Мальцевские чтения
Новосибирск (2006)
Тезисы докладов

Каноническое разложение конкатенации факторных языков

А. Э. Фрид

Пусть u, v — слова над алфавитом Σ. Слово u называется *подсловом*, или *фактором*, слова v, если $v = sut$ для некоторых слов s и t (возможно, пустых). *Факторным замыканием* языка L называется язык $\mathrm{Fac}(L)$ всех подслов слов из L. Язык L называется *факторным*, если если $L =\mathrm{Fac}(L)$.

Напомним, что конкатенация языков определяется естественным образом: $XY = \{xy | x \in X, y \in Y\}$. В данной работе продолжено исследование моноида факторных языков по конкатенации. Заметим, что роль единицы в этом моноиде играет язык $\{\lambda\}$, где λ — пустое слово.

Назовем факторный язык L *неразложимым*, если равенство $L = XY$ для факторных X и Y влечет за собой $L = X$ или $L = Y$. Примерами неразложимых языков являются, например, язык Δ^* для всякого подалфавита $\Delta \subseteq \Sigma$, язык $\{\lambda, a\}$ для всякого $a \in \Sigma$ или язык подслов всякого рекуррентного бесконечного слова.

Разложение $L = L_1 \cdots L_k$, где $L, L_1 \ldots, L_k$ — факторные языки, причем $L \neq \{\lambda\}$, называется *минимальным*, если для всех $i = 1, \ldots, k$ язык L_i непуст, и $L \neq L_1 \cdots L_i' \cdots L_k$ ни для какого подмножества $L_i' \subsetneq L_i$. *Каноническим* называется минимальное разложение на неразложимые сомножители.

В [**1**] были доказаны существование и единственность канонического разложения всякого факторного языка. В [**2**] доказано, что сомножители канонического разложения регулярного факторного языка регулярны.

Обозначим через $\mathcal{C}(X)$ каноническое разложение факторного языка X. В данной работе исследуется следующий вопрос: пусть нам даны факторные языки A и B и их канонические разложения $\mathcal{C}(A)$ и $\mathcal{C}(B)$. Как выглядит $\mathcal{C}(AB)$?

Ответ на этот вопрос может быть весьма далек от последовательно выписанных канонических разложений языков A и B. Пусть, например, $A = a^*$, а $B = (a^* + b^*)^3$ (где $+$ обозначает объединение). Оба эти языка даны в канонических разложениях, однако $\mathcal{C}(AB) = a^*b^*a^*b^*$.

Чтобы сформулировать основной результат, определим подалфавиты $\Pi = \{a \in \Sigma | Aa \subseteq A\}$ и $\Delta = \{a \in \Sigma | aB \subseteq B\}$. От их взаимоотношений зависит вид канонического разложения $\mathcal{C}(AB)$.

Теорема 1. *Пусть A и B — факторные языки, причем $\mathcal{C}(A) = A_1 \cdots A_k$ и $\mathcal{C}(B) = B_1 \cdots B_m$. Тогда верно следующее:*

1. *Если* $\Delta \backslash \Pi \neq \emptyset$ *и* $\Pi \backslash \Delta \neq \emptyset$, *то* $\mathcal{C}(AB) = \mathcal{C}(A)\mathcal{C}(B)$.
2. *Если* $\Delta = \Pi$ *и* $A_k \neq \Delta^*$, $B_1 \neq \Delta^*$, *то* $\mathcal{C}(AB) = \mathcal{C}(A)\mathcal{C}(B)$.
3. *Если* $\Delta = \Pi$ *и* $A_k = \Delta^*$, *то* $\mathcal{C}(AB) = A_1 \cdots A_{k-1}\mathcal{C}(B)$. *Симметрично, если* $\Delta = \Pi$ *и* $B_1 = \Delta^*$, *то* $\mathcal{C}(AB) = \mathcal{C}(A)B_2 \cdots B_m$.
4. *Если* $\Pi \subsetneq \Delta$, *то* $\mathcal{C}(AB) = \mathcal{C}(A')\mathcal{C}(B)$, *где* $A' = Fac(A \backslash A\Delta)$. *Симметрично, если* $\Delta \subsetneq \Pi$, *то* $\mathcal{C}(AB) = \mathcal{C}(A)\mathcal{C}(B')$, *где* $B' = Fac(B \backslash \Pi B)$.

В случае 4 вид канонических разложений языков A' и B' может быть уточнен в терминах $\mathcal{C}(A)$ и $\mathcal{C}(B)$ соответственно.

Следствие 1. *Пусть* A, B *— факторные языки. Тогда* $\mathcal{C}(AB)$ *либо начинается с* $\mathcal{C}(A)$, *либо кончается на* $\mathcal{C}(B)$.

Примеры. Если $A = \{a,b\}^*$ и $B = \{a,c\}^*$, то $\Pi = \{a,b\}$, $\Delta = \{a,c\}$, и $\mathcal{C}(AB) = \{a,b\}^* \cdot \{a,c\}^*$.

Если $A = \mathrm{Fac}\{a,ab\}^*$ и $B = \mathrm{Fac}\{a,ac\}^*$, то $\Pi = \Delta = \{a\}$, и $\mathcal{C}(AB) = \mathrm{Fac}\{a,ab\}^*\mathrm{Fac}\{a,ac\}^*$.

Если $A = a^*$ и $B = \mathrm{Fac}\{a,ab\}^*$, то $\Pi = \Delta = \{a\}$ и $AB = B$.

Если $A = a^*b^*$ и $B = b^*a^*$, то $\mathcal{C}(AB) = a^*b^*a^*$, и здесь не важно, какое именно из вхождений языка b^* было удалено.

Пусть $A = (a^*b^*)^k + (b^*a^*)^k$; тогда $\mathcal{C}(A) = (a^* + b^*)^{2k}$, то есть $A_1 = \cdots = A_{2k} = (a^* + b^*)$ (здесь $+$ обозначает объединение). Конкатенация с $B = a^*$ дает $A' = (a^*b^*)^k$ и $\mathcal{C}(AB) = (a^* \cdot b^*)^k \cdot a^*$.

Полученные результаты дают надежду на решение проблемы коммутации для факторных языков. Рассмотрим уравнение

$$XY = YX.$$

Если X и Y — слова, то $X = Z^n$ и $Y = Z^m$ для некоторого слова Z. Если X и Y — языки общего вида, решение может быть несравненно более сложным: так, М. Кунц [**4**] нашел пример конечного языка X, для которого максимальный язык Y, коммутирующий с ним, не является даже рекурсивно перечислимым. Тем не менее, для некоторых узких семейств языков — в частности, для кодов [**3**] — решение проблемы коммутирования выглядит почти так же просто, как для слов.

Для факторных языков это не так: коммутирующие факторные языки X и Y не обязаны быть степенями одного и того же языка Z. Примером тому служат языки $A = a^*b^*a^*$ и $B = (a^* + b^*)^2$, дающие $AB = BA = a^*b^*a^*b^*a^*$. Этот пример был найден с помощью теоремы 1, которая дает некоторый шанс решить проблему коммутации факторных языков: ведь канонические разложения языков можно сравнивать как слова и применять к ним классические техники работы со словами[1].

Список литературы

[1] S. V. Avgustinovich, A. E. Frid, *A unique decomposition theorem for factorial languages*, Internat. J. Algebra Comput. **15** (2005), 149–160.

[2] S. V. Avgustinovich, A. E. Frid, *Canonical decomposition of a regular factorial language*, in: D. Grigoriev, J. Harrison, E. Hirsch (Eds.), Computer Science - Theory and Applications, Springer, 2006 (LNCS 3967), 18–22.

[3] J. Karhumäki, M. Latteux, I. Petre, *Commutation with codes*, Theoret. Comput. Sci. **340** (2005), 322-333.

[1] При поддержке РФФИ (гранты 05-01-00364 и 06-01-00694).

[4] M. Kunc, *The power of commuting with finite sets of words*, in: Theoretical Aspects of Computer Science (STACS'05), Springer, 2005 (LNCS 3404), 569–580.

Институт математики им. С. Л. Соболева СО РАН
E-mail: frid@math.nsc.ru