Егорушкин О.И., Колбасина И.В., Попов Н.А., Цокин А.В.

Об интегральном представлении синтаксического многочлена программы

*Сибирский государственный университет науки и технологий
имени академика М.Ф. Решетнёва
(Россия, Красноярск)*

doi:10.18411/spc-22-05-2018-03

idsp: 000001:spc-22-05-2018-03

Аннотация

Разрабатывается подход к синтаксическому анализу контекстно-свободных языков, основанный на интегральном анализе синтаксического многочлена. Результаты имеют приложение в теории формальных языков.

Ключевые слова: формальные степенные ряды, контекстно-свободные языки, синтаксический анализ, интегральное представление.

Исходным объектом в теории формальных языков является алфавит, т.е. множество символов $z_1, \dots, z_n, x_1, \dots, x_m$. Над этими символами определена некоммутативная операция умножения (конкатенации) и коммутативная операция формальной суммы; алфавит вместе с операциями является полукольцом. Символы $x_1, \dots, x_m$ называются терминальными символами и интерпретируются как словарь формального языка, а символы $z_1, \dots, z_n$ называются нетерминальными и нужны для задания совокупности грамматических правил (грамматики), которые порождают язык. По этим правилам определяются правильные мономы от терминальных символов $x_1, \dots, x_m$, которые интерпретируются как правильные предложения языка. Определим также коммутативную операцию умножения символов на комплексные числа, тогда можно рассматривать символьные многочлены и формальные степенные ряды (ФСР) с числовыми коэффициентами. Формальным языком является такой ФСР, членами которого являются все правильные мономы, определенные данной грамматикой [1-4].

Одной из важных проблем, связанных с разработкой систем и языков программирования, является проблема синтаксического анализа программ. Как было отмечено выше, большинство языков программирования является кс-языками, которые можно представить в виде ФСР, поэтому каждая программа, написанная на языке программирования, может рассматриваться как моном соответствующего ФСР. В связи с этим рассмотрим проблему синтаксического анализа мономов кс-языка.

Для того, чтобы сформулировать эту проблему, рассмотрим подробнее систему полиномиальных уравнений Хомского-Шутценберге (2), которая определяет кс-язык. Как известно [3, 4], грамматика кс-языка является множеством правил подстановки

$$z_j \rightarrow q_{j1}(z, x), \dots, z_j \rightarrow q_{jp_j}(z, x), j = 1, \dots, n, \quad (1)$$

где $q_{jk}(z, x)$ является мономом от некоммутативных символьных переменных с числовым коэффициентом, равным единице. Правила подстановки можно применять к начальному символу z_1 , а затем к другим мономам в любом порядке неограниченное число раз, что позволяют выводить новые правильные мономы, образующие кс-язык.

Проблема синтаксического анализа мономов состоит в том, чтобы определить, принадлежит ли моном данному кс-языку, т.е. может ли быть получен из начального символа z_1 при помощи правил подстановки (6), а также установить, какие правила подстановки и сколько раз использовались при выводе этого монома; при этом порядок использования правил подстановки не имеет значения.

Информацию о полиномиальных метках монома можно получить в виде $(n + m)$ -кратного интеграла по циклу, где числа n и m не зависят от степени монома и равны

числу нетерминальных и терминальных символов грамматики κ -языка соответственно.

Рассмотрим коммутативный образ ФСР

$$ci(z_1^*(x, t)) = \sum_{\alpha} s_{\alpha}(t) x^{\alpha}, \quad (2)$$

сгруппированный в кратный ряд Гартогса [5].

Имеет место следующее предложение.

Предложение. При всех мультииндексах α голоморфные в нуле коэффициенты ряда Гартогса $s_{\alpha}(t)$ являются многочленами.

Определение. Синтаксическим многочленом монома ν относительно κ -языка $z_1(x) = z_1^*(x, e)$ называется коэффициент $s_{\alpha}(t)$ ряда Гартогса (2), такой, что $x^{\alpha} = ci(\nu)$.

Замечание 1. Мономиальные метки, содержащиеся в некоммутативных мономах κ -языка, не исчезают при переходе от ФСР к его коммутативному образу (2) и сохраняются в виде мономов синтаксических многочленов, поскольку все коэффициенты ФСР являются целыми положительными числами. Следовательно, если синтаксический многочлен монома относительно κ -языка равен нулю, то моном не принадлежит этому языку.

Замечание 2. Для проведения синтаксического анализа монома ν , такого, что $ci(\nu) = x^{\alpha}$, следует найти синтаксический многочлен $s_{\alpha}(t)$. Каждый моном многочлена $s_{\alpha}(t)$ является произведением мономиальных меток правил подстановки, которые позволяют получить некоторые мономы, имеющие тот же коммутативный образ x^{α} , поэтому для завершения синтаксического анализа монома ν остаётся проверить, можно ли получить его с помощью правил подстановки, соответствующих всем мономам синтаксического многочлена $s_{\alpha}(t)$.

Следующее предложение даёт принципиальную возможность получить синтаксические многочлены $s_{\alpha}(t)$ в виде кратного интеграла по циклу, который может быть вычислен с помощью многомерных вычетов.

Предложение. При всех (x, t) , достаточно близких к нулю, голоморфная в нуле алгебраическая функция $ci(z_1^*(x, t))$, представленная рядом Гартогса (2), задаётся формулой

$$ci(z_1^*(x, t)) = \frac{1}{(2\pi i)^n} \int_{\gamma_z} \frac{z_1 \det(\delta_{ij} - \frac{\partial(ci(Q_i^*(z, x, t)))}{\partial z_j})}{(z - ci(Q^*(z, x, t)))} dz. \quad (3)$$

Вычисляя тейлоровские коэффициенты параметрического интеграла (3), можно получить синтаксические многочлены любого монома, осуществив тем самым его синтаксический анализ.

1. Сафонов К.В., Егорушкин О.И. О синтаксическом анализе и проблеме В.М. Глушкова распознавания контекстно-свободных языков Хомского // Вестник Томского государственного университета. 2006. № 17. С. 63–67.
2. Safonov K.V. On conditions for the sum of a power series to be algebraic and rational // Mathematical Notes. 1987. № 3 (41). P. 185–189.
3. Safonov K.V. On Power Series of Algebraic and Rational functions in C^n // Journal of Math. Analysis and Applications. 2000. V. 243. P. 261–277.
4. Salomaa A., Soittola M. Automata-Theoretic Aspects of Formal Power Series. N.-Y.: Springer Verlag. 1978. 171 p.
5. Семёнов А.Л. Алгоритмические проблемы для степенных рядов и контекстно-свободных грамматик // Доклады АН СССР. 1973. Т. 212. С. 50–52.