Труды Петрозаводского государственного университета
Серия *"Прикладная математика и информатика"* Выпуск 7, 1998

УДК 519.7, 511.512, 512.644

Решение одного класса линейных диофантовых уравнений в неотрицательных целых числах методами теории формальных языков

Д. Ж. Корзун

На примере конкретного класса линейных диофантовых уравнений демонстрируется использование метода порождающей грамматики для их решения в неотрицательных целых числах. При этом получены формулы, описывающее множество всех решений уравнений этого класса, независимо от того, является ли оно однородным или неоднородным.

§1. Введение

В данной работе мы хотим на конкретном примере показать технику использования предложенного в [1] и развитого в [2] метода порождающей грамматики. Этот метод предназначен для исследования некоторых систем линейных диофантовых уравнений (ЛДУ) в неотрицательных целых числах.

Решение линейных диофантовых уравнений или их систем в неотрицательных целых числах является NP-полной задачей [3, 4, 5]. Большинство существующих методов для решения линейных диофантовых уравнений основаны на решении однородного уравнения, к которому можно свести любое неоднородное уравнение. Однако после

нахождения решений однородного уравнения необходимо произвести обратный переход к неоднородному, что, как правило, реализуется с помощью переборных алгоритмов, которые могут потребовать значительных затрат времени, если множество решений однородного уравнения велико.

Решение однородного уравнения [3, 5, 8] сводится к поиску конечного множества минимальных решений, которое называют базисным, ибо любое решение этого уравнения может быть выражено в виде неотрицательной целой линейной комбинации базисных и наоборот. Однако известные способы вычисления базиса довольно медленны и число шагов алгоритма экспоненциально зависит от абсолютной величины коэффициентов, входящих в уравнение. Описание и сравнение наиболее популярных среди известных на данный момент методов можно найти в [6–11].

В данной работе предлагается новый метод решения специального класса диофантовых уравнений:

$$\sum_{i=1}^{n} x_i + b_1 = \sum_{i=1}^{n} a_i x_i + b_2 \,, \quad a_i \geqslant 0, \;\; b_1, b_2 \geqslant 0 \,,$$

для которого множество решений может быть описано довольно просто, независимо от того, является ли рассматриваемое уравнение однородным или неоднородным.

Используемая здесь терминология и некоторые факты теории формальных языков заимствованы из [12].

§2. Уравнение вида $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} a_i x_i + b$

Рассматривается уравнение

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} a_i x_i + b \,. \tag{1}$$

Предполагаем, что $a_i \geqslant 0$ $(i = 1, 2, \ldots, n)$, $b \geqslant 0$. Также не будем рассматривать случай, когда $a_k = 1$ для некоторого k, поскольку, вычитая из обеих частей уравнения (1) величину x_k, приходим к новому

уравнению уже с $n-1$ переменной (не содержащему переменной x_k). Решение исходной задачи получается из решения новой задачи, а x_k при этом полагаем быть равным произвольному неотрицательному целому числу.

Таким образом, уравнение, в котором есть $a_k = 1$, всегда можно свести к уравнению вида (1), в котором все $a_i > 1$.

Введем новую переменную y и перепишем уравнение (1) в виде эквивалентной системы:

$$\begin{cases} \sum\limits_{i=1}^{n} x_i = \sum\limits_{i=1}^{n} a_i x_i + by \\ y = 1 \end{cases} \tag{2}$$

Система (2) является ассоциированной[1] для грамматики $G = (\{S,A\},\varnothing,P,S)$ и пустой цепочки $x = e$. При этом множество правил P задается следующим образом:

$$\begin{array}{lcl} y & : & S \longrightarrow A^b \\ x_1 & : & A \longrightarrow A^{a_1} \\ x_2 & : & A \longrightarrow A^{a_2} \\ & \ldots & \\ x_n & : & A \longrightarrow A^{a_n} \end{array} \tag{3}$$

Здесь в левом столбце указаны переменные из (2), соответствующие правилам, указанным в правом столбце (значение переменной равняется числу применений этого правила в выводе цепочки $x = e$).

Если $b > 0$ и $a_i > 0\ \forall i$, то язык грамматики пуст ($L[G] = \varnothing$), поскольку из начального нетерминала S нельзя вывести ни одной терминальной цепочки (в нашем случае — пустой цепочки). Отсюда следует, что в этом случае система (2) неразрешима в неотрицательных целых числах, а значит, и уравнение (1) не имеет решений.

Если $b = 0$ и $a_i > 0\ \forall i$, то $L[G] = \{e\}$, поскольку 1-е правило грамматики при этих условиях имеет вид $S \longrightarrow e$. Из элементарных деревьев для правил $A \longrightarrow A^{a_i}$ нельзя составить уравновешенного леса, ибо тогда обязательно $a_i > 1$ и число появлений нетерминала A в кроне всегда будет превосходить число его появлений среди корней.

[1] Иногда мы будем использовать для выражения этого факта другую фразу: “грамматика порождает данную систему ЛДУ”

По теореме об общем решении для ассоциированной системы (теорема 5 из [2]) любое решение системы (2) имеет вид:

$$\vec{\xi} = (y, x_1, \ldots, x_n)^\top = \vec{\xi}_0 \ ,$$

где $\vec{\xi}_0$ — вектор, определяемый деревом разбора пустой цепочки в грамматике G. Поскольку такой вывод всего один, а именно $S \Rightarrow e$ (применяется только 1-е правило $S \longrightarrow e$), то вектор $\vec{\xi}_0$ находится единственным образом:

$$\vec{\xi}_0 = (1, 0, \ldots, 0)^\top \ ,$$

а значит, уравнение (1) в этом случае имеет лишь нулевое решение $x_i = 0$, $i = 1, 2, \ldots, n$.

Пусть теперь $b > 0$ и некоторые (но не все) a_i равны 0. Ясно, что в этом случае всегда можно соответствующей перенумерацией переменных получить, что $a_1 = a_2 = \ldots = a_m = 0$, $a_{m+1} > 1$, $\ldots$, $a_n > 1$ ($0 < m < n$). Тогда множество правил порождающей грамматики G имеет вид:

$$\begin{array}{lcl} y & : & S \longrightarrow A^b \\ x_1 & : & A \longrightarrow e \\ & \ldots & \\ x_m & : & A \longrightarrow e \\ x_{m+1} & : & A \longrightarrow A^{a_{m+1}} \\ & \ldots & \\ x_n & : & A \longrightarrow A^{a_n} \end{array} \tag{4}$$

Заметим, что несмотря на то, что здесь имеется по меньшей мере m одинаковых правил (соответствуют переменным $x_1, \ldots x_m$), их нельзя объединять в одно, поскольку они соответствуют различным переменным.

Любой вывод пустой цепочки $x = e$ имеет дерево разбора вида, показанного на рис. 1. При развертывании нетерминала A мы можем применить любое из правил, соответствующих переменным x_1, $\ldots$, x_m. Всего нетерминалов A в этом дереве — b штук, следовательно, число всех возможных таких деревьев равняется количеству неотрицательных целых решений уравнения

$$k_1 + k_2 + \ldots + k_m = b \ . \tag{5}$$

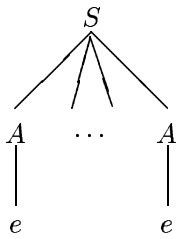


Рис. 1. Общий вид дерева разбора пустой цепочки в порождающей грамматике G

Обозначим это число как M_0.

Этим деревьям соответствует M_0 векторов:

$$\vec{\xi}_0^{\,j} = (1, k_1, \ldots, k_m, 0, \ldots, 0)^\top \quad , \quad k_1 + \ldots + k_m = b,\ j = 1, \ldots, M_0 \ .$$

Далее построим все возможные чисто циклические деревья для этой грамматики. Заметим, что уравновешенных лесов циклических деревьев с непустой кроной здесь не может быть, ибо $\Sigma = \varnothing$. Любое простейшее циклическое дерево в данном случае имеет вид, показанный на рис. 2. В качестве первого правила может быть выбрано лю-

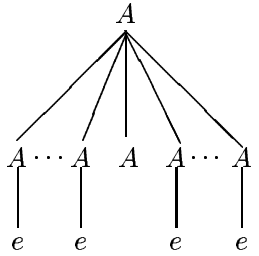


Рис. 2. Простейшее дерево разбора пустой цепочки

бое из соответствующих переменным x_{m+1}, ..., x_n. Всего — $(n-m)$ вариантов. В результате появляются a_l нетерминалов (l — номер выбранного первого правила для дерева), которые затем раскрываются в

пустую цепочку, кроме одного (дерево циклическое). Для этого можно воспользоваться лишь правилами, соответствующими переменным x_1, ..., x_m. Таким образом, если первое правило в дереве зафиксировано (пусть, для определенности, соответствующее x_l), то при этом условии число таких чисто циклических деревьев будет равняться числу неотрицательных целых решений уравнения

$$k_1 + k_2 + \ldots + k_m = a_l - 1 \ . \tag{6}$$

Обозначим это число через M_l.

Этим чисто циклическим деревьям соответствуют векторы

$$\vec{\xi}_c^{\,lr} = (0, k_1, \ldots, k_m, 0, \ldots, 1, \ldots, 0)^\top \ , \quad \begin{array}{l} k_1 + \ldots + k_m = a_l - 1 \ , \\ l = m+1, \ldots, n \ , \\ r = 1, \ldots, M_l \ , \end{array}$$

где 1 стоит на месте, соответствующем переменной x_l.

Согласно теореме об общем решении ассоциированной системы, получаем формулу для общего решения системы (2):

$$\vec{\xi} = \vec{\xi}_0^{\,j} + \sum_{l=m+1}^{n} \sum_{r=1}^{M_l} p_{lr} \vec{\xi}_c^{\,lr} \quad , \quad p_{lr} \geqslant 0 \ . \tag{7}$$

Формула (7) и определяет все решения системы (2), а значит, и уравнения (1).

Пример 1. Рассмотрим уравнение $x+y-2z=2$. Оно эквивалентно уравнению $x+y+z=3z+2$ типа (1). Здесь $n=3$, $m=2$, $a_3=3$, $b=2$. $M_0=3$ и векторы $\vec{\xi}_0^{\,j}$ $(j=1,2,3)$ имеют вид:

$$\vec{\xi}_0^{\,1} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 0 \end{pmatrix} \ , \quad \vec{\xi}_0^{\,2} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \ , \quad \vec{\xi}_0^{\,3} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix} .$$

$M_3=3$ и векторы $\vec{\xi}_c^{\,3r}$ $(r=1,2,3)$ имеют вид:

$$\vec{\xi}_0^{\,31} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \ , \quad \vec{\xi}_0^{\,32} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \ , \quad \vec{\xi}_0^{\,33} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix} .$$

Следовательно, любое решение обязательно содержится в одной из следующих последовательностей:

$$\begin{pmatrix} 2+2p_1+p_2 \\ p_2+2p_3 \\ p_1+p_2+p_3 \end{pmatrix} \ , \quad \begin{pmatrix} 1+2p_1+p_2 \\ 1+p_2+2p_3 \\ p_1+p_2+p_3 \end{pmatrix} \ , \quad \begin{pmatrix} 2p_1+p_2 \\ 2+p_2+2p_3 \\ p_1+p_2+p_3 \end{pmatrix} \ ,$$

$$p_r \geqslant 0 \ , \ r = 1,2,3.$$

В силу предыдущих рассуждений верна следующая теорема.

Теорема 1. *Вектор $\vec{x}$ является решением уравнения (1) тогда и только тогда, когда он имеет вид*

$$\vec{x} = \vec{x}_0^j + \vec{f}(\mathbb{P}) \quad \textit{для некоторого } j \in \{1,\ldots,M_0\} \ , \tag{8}$$

где

$$\vec{x}_0^j = (k_1,k_2,\ldots,k_m,0,\ldots,0)^\top \ , \quad \begin{array}{l} k_1+k_2+\ldots+k_m = b \ , \\ j = 1,2,\ldots,M_0 \ , \end{array} \tag{9}$$

$$\vec{f}(\mathbb{P}) = \sum_{l=m+1}^{n} \sum_{r=1}^{M_l} p_{lr}\vec{x}_c^{lr} \ , \quad \mathbb{P} = (p_{lr}) \ , \ p_{lr} \geqslant 0 \ , \tag{10}$$

и

$$\vec{x}_c^{lr} = (k_1,\ldots,k_m,0,\ldots,1,\ldots,0)^\top \ , \quad \begin{array}{l} k_1+\ldots+k_m = a_l - 1 \ , \\ l = m+1,\ldots,n \ , \\ r = 1,\ldots,M_l \\ \textit{и } 1 \textit{ стоит на } l\textit{-м месте.} \end{array} \tag{11}$$

Замечание. Между множеством решений уравнения (1) и множеством пар вида $(j,\mathbb{P})$, вообще говоря, нет взаимно однозначного соответствия. Это означает, что могут существовать две различные пары $(j_1,\mathbb{P}_1)$ и $(j_2,\mathbb{P}_2)$, которые порождают одно и то же решение по формуле (8). Например, одно из решений $(x,y,z)^\top = (2,2,2)^\top$ однородного уравнения $x+y=2z$ может быть представлено двумя способами:

$$\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$$

Согласно теореме 1 решение уравнения (1) складывается из двух компонент. Первая — это вектор $\vec{x}_0^j$ для некоторого j из $\{1,\ldots,M_0\}$. Число таких векторов конечно и равняется M_0. Вторая компонента решения — $\vec{f}(\mathbb{P})$, где $\mathbb{P}$ — произвольная матрица с неотрицательными целыми компонентами p_{lr} $(r=1,\ldots,M_l,\ l=m+1,\ldots,n)$, а поэтому может принимать бесконечное число значений. Однако $\vec{f}(\mathbb{P})$ имеет вид неотрицательной линейной комбинации конечного числа векторов $\vec{x}_c^{lr}$ с коэффициентами из элементов матрицы $\mathbb{P}$. Отсюда сразу получаем, что $\vec{f}(\mathbb{P}) = \vec{0} \Leftrightarrow \mathbb{P} = \mathbb{O}$ и $\vec{f}(\mathbb{P}) > \vec{0} \Leftrightarrow \mathbb{P} \neq \mathbb{O}$.

Обозначим $H = \{\vec{x}_0^j \mid j = 1,\ldots,M_0\}$ и $B = \{\vec{x}_c^{lr} \mid r = 1,\ldots,M_l, l = m+1,\ldots,n\}$.

Теорема 2. *Пусть $\vec{x}$ — произвольный вектор из H. Тогда $\vec{x}$ есть решение уравнения (1).*

Доказательство. Достаточно положить в (8) $\mathbb{P} = \mathbb{O}$. □

Рассмотрим однородное уравнение для (1):

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} a_i x_i \ . \tag{12}$$

Теорема 3. *Пусть $\vec{x}$ — произвольный вектор из B. Тогда $\vec{x}$ есть решение однородного уравнения (12). Более того, любая неотрицательная линейная комбинация векторов из B является решением уравнения (12).*

Доказательство. Достаточно посмотреть на формулу (8) при $b=0$, поскольку в этом случае $H = \varnothing$. □

Пусть S — множество всех решений уравнения (1). Обозначим как $L(B)$ неотрицательную линейную комбинацию векторов множества B. Суммируя результаты предыдущих теорем, получаем:

Теорема 4. *Общее решение уравнения (1) определяется следующей формулой:*

$$S = H + L(B) \ . \tag{13}$$

Замечания

1. Если бы решения искались в кольце (например, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$), то для определения общего решения неоднородного уравнения по известному общему решению однородного уравнения достаточно всего лишь одного частного решения. В случае неотрицательных целых требуется, вообще говоря, большее, хотя и конечное, их число.

2. Если $m = n$, то есть $a_i = 0$ $(i = 1, \ldots, n)$, то уравнение (1) принимает вид
$$\sum_{i=1}^{n} x_i = b\,. \tag{14}$$
В этом случае $B = \varnothing$, то есть базисных решений однородного уравнения нет, и множество H содержит в себе все решения, число которых в этом случае конечно.

3. Если $m < n$, то базисное множество для однородного уравнения всегда непусто и уравнение имеет бесконечное число решений.

Элементы множеств H и B удовлетворяют так называемому свойству минимальности, которое может быть сформулировано следующим образом:

Теорема 5. *Пусть вектор $\vec{x}$ — произвольное решение неоднородного уравнения (1), а $\vec{x}_c$ — произвольное ненулевое решение однородного уравнения (12). Тогда*[2]

(a) Найдется $\vec{y} \in H$ такой, что $\vec{y} \leqslant \vec{x}$.

(b) Для всех $\vec{y} \in H$ неверно, что $\vec{y} \geqslant \vec{x}$, если только $\vec{y} \neq \vec{x}$.

(c) Найдется $\vec{y} \in B$ такой, что $\vec{y} \leqslant \vec{x}_c$.

(d) Для всех $\vec{y} \in B$ неверно, что $\vec{y} \geqslant \vec{x}_0$, если только $\vec{y} \neq \vec{x}_0$.

Доказательство. Во-первых, если $\vec{y}^1$ и $\vec{y}^2$ — элементы из H, то они несравнимы между собой, поскольку если допустить, например, что $\vec{y}^1 \geqslant \vec{y}^2$, то это значит, согласно (9), что существует 2 различных

[2]Здесь отношения $\geqslant$ и $\leqslant$ для векторов определяются через соответствующее сравнение всех их компонент: $\vec{x} \geqslant \vec{y}$, если $x_i \geqslant y_i$ $\forall i$. Эти отношения порождают частичный порядок, а поэтому не все элементы сравнимы друг с другом.

набора неотрицательных целых чисел $(k_1^1, \ldots, k_m^1)$ и $(k_1^2, \ldots, k_m^2)$ таких, что $k_1^1 + \ldots + k_m^1 = k_1^2 + \ldots + k_m^2 = b$ и $k_1^1 \geqslant k_1^2$, $\ldots$, $k_m^1 \geqslant k_m^2$. Но это невозможно. Аналогичным образом доказывается, что элементы из B также несравнимы между собой.

Далее, $\vec{x} = \vec{y}' + \vec{f}$, где $\vec{y}' \in H$. Ясно, что $\vec{x} \geqslant \vec{y}'$, тем самым доказано утверждение (a).

Предположим теперь, что существует $\vec{y} \in H$ такой, что $\vec{y} > \vec{x}$. Но тогда $\vec{y} > \vec{y}' + \vec{f}$ и тем более $\vec{y} > \vec{y}'$, что невозможно в силу несравнимости векторов из H, что доказывает справедливость утверждения (b).

Так как $\vec{x}_c$ — решение (12), то $\vec{x}_c = \sum\limits_{\vec{y} \in B} p_y \vec{y}$. Поскольку $\vec{x}_c \neq \vec{0}$, то среди неотрицательных чисел p_y есть хотя бы одно отличное от 0. Соответствующий этому числу вектор $\vec{y} \in B$ и есть искомый для утверждения (c).

Для доказательства утверждения (d) предположим существование вектора $\vec{y} \in B$ такого, что $\vec{y} > \vec{x}_0$. Это означает, что существует $\vec{y}' \in B$ и $\vec{y} > \vec{y}'$, поскольку $\vec{x}_0 = \sum\limits_{\vec{y}' \in B} p_{y'} \vec{y}'$, что противоречит несравнимости элементов множества B. □

Следующая теорема позволяет нам не беспокоиться о том, каким способом получены множества H и B, определяющие общее решение уравнения (1). В частности, всегда можно получать их с помощью формул (9) и (11).

Теорема 6. *Пусть множества H и B удовлетворяют свойству минимальности и удовлетворяют равенству $S = H + L(B)$, где S — общее решение (1). Тогда эти множества определяются единственным образом.*

Доказательство. Предположим, что существуют H_1, B_1 и H_2, B_2 такие, что для них выполнены условия теоремы.

Пусть $\vec{h} \in H_1$ и $\vec{h} \notin H_2$. Для вектора $\vec{h}$ имеем разложение:
$$\vec{h} = \vec{h}_2 + \sum_{y \in B_2} p_y \vec{y} \quad , \quad \vec{h}_2 \in H_2\,.$$

Отсюда получаем, что $\vec{h} \geqslant \vec{h}_2$. А это противоречит свойству минимальности элементов из H_1 — существует еще меньшее решение $\vec{h}_2$

(утверждение (b) предыдущей теоремы). Таким образом, $H_1 \subset H_2$. Аналогично устанавливается, что $H_2 \subset H_1$. Тем самым доказано совпадение множеств H_1 и H_2.

Пусть теперь $\vec{y} \in B_1$ и $\vec{y} \notin B_2$. Тогда $\vec{y} = \sum\limits_{\vec{y}' \in B_2} p_{y'} \vec{y}'$, поскольку $\vec{y}$ — решение однородного уравнения (12). Так как $\vec{y} \neq \vec{0}$, то хотя бы одно $p_{y'} > 0$. Отсюда имеем неравенство $\vec{y} \geqslant \vec{y}'$ для некоторого $\vec{y}' \in B_2$. А это противоречит минимальности элементов из B_1 (утверждение (d) теоремы 5). Следовательно, $B_1 \in B_2$. Аналогично можно получить, что $B_2 \in B_1$, а значит, $B_1 = B_2$. □

Введем, следуя [4], следующее определение:

Определение 1. *Положительное целочисленное решение однородного уравнения (12) называется* **неприводимым**, *если его нельзя представить в виде суммы других положительных целочисленных решений.*

Утверждение 1. *Множество B есть множество всех неприводимых решений уравнения (12).*

Доказательство. В самом деле, если $\vec{y} \in B$ и $\vec{y} = \vec{y}_1 + \vec{y}_2$, где $\vec{y}_1$ и $\vec{y}_2$ — ненулевые решения (12), то это нарушает условие минимальности элементов из B. Поскольку множество всех решений уравнения (12) есть $L(B)$, то B содержит все неприводимые решения. □

Таким образом, множество B есть базис Гильберта [4] для множества всех неотрицательных целочисленных решений однородного уравнения (12).

Определим теперь число векторов в множествах H и B.

Теорема 7. *Число минимальных решений неоднородного уравнения (1) определяется формулой:*

$$M_0 = |H| = C^b_{m+b-1} = \frac{(m+b-1)!}{b!(m-1)!} \,. \tag{15}$$

Доказательство. Число M_0 есть число неотрицательных целых решений уравнения (5). Согласно [13] это число легко находится из комбинаторных соображений и равно величине C^b_{m+b-1}. □

Теорема 8. *Число минимальных базисных решений однородного уравнения (12) определяется формулой:*

$$|B| = \sum_{l=m+1}^{n} C^{a_l-1}_{m+a_l-2} = \sum_{l=m+1}^{n} \frac{(m+a_l-2)!}{(a_l-1)!(m+a_l-2)!} \,. \tag{16}$$

Доказательство. В самом деле, $|B| = \sum\limits_{l=m+1}^{n} M_l$, но M_l — это число неотрицательных целых решений уравнения $k_1 + \ldots + k_m = a_l - 1$, а значит, $M_l = C^{a_l-1}_{m+a_l-2}$. □

§3. Уравнение вида $\sum\limits_{i=1}^{n} x_i + b = \sum\limits_{i=1}^{n} a_i x_i$

Рассмотрим теперь уравнение вида

$$\sum_{i=1}^{n} x_i + b = \sum_{i=1}^{n} a_i x_i \,. \tag{17}$$

Как и раньше предполагаем, что $b \geqslant 0$, $a_i \geqslant 0$ и $a_i \neq 1$.

Уравнение (17) можно переписать в виде эквивалентной системы, введя дополнительные переменные y_1 и y_2:

$$\begin{cases} y_2 + \sum\limits_{i=1}^{n} x_i = y_1 + \sum\limits_{i=1}^{n} a_i x_i \\ y_2 = b + 1 \\ y_1 = 1 \end{cases} \tag{18}$$

Система (18) является ассоциированной для грамматики $G = (\{S, A\}, \{a\}, P, S)$ и цепочки $x = a^{b+1}$, где множество правил P задается следующим образом:

$$\begin{array}{lcl} y_1 & : & S \longrightarrow A \\ y_2 & : & A \longrightarrow a \\ x_1 & : & A \longrightarrow A^{a_1} \\ x_2 & : & A \longrightarrow A^{a_2} \\ & & \ldots \\ x_n & : & A \longrightarrow A^{a_n} \end{array} \tag{19}$$

Здесь в левом столбце указаны переменные из (18), соответствующие правилам, указанным в правом столбце (значение переменной равняется числу применений этого правила в выводе некоторой терминальной цепочки).

В зависимости от вида коэффициентов a_i $(i = 1, \ldots, n)$ возможны ситуации:

1) $a_i = 0$ для $i = 1, 2, \ldots, n$;

2) $a_i > 0$ для $i = 1, 2, \ldots, n$;

3) $a_1 = a_2 = \ldots = a_m = 0$ $(0 < m < n)$ и $a_l > 0$ $(l = m+1, \ldots, n)$.

В первом случае, если $b > 0$, то решений нет. Если $b = 0$, то единственное решение уравнения (17) — нулевое.

Во втором случае приходим к уравнению

$$\sum_{i=1}^{n} c_i x_i = b \,, \tag{20}$$

где $c_i = a_i - 1 > 0$ $(i = 1, \ldots, n)$.

С точки зрения порождающей грамматики (18) любому решению (20) соответствует лес решения $F_{a^{b+1}} = \{T_{a^s}\} \cup F_c^1$, где T_{a^s} — некоторое дерево разбора цепочки $y = a^s$, а F_c^1 — уравновешенный лес циклических деревьев (чисто циклических деревьев для этой грамматики не существует), крона которого совпадает с цепочкой a^{b+1-s}. На рис. 3 показаны структурные элементы леса $F_{a^{b+1}}$.

Однако ясно, что любой такой лес $F_{a^{b+1}}$ легко преобразовать в некоторое дерево разбора $T_{a^{b+1}}$ так, чтобы равновесие не было нарушено. Таким образом, любое решение (20) соответствует некоторому дереву разбора $T_{a^{b+1}}$ и наоборот.

Рассмотрим, наконец, наиболее интересный — последний случай из приведенных выше. Правила порождающей грамматики теперь вы-

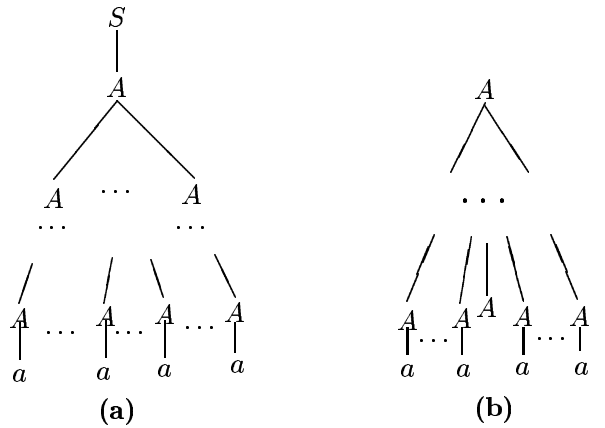


Рис. 3. (a) — дерево T_{a^s} разбора цепочки a^s; (b) — циклическое дерево из леса F_c^1

глядят следующим образом:

$$\begin{array}{lcl} y_1 & : & S \longrightarrow A \\ y_2 & : & A \longrightarrow a \\ x_1 & : & A \longrightarrow e \\ & \ldots & \\ x_m & : & A \longrightarrow e \\ x_{m+1} & : & A \longrightarrow A^{a_{m+1}} \\ & \ldots & \\ x_n & : & A \longrightarrow A^{a_n} \end{array} \tag{21}$$

Легко заметить, что лес решения системы имеет вид:

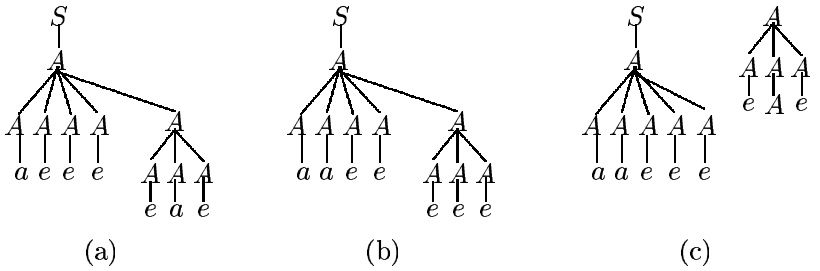
$$F = \{T_{a^{b+1}, e^s}\} \cup F_c^0 \,,$$

где T_{a^{b+1}, e^s} — дерево разбора цепочки $x = a^{b+1}$, крона которого состоит из $b+1$ терминала a и s пустых цепочек e. F_c^0 — лес чисто циклических деревьев, поскольку, как и ранее, циклические деревья с непустой кроной могут быть включены в состав дерева разбора без изменения равновесия леса решения.

Всегда можно предполагать, что $s < \max\limits_{m+1<l<n} \{a_l\}$, ибо в противном случае в дереве T_{a^{b+1}, e^s} есть поддерево, соответствующее вы-

воду $A \Rightarrow^+ A^{a_{l_0}}$ ($l_0 \leqslant s$). Тогда путем перестановки его потомков можно добиться, чтобы в T_{a^{b+1},e^s} появилось поддерево для вывода $A \Rightarrow A^{a_{l_0}} \Rightarrow^+ e$, поскольку число поддеревьев $A \Rightarrow^+ e$ в T_{a^{b+1},e^s} равно $s \geqslant a_{l_0}$. Полученное дерево легко расщепить на два: $T_{a^{b+1},e^{s-a_{l_0}}}$ и чисто циклическое дерево. Продолжая этот процесс, получим в результате дерево разбора T_{a^{b+1},e^s} такое, что $0 \leqslant s < \max\limits_{m+1<l<n}\{a_l\}$, и некоторый лес чисто циклических деревьев.

Пример 2. Рассмотрим уравнение $x_1 + x_2 + x_3 + 1 = 3x_2 + 5x_3$. Одному из его решений $x = (3,1,1)^\top$ можно сопоставить лес решений $\{T_{a^2,e^5}\}$, изображенный на рис. 4 (a). Дерево разбора цепочки a^2 из этого леса можно преобразовать так, чтобы получить дерево с поддеревом $A \Rightarrow A \Rightarrow^+ e$, что показано на рис. 4 (b). В свою очередь, последнее дерево может быть расщеплено на два, что демонстрируется на рис. 4 (c).



(a) (b) (c)

Рис. 4. Преобразование леса $\{T_{a^2,e^5}\}$ в лес $\{T_{a^2,e^3}\} \cup F_c^0$

Дерево T_{a^{b+1},e^s} может быть получено из решения следующих систем (их количество равняется $\max\limits_{m+1\leqslant l\leqslant n}\{a_l\}$):

$$\begin{cases} \sum\limits_{l=m+1}^{n} (a_l - 1)x_i = b + s\,, \\ \sum\limits_{i=1}^{m} x_i = s\,, \quad s = 0, 1, \ldots, \max\limits_{m+1\leqslant l\leqslant n}\{a_l\} - 1\,, \end{cases} \tag{22}$$

при этом $y_1 = 1$, $y_2 = b + 1$.

Данная система получена из тех соображений, что в дереве T_{a^{b+1},e^s} нетерминалов A, которые непосредственно порождают лист (терминал a или пустую цепочку e), ровно $\sum\limits_{l=m+1}^{n} (a_l - 1)x_l + 1$ и они должны быть раскрыты в $b + 1$ терминалов a и s пустых цепочек.

Замечание. Системы уравнений (22) не эквивалентны (17), поскольку s пробегает конечное число значений. Число решений (22) также конечно. Кроме того, полученное таким образом множество векторов может не обладать свойством минимальности.

Пусть

$$H = \left\{ \vec{x} \;\middle|\; \begin{array}{l} \vec{x} \text{ — решение системы (22) для некоторого } s, \\ \text{несравнимое } (\leqslant, \geqslant) \text{ с другими решениями этой} \\ \text{системы} \end{array} \right\} \tag{23}$$

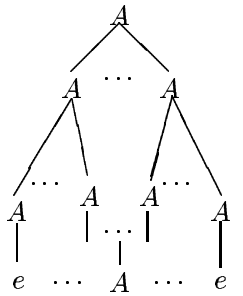


Рис. 5. Общий вид чисто циклического дерева

Чисто циклические деревья, составляющие лес F_c^0, имеют вид, показанный на рис. 5. Пусть при построении такого чисто циклического дерева правило для переменной x_l применено k_l раз ($l = m+1, \ldots, n$) и k_i раз для правила, соответствующего x_i ($i = 1, \ldots, m$). Тогда в дереве будет всего $\sum\limits_{l=m+1}^{n} a_l k_l + 1$ нетерминалов A. Из них не являются непосредственными предками пустой цепочки e ровно $\sum\limits_{l=m+1}^{n} k_l + 1$. Вычитая из первого второе, получаем, что в дереве $\sum\limits_{l=m+1}^{n} (a_l - 1)k_l$ нетерминалов A, которые раскрываются в пустую цепочку. Далее,

поскольку k_i $(i = 1, \ldots, m)$ определяют, сколько раз соответствующее правило для раскрытия A в пустую цепочку должно быть применено, то получаем следующее уравнение для определения k_i и k_l:

$$\sum_{i=1}^{m} k_i = \sum_{l=m+1}^{n} (a_l - 1)k_l$$

или в другой форме:

$$\sum_{i=1}^{n} x_i = \sum_{l=m+1}^{n} a_l x_l \ , \tag{24}$$

где $x_i = k_i$ $(i = 1, \ldots, n)$.

Но уравнение (24) есть однородное уравнение типа (12) из предыдущего параграфа. Его общее решение есть множество $L(B)$, где

$$B = \left\{ \vec{x}^{lr} = (k_1, \ldots, k_m, 0, \ldots, 1, \ldots, 0)^\top \;\middle|\; \begin{array}{l} k_1 + \ldots + k_m = a_l - 1, \\ l = m+1, \ldots, n, \\ 1 \text{ стоит на } l\text{-м месте} \end{array} \right\} \tag{25}$$

Для уравнения (17) и множеств H и B, определенных согласно (23), (25), остаются справедливы теоремы 3, 4, 8 и утверждение 1 из предыдущего параграфа. Теорема 2 здесь имеет аналог в виде:

Теорема 9. *Пусть $\vec{x}$ — произвольный вектор из H. Тогда $\vec{x}$ есть решение уравнения (17).*

Теорема 5 (свойство минимальности) переформулируется следующим образом:

Теорема 10. *Пусть вектор $\vec{x}$ — произвольное решение неоднородного уравнения (17), а $\vec{x}_c$ — произвольное решение однородного уравнения (12). Тогда:*

(a) Найдется $\vec{y} \in H$ такой, что $\vec{y} \leqslant \vec{x}$.

(b) Для всех $\vec{y} \in H$ неверно, что $\vec{y} \geqslant \vec{x}$, если только $\vec{y} \neq \vec{x}$.

(c) Найдется $\vec{y} \in B$ такой, что $\vec{y} \leqslant \vec{x}_c$.

(d) Для всех $\vec{y} \in B$ неверно, что $\vec{y} \geqslant \vec{x}_0$, если только $\vec{y} \neq \vec{x}_0$.

Сохраняется свойство единственности для множеств H и B.

Теорема 11. *Пусть множества H и B удовлетворяют свойству минимальности и удовлетворяют равенству $S = H + L(B)$, где S — общее решение (17). Тогда эти множества определяются единственным образом.*

Множество H для уравнения (17) имеет несколько другую структуру, нежели аналогичное множество для уравнения (1), и точное вычисление его элементов довольно затруднительно. Однако имеет место следующая оценка:

Теорема 12. *Пусть $N = \max\limits_{m+1 \leqslant l \leqslant n} \{a_l\} - 1$. Тогда число минимальных решений неоднородного уравнения (17) не превосходит величины M_0:*

$$|H| \leqslant M_0 = \sum_{s=0}^{N} C_{m+s-1}^{s} C_{n-m+b+s-1}^{b+s} \tag{26}$$

Доказательство. В самом деле, число решений первого из уравнений системы (22) не превосходит числа решений уравнения

$$\sum_{l=m+1}^{n} x_l = b + s \ ,$$

а оно имеет ровно $C_{n-m+b+s-1}^{b+s}$ решений. Число решений второго уравнения системы (22) равно C_{m+s-1}^{s}. Таким образом, при фиксированном s число решений системы не превосходит величины $C_{m+s-1}^{s} C_{n-m+b+s-1}^{b+s}$. Суммируя по всем s от 0 до N, получаем верхнюю оценку общего числа минимальных решений. □

Пример 3. Рассмотрим уравнение $x+5 = y+2z$. Оно эквивалентно уравнению $x + y + z + 5 = 2y + 3z$. Здесь $n = 3$, $m = 1$, $a_2 = 2$, $a_3 = 3$.

Согласно (23) и (25) вычисляем H (отсеивая неминимальные элементы):

$$H = \left\{ \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \right\} \ , \quad B = \left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Отсюда получаем, что любое решение принадлежит множеству:

$$\left\{ \begin{pmatrix} p_1+2p_2 \\ 1+p_1 \\ 2+p_2 \end{pmatrix}, \begin{pmatrix} p_1+2p_2 \\ 3+p_1 \\ 1+p_2 \end{pmatrix}, \begin{pmatrix} p_1+2p_2 \\ 5+p_1 \\ p_2 \end{pmatrix}, \begin{pmatrix} 1+p_1+2p_2 \\ p_1 \\ 3+p_2 \end{pmatrix} \right\}$$

§4. Общий случай

Рассмотрим уравнение

$$\sum_{i=1}^{n} x_i + b_1 = \sum_{i=1}^{n} a_i x_i + b_2 \tag{27}$$

и соответствующее ему однородное уравнение

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} a_i x_i \tag{28}$$

Полагая $b = b_2 - b_1$, если $b_2 - b_1 > 0$, можно уравнение (27) свести к уравнению (1). В противном случае, если $b_1 - b_2 > 0$, то полагаем $b = b_1 - b_2$ и приходим к уравнению (17). Таким образом, результаты, полученные в предыдущих двух параграфах, позволяют успешно решать уравнение (27).

Теорема 13. *Множество S всех решений уравнения (27) определяется формулой*

$$S = H + L(B)\,, \tag{29}$$

где множество H есть конечное множество всех минимальных решений неоднородного уравнения (27), а множество B — конечное множество всех минимальных (неприводимых) решений однородного уравнения (28). Причем множества H и B определяются единственным образом.

Как определить множества H и B, было описано в § 2 и § 3.

Теорема 14. *Число элементов множеств H и B конечно.*

Конкретные оценки мощности этих множеств даются формулами (15), (16) и (26).

Задача нахождения множеств H и B сводится к нахождению всех решений уравнений типа

$$\sum_{i=1}^{m} x_i = d \tag{30}$$

или

$$\sum_{i=1}^{m} c_i x_i = d\,. \tag{31}$$

Обе задачи имеют довольно простой рекурсивный алгоритм решения. Обозначим за $\mathbf{P}(x_1, \ldots, x_m, d)$ исходную задачу ((30) или (31)), и тогда она сводится к решению нескольких задач меньшего размера:

$$\Big\{ x_1 = q, \mathbf{P}(x_2, \ldots, x_m, d-q) \Big\}\,,\quad q = 0, 1, \ldots, d\,,$$

или

$$\Big\{ x_1 = q, \mathbf{P}(x_2, \ldots, x_m, d-c_1 q) \Big\}\,,\quad q = 0, 1, \ldots, \left[\frac{d}{c_1}\right]\,,$$

в зависимости от того, какое из уравнений (30) или (31) решается.

В случае уравнения (30) число шагов данного алгоритма пропорционально числу решений. Для уравнения (31) алгоритм работает хуже, поскольку некоторые из подзадач, к которым сводится исходная, могут не иметь решения, а значит, возникают излишние ветвления. Однако алгоритм может быть частично улучшен за счет введения различных дополнительных проверок, имеющих цель пресекать ненужные ветвления на как можно более раннем этапе.

§5. Интерпретация описанного метода

Используемый нами метод порождающей грамматики для решения класса диофантовых уравнений вида (27) в неотрицательных целых числах имеет довольно прозрачную интерпретацию с точки зрения обычной техники. В этом методе уравнение (27) расщепляется на несколько уравнений, каждое из которых решается по отдельности. После этого любое решение получается как некоторая сумма найденных решений расщепленных уравнений.

Продемонстрируем вышесказанное на примере уравнения вида (1). Первоначально оно расщепляется на два уравнения:

$$\sum_{i=1}^{m} x_i = b \quad \text{и} \quad \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} a_i x_i \ ,$$

то есть на "усеченное" неоднородное уравнение, которое имеет вид (30), и однородное уравнение. Первое из них определяет некоторое дерево разбора пустой цепочки, а второе — некоторый уравновешенный лес чисто циклических деревьев. Все решения первого расщепленного уравнения, в которых компоненты $x_l = 0$ $(l = m+1, \ldots, n)$, составляют множество H, а решения второго — множество $L(B)$.

Для нахождения B продолжается расщепление последнего уравнения на $n - m$ уравнений вида:

$$\sum_{i=1}^{m} x_i + x_l = a_l x_l \ , \quad l = m+1, \ldots, n \ ,$$

или в другой форме:

$$\sum_{i=1}^{m} x_i = (a_l - 1) x_l \ , \quad l = m+1, \ldots, n \ .$$

Каждое из этих уравнений определяет некоторое чисто циклическое дерево грамматики.

Поскольку уравнения однородные, а значит, если $\vec{x}^*$ — решение, то решением является и любой вектор $p\vec{x}^*$ для произвольного числа $p \geqslant 0$. Следовательно, мы можем зафиксировать компоненту $x_l = 1$ и получить уже знакомые нам уравнения для определения B:

$$\sum_{i=1}^{m} x_i = a_l - 1 \ , \quad l = m+1, \ldots, n \ .$$

Легко показать, что любая сумма решений расщепленных уравнений дает решение и исходного уравнения. Однако обратное утверждение о том, что любое решение уравнения (1) всегда может быть представлено в виде суммы решений расщепленных уравнений, вообще говоря, не очевидно. Развитая в [2] теория позволяет сказать, что это верно. Прямой же метод доказательства этого факта, напротив, может оказаться довольно трудным.

§6. Заключение

В данной работе демонстрируется, насколько успешно может быть использована теория, основы которой заложены в [1] и развиты в [2]. Показано, что теория формальных языков полезна не только в таких стандартных областях ее применения, как теория компиляции, синтаксического и лексического анализа, но и для решения других, совершенно отличных от этих, проблем математики.

В работе получены следующие результаты:

1) Показано, как задача решения уравнения вида

$$\sum_{i=1}^{n} x_i + b_1 = \sum_{i=1}^{n} a_i x_i + b_2$$

с неотрицательными целыми коэффициентами в неотрицательных целых числах сводится к нахождению решений уравнений вида $\sum_{i=1}^{m} c_i x_i = d \ \ (c_i > 0 \ \ \forall i = 1, 2, \ldots, m)$, которые, в свою очередь, относительно легко решаются при помощи ЭВМ.

2) Доказано, что множество всех решений уравнения (27) имеет вид $S = H + L(B)$, где H — множество минимальных решений соответствующего неоднородного уравнения, а B — множество минимальных (неприводимых) решений соответствующего однородного уравнения. При этом множества H и B определяются единственным образом.

3) Доказано, что число элементов множеств H и B всегда конечно, и предложен алгоритм их построения по исходным данным.

4) Для числа элементов множества B получена точная оценка.

5) Для числа элементов множества H получена оценка, которая является точной для уравнения (1) и оценкой сверху, если рассматривается уравнение вида (17).

Заметим, что факт представления множества решений в виде $S = H + L(B)$, а также конечность множеств H и B были известны

и ранее, причем для более общего случая — систем линейных диофантовых уравнений. Однако в данной работе мы специально приводим доказательства для частного случая, поскольку они опираются на развитую нами теорию и тем самым демонстрируют ее достоинства.

Литература

1. *Filgueiras M., Tomás A.* Solving Linear Constraints on Finite Domains through Parsing // Proc. of EPTA'91, 1991.

2. *Богоявленский Ю. А., Корзун Д. Ж.* Общий вид решения системы линейных диофантовых уравнений, ассоциированной с контекстно-свободной грамматикой // Труды Петрозаводского государственного университета. Сер. "Прикладная математика и информатика". Вып. 6. Петрозаводск: Изд-во ПетрГУ, 1997. С. 79–94.

3. *Comon H., Dincbas M.* A Methodological View of Constraint Solving, 1995. (`ftp://ftp.lri.fr/LRI/articles/comon/casm.ps.Z`)

4. *Схрейвер А.* Теория линейного и целочисленного программирования. Т. 1, 2. М.:Мир, 1991.

5. *Tomás A., Filgueiras M.* A New Method for Solving Linear Constraints on the Natural Numbers // Proc. of EPTA'91, 1991.

6. *Huet G.* An algorithm to generate the basis of solutions to homogeneous linear diophantine equations // Informational Processing Letters. Vol. 3. No. 7. 1978. P. 144–147.

7. *Contejean E., Devie H.* An Efficient Incremental Algorithm for Solving Systems of Linear Diophantine Equations // Information and Computation. Vol. 13. No. 1. 1994. P. 143–172.

8. *Ajili F., Contejean E.* Avoiding Slack Variables in the Solving of Linear Diophantine Equations and Inequations // Technical Report, 1996. (`ftp://ftp.lri.fr/LRI/articles/contejean/tcs97.ps.gz`).

9. *Pottier L.* Minimal solutions of linear diophantine systems: bounds and algorithms // Proceedings of the 4th International Conference on Rewriting Techiques and Applications (RTA'91). Como (Italy), 1991. P. 162–173.

10. *Clausen M., Fortenbacher A.* Efficient Solution of Linear Diophantine Equations // J. Symbolic Computations. Vol. 8 (1 & 2). 1989. P. 201–216.

11. *Boudet A., Comon H.* Diophantine Equations, Presburger Arithmetic and Finite Automata // Proceedings of the 21st International Colloquim on Trees in Algebra and Programming (CAAP'96). 1996. P. 30–43.

12. *Ахо А. В., Ульман Д. Д.* Теория синтаксического анализа, перевода и компиляции. Т. 1, 2. М.:Мир, 1978.

13. *Нефёдов В. Н., Осипова В. А.* Курс дискретной математики. М.: Изд-во МАИ, 1992.