Доклады Академии наук СССР
1973. Том 212, № 1

УДК 519.48

МАТЕМАТИКА

А. Л. СЕМЕНОВ

АЛГОРИТМИЧЕСКИЕ ПРОБЛЕМЫ ДЛЯ СТЕПЕННЫХ РЯДОВ И КОНТЕКСТНО-СВОБОДНЫХ ГРАММАТИК

(Представлено академиком П. С. Новиковым 10 I 1973)

Настоящая заметка посвящена использованию степенных рядов для решения некоторых алгоритмических проблем для контекстно-свободных (к.с.) грамматик.

Понятия контекстно-свободной, нормальной (в смысле Хомского) и однозначной грамматики считаются известными ([1]). В дальнейшем под грамматикой понимается нормальная к. с. грамматика. Буквами $A_1, A_2, \ldots, A_n$, $B_1, B_2, \ldots, B_m$ мы будем обозначать нетерминальные символы, A_1, B_1 — всегда начальные символы, $a_1, a_2, \ldots, a_k$ — терминальные символы. Грамматике $\langle V, V_T, A_1, P\rangle$ можно поставить в соответствие систему уравнений относительно формальных степенных рядов $A_1, A_2, \ldots, A_n$ от некоммутирующих переменных $a_1, \ldots, a_k$ с коэффициентами из поля действительных чисел. Одним из решений этой системы будет n-ка степенных рядов, для которой коэффициент при цепочке в ряде A_i равен числу выводов этой цепочки в грамматике $\langle V, V_T, A_i, P\rangle$ ([1]).

Известно следующее обобщение к. с. грамматик ([2]). Каждому правилу сопоставлено некоторое действительное число — его вес, которое является коэффициентом при правой части этого правила в правой части соответствующего уравнения. Вес выводимой цепочки определяется как сумма по всем выводам данной цепочки произведений весов, примененных в выводе правил. Дальнейшие рассуждения в равной степени применимы к весовым и обычным грамматикам, т.е. весовым грамматикам, у которых все веса равны 1.

Для системы уравнений, соответствующей грамматике, существует только одно решение с нулевыми свободными членами степенных рядов (отсутствие свободного члена соответствует невыводимости пустой цепочки из A_i). Действительно, множество слов длины 1, выводимых из A_i, определяется однозначно в силу нормальности грамматики. Коэффициенты при цепочках длины $k+1$ в левых частях уравнений определяются коэффициентами при цепочках длины не более k в A_i, стоящих в правых частях уравнений.

Мы будем использовать степенные ряды над полем действительных чисел с коммутирующими переменными. Гомоморфизм, отображающий кольцо степенных рядов, задается отображением V_T на множество образующих свободной коммутативной полугруппы и считается фиксированным. Очевидно, коммутативный образ системы уравнений, соответствующей грамматике, имеет единственное решение без свободных членов, которое является коммутативным образом решения без свободных членов некоммутативной системы.

Укажем некоторую оценку для роста коэффициентов степенного ряда, задаваемого грамматикой.

Лемма 1. *Для всякой грамматики с p правилами число выводов слов длины n не превосходит p^{2n}.*

Доказательство. Каждому дереву вывода можно однозначно сопоставить последовательность применения правил, задающую левосторонний вывод. Длина этой последовательности не превосходит $2n$.

Из леммы 1 непосредственно вытекает

Лемма 2. *Коммутативный образ ряда, задаваемого грамматикой, имеет ненулевой радиус сходимости.*

Фиксируем алфавит $\{a_1, \ldots, a_k\}$ коммутирующих переменных коммутативных сходящихся степенных рядов. Для заданных рядов $r_1, \ldots, r_n$ существует ε такое, что при подстановке действительных чисел $\tilde{a}_1, \ldots, \tilde{a}_k$, $|\tilde{a}_i| < \varepsilon$, вместо $a_1, \ldots, a_k$ в ряды $r_1, \ldots, r_n$ получим некоторые сходящиеся ряды с суммами $\tilde{r}_1, \ldots, \tilde{r}_n$ соответственно. Такая подстановка задает гомоморфизм подкольца с образующими $r_1, \ldots, r_n$ кольца степенных рядов в кольцо действительных чисел.

Следующая лемма известна из анализа ([3]).

Лемма 3. *Если два ряда сходятся в некоторой окрестности $|\tilde{a}_i| < \varepsilon$ и задают в этой окрестности одну и ту же функцию, то они равны.*

Лемма 4. *В некоторой окрестности $|\tilde{a}_i| < \varepsilon$, $i = 1, \ldots, k$; $|\tilde{A}_j| < \varepsilon$, $j = 1, \ldots, n$, система уравнений, соответствующая грамматике, задает $(\tilde{A}_1, \ldots, \tilde{A}_n)$ как однозначную функцию от $(\tilde{a}_1, \ldots, \tilde{a}_k)$.*

Доказательство. Утверждение леммы непосредственно вытекает из теоремы о неявной функции ([3]) и нормальности грамматики.

Теорема. *Существует алгоритм, позволяющий по двум заданным грамматикам установить, совпадают ли коммутативные образы порождаемых ими степенных рядов.*

Доказательство. Пусть общий терминальный алфавит грамматик есть $\{a_1, \ldots, a_k\}$, нетерминальные алфавиты $\{A_1, \ldots, A_n\}$ и $\{B_1, \ldots, B_m\}$. Предикат $P_1(A_1, \ldots, A_n, a_1, \ldots, a_k)$ истинный, когда коммутативные ряды $A_1, \ldots, A_n$ являются решениями системы уравнений, задаваемой первой грамматикой, т.е. P_1 есть конъюнкция равенств системы; $\tilde{P}_1$ есть конъюнкция тех же равенств, но A_i и a_j везде заменены на $\tilde{A}_i$ и $\tilde{a}_j$. Аналогично определим P_2 и $\tilde{P}_2$ для второй системы. Равенство рядов, порождаемых грамматиками, есть равенство рядов без свободного члена A_1 и B_1, являющихся первыми компонентами решений соответствующих систем. По лемме 3 равенство этих рядов эквивалентно совпадению соответствующих им функций $\tilde{A}_1$ и $\tilde{B}_1$. По лемме 4 решение, задаваемое такой функцией, существует и единственно при $|\tilde{a}_i| < \varepsilon$, $|\tilde{A}_j| < \varepsilon$, $|\tilde{B}_l| < \varepsilon$ для некоторого ε. Следовательно, равенство функций эквивалентно равенству первых компонент решений системы (как функций от $\tilde{a}_i$). Итак, равенство коммутативных образов рядов, задаваемых грамматиками, эквивалентно истинности следующего высказывания:

$$\exists \varepsilon > 0 \ \forall \tilde{A}_1, \ldots, \tilde{A}_n, \tilde{B}_1, \ldots, \tilde{B}_n, \ \tilde{a}_1, \ldots, \tilde{a}_k \Big(\Big(\bigwedge_1^n |\tilde{A}_i| < \varepsilon \wedge \bigwedge_1^m |\tilde{B}_i| < \varepsilon \wedge$$

$$\wedge \bigwedge_1^k |\tilde{a}_i| < \varepsilon \wedge P_1(\tilde{A}_1, \ldots, \tilde{A}_n, \tilde{a}_1, \ldots, \tilde{a}_k) \wedge P_2(\tilde{B}_1, \ldots, \tilde{B}_m, \tilde{a}_1, \ldots, \tilde{a}_k)\Big) \supset$$

$$\supset \tilde{A}_1 = \tilde{B}_1\Big).$$

Из разрешимости элементарной теории действительных чисел ([4]) вытекает существование алгоритма, позволяющего проверить истинность этого высказывания. Теорема доказана.

Следствие 1. *Если для некоторого языка имеется однозначная грамматика, то можно алгоритмически решить вопрос, является ли другая грамматика, порождающая тот же язык, однозначной.*

Следствие 2. *Пусть заданы две однозначные грамматики, порождающие языки, про которые известно, что один из них содержится в другом. Тогда разрешима проблема равенства этих языков.*

Следствия 1 и 2 непосредственно вытекают из теоремы. Заметим, что в условиях следствия 2, если языки не совпадают, можно, исходя из знака разности, алгоритмически решить, какое из включений имеет место.

Следствие 3. *Разрешима проблема равенства языка, порождаемого однозначной грамматикой, и автоматного языка.*

Доказательство. Можно эффективно построить пересечение к. с. языка с (автоматным) дополнением к автоматному и проверить пустоту этого пересечения; следовательно, проблема включения разрешима. По автоматной грамматике можно построить однозначную грамматику, порождающую тот же язык ([1]). Остается применить следствие 2. Последний факт представляет интерес в связи с тем, что проблема равенства к. с. языка автоматному разрешима только в том случае, когда последний ограничен ([5]).

Теорема и следствия применимы не только к однозначным грамматикам, но и в случаях, когда все выводимые цепочки имеют равную степень неоднозначности или все цепочки имеют степень неоднозначности, не превосходящую двух, а множество цепочек с неоднозначностью два или его дополнение задается некоторой однозначной грамматикой, и т.д. Практический интерес представляет также следующая задача, ответ на которую дает доказанная теорема: по двум заданным вероятностным к. с. грамматикам ([2]) установить, совпадают ли для этих грамматик вероятности вывода цепочки длины n для всех n.

Автор благодарит А. А. Мучника за постановку задач и помощь в написании настоящей работы.

Московский государственный университет
им. М. В. Ломоносова

Поступило
4 I 1973

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

[1] Н. Хомский, М. П. Шютценберже, Кибернетич. сборн., нов. сер., в. 3, 1966. [2] K. S. Fu, T. Huang, Int. J. Comput. and Inform Sci., 1, № 2 (1972). [3] У. Рудин, Основы математического анализа, М., 1966. [4] P. Cohen, Comm. on Pure and Appl. Math., 22, № 1 (1969). [5] J. Hopcroft, Math. Syst. Theory, 3, № 2 (1969).