

УДК 519.7, 511.512

О СУЩЕСТВОВАНИИ ПОРОЖДАЮЩЕЙ КС-ГРАММАТИКИ ДЛЯ ПРОИЗВОЛЬНОЙ ЛИНЕЙНОЙ ДИОФАНТОВОЙ СИСТЕМЫ

Д. Г. Корзун

В работе доказывается существование контекстно-свободной грамматики и терминальной цепочки, порождающих произвольно заданную систему линейных диофантовых уравнений (ЛДУ).

§ 1. Введение

В данной статье под системой ЛДУ мы понимаем систему линейных уравнений с целыми коэффициентами, решения которой ищутся в множестве неотрицательных целых чисел $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$. Две системы будем называть эквивалентными, если по решениям одной системы однозначно определяются решения второй, и наоборот.

В работе [1] описывается метод построения некоторой системы ЛДУ по произвольной КС-грамматике $G = (N, \Sigma, P, S)$ и цепочке $x \in \Sigma^*$. Такая система имеет вид:

$$\left\{ \begin{array}{l} \sum_{i \in H_S} \xi_i - \sum_{k=1}^n c_{Sk} \xi_k = 1, \\ \sum_{i \in H_A} \xi_i - \sum_{k=1}^n c_{Ak} \xi_k = 0 \quad \forall A \in N \setminus \{S\}, \\ \sum_{k=1}^n c_{ak} \xi_k = b_a \quad \forall a \in \Sigma. \end{array} \right. \quad (1)$$

Здесь S — начальный нетерминал грамматики G , H_A — множество номеров тех правил грамматики, в левой части которых находится

нетерминал A (формально $H_A = \{i \mid p_i = (A \rightarrow \alpha), p_i \in P\}$), c_{U_k} — число вхождений символа U (терминал или нетерминал) в правую часть правила p_k , а b_a — число вхождений a в цепочку x .

В [2] этот метод получил формальное обоснование на основе балансовых соотношений в дереве разбора цепочки x в грамматике G . При этом грамматика G называется *порождающей*, а построенная система $D(G, x)$ — *ассоциированной*.

Если известно, как по КС-грамматике построить ассоциированную систему ЛДУ, то весьма естественным становится вопрос об анализе обратной задачи — можно ли найти такую грамматику G и цепочку x , для которых произвольно заданная система $D : A\xi = d$ является ассоциированной.

В этой работе мы исследуем следующие вопросы подобной задачи:

- 1) существует ли порождающая грамматика;
- 2) единственным ли образом она определяется;
- 3) как построить порождающую грамматику.

Оказывается, что для произвольной системы ЛДУ порождающая КС-грамматика всегда существует, точнее всегда можно найти такую пару (G, x) , которая порождает эквивалентную данной системе ЛДУ. Однако единственность не гарантируется. При этом соответствующую грамматику и цепочку (G, x) всегда можно построить явным образом.

Терминология и обозначения, применяемые в статье относительно объектов теории формальных языков, соответствуют [3].

§ 2. Подготовительные соображения

Пусть дана произвольная система D линейных диофантовых уравнений:

$$A\xi = d. \quad (2)$$

Здесь A — целочисленная $m \times n$ -матрица коэффициентов, $\xi \in \mathbb{Z}_+^n$ — вектор неизвестных, $d \in \mathbb{Z}^m$ — вектор правых частей.

Обозначим через $S = S(D)$ множество решений этой системы, т. е. $S(D) = \{\xi \mid A\xi = d\}$.

Как известно [4], множество S можно описать с помощью базиса Гильберта:

$$S = H + L(B), \quad (3)$$

где H — конечное множество минимальных частных решений системы (2), B — конечный базис Гильберта, который определяется единственным образом как множество минимальных решений однородной системы для (2), а $L(B)$ — линейная неотрицательная целочисленная комбинация элементов из B .

В силу этого возможны 3 ситуации: либо система D несовместна ($H = \emptyset$), либо имеет конечное число решений ($H \neq \emptyset, B = \emptyset$), либо множество решений бесконечно ($H \neq \emptyset, B \neq \emptyset$).

§ 3. Основная теорема

ЛЕММА 1. Если система (2) несовместна, то в качестве порождающей грамматики можно взять $G = (N, \Sigma, P, S)$, где множество нетерминалов — $N = \{S\}$, множество терминальных символов — $\Sigma = \emptyset$, множество правил P грамматики определяется как¹:

$$\begin{aligned} \xi_1 &: S \longrightarrow S \\ \xi_2 &: S \longrightarrow S \\ &\dots \\ \xi_n &: S \longrightarrow S, \end{aligned} \tag{4}$$

нетерминал S — начальный символ. В качестве цепочки x берется пустая цепочка ϵ .

ДОКАЗАТЕЛЬСТВО. Если система несовместна, то она эквивалентна уравнению $0 = 1$. Последнее можно переписать следующим образом:

$$\sum_{i=1}^n \xi_i - \sum_{i=1}^n \xi_i = 1.$$

Согласно (1) построенная в условии леммы грамматика G и цепочка x порождают данное уравнение. \square

ЛЕММА 2. Пусть система (2) имеет конечное число решений, т. е. $B = \emptyset, H = \{h^1, h^2, \dots, h^m\}$. Тогда в качестве порождающей грамматики можно взять $G = (N, \Sigma, P, S)$, где множество нетерминалов — $N = \{S, A_1, \dots, A_n\}$, множество терминальных символов — $\Sigma = \emptyset$,

¹Здесь и далее при описании правил порождающей грамматики слева через двоеточие указываются соответствующие переменные ассоциированной системы.

множество правил P определяется как:

$$\begin{aligned}
 y_1 & : S \longrightarrow A_1^{h_1^1} A_2^{h_2^1} \dots A_n^{h_n^1} \\
 y_2 & : S \longrightarrow A_1^{c_1^2} A_2^{h_2^2} \dots A_n^{h_n^2} \\
 & \dots \\
 y_m & : S \longrightarrow A_1^{h_1^m} A_2^{h_2^m} \dots A_n^{h_n^m} \\
 \xi_1 & : A_1 \longrightarrow e \\
 \xi_2 & : A_2 \longrightarrow e \\
 & \dots \\
 \xi_n & : A_n \longrightarrow e,
 \end{aligned} \tag{5}$$

нетерминал S — начальный символ. В качестве цепочки x берется пустая цепочка e .

ДОКАЗАТЕЛЬСТВО. Пусть G — грамматика из условия леммы. Согласно (1) ассоциированная система имеет вид:

$$\begin{cases} \sum_{j=1}^m y_j = 1 \\ \xi_i = \sum_{j=1}^m h_i^j y_j, \quad i = 1, 2, \dots, n \end{cases}$$

Так как решение ищется в неотрицательных целых, то в силу 1-го уравнения этой системы получаем, что $y_j \in \{0, 1\}$. Более того, если $y_{j_0} = 1$ для некоторого j_0 , то $y_j = 0 \forall j \neq j_0$. Таким образом, любое решение ассоциированной системы имеет вид:

$$(y_1, \dots, y_{j_0}, \dots, y_m, \xi_1, \dots, \xi_n) = (0, \dots, 1, \dots, 0, h_1^{j_0}, \dots, h_n^{j_0}),$$

и наоборот. Таким образом, $\xi_i = h_i^{j_0}$. А это совпадает с (3) при $B = \emptyset$. \square

ЛЕММА 3. Пусть система (2) имеет бесконечное число решений, т. е. $H = \{h^1, h^2, \dots, h^m\}$, $B = \{b^1, b^2, \dots, b^l\}$. Тогда в качестве порождающей грамматики можно взять $G = (N, \Sigma, P, S)$, где множество нетерминалов — $N = \{S, A_1, \dots, A_n\}$, множество терминальных символов

— $\Sigma = \emptyset$, множество правил P определяется как:

$$\begin{aligned}
 y_1 & : S \longrightarrow A_1^{h_1^1} A_2^{h_2^1} \dots A_n^{h_n^1} \\
 y_2 & : S \longrightarrow A_1^{c_1^2} A_2^{h_2^2} \dots A_n^{h_n^2} \\
 & \dots \\
 y_m & : S \longrightarrow A_1^{h_1^m} A_2^{h_2^m} \dots A_n^{h_n^m} \\
 z_1 & : S \longrightarrow S A_1^{b_1^1} A_2^{b_2^1} \dots A_n^{b_n^1} \\
 z_2 & : S \longrightarrow S A_1^{c_1^2} A_2^{h_2^2} \dots A_n^{b_n^2} \\
 & \dots \\
 z_l & : S \longrightarrow S A_1^{b_1^l} A_2^{b_2^l} \dots A_n^{b_n^l} \\
 \xi_1 & : A_1 \longrightarrow e \\
 \xi_2 & : A_2 \longrightarrow e \\
 & \dots \\
 \xi_n & : A_n \longrightarrow e,
 \end{aligned} \tag{6}$$

нетерминал S — начальный символ. В качестве цепочки x берется пустая цепочка e .

ДОКАЗАТЕЛЬСТВО. Ассоциированная система для этой грамматики имеет вид:

$$\left\{ \begin{array}{l} \sum_{j=1}^m y_j + \sum_{k=1}^l z_k - \sum_{k=1}^l z_k = 1 \quad \Rightarrow \quad \sum_{j=1}^m y_j = 1 \\ \xi_i = \sum_{j=1}^m h_i^j y_j + \sum_{k=1}^l b_i^k z_k, \quad i = 1, 2, \dots, n \end{array} \right.$$

Из первого уравнения получаем, что $y_j \in \{0, 1\}$ и, если $y_{j_0} = 1$ для некоторого j_0 , то $y_j = 0 \forall j \neq j_0$. Далее, поскольку z_k — произвольное неотрицательное целое, то любое решение системы имеет вид:

$$\begin{aligned}
 & (y_1, \dots, y_{j_0}, \dots, y_m, z_1, \dots, z_l, \xi_1, \dots, \xi_n) = \\
 & = (0, \dots, 1, \dots, 0, z_1, \dots, z_l, h_1^{j_0}, \dots, h_n^{j_0})
 \end{aligned}$$

и наоборот. Таким образом, $\xi_i = h_i^{j_0} + \sum_{k=1}^l b_i^k z_k$, что совпадает с (3). \square

Из лемм 1–3 и замечания в конце § 2 следует основная теорема данной работы:

ТЕОРЕМА 1. Для произвольной системы ЛДУ вида (2) существует эквивалентная система, являющаяся ассоциированной для некоторой

КС-грамматики G и терминальной цепочки x . При этом можно выбрать множество терминальных символов (алфавит Σ) пустым, а $x = e$.

СЛЕДСТВИЕ. . Порождающая грамматика G и терминальная цепочка x для произвольной системы ЛДУ определяются, вообще говоря, не единственным образом.

ДОКАЗАТЕЛЬСТВО. Пусть $G = (N, \Sigma, P, S)$ и $x \in \Sigma^*$. Допустим также, что или $\Sigma \neq \emptyset$, или $x \neq e$. Тогда можно построить ассоциированную систему $D(G, x)$ по (1). Согласно теореме для этой системы существует порождающая грамматика $G' = (N, \emptyset, P, S)$ и $x' = e$. Таким образом, есть, по крайней мере, две различные пары (G, x) и (G', x') , порождающие эквивалентные системы. \square

Заметим, что леммы 1–3 не только доказывают существование порождающей грамматики, но и позволяют явно ее построить.

Заключение

Предлагаемые в леммах 1–3 варианты порождающей грамматики требуют знания множеств H и B , т. е., по сути дела, перед построением грамматики необходимо решить соответствующую систему ЛДУ. Однако следствие позволяет надеяться, что для заданной системы существует порождающая грамматика, построение которой не требует явного нахождения решений. Этого, как кажется, можно добиться, во-первых, за счет увеличения терминального алфавита и более сложной терминальной цепочки x , а во-вторых, за счет более “изощренных” правил грамматики, позволяющих строить более сложные деревья разбора.

Большой интерес представляет собой задача построения порождающей грамматики для специальных частных видов систем ЛДУ, особенно если при этом не требуется предварительно находить решения. Так, например, в работе автора [5] такие грамматики были найдены для уравнения вида: $\sum_{i=1}^n x_i + b_1 = \sum_{j=1}^m a_j y_j + b_2$, что в свою очередь позволило свести задачу его решения к решению более простого уравнения: $\sum_{i=1}^k z_i = d$ или $\sum_{i=1}^k c_i z_i = d$ ($c_i > 0$).

In this paper we prove the existence of a context-free grammar and a terminal string, which generate a given system of linear diophantine equations.

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. Г. *Общий вид решения системы линейных диофантовых уравнений, ассоциированной с контекстно-свободной грамматикой* // Труды Петрозаводского государственного университета. Сер. Прикладная математика и информатика. Вып. 6. 1997. С. 79–94.
- [3] Ахо А. В., Ульман Д. Д. *Теория синтаксического анализа, перевода и компиляции*. Т. 1, 2. М.: Мир, 1978.
- [4] Схрейвер А. *Теория линейного и целочисленного программирования*. Т. 1, 2. М.: Мир, 1991.
- [5] Корзун Д. Г. *Решение одного класса линейных диофантовых уравнений в неотрицательных целых числах методами теории формальных языков* // Труды Петрозаводского государственного университета. Сер. Прикладная математика и информатика. Вып. 7. 1998. С. 101–124.

Петрозаводский государственный университет,
математический факультет,
185640, Петрозаводск, пр. Ленина, 33