Б. Ю. Конев

УТОЧНЕНИЕ ОЦЕНОК ВЫСОТЫ ТЕРМОВ В НАИБОЛЕЕ ОБЩЕМ УНИФИКАТОРЕ

Введение

Фиксируем произвольную алгебраическую сигнатуру – язык, в который входят лишь функциональные символы (произвольной арности) и счетное множество переменных Var. Как обычно, множество Tm термов языка определяется следующим образом:

(1) каждая переменная или 0-местый функциональный знак – это терм,

(2) если $t_1, t_2, \ldots, t_n \in \mathrm{Tm}$, f^n – функциональный знак арности $n \geqslant 0$, то $f^n(t_1, t_2, \ldots, t_n) \in \mathrm{Tm}$,

(3) выражение является термом только в том случае, если оно получено исходя из правил (1) и (2).

Пусть $\mathrm{Var}(t)$ обозначает совокупность всех переменных, входящих в терм t. Под *областью задания* произвольного отображения $\sigma : \mathrm{Var} \to \mathrm{Tm}$ будем понимать множество $\mathrm{Dom}\,\sigma = \{x \in \mathrm{Var} \mid \sigma x \neq x\}$. Значение отображения σ на терме $t \in \mathrm{Tm}$ – это терм σt, получающийся из t одновременным замещением каждой переменной $x \in \mathrm{Var}(t)$ на терм σx. Для двух отображений $\sigma_1, \sigma_2 : \mathrm{Var} \to \mathrm{Tm}$ всегда определена их композиция, характеризующаяся равенством $(\sigma_1 \circ \sigma_2)t = \sigma_1(\sigma_2 t)$, справедливым при всех $t \in \mathrm{Tm}$.

Отображение $\sigma : \mathrm{Var} \to \mathrm{Tm}$ с конечной областью задания называется *подстановкой*. Поскольку $\mathrm{Dom}(\sigma_1 \circ \sigma_2) \subseteq \mathrm{Dom}\,\sigma_1 \cup \mathrm{Dom}\,\sigma_2$, композиция произвольных подстановок является подстановкой. Две подстановки равны, если они совпадают как отображения.

Пусть $\Gamma = \{\langle T_i, \Theta_i \rangle \mid i \in I\}$ – некоторое множество пар термов. Подстановка σ называется *унификатором* множества Γ, если $\sigma(T_i) = \sigma(\Theta_i)$ для всех $i \in I$. Подстановка σ называется *наиболее общим унификатором*, если для любого другого унификатора τ существует такая подстановка λ, что $\tau = \lambda \circ \sigma$.

Основные результаты доложены на заседании семинара 24 февраля 1997г.

Замечание 1. Иногда исходные пары термов рассматривают в виде системы равенств, и вместо задачи поиска унификатора ставят задачу поиска решения системы уравнений в термах

$$\begin{cases} T_1 = \Theta_1 \\ \quad \vdots \\ T_n = \Theta_n \end{cases}$$

В этом случае унификатор называют *решением* этой системы, а наиболее общий унификатор – *универсальным решением*.

Пусть v – вхождение подтерма t' в терм t. *Глубиной* v в t будем называть число таких вхождений функциональных знаков в терм t, в области действия которых находится данное вхождение терма t' (обозначим её $h_v[t]$). Если имеется единственное вхождение терма t' в t, или это вхождение фиксировано контекстом, мы будем просто писать $h_{t'}[t]$. *Разностью глубин* переменной x в терме t будем называть разность между наибольшей и наименьшей глубинами вхождений переменной x в t, обозначим ее δ_x. Высотой терма t называется максимум глубин входящих в него переменных и констант, обозначим ее $h[t]$.

Нас интересуют оценки высоты термов в наиболее общем унификаторе σ термов $\{\langle T_i, \Theta_i \rangle\}$ и их применение к оценкам высоты термов в доказательствах в исчислениях генценовского типа. Впервые эти оценки были применены к задаче определения увеличения сложности доказательства при устранении сечения в [5].

Наиболее простые оценки высоты термов в наиболее общем унификаторе могут быть найдены в [4]. В этой работе высоты термов в наиболее общем унификаторе оценивались в терминах высот исходных термов T_i, Θ_i и числа входящий в них переменных, а именно $h[\sigma(x)] \leqslant \mathrm{Var}[T_i, \Theta_i] \cdot h_0$, где h_0 – максимум высот термов T_i, Θ_i. Здесь и далее запись $\mathrm{Var}[T_i, \Theta_i]$ означает совокупное количество переменных термов T_i, Θ_i. Аналогичные оценки были независимо получены Крайчеком и Пудлаком [6].

Эти оценки можно улучшить, исключив из них те переменные, которые входят в множество термов Γ только на “нулевом уровне”, то есть только в пары вида $\langle x, t \rangle$, причем x не входит как подтерм ни в какой терм. Тогда оценка имеет вид: $h[\sigma(x)] \leqslant \mathrm{Var}_0[T_i, \Theta_i] \cdot h_0$, где Var_0 и обозначает число переменных термов T_i и Θ_i, входящих на ненулевом уровне (см. [5]).

После этого возникает вопрос, о возможности исключения из оценки переменных, входящих только на одном уровне, но это не нулевой уровень, то есть переменных, разность глубин для которых равна нулю.

Известны следующие оценки ([2, 3]):

$$(1)\ \ h[\sigma(x)] \leqslant \delta \cdot \rho^{\,|\,\mathrm{Var}\,|} \cdot |\,\mathrm{Var}\,| + h_0$$
$$(2)\ \ h[\sigma(x)] \leqslant \sum_{v\,\in\,\mathrm{Var}} h_v[T_i, \Theta_i] + h_0,$$

где δ – наибольшая разность глубин переменных в T_i и Θ_i, $h_v[T_i, \Theta_i]$ – максимум глубин вхождений переменной v в термы T_i и Θ_i, ρ – максимум числа вхождений переменных в T_i и Θ_i.

Замечание 2. Пусть нам дано множество Γ, состоящее из n пар термов. Пусть G – n-местный функциональный знак. Рассмотрим термы $< G(T_1, T_2, \ldots, T_n),\ G(\Theta_1, \Theta_2, \ldots, \Theta_n) >$. Они унифицируемы тогда и только тогда, когда унфицируемы все пары термов из множества Γ. Таким образом, если мы не накладываем ограничений на арность функциональных символов, можно ограничиться рассмотрением только двух термов T и Θ.

Целью работы являлось выяснение возможности улучшения оценки (1) и исключения из суммы в оценке (2) слагаемых с нулевой разностью глубин. В случае систем уравнений в термах с одноместными функциональными знаками оценка в терминах разностей глубин является линейной, $h[\sigma(x)] \leqslant \sum_{x \in \mathrm{Var}} \delta_x + h_0$. Кроме того, в задаче унификации рекурсивные выражения типа “пройтись по этой ветке несколько раз” не реализуются (x и $f(x)$ не унифицируются), поэтому казалось, что введение двуместных функциональных знаков не должно приводить к экспоненциальной зависимости высоты унификатора от разностей глубин неизвестных исходных термов. Для многих примеров, в оценке (2) суммирование достаточно производить только по переменным с ненулевой разностью глубин.

В данной статье строится пример, показывающий что оценка (1) является экспоненциальной и из суммы оценки (2) нельзя исключить члены с нулевой разностью глубин. Поскольку оказалось, что оценка (1) имеет экспоненциальный характер, она была несколько уточнена, а именно, из этой оценки была исключена зависимость от числа вхождений одной и той же переменной.

§1. Нижняя оценка

Рассмотрим термы

$$A_n(t) \rightleftharpoons g(\underbrace{f(f(\dots f}_{2^n \text{раз}}(t)\dots), x_n)),$$

$$B_n(t) \rightleftharpoons g(\underbrace{f(f(\dots f}_{2^n \text{раз}}(x_n)\dots), t)).$$

Здесь и далее

g – двуместный функциональный символ,
f – одноместный, функциональный символ,
x_n – переменная,
t – некоторый терм.

Лемма 1.1. *Если x_n не входит в термы t, t', то $A_n(t)$ и $B_n(t')$ унифицируемы тогда и только тогда, когда унифицируемы t и t'.*

Доказательство.

Тогда: пусть $A_n(t)$ и $B_n(t')$ унифицируемы, тогда существует подстановка σ такая, что

$$\sigma(g(\underbrace{f(f(\dots f}_{2^n \text{раз}}(t)\dots), x_n))) = \sigma(g(\underbrace{f(f(\dots f}_{2^n \text{ раз}}(x_n)\dots), t'))),$$

откуда $\sigma(f(f(\dots(t)\dots))) = \sigma(f(f(\dots(x_n)\dots)))$, следовательно, $\sigma(t) = \sigma(x_n)$. Аналогично, $\sigma(x_n) = \sigma(t')$. Окончательно, $\sigma(t) = \sigma(t')$.

Только тогда: пусть t, t' – унифицируемы. Тогда существует наиболее общий унификатор σ, такой что $\sigma(t) = \sigma(t')$. Так как x_n не входит в термы t, t', а σ – наиболее общий унификатор, $\sigma(x_n) = x_n$. Тогда доопределим $\sigma'(x_n) = \sigma(t)$. Из определений $A_n(t)$, $B_n(t)$ видно, что $\sigma'(A_n(t)) = \sigma'(B_n(t'))$. •

Непосредственно из определения $A_n(t)$, $B_n(t)$ мы имеем следующие значения глубин t, x_n в термах $A_n(t)$, $B_n(t)$:

$$\begin{aligned} h_t\ &[A_n(t)] = 2^n + 1, \\ h_{x_n}&[A_n(t)] = 1, \\ h_t\ &[B_n(t)] = 1, \\ h_{x_n}&[B_n(t)] = 2^n + 1. \end{aligned}$$

Для дальнейшего рассмотрения сформулируем следующее полезное наблюдение.

Лемма 1.2. *Пусть t' – подтерм терма t, а t'' – подтерм t'. Зафиксируем вхождение v терма t' в терм t и вхождение v' терма t'' в t', что фиксирует вхождение w терма t'' в t. Тогда $h_w[t] = h_v[t] + h_{v'}[t']$.*

Доказательство следует из определения глубины вхождения подтерма. $\bullet$

Обозначим через $t{\downarrow}^s_{t'}$ терм, полученный из терма t подстановкой $\sigma : s \to t'$.

Пусть s – некоторая неизвестная. Рассмотрим термы $\mathcal{A}_n$, $\mathcal{B}_n$

$$\begin{aligned} \mathcal{A}_0 &\leftrightharpoons A_0(s), & \mathcal{B}_0 &\leftrightharpoons B_0(s),\\ \mathcal{A}_{n+1} &\leftrightharpoons \mathcal{A}_n{\downarrow}^s_{A_{n+1}(s)}, & \mathcal{B}_{n+1} &\leftrightharpoons \mathcal{B}_n{\downarrow}^s_{B_{n+1}(s)}. \end{aligned}$$

По сути, выше приведена рекуррентная запись того факта, что

$$\mathcal{A}_n = A_0(A_1(\ldots(A_n(s)\ldots))) \text{ и } \mathcal{B}_n = B_0(B_1(\ldots(B_n(s)\ldots))).$$

Пусть t – некоторый терм. Обозначим тогда

$$\begin{aligned} \mathcal{A}_n[t] &\leftrightharpoons \mathcal{A}_n{\downarrow}^s_t,\\ \mathcal{B}_n[t] &\leftrightharpoons \mathcal{B}_n{\downarrow}^s_t. \end{aligned}$$

Лемма 1.3.

(1) *Если переменные $x_0, x_1, \ldots, x_n$ не входят в термы t, t', то терм $\mathcal{A}_n[t]$ унифицируем с термом $\mathcal{B}_n[t']$ тогда и только тогда, когда унифицируемы термы t и t'.*

(2) *Справедливы следующие равенства:*

$$\begin{aligned} h_{x_i}[\mathcal{A}_n[t]] &= 2^i + i, & 0 \leqslant i \leqslant n\\ h_t\ [\mathcal{A}_n[t]] &= 2^{n+1} + n, &\\ h_{x_i}[\mathcal{B}_n[t']] &= 2^i + i + 1, & 0 \leqslant i \leqslant n\\ h_{t'}[\mathcal{B}_n[t']] &= n + 1. & \end{aligned}$$

Доказательство.

(1) Утверждение доказывается индукцией по n:

База индукции: $n = 0$

Раскроем определение термов $\mathcal{A}_0[t]$ и $\mathcal{B}_0[t]$:

$$\mathcal{A}_0[t] = A_0(s)\downarrow^s_t = A(t),$$
$$\mathcal{B}_0[t'] = B_0(s)\downarrow^s_{t'} = B_0(t').$$

По Лемме 1.1, термы $\mathcal{A}_0[t]$ и $\mathcal{B}_0[t']$ унифицируемы тогда и только тогда, когда унифицируемы термы t и t'.

Индукционный переход:

по определению

$$\mathcal{A}_{n+1}[t] = \mathcal{A}_{n+1}\downarrow^s_t = (\mathcal{A}_n\downarrow^s_{A_{n+1}(s)})\downarrow^s_t =$$
$$= \mathcal{A}_n\downarrow^s_{A_{n+1}(t)} = \mathcal{A}_n[A_{n+1}(t)].$$

Аналогично, $\mathcal{B}_{n+1}[t'] = \mathcal{B}_n[B_n(t')]$.

Пусть термы $\mathcal{A}_{n+1}[t]$ и $\mathcal{B}_{n+1}[t]$ унифицируемы. Как было только что показано, это означает, что унифицируемы термы $\mathcal{A}_n[A_{n+1}(t)]$ и $\mathcal{B}_n[B_{n+1}(t')]$. Тогда, по индукционному предположению, унифицируемы термы $A_{n+1}(t)$ и $\mathcal{B}_{n+1}(t')$. Отсюда по Лемме 1.1 следует, что термы t и t' унифицируемы. Ясно, что это рассуждение обращается, то есть из унифицируемости термов t и t' следует унифицируемость термов $\mathcal{A}_{n+1}[t]$ и $\mathcal{B}_{n+1}[t']$.

(2) Индукция по n:

$$\mathcal{A}_0[t] = A_0(t) \implies \begin{cases} h_{x_0}[\mathcal{A}_0[t]] = 1 \\ h_t[\mathcal{A}_0[t]] = 2^0 + 1 = 2 \end{cases}$$

Так как $\mathcal{A}_{n+1}$ получается из $\mathcal{A}_n$ подстановкой вместо переменной терма, не содержащего $x_0, x_1, \ldots, x_n$, то $h_{x_i}[\mathcal{A}_{n+1}] = h_{x_i}[\mathcal{A}_n], \quad 0 \leqslant i \leqslant n$. По индукционному предположению и из свойств $A_n(t)$:

$$\begin{cases} h_s[\mathcal{A}_n] = h_t[\mathcal{A}_n[t]] = 2^{n+1} + n \\ h_t[A_{n+1}(t)] = 2^{n+1} + 1 \\ h_{x_n}[A_{n+1}(t)] = 1 \end{cases}$$

По определению,

$$\mathcal{A}_{n+1}[t] = \mathcal{A}_n\downarrow^s_{A_{n+1}(t)}.$$

По Лемме 1.2 мы имеем:

$$h_t[\mathcal{A}_{n+1}[t]] = 2^{n+1} + n + 2^{n+1} + 1 = 2^{(n+1)+1} + (n+1),$$
$$h_{x_n}[\mathcal{A}_{n+1}[t]] = 2^{n+1} + (n+1).$$

Доказательство утверждения для $\mathcal{B}_n[t']$ – аналогично. •

Пусть $C_n[t] \rightleftharpoons \underbrace{f(f(\ldots(f}_{n}(t)\ldots)))$.

Замечание. Из вида термов $C_n[x]$ и $C_{n+k}[y]$ ясно, что они унифицируемы, их унификатором является $\sigma : x \to C_k[y]$.

Окончательно, рассмотрим термы:

$$T = G(C_{2^{n+1}+n+2}[c], C_{2^{n+1}+n+2}[y_n], C_{2^{n+1}+n+2}[y_{n-1}], ..., C_{2^{n+1}+n+2}[y_1],$$
$$f(\mathcal{A}_n[F(y_1,\ldots,y_n)]), \mathcal{B}_n[F(z_1,\ldots,z_n)])$$
$$\Theta = G(C_{n+2}[z_n], C_{n+2}[z_{n-1}], \ldots, C_{n+2}[z_0], f(u), u),$$

где

c – константа,
f – одноместный функциональный знак,
g – двуместный,
F – n-местный,
G – $(n+3)$- местный,
u, $y_1, y_2, \ldots, y_n$, $z_0, z_1, \ldots, z_n$ – новые переменные.

Пусть v_i – вхождение неизвестного x_i в терм T, задаваемое вхождением x_i в $\mathcal{A}_n[F(y_1,\ldots,y_n)]$, w_i – вхождение x_i в терм T, задаваемое вхождением x_i в $\mathcal{B}_n[F(z_1,z_2,\ldots,z_n)]$. По Лемме 1.3, $h_{v_i}[T] = 2 + (2^i + i) = 1 + (2^i + i + 1) = h_{w_i}[T]$. Аналогично, из Леммы 1.3 и структуры термов T, Θ мы видим, что разности глубин всех переменных, кроме u, равны нулю. Разность глубин вхождений переменной u равна единице.

Замечание. Мы будем доказывать, что унифицируемость термов T и Θ эквивалентна унифицируемости некоторого множества пар термов Γ. Для упрощения записи, мы перейдем к записи множества пар унифицируемых термов как системы равенств.

Равенство $T = \Theta$ дает:

$$
\text{(i)} \begin{cases} z_n = C_{2^{n+1}}[c] \\ z_{n-1} = C_{2^{n+2}}[y_n] \\ \vdots \\ z_0 = C_{2^{n+2}}[y_1] \end{cases}
$$

$$
\mathcal{A}_n[F(y_1, y_2, \ldots, y_n)] = \mathcal{B}_n[F(z_1, z_2 \ldots, z_n)].
$$

Раскрывая последнее равенство, получим $F(y_1, y_2, \ldots, y_n) = F(z_1, z_2, \ldots, z_n)$. Откуда

$$
\text{(ii)} \begin{cases} y_1 = z_1 \\ y_2 = z_2 \\ \vdots \\ y_n = z_n \end{cases}
$$

Объединяя (i) и (ii), имеем:

$$
\begin{aligned}
z_n &= C_{2^{n+1}}[c] \\
&\vdots \\
z_1 &= \underbrace{C_{2^{n+1}}[C_{2^{n+1}}[\ldots C_{2^{n+1}}}_{n}[c]]] \\
z_0 &= \underbrace{C_{2^{n+1}}[C_{2^{n+1}}[\ldots C_{2^{n+1}}}_{n+1}[c]]].
\end{aligned}
$$

Таким образом,

$$
\begin{aligned}
h[\sigma(z_n)] &= 2^{n+1} \\
&\vdots \\
h[\sigma(z_1)] &= n \cdot 2^{n+1} \\
h[\sigma(z_0)] &= (n+1) \cdot 2^{n+1}.
\end{aligned}
$$

Поскольку, в терм T входят лишь переменные $x_0, x_1, \ldots, x_n$, $y_1, y_2, \ldots, y_n$, $z_0, z_1, \ldots, z_n$ – их глубины мы уже оценили и единственная константа c – её глубина равна $2^{n+1} + n + 3$, то высота терма T равна наибольшей из глубин входящих в T переменных и констант, $h_0 = h[y_i] = h[c] = 2^{n+1} + n + 3$. Поскольку $\sigma(u) = \sigma(\mathcal{A}_n[F(y_1, y_2, \ldots, y_n)])$, по Лемме 1.2 $h[\sigma(u)] = h[\sigma(y_1)] + h_{y_1}[T] = h[\sigma(z_1)] + h_0 = n \cdot 2^{n+1} + h_0$.

Тем самым доказана следующая теорема.

Теорема 1. *Существуют термы T, Θ, содержащие $3 \cdot n + 3$ переменных, такие что*

(1) *для $3 \cdot n + 2$ переменных разности глубин равны нулю,*
(2) $h[\sigma] > \delta \cdot n \cdot 2^{n+1} + h_0$,
(3) $h[\sigma] > \sum\limits_{\substack{v \in \mathrm{Var} \\ \delta_v \neq 0}} h_v[T, \Theta] + h_0$,

где $h[\sigma]$ – наибольшая высота термов, входящих в наиболее общий унификатор σ термов T, Θ.

§2. Верхняя оценка

Для удобства выберем древовидное представление термов, в котором каждому вхождению функциональных знаков и переменных соответствует вершина дерева, причем подтермы представляются соответствующими поддеревьями. При таком представлении глубина вхождения v терма t' в терм t есть длина пути от корня до представляющей t' вершины в дереве, представляющем t.

В [7] Патерсон и Вегман[1] показали, что понятие унификации тесно связано с понятием эквивалентности на вершинах графовых структур, представляющих термы. Для построения эффективного алгоритма унификации они представляли термы ориентированными мультиграфами без циклов. Единственное отличие таких мультиграфов от деревьев состоит в том, что переменные представлены в единственном экземпляре. Для ясности, рассмотрим на примере древовидное представление и представление мультиграфами терма $g(f(x, x), h(x, y))$.

[1]Аналогичные идеи можно найти в совместной работе Г.В. Давыдов, С.Ю. Маслов, Г.Е. Минц, В.П. Оревков и А.О. Слисенко [1].

Древовидное представление

Представление ациклическими мультиграфами

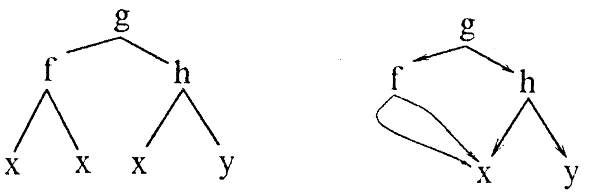

Рис. 1.

Пусть на вершинах графа задано некоторое отношение эквивалентности. Определим отношение $<^*$ на классах эквивалентности. Рассмотрим отношение $<'$, заданное на классах эквивалентности вершин следующим образом: если $x < y$ (то есть вершина x является предком вершины y в графе G), то $[x] <' [y]$ (как обычно, $[x]$ – класс эквивалентности вершины x). Пусть $<^*$ – транзитивное замыкание $<'$.

Отношение эквивалентности на вершинах мультиграфа называется *верным*, если оно удовлетворяет следующим условиям:

(1) если две функциональные вершины эквивалентны, то их соответствующие сыновья также попарно эквивалентны;

(2) каждый класс эквивалентности *однороден*, то есть не содержит функциональных вершин, помеченных различными функциональными знаками;

(3) Отношение $<^*$ является отношением строгого порядка на множестве классов эквивалентности вершин графа G.

Лемма 2.1. *Вершины u, v унифицируемы тогда и только тогда, когда существует верное отношение эквивалентности, такое что $u \sim v$. Кроме того, в случае унифицируемости, существует единственное минимальное верное отношение эквивалентности*

Доказательство. См. [7, Лемма 1, с. 160]. •

Построенное по унификотору σ верное отношение отношение эквивалентности Σ обладает следующим свойством: вершины, эквивалентные по отношению Σ переходят под действием σ в графически равные термы, причем наиболее общему унификатору соответствует минимальное отношение эквивалентности.

Как было сказано, для наших целей более подходят древовидные формы представления термов. Для того чтобы Лемма 2.1 и указанное свойство верного отношения остались верными для древесных структур, помимо условий (1), (2), (3) небходимо добавить условие

(0) вершины, помеченные одинаковыми переменными эквивалентны.

Докажем некоторое вспомогательное утверждение.

Лемма 2.2. *Пусть термы T и Θ унифицируемы. Рассмотрим минимальное верное отношение эквивалентности Σ на вершинах соответствующих этим термам деревьев, C – класс эквивалентных вершин по отношению Σ. Пусть δ_C есть разность глубин класса эквивалентости C, то есть наибольшая разность глубин входящих в класс C вершин. Тогда*

$$\delta_C \leqslant (2^{k+1} - 1)\delta, \tag{$*$}$$

где k – число переменных, таких что помеченные ими вершины либо лежат в C, либо принадлежат классам меньшим C по отношению $<^*$, то есть $k = \#\{z \in \mathrm{Var}(T, \Theta) | v \leqslant^* C,\ v$ – вхождение z в T или $\Theta\}$, δ – наибольшая разность глубин переменных, входящий в T и Θ.

Доказательство. Проведем индукцию по отношению порядка $<^*$ на классах эквивалентности. Разность глубин корней, представляющих T и Θ, равна нулю, $k \geqslant 0$, следовательно, для минимального класса условие $(*)$ выполнено.

Пусть оценка $(*)$ выполнена для всех классов эквивалентности, меньших класса C по отношению $<^*$. Разберем для начала несколько случаев.

Если класс C не содержит переменных (см. Рис. 2), то вершины этого класса могли оказаться эквивалентными исключительно в силу условия (1) верного отношения эквивалентности, то есть имеется такой класс C', что каждая вершина класса C есть сын некоторой вершины класса C'. То есть, разность глубин класса C равна разности глубин класса C'. Число переменных, вхождения которых меньше либо равны класса S, не изменилось.

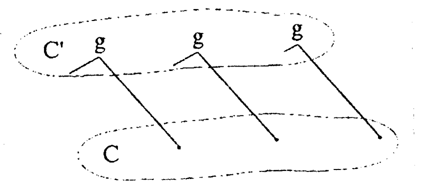

Рис. 2.

Рассмотрим теперь класс C, в который входит лишь одна переменная x (рис. 3). В силу условия (0) такой класс может состоять из потомков нескольких классов эквивалентности. Пусть C_1', $C_2', \ldots$, C_n' – его *непосредственные* по отношению $<^*$ предшественники, k_i число переменных, меньших класса C_i'. Пусть C_i – множество вершин класса C, чьи отцы принадлежат C_i', тогда $\cup_i C_i = C$.

Вершины множеств $C_1, \ldots, C_n$ могут оказаться эквивалентными только в силу условия (0) верного отношения эквивалентности, но это означает, что каждое множество вершин C_i содержит вхождение переменной x.

Пусть u, v – две вершины класса C. Если u, $v \in C_i$ для некоторого i, то утверждение (*) доказывается, также как в предыдущем случае, поскольку $k \geqslant k_i$.

Рассмотрим случай, когда $u \in C_l$, $v \in C_m$.

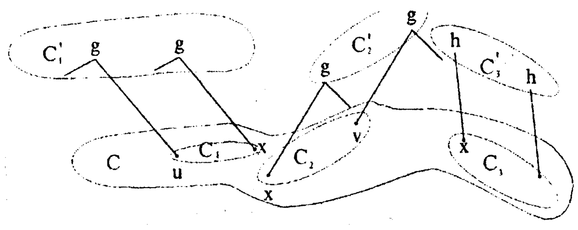

Рис. 3.

Разность высот u и v не больше чем разность высот u и вхождения x в C_l плюс разность высот v и вхождения x в C_m плюс разность высот вхождений x в C_l и C_m. Разность высот u и вхождения x в C_l равняется разности высот вхождений их родителей в класс C'_l, аналогично для класса C_m. Таким образом,

$$|h_u - h_v| \leqslant \delta_{C'_l} + \delta_{C'_m} + \delta_x \leqslant ((2^{k_l+1} - 1) + (2^{k_m+1} - 1) + 1)\delta.$$

Но так как $<^*$ – отношение строгого порядка, то x не меньше C_i, следовательно, $k \geqslant k_i + 1$. Поэтому,

$$|h_u - h_v| \leqslant (2^k - 1 + 2^k - 1 + 1)\delta = (2^{k+1} - 1)\delta.$$

Пусть класс C содержит две переменных x и y (Рис. 4). Аналогично предыдущему случаю рассмотрим классы непосредственных предшественников $C'_1, C'_2, \ldots, C'_n$ и разбиение C на множества вершин $C_1, C_2, \ldots, C_n$. Для того, чтобы вершины из множеств C_1, $C_2, \ldots, C_n$ оказались эквивалентными, необходимо, чтобы все C_i содержали вхождения x или y и, кроме того, чтобы было хотя бы одно множество C_p, содержащее и вхождение x, и вхождение y.

Пусть $u \in C_l$, $x \in C_l$, $v \in C_m$, $y \in C_m$, тогда

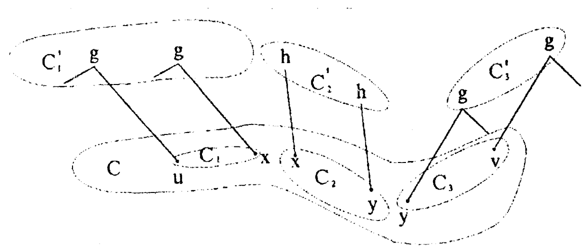

Рис. 4.

$$|h_u - h_v| \leqslant \delta_{C'_l} + \delta_{C'_p} + \delta_{C'_m} + \delta_x + \delta_y \leqslant$$
$$\leqslant (2^{k_l+1} - 1 + 2^{k_p+1} - 1 + 2^{k_m+1} - 1)\delta + \delta_x + \delta_y.$$

Так как x и y не меньше C_i, то $k \geqslant k_i + 2$. Тогда,

$$|h_u - h_v| \leqslant 3(2^{k-1} - 1)\delta + 2\delta \leqslant 4(2^{k-1} - 1)\delta + 2\delta \leqslant$$
$$\leqslant (2^{k+1} - 1)\delta.$$

Случаи, когда и u, и v принадлежат некоторому C_j, либо $u \in C_l$, $v \in C_m$, и множества C_l, C_m оба содержат вхождение x соответствуют разобраным ранее.

Рассуждая аналогично для случая, когда C содержит r переменных, мы имеем:

$$|h_u - h_v| \leqslant (r+1)(2^{k-r+1}-1)\delta + r\delta.$$

Так как $r+1 \leqslant 2^r$,

$$|h_u - h_v| \leqslant 2^r(2^{k-r+1}-1)\delta + r\delta = (2^{k+1} - 2^r + r)\delta \leqslant (2^{k+1}-1)\delta.$$

•

Теорема 2. *Каковы бы ни были термы T и Θ, если σ является наиболее общим унификатором термов T и Θ, то*

$$h[\sigma[x]] \leqslant (2^{|\operatorname{Var}[T,\Theta]|+2} - |\operatorname{Var}[T,\Theta]| - 3)\delta + h_0$$

Доказательство. Построим по наиболее общему унификатору σ минимальное верное отношение эквивалентности на вершинах деревьев, представляющих термы T и Θ. Это отношение обладает тем свойством, что термы, соответствующие эквивалентным вершинам, переходят под действием унификатора σ в один и тот же терм. Это свойство позволяет написать унификатор в его компактной записи,

$$\begin{array}{l} x_1 \to t_1[x_{j_{1,1}}, x_{j_{1,2}}, \ldots, x_{j_{1,i_1}}] \\ x_2 \to t_2[x_{j_{2,1}}, x_{j_{2,2}}, \ldots, x_{j_{2,i_2}}] \\ \quad\vdots \\ x_n \to t_n[x_{j_{n,1}}, x_{j_{n,2}}, \ldots, x_{j_{n,i_n}}]. \end{array}$$

где все t_i, – либо сами x_i, либо некоторые подтермы термов T или Θ, соответствующие эквивалентной x_i функциональной вершине v_i.

Для того, чтобы получить унификатор в его стандартной записи, например, для x_1, надо последовательно подставить вместо $x_{j_{1,1}}, x_{j_{1,2}}, \ldots, x_{j_{1,i_1}}$ их значения в указанной записи. Мы получим терм, в который входят новые переменные. Продолжим подставлять вместо переменных значение до тех пор, пока мы не получим такой терм, что в него не входят переменные, эквивалентные функциональным вершинам. Свойство (3) верного отношения эквивалентности гарантирует то, что у нас не возникнет цикла. В итоге мы получим унификатор в стандартной записи. Эту идею поясняет рис. 5. На нем пунктирные линии соединяют эквивалентные вершины представляющий термы T, Θ деревьев.

Компактное представление унификатора — Стандартное представление

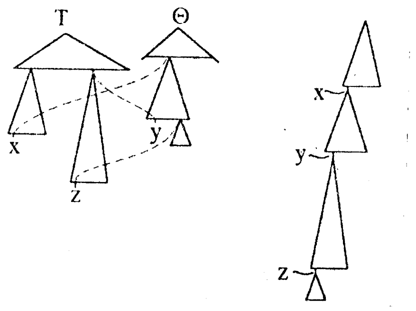

Рис. 5.

Рассмотрим путь от корня до вершины в дереве, представляющем $\sigma(x_i)$, длина которого равна высоте терма $\sigma(x_i)$. Как было сказано, этот путь складывается из пути в терме t_i от вершины до листа, помеченного некоторым x_i', затем из пути в терме t_i' от корня до некоторой вершины x_i'' и т.д. до тех пор, пока мы не придем в вершину, помеченную 0-местным функциональным знаком или переменной, не эквивалентной никакой функциональной вершине. Для краткости, назовем такую вершину *окончательным листом*.

Заметим теперь, что термы t_i, t_i', t_i'' и т.д. являются подтермами термов T или Θ, а значит они все представлены вершинами деревьев, представляющих термы T, Θ. Терму t_i соответствует вершина v_i, мы проходим от нее путь до листа помеченного x_i'. Терму t_i' соответствует вершина v_i'. Мы как бы перескакиваем от x_i' до эквивалентной вершины v_i' и продолжаем путь до листа, помеченного x_i''. Такие прыжки как раз соответствуют пунктирным линиям на Рис. 5. Наши перемещения имеют следующий вид (см. Рис. 6).

Отсюда ясно, что длина пройденного пути не превышает высоты исходных деревьев плюс суммарную длину прыжков.

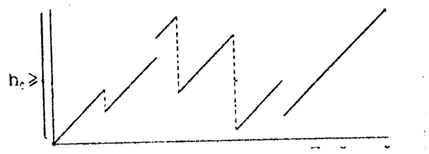

Рис. 6.

Для формального доказательства проведем индукцию по числу прыжков. Рассмотрим произвольную вершину v деревьев, представляющих термы T, Θ. Этой вершине соответствует некий терм t. Обозначим Δ_t суммарную длину прыжков на пути от v к окончательному листу с самой большой глубиной. Терм t переходит под действием σ в терм $\sigma(t)$. Индукционное предположение состоит в том, что высота терма $\sigma(t)$ не превосходит слудующего выражения:

$$h[\sigma(t)] \leqslant (h_0 - h_t) + \Delta_t.$$

Напомним, что h_t есть глубина вхождения терма t в наши деревья, $h[\sigma(t)]$ – высота терма $\sigma(t)$, h_0 – максимум высот термов T, Θ.

Если на пути от вершины v в окончательный лист нет прыжков, утверждение очевидно верно. Пусть мы уже доказали утверждение для случая n прыжков, на пути от v к окончательному листу имеется $n+1$ прыжок. Рассмотрим первый из них. Мы проходим путь от вершины v до некоторого листа, помеченного переменной x, а затем совершаем прыжок к некоторой вершине u, помеченной функциональным знаком. Этой вершине соответствует терм t'. Поскольку на пути от вершины v к окончательному листу мы совершаем $n+1$ прыжок, на пути от u прыжков будет на единицу меньше, то есть n. Путь от v к окончательному листу складывается из кусочка пути от v к листу, помеченному x и пути от u до окончательного листа, $h[\sigma(t)] = h_x - h_t + h[\sigma(t')]$. По индукционному предположению, $h[\sigma(t') \leqslant (h_{t'} - h_0) + \Delta_{t'}$. Отсюда,

$$\begin{aligned} h[\sigma(t)] &\leqslant h_x - h_t + (h_{t'} - h_0) + \Delta_{t'} = (h_x - h_0) + (h_{t'} - h_t) + \Delta_{t'} \\ &\leqslant (h_x - h_0) + |h_{t'} - h_t| + \Delta_{t'} = (h_x - h_0) + \Delta_t \end{aligned}$$

Остается оценить длину прыжков. Исходя из Леммы 2.2 прыжок от вершины, помеченной x_i' к вершине, помеченной t_i' оценивается так: $\delta_{[x]} \leqslant (2^{k_x+1} - 1)\delta$, где k_x обозначает число переменных меньших или равных класса эквивалентности $[x]$. Прыжок от вершины, помеченной x_i'' до вершины, помеченной t_i'' аналогично оценивается исходя из $k_{x'}$ – числа переменных меньших или равных класса эквивалентности $[x']$. Так как класс $[x_i']$ меньше класса $[x_i'']$, $k_{x'} \geqslant k_x + 1$. То есть оценивая прыжки по Лемме 2.2, мы получаем, что оценка слудующего прыжка по крайней мере в два раза больше предыдущего. Но так как циклов нет, мы не можем совершить более $\mathrm{Var}[T, \Theta]$ прыжков. Таким образом, прыжки составляют фрагмент геометрической прогресии. Оценим их суммарную длину, как сумму всей геометрический прогрессии от 1-го до $|\mathrm{Var}|$-того члена. Максимальный прыжок составляет $(2^{|\mathrm{Var}[T,\Theta]|+1} - 1)\delta$, тогда суммарная длина всех прыжков оценивается как

$$(2^1 - 1)\delta + \cdots + (2^{|\mathrm{Var}[T,\Theta]|+1} - 1)\delta \leqslant (2^{|\mathrm{Var}[T,\Theta]|+2} - |\mathrm{Var}[T, \Theta]| - 3)\delta$$

Таким образом мы получаем требуемое утверждение. •

Литература

1. Г. В. Давыдов, С. Ю. Маслов, Г. Е. Минц, В. П. Оревков и А. О. Слисенко, *Машинный алгоритм установления выводимости на основе обратного метода.* — Зап. научн. семин. ЛОМИ **16** (1969), 8–15.
2. В. П. Оревков, *Одна оценка сложности термов в доказательствах.* — В кн.: Семиотические аспекты формализации интелектуальной деятельности. Всесоюзная школа-семинар “Боржоми – 88”. Тезисы докладов и сообщений.. 1988, с. 108–110.
3. В. П. Оревков, *Complexity of proofs and their transformations in axiomatic theories.* — Amer. Math. Soc., translations of mathematical monographs **128** (1991).
4. В. П. Оревков, *Алгоритм Британского музея может быть эффективнее метода резолюций..* — В кн.: Ч. Чень, Р. Ли, Математическая логика и автоматическое доказательство теорем, дополнение Б.. 1983, с. 314–325.
5. В. П. Оревков, *Нижние оценки увеличения сложности выводов после устранения сечений.* — Зап. научн. семин. ЛОМИ **88** (1979), 137–162.
6. J. Krajíček and P. Pudlák, *The number of proof lines and size of proofs in first order logic.* — Arch. Math. Logic **27** (1988), 69–84.
7. M. S. Paterson and M. N. Wegman, *Linear unification.* — Journal of computer and system science **16** (1978), 158–167.

Konev B. Yu. Refinement of bounds of the heights of terms in the most general unifer.

We consider the most general unifer of a pair of terms. Our goal is to estimate the heights of terms in the unifer respect to the depth differences of variables in given terms. The depth difference of a variable is the maximum of the differences between the depths of occurrences of the variable in the terms. Such bounds are useful in finding upper bounds of the heights of terms in proofs in Gentzen-type sequential calculi. We construct a series of examples proving an exponential lower bound on the heights of terms in the most general unifer. We also improve the previously known upper bound of heights in the most general unifer.

Санкт-Петербургское отделение
Математичекого института
им. В. А. Стеклова РАН

Поступило 20 апреля 1997 г.