$ пути на отрезке входа u_t и минимальной длиной выхода среди выходов всех путей со входом u_t из q_1 во множество Q_2'. Чтобы следить за этой минимальной длиной, в каждой не текущей паре $\langle q', m\rangle$ в m будет вычисляться разность между минимальной длиной выхода среди выходов всех путей из q_1 в q' со входом u_t и минимальной длиной выхода среди выходов всех путей со входом u_t из q_1 во множество Q_2'. Начальными состояниями $\mathfrak{A}(D)$ являются такие, в которых Q_1 — множество начальных состояний, Q_2 и q — те, что указаны в начале D;

если первый отрезок диаграммы имеет третий тип, то L непусто и для первого подтипа L=0, а для второго $Q'_1 = Q'_2 = \{q\}$, i=0 и во всех парах $m = 0$. Текущая пара: $\langle q, 0\rangle$. Заключительными состояниями $\mathfrak{A}(D)$ являются такие, где Q_2 и q — те, которые указаны в конце D, и множество $Q_1 \setminus Q_2$ не содержит заключительных состояний: если последний отрезок диаграммы имеет третий тип, то L непусто и для первого подтипа равно длине указанного при последнем отрезке D слова, а для второго $Q'_2 = \{q\}$, $i = n^3 + 1$, и в текущей паре m равно числу, указанному в D.

Опишем переходы $\mathfrak{A}(D)$. Скажем, что состояние q'' из $\mathfrak{A}$ является последователем состояния q' по букве a, если есть переход из q' в q'' со входом a. Переход из состояния s_1 в состояние s_2 со входной буквой a и выходным словом v существует тогда и только тогда, когда выполнены все следующие условия.

1. $Q_1(s_2)$ (так обозначаем компоненту Q_1 в состоянии s_2) есть множество всех последователей состояний из $Q_1(s_1)$ по a. Таким образом, компонента Q_1 вычисляется детерминированно.

2. Существует переход в $\mathfrak{A}$ из $q(s_1)$ в $q(s_2)$ со входом a и выходом v.

3. Для каждого состояния q' из $Q_2(s_1)$ хотя бы один из его последователей по a принадлежит $Q_2(s_2)$. Для каждого состояния q' из $Q_1(s_1) \setminus Q_2(s_1)$ все его последователи по a не принадлежат $Q_2(s_2)$.

4. В D существует отрезок R: $[(q_1, M_1), (q_2, M_2)]$ либо типа 1 или 3 такой, что $M_1 \geq Q_2(s_1) \geq M_2$, $M_1 \geq Q_2(s_2) \geq M_2$, (очевидно, такой отрезок может быть лишь один), либо типа 2 такой, что $M_1 = Q_2(s_1)$, $M_2 = Q_2(s_2)$, $q_1 = q(s_1)$, $q_2 = q(s_2)$, v равно выходу, указанному при R. В этом последнем случае, если предыдущий отрезок D (до (q_1, M_1)) имеет третий тип, то должно выполняться: для первого подтипа $L(s_1)$ равно указанному в D, а для второго $Q'_2(s_1) = \{q_1\}$, m в текущей паре в s_1 равно указанному в D, $i(s_1) = n^3 + 1$. Аналогично, если последующий отрезок D (от (q_2, M_2)) имеет третий тип, то должны выполняться очевидные начальные условия. В случае, если R третьего типа, то должны выполняться перечисленные в следующем абзаце условия.

Для первого подтипа v не длиннее продолжения указанного в D выхода после $L(s_1)$ первых букв и согласованно с ним, $L(s_2) = L(s_1) + |v|$. Для второго подтипа выполняется следующее. $Q'_1(s_2)$ состоит из всех последователей состояний из $Q'_1(s_1)$ по a. Для каждого состояния q' из $Q'_2(s_1)$ хотя бы один из его последователей по a принадлежит $Q'_2(s_2)$. Для каждого состояния q' из $Q'_1(s_1) \setminus Q'_2(s_1)$ все его последователи не принадлежат $Q'_2(s_2)$, $i(s_2) = \min(n^3 + 1, i(s_1) + |v|)$. $P(s_2)$ получается по следующим правилам. Сначала сопоставим каждому $q' \in Q'_2(s_2)$ минимальное число среди всех сумм $m + |w|$ таких, что для некоторого q'' $\langle m, q''\rangle \in P(s_1)$ и существует переход из q'' в q' со входом a и выходом w.

Среди всех сопоставленных чисел возьмем минимальное число k. Если $k \neq 0$, уменьшим все эти числа на k. При этом все получившиеся числа не превышают $2n^3$. Эти числа в парах с соответствующими им состояниями войдут в $P(s_2)$. В текущей паре $m(s_2) = m(s_1) + |v| - k \leq 2n^3$.

Очевидно, что если в $\mathfrak{A}$ есть допускающий путь l с диаграммой D, то в $\mathfrak{A}(D)$ существует допускающий путь l' такой, что $вх(l) = вх(l')$, $вых(l) = вых(l')$. Пусть, наоборот, в $\mathfrak{A}(D)$ есть допускающий путь l'. Покажем, что соответствующий ему угаданный путь l в $\mathfrak{A}$ имеет диаграмму D. Компонента Q_1 вычисляется верно, так как она детерминированна. Покажем, что Q_2 тоже вычисляется верно. Пусть это не так. Пусть Q_2 в некоторый момент имеет лишнее состояние. Тогда в силу того, что элементы Q_2 имеют хотя бы одного последователя в Q_2 на каждом шагу, в конце Q_2 будет содержать незаключительное состояние, что противоречит тому, что путь l' допускающий. Пусть Q_2 не содержит некоторого состояния, которое на самом деле входит в множество двусторонней достижимости. Так как все последователи состояний из $Q_1 \setminus Q_2$ не принадлежат Q_2, то в конце некоторое заключительное состояние будет принадлежать $Q_1 \setminus Q_2$, что невозможно. Аналогично доказывается правильность вычисления Q_1' и Q_2' на отрезках. Переходы $\mathfrak{A}(D)$ устроены так, что находясь на каком-либо отрезке D первого или третьего типа, можно уйти с него только по переходу, соответствующему следующему отрезку второго типа, и лишь тогда, когда информация о первом отрезке соответствует D. Правильность вычисления этой информации очевидна. Построение $\mathfrak{A}(D)$, очевидно, производится за экспоненциальное время. Теорема 1 и достаточность критерия конечнозначности доказаны.

Можно построить $\mathfrak{A}(D)$ так, чтобы он обладал следующим дополнительным свойством: для любого слова u существует не более одного допускающего пути со входом u. Для этого нужно с самого начала перенумеровать все переходы $\mathfrak{A}$. В определении состояний $\mathfrak{A}(D)$ следует добавить для отрезков первого типа — компоненты, вычисляющие множества локальной односторонней и двусторонней достижимости, для отрезков третьего типа первого подтипа — множество "плохих" пар вида $\langle q, i\rangle$, а для второго подтипа — множество "плохих" троек вида $\langle q, m, i\rangle$. Здесь q — состояние $\mathfrak{A}$, в котором по предположению оканчивается начало некоторого возможного пути, последовательность переходов которого лексикографически меньше, чем у угадываемого пути. В числах m, i вычисляются те же характеристики возможного пути, которые раньше вычислялись только для угадываемого пути. В устройстве переходов $\mathfrak{A}(D)$ добавляется правило о том, что все переходы из текущего состояния, которые меньше угаданного, либо ведут в локально двусторонне

недостижимое состояние (на отрезках типа 1, 3), либо не соответствуют указанному в D переходу (тип 2), либо (отрезок типа 3) ведут в состояние, которое входит в "плохую" пару или тройку. Если в $\mathfrak{A}(D)$ есть переход из s_1 в s_2 со входом a и выходом v, то для каждого состояния q'', являющегося последователем по a состояния q', входящего в "плохую" пару $\langle q', i\rangle \in P(s_1)$ (первый подтип) или в "плохую" тройку $\langle q', m, i\rangle \in P(s_1)$ (второй подтип), должно выполняться хотя бы одно из следующих условий.

1. $q'' \notin Q'_2(s_2)$.

2. В $P(s_2)$ существует "плохая" пара $\langle q'', i_1\rangle$ или "плохая" тройка $\langle q'', m_1, i_1\rangle$, где m_1 и i_1 вычислены по m, i, v так же, как вычислялись для угадываемого пути.

3. Для первого подтипа v не согласовано с продолжением указанного в D слова или выходит за его пределы. В конце отрезка для любой "плохой" пары число i не должно быть равно длине указанного в D слова, а для любой "плохой" тройки должно быть либо m не равно числу из D, либо $i < n^3 + 1$.

Легко видеть, что описанный преобразователь для каждого входа допускает лишь тот путь с диаграммой D, последовательность переходов которого наименьшая в лексикографическом смысле. Укажем одно следствие наших построений.

ТЕОРЕМА 3. *Для любого слова u в $\mathfrak{A}$ существует множество $M(u)$ из* $\exp(\mathrm{poly}(n))$ *допускающих путей со входом u такое, что для любого допускающего пути l со входом u существует путь $l' \in M(u)$ такой, что вых(l') = вых(l) и для любого $u' \subseteq u$ выполняется: $|d(l, l', u')| \le 2n^3$.*

В частности, как доказал А.Вебер в [1], значность конечнозначного преобразователя размера n не превосходит $\exp(\mathrm{poly}(n))$.

Для доказательства теоремы 3 возьмем в качестве $M(u)$ по одному пути с каждой возможной диаграммой. Утверждение теоремы 3 легко следует из лемм 2, 3, 4.

Нижняя оценка значности конечнозначного преобразователя экспоненциальна. Легко показать, что даже число различных длин выходов на одном входе может быть экспоненциальным. Для этого рассмотрим преобразователь $\mathfrak{B}$, имеющий $m+1$ пару состояний: (q_1, q'_1), $(q_2, q'_2), \ldots, (q_{m+1}, q'_{m+1})$. Начальные состояния q_1, q'_1, заключительные q_{m+1}, q'_{m+1}. Переходы следующие:

$$\langle q_i, u_i, \Lambda, q_{i+1}\rangle,\ \langle q_i, u_i, \Lambda, q'_{i+1}\rangle,\ \langle q'_i, u_i, \Lambda, q_{i+1}\rangle,$$
$$\langle q'_i, u_i, \Lambda, q'_{i+1}\rangle,\ \langle q_i, u_i, c, q_i\rangle,\ \langle q'_i, u_i, \Lambda, q'_i\rangle,$$

где $u_i = a$ при нечетном i и $u_i = b$ при четном, a, b, c — буквы. Очевидно, что $\mathfrak{B}$ конечнозначный. Пусть $f(k) = 2^{m-k}$. Тогда для входа

$a^{f(0)}a\,b^{f(1)}b\,a^{f(2)}a\,b^{f(3)}b\ldots a^2a\,bb$ (считаем m четным) преобразователь $\mathfrak{B}$ может, очевидно, выдать любой выход c^k где число k записывается в двоичной системе числом из не более m цифр. Так как таких чисел всего 2^m, то нижняя экспоненциальная оценка доказана.

В общем случае дополнение до множества, распознаваемого недетерминированным автоматом размера m распознается автоматом размера $\exp(\mathrm{poly}(m))$. Однако, в данном случае имеет место

ТЕОРЕМА 4. *Для любой диаграммы D существует автомат $\mathfrak{A}'(D)$ размера $\exp(\mathrm{poly}(n))$, который распознает множество таких входов u, что не существует допускающего пути в $\mathfrak{A}(D)$ со входом u. Для каждого u в $\mathfrak{A}'(D)$ существует не более одного допускающего пути со входом u.*

ДОКАЗАТЕЛЬСТВО. Состояния $\mathfrak{A}'(D)$ включают множества Q_1, Q_2, в которых вычисляются множества односторонней и двусторонней достижимости. Кроме того, они содержат множества Q'_1 и Q'_2 локальной односторонней и двусторонней достижимости, а для отрезков третьего типа еще множество наборов — троек $\langle q, m, i\rangle$ или пар $\langle q, i\rangle$. Здесь q — состояние из Q'_2, в котором оканчивается возможный путь, т.е. путь с прочитанной частью входа, который может оказаться началом некоторого пути с диаграммой D, m и i — те же характеристики этого пути на отрезке, как и в $\mathfrak{A}(D)$. Множество наборов полное, т.е. если существует возможный путь с некоторыми характеристиками, то существует и соответствующий им набор. Q_1, Q_2, Q'_1, Q'_2 вычисляются так же, как в $\mathfrak{A}(D)$. Если при очередном переходе новое Q_2 остается на том же отрезке D, что и старое, то новые Q'_1, Q'_2 вычисляются так же, как в $\mathfrak{A}(D)$, а новое множество наборов вычисляется по старому естественным образом с сохранением свойства полноты. Если очередной переход во множествах соответствует отрезку второго типа, а перед этим был отрезок первого или третьего типа, то переход на следующий отрезок D осуществляется лишь в том случае, если Q'_2 состоит из конечного состояния отрезка, а для третьего типа среди наборов есть хотя бы один, который соответствует информации D для пройденного отрезка. Если же Q'_2 пусто или нет соответствующего набора, а также в случае, когда очередной переход во множествах противоречит D, совершается переход в состояние $\langle Q_1, Q_2, \text{“сбой”}\rangle$, и дальше продолжают вычисляться только Q_1 и Q_2, слово “сбой” сохраняется. В заключительных состояниях $\mathfrak{A}'(D)$ все состояния из Q_2 заключительные, а из $Q_1 \setminus Q_2$ — незаключительные. Среди состояний $\mathfrak{A}'(D)$, в которых выполняется это свойство, незаключительными являются лишь такие, где нет сбоя, Q_2 равно последнему множеству из D, и существует набор, соответствующий информации, приписанной к последнему отрезку D.

Легко видеть, что если в $\mathfrak{A}$ нет допускающего пути со входом u и диаграммой D, то $\mathfrak{A}'(D)$ допускает u. Пусть наоборот, в $\mathfrak{A}'(D)$ есть допускающий путь со входом u. Так же, как в доказательстве теоремы 1 показывается, что Q_1, Q_2, Q_1', Q_2' вычисляются верно. Из устройства переходов и полноты множества наборов следует, что учитываются все возможные начала путей, претендующих на диаграмму D и верно определяется их отсутствие. Поэтому, в $\mathfrak{A}$ нет допускающего пути со входом u и диаграммой D. Из того, что входная буква и фиксированный вариант угадывания Q_2 и Q_2' однозначно определяют переход из любого состояния $\mathfrak{A}'(D)$, следует единственность допускающего пути со входом u. Теорема 4 доказана. Отметим, что подобную теорему для двухэкспоненциальной оценки доказал в [2] А.Вебер.

А.Вебер доказал в [1] следующую теорему.

ТЕОРЕМА 5. *Существует полиномиальный алгоритм, который по произвольному преобразователю решает, конечнозначный он или нет.*

ДОКАЗАТЕЛЬСТВО. Этот результат следует из леммы 1, критерия конечнозначности (теорема 2) и того, что по произвольному преобразователю $\mathfrak{A}$ без пустых входов можно за полиномиальное время решить, удовлетворяет ли он критерию. Покажем, как это сделать. Перебираем пары состояний $\langle s_1, s_2\rangle$. Для каждой пары сначала проверим, выполняется ли условие (2). Для этого, построим следующий недетерминированный автомат A_1. Его состояния — упорядоченные четверки $\langle q_1, q_2, q_3, m\rangle$, где q_1, q_2, q_3 — состояния $\mathfrak{A}$, а m — либо целое число такое, что $|m| \le n^3$, либо $*$. Переход из состояния $\langle q_1, q_2, q_3, m\rangle$ в состояние $\langle q_1', q_2', q_3', m'\rangle$ со входом a существует тогда и только тогда, когда во-первых для каждого $i = 1, 2, 3$ существует переход в $\mathfrak{A}$ из q_i в q_i' со входом a (обозначим его выход через v_i) и во-вторых если m — число и $|m + |v_1| - |v_2|| \le n^3$, то $m' = m + |v_1| - |v_2|$, иначе $m' = *$. Начальное состояние A_1: $\langle s_1, s_1, s_2, 0\rangle$, заключительное: $\langle s_1, s_2, s_2, *\rangle$. В m вычисляется разность между длинами выходов p_1 и p_2, а $*$ указывает, что ее модуль превысил n^3.

Очевидно, что из существования допускающего пути в A_1 следует существование указанных в критерии путей p_1, p_2, p_3 таких, что на некотором начале u' их общего входа $|d(p_1, p_2, u')| > n^3$. Обратно, из существования таких путей вытекает, очевидно, существование допускающего пути в A_1. Таким образом, выполненность условия (2) для s_1, s_2 эквивалентна отсутствию в A_1 допускающего пути, что, очевидно, проверяется за полиномиальное время.

Теперь проверим для s_1, s_2 выполненность условия (1), предполагая, что условие (2) для них выполнено. Сначала заметим, что условие (1) эквивалентно следующему условию (1)$'$: слова *вых*(p_1) и *вых*(p_2) согла-

сованы. Действительно, если (1) нарушается для путей p_1, p_2, p_3, то для путей $p'_1 = p_1^k$, $p'_2 = p_2 p_3^{k-1}$, $p'_3 = p_3^k$, где k достаточно велико, нарушается условие $(1)'$. Построим следующий автомат A_2. Его состояния делятся на три множества: Q_1, Q_2, Q_3. В Q_1 это пятерки $\langle q_1, q_2, q_3, m, a\rangle$, здесь q_1, q_2, q_3 так же, как в A_1, соответствуют последним состояниям угадываемых путей p_1, p_2, p_3: в m вычисляется разность длин выходов p_1 и p_2. При этом требуется, чтобы $|m| \le n^3$, если это неравенство нарушается, переход отсутствует. В a естественным образом вычисляется последняя буква того пути из p_1, p_2, выход которого в данный момент строго длиннее (если $|m| > 0$), а если $m = 0$, то $a = \Lambda$. Начальное состояние: $\langle s_1, s_1, s_2, 0, \Lambda\rangle$, заключительных состояний в Q_1 нет.

Находясь в произвольном состоянии $\langle q_1, q_2, q_3, m, a\rangle \in Q_1$ $(m \neq 0)$, A_2 может "предположить", что именно в указанной букве a произойдет рассогласованность *вых*(p_1) и *вых*(p_2) (назовем ее сбоем), и перейти (с пустым входом и выходом) в состояние из Q_2, имеющее вид $\langle q_1, q_2, q_3, a, k\rangle$, где k сначала равно m. Здесь в q_1, q_2, q_3 вычисляется обычная информация, a — буква предполагаемого сбоя, k указывает на сколько букв надо нарастить "короткий" выход, чтобы проверить, действительно ли в указанном месте произойдет сбой. Переходы между состояниями в Q_2 очевидным образом уменьшают $|k|$ по мере нарастания "короткого" выхода, знак k определяет путь с "коротким" выходом. Когда на "коротком" выходе появляется буква, в которой по предположению должен быть сбой, проверяется, действительно ли эта буква не равна a. Если это так, A_2 может перейти в состояние из Q_3, которое имеет вид $\langle q_1, q_2, q_3\rangle$. Если, находясь в состоянии $\langle q_1, q_2, q_3, 0, \Lambda\rangle \in Q_1$, A_2 обнаруживает при очередном переходе несогласованность выходов p_1 и p_2, то он тоже может перейти в $\langle q'_1, q'_2, q'_3\rangle \in Q_3$. Нахождение в состоянии из Q_3 означает, что сбой произошел и теперь осталось достроить пути p_1, p_2, p_3 до конца. Переходы в Q_3 определяются естественным образом, заключительное состояние: $\langle s_1, s_2, s_2\rangle$. Очевидно, что условие $(1)'$ выполнено для s_1, s_2 тогда и только тогда, когда не существует допускающего пути в A_2. Полиномиальность времени проверки критерия конечнозначности и теорема 5 доказаны.

Перейдем к вопросам, связанным с вложенностью одного преобразователя в другой. В [3] доказана разрешимость выяснения вопроса о вложенности произвольного преобразователя в конечнозначный без оценки на время алгоритма. А.Вебер в [2] доказал разрешимость этого вопроса за время $\exp(\exp(\mathrm{poly}(n)))$, где n — сумма размеров преобразователей. Размер зоны у этого алгоритма тоже двойная экспонента. Мы докажем теорему, усиливающую этот результат, построив алгоритм с экспоненциальной зоной.

ТЕОРЕМА 6. *Существует детерминированный алгоритм, работающий на зоне* $\exp(\mathrm{poly}(n))$, *который по произвольному преобразователю* $\mathfrak{A}_1$ *и конечнозначному преобразователю* $\mathfrak{A}_2$ *решает вопрос о вложенности* $\mathfrak{A}_1$ *в* $\mathfrak{A}_2$.

ДОКАЗАТЕЛЬСТВО. В силу леммы 1 можно считать, что $\mathfrak{A}_1$ и $\mathfrak{A}_2$ не имеют пустых входов. Следующая лемма утверждает, что если $\mathfrak{A}_1$ не вложен в $\mathfrak{A}_2$, то это проявляется на выходах экспоненциальной длины.

ЛЕММА 5. *Если преобразователь* $\mathfrak{A}_1$ *не вложен в конечнозначный преобразователь* $\mathfrak{A}_2$, *то существует пара* $\langle u, v\rangle$ *такая, что* $\langle u, v\rangle \in \Gamma(\mathfrak{A}_1)$, $\langle u, v\rangle \notin \Gamma(\mathfrak{A}_2)$ *и* $|v| \leq \exp(p_1(n))$, *где* $p_1(n)$ — *некоторый полином.*

ДОКАЗАТЕЛЬСТВО. Рассмотрим допускающий путь l_1 в $\mathfrak{A}_1$ с выходом минимальной длины такой, что $\langle u, v_1\rangle \in \Gamma(\mathfrak{A}_1)$, $\langle u, v_1\rangle \notin \Gamma(\mathfrak{A}_2)$, где $u = \mathit{вх}(l_1)$, $v_1 = \mathit{вых}(l_1)$. Предположим, что $|v_1| > \exp(p_1(n))$, где степень $p_1(n)$ достаточно велика. Будем далее подразумевать под словосочетаниями типа "достаточно много" $\exp(\mathrm{poly}(n))$, где степень полинома настолько велика, чтобы можно было выполнить все описываемые действия. Если $|v_1| = |v_2|$, но $v_1 \neq v_2$, то назовем левым (правым) сбоем между v_1 и v_2 первую слева (справа) пару неравных букв слов v_1 и v_2, равноотстоящих от начала (конца). По теореме 3, количество различных v_2 таких, что $\langle v_1, v_2\rangle \in \Gamma(\mathfrak{A}_2)$ — фиксированная экспонента, поэтому и букв на v_1, в которых имеет место левый или правый сбой между v_1 и каким-нибудь из таких v_2, тоже мало. Возьмем в слове v_1 достаточно большое подслово r, в которое не попадает ни один сбой и которое лежит от ближайшего сбоя на расстоянии, в достаточно большое число раз превышающем $|r|$.

Будем называть начала входа u точками. Иногда под точкой будем понимать также то место на пути, где заканчивается соответствующее начало входа. Выделим достаточно много точек, так чтобы выполнялись следующие три условия:

1. Слова $\mathit{вых}(l_1(u'))$ для всех выделенных точек u' заканчиваются внутри r и различны.

2. Во всех выделенных точках l_1 проходит через одно и то же состояние.

3. Во всех выделенных точках множества левосторонней и правосторонней достижимости в $\mathfrak{A}_2$ одни и те же (множеством правосторонней достижимости в точке u' мы называем множество состояний, из которых существует путь в заключительное состояние со входом u'', где $u'u'' = u$).

Будем следующим образом сужать по шагам множество выделенных точек и одновременно строить в $\mathfrak{A}_2$ множество отмеченных состояний.

На очередном шаге для каждого неотмеченного состояния q из $\mathfrak{A}_2$ проверяем, существует ли допускающий путь в $\mathfrak{A}_2$ со входом u, который находится в q в не менее чем $m/\exp(n)$ выделенных точках, где m — количество выделенных точек на данный момент. Если существует, то отмечаем q, ставим ему в соответствие один из описанных путей, а множество выделенных точек уменьшаем в $\exp(n)$ раз, так чтобы этот путь во всех выделенных точках проходил через q. Путь, поставленный в соответствие отмеченному состоянию q, будем обозначать $l(q)$ и называть отмеченным. Когда на очередном шаге ни одно неотмеченное состояние не будет отмечено, процесс останавливается. Так как шагов не более n, в конце процесса количество выделеннных точек достаточно велико. Каждый путь $l(q)$ во всех выделенных точках находится в состоянии q.

Разделим все выделенные точки на три примерно равные части идущих подряд точек, а между левой и средней частью зафиксируем одну из выделенных точек u_b, которую будем называть граничной. Каждой выделенной точке u' из средней части сопоставим множество, состоящее из всех таких наборов $\langle q, q_1, q_2, d\rangle$, что в $\mathfrak{A}_2$ существует путь l из q_1 в q_2 со входом u'', где $u' = u_b u''$, q — отмеченное состояние и разность между $|\,\mathit{вых}(l)|$ и длиной выхода $l(q)$ на отрезке входа u'' равна d, причем $|d| \leq 2n^3$. Число описанных множеств — фиксированная экспонента, поэтому существует много выделенных средних точек, которым соответствует одно и то же множество. В средней части будем считать выделенными только эти точки.

Так как во всех выделенных точках путь l_1 находится в одном состоянии, мы можем выбросить любое множество отрезков между выделенными точками и получить укороченный допускающий путь l_1' в $\mathfrak{A}_1$. В силу выбора l_1, существует допускающий путь l_2' в $\mathfrak{A}_2$ такой, что $\mathit{вх}(l_2') = \mathit{вх}(l_1')$, $\mathit{вых}(l_2') = \mathit{вых}(l_1')$. Пусть выбросы отрезков произведены в средней трети. Будем понимать под выделенными точками пути с выбросами точки, получившиеся из точек первоначального пути после выбросов из них. Покажем, что l_2' хотя бы в одной из выделенных точек первой и последней трети проходит через отмеченные состояния. Будем обозначать через $q(t)$ состояние, в котором находится l_2' в точке t. Рассмотрим на $\mathit{вх}(l_2')$ точку t, соответствующую самому правому выброшенному отрезку. В силу совпадения множеств правосторонней достижимости во всех выделенных точках, существует путь из $q(t)$ в заключительное состояние, вход которого равен $u'u''$, где u' — вход самого правого выброшенного отрезка, а u'' — вход l_2' от t до конца. Соединив старое начало с новым концом, получим новый допускающий путь, на входе которого одним выброшенным отрезком меньше. Таким же образом вставим остальные выброшенные отрезки и получим допускающий

путь со входом u. В первой трети выделенных точек он проходит через некоторое состояние q_1, которое встречается на примерно $1/3n$ от количества всех выделенных точек. Если бы q_1 не было отмеченным, это бы противоречило завершенности построения отмеченных состояний. Аналогично показывается, что на последней трети существует выделенная точка, в которой l'_2 находится в отмеченном состоянии q_2.

Так как пути $l(q_1)$ и $l(q_2)$ во всех выделенных точках проходят через q_1 и q_2, то из этих путей можно сделать выбросы на тех же отрезках входа, что и для l_1. Будем обозначать пути с такими выбросами через $l'(q_1)$ и $l'(q_2)$. Из условия (2) (при $s_1 = q_1$, $s_2 = q_2$, p_1, p_2, p_3 — участки путей $l'(q_1)$, l'_2, $l'(q_2)$ соответственно) следует, что разность длин выходов пути l'_2 на отрезке от u_b до любой средней точки u' и пути $l'(q_1)$ на том же участке входа, не превосходит по модулю $2n^3$. Учитывая это, рассмотрим следующий процесс "восстановления" l'_2 до пути с первоначальным входом. Пусть самый левый выброшенный отрезок расположен между точками t_1 и t_2 и его вход есть u_1. Обозначим через l' участок пути l'_2 от u_b до t_1. По построению (совпадение множеств наборов в средних точках) существует путь l'' из $q(u_b)$ в $q(t_1)$ такой, что $вх(l'') = вх(l')u_1$, а разность $|\,вых(l'')| - |\,вых(l')|$ равна длине выхода пути $l(q_1)$ на отрезке входа u_1. Вход нового допускающего пути, получающегося заменой в l'_2 участка l' на l'', содержит на один выброшенный отрезок меньше. Можно сказать, что мы вставили отрезок на входе и при этом выход удлинился на длину выхода пути $l(q_1)$ на вставленном отрезке входа. После этого, таким же образом вставляем второй слева отрезок и так далее. В конце получим "восстановленный" путь со входом u.

Будем обозначать для выброса α через $l'_2(\alpha)$ некоторый допускающий путь с выбросом α в $\mathfrak{A}_2$; через $l_2(\alpha)$ — некоторый "восстановленный" путь со входом u, построенный описанным выше процессом; $q_1(\alpha)$ — некоторое отмеченное состояние, через которое проходит в первой трети выделенных точек путь $l'_2(\alpha)$; $t_1(\alpha)$ — некоторая выделенная точка первой трети в которой $l_2(\alpha)$ находится в $q_1(\alpha)$; $q_2(\alpha)$ и $t_2(\alpha)$ — то же самое для последней трети; $l'_\alpha(q_1)$ — путь $l(q_1)$ с выбросом α; $|\alpha|$ — сумма длин выходов выбрасываемых из l_1 отрезков.

Любой отрезок между двумя средними выделенными точками имеет относительно каждого отмеченного состояния q один из трех типов: длина выхода $l(q)$ на этом отрезке может быть больше длины выхода пути l_1 на нем (положительный тип), меньше (отрицательный тип) или равна (нулевой тип). Таким образом, каждому такому отрезку сопоставляется набор пар $\langle q, \text{тип}\rangle$. Всего таких наборов — фиксированная экспонента. Выберем достаточно много непересекающихся отрезков, которым соответствует один и тот же набор. Упорядочим их и рассмо-

трим последовательность S выбросов, в которой m-тый выброс состоит из первых m отрезков в нашем упорядочении. Каждому выбросу α из S поставим в соответствие пару $\langle q, v\rangle$, где $q = q_1(\alpha)$, $v = \mathit{вых}(l_2(\alpha))$. По теореме 3 количество таких пар — фиксированная экспонента. Пусть α_1 и α_2, — два различных выброса из S, которым соответствует одна и та же пара $\langle q, v\rangle$. Если бы тип всех отрезков относительно q был ненулевым, то, очевидно, разности между длинами выходов "восстановленных" путей и $|v_1|$ были бы различны для α_1 и α_2, значит эти выходы тоже были бы различны. Полученное противоречие показывает, что тип всех отрезков относительно q нулевой, что с учетом конструкции вставки дает для всех $\alpha \in S$: $|\mathit{вых}(l_2(\alpha))| = |v_1|$.

Скажем, что выброс α является выбросом нулевого типа, если тип всех отрезков из α относительно $q_1(\alpha)$ нулевой. Таким образом, на любом участке средней части с достаточно большим количеством выделенных точек существует достаточно много выбросов нулевого типа: $\alpha_1, \alpha_2, \ldots, \alpha_k$, где все $|\alpha_i|$ различны. Поэтому, можно взять достаточно много выбросов на средней трети так, чтобы выполнялись следующие условия:

1. Выбросы упорядочены, т.е. каждый отрезок одного из любых двух выбросов лежит строго левее каждого отрезка другого и не пересекается с ним.

2. Все выбросы являются выбросами нулевого типа.

3. Для любых двух выбросов α_1, α_2: $|\alpha_1| \neq |\alpha_2|$.

4. Для всех выбросов α $q_1(\alpha)$ одно и то же (обозначим его q_1).

5. Для всех α разности между длиной выхода $l'_2(\alpha)$ на отрезке от $t_1(\alpha)$ до u_b и длиной выхода $l'_\alpha(q_1)$ на том же отрезке одинаковы (из (2) следует, что эта разность по модулю не превосходит n^3).

6. У всех $l_2(\alpha)$ (а значит и $l'_2(\alpha)$) одинаковый выход от начала до u_b.

7. Все $\mathit{вых}(l_2(\alpha))$ одинаковы (обозначим этот выход v_2).

Из условия 2 вытекает, что выход пути $l(q_1)$ не пуст на любом отрезке, входящем в выбросы. Тогда, из условия (1) и того, что разность выходов путей $l'_2(\alpha)$ и $l'_\alpha(q_1)$ на отрезке от $t_1(\alpha)$ до $t_2(\alpha)$ не меняется после вставки, следует, что для всех α (кроме, возможно, n^3 самых правых) $\mathit{вых}(l_2(\alpha))$ фактически получается из $\mathit{вых}(l'_2(\alpha))$ вставкой выходов пути $l(q_1)$ на отрезках из α. Из условия 5 следует, что для всех α (кроме, возможно, n^3 самых левых) выходы всех $l'_2(\alpha)$ после точки u_b повторяют выходы путей $l'_\alpha(q_1)$ с одного и того же места на общем начале выходов $l'_\alpha(q_1)$. Учитывая условие 6, заключаем, что если выброс α_1 строго левее чем α_2, то и вставка на выходе для α_1 происходит строго левее (считая по длине всего выхода) вставки для α_2. Назовем это свойство монотонностью. Напомним, что $v_1 \neq v_2$, $|v_1| = |v_2|$ (в силу условия 2)

и для любого выброса если из v_1 и v_2 выбросить соответствующие отрезки, они станут равными. В силу монотонности, существует много таких выбросов, для которых участок выбросов из v_2 не содержит ни левого ни правого сбоя между v_1 и v_2. Рассмотрим эти выбросы. Легко видеть, что для того, чтобы после них сбои исчезли, необходимо, чтобы выбросы из v_1 и v_2 лежали по разные стороны от обоих сбоев. Поэтому, из расположения участка r следует, что выбросы из v_1 лежат по одну сторону от обоих сбоев. Пусть они лежат слева от левого сбоя, тогда выбросы из v_2 лежат справа от правого сбоя. Рассмотрим два выброса α_1 и α_2, где α_1 левее α_2. Рассмотрим подслово w из v_1 от правого конца α_2 до левого сбоя. При совершении выброса α_2, w сдвигается влево и отождествляется само с собой. Следовательно, w периодично с периодом длины $|\alpha_2|$. Легко видеть, что если продолжить период вправо за w, то нарушение его должно произойти в точности на расстоянии $|\alpha_2|$ от левого сбоя, иначе сбой бы не исчез. Рассуждая аналогично по отношению к α_1, легко показать, что w периодично с периодом длины $|\alpha_1|$ и что его нарушение происходит в точности на $|\alpha_1|$ правее левого сбоя. По построению подслова r, $|w| \gg |\alpha_1| \cdot |\alpha_2|$, поэтому w периодично с периодом длины $|\alpha_1| \cdot |\alpha_2|$ и нарушение этого большого периода справа от левого сбоя является, очевидно, нарушением обоих малых периодов. Так как $|\alpha_1| \neq |\alpha_2|$ (условие 3), мы получили противоречие с тем, что нарушения этих двух периодов произошли в разных местах. Случай, когда выбросы из v_1 лежат правее правого сбоя, рассматривается симметричным образом. Лемма 5 доказана.

Для полноты рассмотрим также оценку на $|u|$ в условиях доказанной леммы. А.Вебер в [2] доказал, что если преобразователь $\mathfrak{A}_1$ не вложен в конечнозначный преобразователь $\mathfrak{A}_2$, то существует пара $\langle u, v\rangle \in \Gamma(\mathfrak{A}_1)$, $\langle u, v\rangle \notin \Gamma(\mathfrak{A}_2)$, где $|u| \leq \exp(\exp(\mathrm{poly}(n)))$. Покажем, как это утверждение можно вывести из леммы 5. Рассмотрим пару $\langle u, v\rangle$, существование которой доказано в лемме 5, с минимальной $|u|$ при данном v. Предположим, что $|u| > \exp(\exp(\mathrm{poly}(n)))$, где степень полинома достаточно велика. Так как $|v| \leq \exp(p_1(n))$, то на допускающем пути l_1 из $\mathfrak{A}_1$ со входом u и выходом v существует достаточно длинный (две экспоненты) отрезок с пустым выходом. Выделим на этом отрезке много точек, в которых l_1 находится в одном состоянии. Каждой точке u' из них поставим в соответствие множество пар вида $\langle q, v'\rangle$, где $v' \subseteq v$, q — состояние из $\mathfrak{A}_2$ такое, что в $\mathfrak{A}_2$ существует путь из начала в q со входом u' и выходом v'. Из теоремы 3 следует, что количество таких множеств — фиксированная двойная экспонента, поэтому можно выбрать две точки u_1, u_2, которым соответствует одно и то же множество. По построению, в $\mathfrak{A}_2$ существует допускающий путь l_2 со входом, получающимся из

u выбрасыванием отрезка от u_1 до u_2, и выходом v. Пусть q — состояние, в котором кончается $l_2(u_1)$. В силу совпадения соответствующих множеств, в $\mathfrak{A}_2$ существует путь l_2' из начала в q такой, что $вх(l_2') = u_2$, $вых(l_2') = вых(l_2(u_1))$. Соединяя путь l_2' с продолжением пути l_2, получим допускающий путь в $\mathfrak{A}_2$ со входом u и выходом v. Это противоречит тому, что $\langle u, v\rangle \notin \Gamma(\mathfrak{A}_2)$. Утверждение доказано.

Легко показать, что нижняя оценка на длину выхода v в формулировке леммы 5 экспоненциальна. Эта оценка, очевидно, следует из того, что для любого n существует недетерминированный автомат A размера $\mathrm{poly}(n)$, который не допускает некоторое слово w, где $|w| = \exp(\mathrm{poly}(n))$, но допускает все слова длины меньшей $|w|$. В качестве w возьмем конкатенацию записей всех n-разрядных двоичных чисел (младшие разряды слева), расположенных в порядке возрастания, начиная от слова из одних нулей и кончая словом из одних единиц. Автомат A в процессе работы угадывает причину, по которой читаемое слово не равно w. Причины могут быть такими: первое слово не нулевое; некоторое последующее число не равно предыдущему, увеличенному на 1 (в этом случае угадывается неверный разряд); длина всего слова не кратна n; последнее слово не единичное. В состояниях A (до угадывания) хранится номер последнего прочитанного разряда и информация о том, есть ли левее него нули в текущем числе. Дальнейшие детали очевидны.

Автору неизвестен ответ на следующий

ВОПРОС. Существует ли двухэкспоненциальная нижняя оценка на длину входа u в формулировке леммы 5?

Докажем теорему 6. Опишем недетерминированный алгоритм, подтверждающий невложенность $\mathfrak{A}_1$ в $\mathfrak{A}_2$ и работающий на экспоненциальной зоне. Сначала алгоритм угадывает выход v, где $|v| \leq \exp(p_1(n))$ и записывает v на ленте. После этого по шагам угадывается вход и путь в $\mathfrak{A}_1$. В каждый момент на ленте указано $v_1 \subseteq v$ — выход угаданного начала пути в $\mathfrak{A}_1$ и множество пар $\langle v', q\rangle$, где $v' \subseteq v$, q — состояние из $\mathfrak{A}_2$ такое, что существует путь из начала в q с угаданным к данному моменту входом и выходом v'. На очередном шаге угадывается очередная буква a входа и очередной переход в $\mathfrak{A}_1$ со входом a. Алгоритм проверяет, что выход у этого перехода продолжает v_1 вдоль v, и дописывает его к v_1. Затем для каждой пары $\langle v', q\rangle$ и каждого перехода из q со входом a и выходом, который продолжает v' вдоль v и не выводит за его пределы, естественным образом строится новая пара. Повторяющиеся пары удаляются из получившегося нового множества. Алгоритм работает до тех пор, пока v_1 — начало v и множество пар непусто. Если $v_1 = v$ и во множестве пар нет пары $\langle v, q\rangle$, где q — заключительное состояние, то алгоритм обнаруживает невложенность $\mathfrak{A}_1$ в $\mathfrak{A}_2$. Правильность работы и

экспоненциальная зона этого алгоритма очевидны. По теореме Сэвича, недетерминированный алгоритм, работающий на зоне S, может быть переделан в детерминированный, распознающий тот же язык и имеющий зону S^2. Отсюда следует существование искомого алгоритма. Теорема 6 доказана.

Конструкция выбросов, использованная для доказательства теоремы 6, позволяет получить следующий результат.

ТЕОРЕМА 7. *Пусть* $\mathfrak{A}_1$ *и* $\mathfrak{A}_2$ — *конечнозначные преобразователи без пустых входов и* $\mathfrak{A}_1$ *вложен в* $\mathfrak{A}_2$. *Тогда для любого допускающего пути* l_1 *в* $\mathfrak{A}_1$ *со входом* u *и выходом* v *существует допускающий путь* l_2 *в* $\mathfrak{A}_2$ *с теми же входом и выходом такой, что для любого* $u' \subseteq u$ *выполняется:* $|d(l_1, l_2, u')| \leq \exp(\mathrm{poly}(n))$.

ДОКАЗАТЕЛЬСТВО. Обозначим через $M_1(u, v)$ и $M_2(u, v)$ множества допускающих путей со входом u и выходом v в $\mathfrak{A}_1$ и $\mathfrak{A}_2$ соответственно. Предположим вопреки утверждению теоремы, что существует путь $l_0 \in M_1(u, v)$ такой, что для любого пути $l_2 \in M_2(u, v)$ существует $u' \subseteq u$ такое, что $|d(l_0, l_2, u')| > k \geq \exp(\mathrm{poly}(n))$, где степень полинома достаточно велика. По теореме 3 в $\mathfrak{A}_2$ существует не более чем экспоненциальное множество M путей из $M_2(u, v)$ такое, что для любого пути $l \in M_2(u, v)$ существует $l' \in M$ такой, что для любой точки u' $|d(l, l', u')| \leq 2n^3$. Рассмотрим множество P допускающих путей l в $\mathfrak{A}_1$, обладающих следующим свойством: на входе l существует множество T из не более $|M|$ точек такое, что для любого пути $l' \in M_2(\mathit{вх}(l), \mathit{вых}(l))$ существует точка $u' \in T$ такая, что $|d(l, l', u')| > k_1 = k - 2n^3$. Легко видеть, что $l_0 \in P$, поэтому P не пусто. Пусть l_1 — путь из P с минимальной длиной выхода, обозначим: $u = \mathit{вх}(l_1)$, $v_1 = \mathit{вых}(l_1)$. Соответствующее ему множество точек обозначим T_1. Из наших предположений следует, что $|v_1| > k_1$. Поэтому, существует длинный отрезок r пути l_1 с большим выходом, не содержащий точек из T_1. Для r повторим всю конструкцию выбросов, описанную в доказательстве леммы 5. Единственное отличие будет состоять в том, что в качестве пути l_2' в $\mathfrak{A}_2$ для пути l_1' в $\mathfrak{A}_1$ с выбросами, мы будем брать не произвольный путь, а такой, что для любой точки $u' \in T_1$ выполняется: $|d(l_1', l_2', u')| \leq k_1$. Этот путь существует в силу условия выбора l_1 и того, что $|\mathit{вых}(l_1')| < |\mathit{вых}(l_1)|$. Повторив соответствующее рассуждение, легко доказать, что существует такой выброс нулевого типа из пути l, что для "восстановленного" пути l_2 в $\mathfrak{A}_2$ $\mathit{вх}(l_2) = u$, $\mathit{вых}(l_2) = v_1$. Последнее равенство следует из того, что при доказательстве леммы 5 мы получили противоречие, предполагая, что для всех выбросов α $\mathit{вых}(l_2(\alpha)) \neq v_1$. Легко видеть, что в силу нулевого типа и конструкции вставки, при вставке отрезка отклонение $d(l_1', l_2', u')$ пути в $\mathfrak{A}_1$ от пути в $\mathfrak{A}_2$ может меняться только для u',

лежащих между граничной точкой u_b и вставляемым на входе отрезком. Следовательно, за пределами r и в частности во всех точках $u' \in T$ выполняется: $|d(l_1, l_2, u')| \leq k_1$. Получили противоречие с тем, что $l_1 \in P$. Теорема 7 доказана.

В одном частном случае мы можем усилить результат теоремы 6. Будем говорить, что преобразователь $\mathfrak{A}$ *имеет конечную задержку*, если существует такое натуральное c, что для любого пути l, если $|\ вх(l)| \geq c$, то $|\ вых(l)| > 0$. Очевидно, что $c \leq \mathrm{poly}(n)$, где $n = |\mathfrak{A}|$.

ТЕОРЕМА 8. *Существует недетерминированный алгоритм, который за недетерминированное время* $\exp(\mathrm{poly}(n))$ *подтверждает невложенность произвольного преобразователя* $\mathfrak{A}_1$ *в конечнозначный преобразователь* $\mathfrak{A}_2$*, где* $\mathfrak{A}_2$ *имеет конечную задержку.*

ДОКАЗАТЕЛЬСТВО. Следующая лемма усиливает в этом частном случае лемму 5.

ЛЕММА 6. *Если преобразователь* $\mathfrak{A}_1$ *не вложен в конечнозначный преобразователь* $\mathfrak{A}_2$ *с конечной задежкой, то существует пара* $\langle u, v\rangle \in \Gamma(\mathfrak{A}_1)$, $\langle u, v\rangle \notin \Gamma(\mathfrak{A}_2)$ *и* $|u| \leq \exp(p_2(n))$ ($p_2(n)$ — полином).

ДОКАЗАТЕЛЬСТВО. Пусть l_1 — допускающий путь в $\mathfrak{A}_1$ минимальной длины такой, что $\langle вх(l_1), вых(l_1)\rangle \notin \Gamma(\mathfrak{A}_2)$. Обозначим: $u = вх(l_1)$, $v_1 = вых(l_1)$. Предположим, что $|u|$ достаточно велико. Возможны два случая. Случай 1: $|v_1| > \exp(p_1(n))$ ($p_1(n)$ — полином из леммы 5). В этом случае приводим это предположение к противоречию точно так же, как в доказательстве леммы 5. Случай 2: $|v_1| \leq \exp(p_1(n))$. В этом случае существует достаточно длинный участок r пути l_1 с пустым выходом. На нем будем делать выбросы так же, как в доказательстве леммы 5, только вместо требования чтобы выход выбрасываемых отрезков был непуст, потребуем, чтобы длина входа каждого выбрасываемого отрезка была больше c. Повторив соответствующее рассуждение, докажем существование выброса α нулевого типа. Но в силу свойства конечной задержки, выход любого отмеченного пути в $\mathfrak{A}_2$ непустой на выбрасываемых отрезках входа. Это противоречит тому, что выход участка r пустой. Лемма 6 доказана.

Докажем теорему 8. Опишем требуемый алгоритм. Он угадывает вход u и выход v, где $|u| \leq \exp(p_2(n))$, $|v| \leq n|u|$. После этого детерминированно определяется, есть ли в $\mathfrak{A}_1$ и $\mathfrak{A}_2$ допускающий путь со входом u и выходом v. Для этого, для каждого $u' \subseteq u$, где u' увеличивается побуквенно, находится множество пар $\langle v', q\rangle$, где $v' \subseteq v$, q — состояние такое, что существует путь из начала в q со входом u' и выходом v'. Подробности очевидны. Алгоритм обнаруживает невложенность $\mathfrak{A}_1$ в $\mathfrak{A}_2$, если $\langle u, v\rangle \in \Gamma(\mathfrak{A}_1)$, $\langle u, v\rangle \notin \Gamma(\mathfrak{A}_2)$. Очевидно, время работы этого алгоритма примерно $|u| \cdot |v|$. Теорема 8 доказана.

ЛИТЕРАТУРА

[1]. A.Weber: Über die Mehrdeutigkeit und Wertigkeit von endlichen Automaten und Transducern. Dissertation, Goethe-Universitat Frankfurt am Main, 1987.

[2]. A.Weber. A Decomposition Theorem for Finite Valued Transducers and an Application to the Equivalence Problem. Manuscript, 1987.

[3]. K.Culik II, J.Karhumaki: The Equivalence of Finite Valued Transducers (on HDTOL Languages) is Decidable. TCS 47 (1986), pp.71-84.