Секция «Математическая логика, алгебра и теория чисел»

Групповые грамматики и контекстно-свободные грамматики

Научный руководитель – Пентус Мати Рейнович

Пшеницын Тихон Григорьевич

Студент (специалист)

Московский государственный университет имени М.В.Ломоносова,

Механико-математический факультет, Москва, Россия

E-mail: ptihon@yandex.ru

В работе исследуется соотношение классов языков, задаваемых групповыми грамматиками и контекстно-свободными грамматиками.

Необходимым понятием для определения групповых грамматик является свободная группа. Под этим термином мы будем понимать множество FG конечных строк, состоящих из символов p_i, $i \in \mathbb{N}$, и формальных обратных к ним p_i^{-1}, с операцией конкатенации, в котором при этом имеют место равенства вида $p_i p_i^{-1} = p_i^{-1} p_i = \varepsilon$, где ε обозначает пустую строку. Для элементов свободной группы определяется понятие длины: длиной элемента $w \in FG$ (обозначается $|w|$) называется длина кратчайшей строки, представляющей данный элемент (например, $|p_1 p_1^{-1} p_2| = |p_2| = 1$). Необходимость рассмотрения свободной группы возникает в исчислении Ламбека, где естественным образом определяется перевод типов в ее элементы.

Групповая грамматика (далее ГГ) — это тройка $\langle \Sigma, H, \triangleright \rangle$, где Σ — конечный алфавит, $H \in FG$, $\triangleright \subseteq \Sigma \times FG$ — конечное бинарное отношение, которое понимается как сопоставление символам алфавита некоторого числа типов — элементов свободной группы. По определению, такая грамматика задает язык (будем называть его *групповым*)

$$\{a_1 \ldots a_n \in \Sigma^+ \colon \exists g_1, \ldots, g_n \in FG : a_i \triangleright g_i \wedge g_1 \ldots g_n = H\}.$$

Определение контекстно-свободных грамматик и задаваемых ими языков (далее КСГ) см. в [2].

Основное утверждение работы заключается в том, что классы групповых и контекстно-свободных языков совпадают.

Для доказательства включения групповых языков в контекстно-свободные используется конструкция, близкая представленной в [1] для предгрупповых грамматик. Используя свойства свободных групп, мы строим по ГГ $G = \langle \Sigma, H, \triangleright \rangle$ грамматику $G' = \langle Tp, \Sigma, P, H \rangle$, где Tp — это те элементы FG, a) длина которых не превосходит максимума из длин типов грамматики G и длины H и b) в кратчайшую запись которых входят только те примитивные типы, которые встречаются в кратчайших записях типов из G. Набор правил P имеет вид:

$$P = \{h \to h_1 \ldots h_k \colon h, h_1, \ldots, h_k \in Tp, 1 < k \leq \max\{2, |H|\}, h = h_1 \ldots h_k\} \cup \\ \cup \{t \to a : t \in Tp, a \in \Sigma, a \triangleright t\}.$$

Доказывается, что G и G' задают один и тот же язык.

Для доказательства обратного включения мы рассматриваем КС-грамматику в нормальной форме Розенкрантца (см. [3]), после чего преобразуем её в групповую грамматику и доказываем, что последняя эквивалентна исходной, для чего используем преимущества, которые даёт вышеупомянутая нормальная форма.

Список литературы

[1] Bechet D. *Parsing Pregroup Grammars and Lambek Calculus using Partial Composition*. Studia Logica, 87.2/3, 2007.

[2] А. Е. Пентус, М.Р. Пентус. *Теория формальных языков*. Издательство ЦПИ при механико-математическом факультете МГУ, 2004.

[3] D. J. Rosenkrantz. Matrix equations and normal forms for context-free grammars. Journal of the ACM, 14:3 (1967), 501–507.