MR2030508 (2004j:68101) 68Q45 68R15

Mateescu, Alexandru (R-BUCHM);
Salomaa, Arto (FIN-TURK-TC); **Yu, Sheng** (3-WON-C)

Subword histories and Parikh matrices. (English. English summary)

J. Comput. System Sci. **68** (2004), *no.* 1, 1–21.

Let u and w be words over a linearly ordered alphabet $\Sigma = \{a_1, \ldots, a_n\}$ with $a_1 < \cdots < a_n$. Then u is a "subword" of w if there exist words $x_1, \ldots, x_k, y_0, \ldots, y_k \in \Sigma^\star$ such that $u = x_1 \ldots x_k$ and $w = y_0 x_1 y_1 \ldots x_k y_k$. Let $|w|_u$ denote the number of times that u ($u \neq \lambda$; λ is the empty word) occurs as a subword of w, and $|w|_\lambda = 1$.

If $\mathcal{M}_{n+1}$ denotes the set of triangular integer matrices of dimension $n+1$, then the "Parikh matrix mapping" Ψ_n is the morphism $\Psi_n\colon \Sigma^\star \to \mathcal{M}_{n+1}$ defined by: if $\Psi_n(a_k) = (m_{i,j})_{1 \le i,j \le n+1}$, then $m_{i,i} = 1$ $(1 \le i \le n+1)$, $m_{k,k+1} = 1$, and $m_{i,j} = 0$ in all other cases. Then for a word w, the superdiagonal $(m_{i,i+1})_{1 \le i \le n}$ of $\Psi_n(w)$ equals the Parikh vector $(|w|_a)_{a \in \Sigma}$ of w. The next diagonals provide information about the order of symbols in w.

Consider a word w over Σ. A "subword history" (SH) in Σ and its "value" are defined recursively by: (i) Each $u \in \Sigma^\star$ is a subword history in Σ and its value in w is $|w|_u$. (ii) If S_1 and S_2 are subword histories with values α_1 and α_2 respectively, then so are $-S_1$, $S_1 + S_2$ and $S_1 \times S_2$ with values $-\alpha_1$, $\alpha_1 + \alpha_2$ and $\alpha_1 \alpha_2$, respectively.

The authors establish a normal form (involving $-$ and $+$ only) for SHs and show the decidability of the equivalence problem for SHs; they use minors of Parikh matrices to prove some inequalities between SHs, among which are special cases of the general Cauchy inequality. Finally, some observations about the structure of Parikh matrices are made. *Peter R. J. Asveld* (NL-TWEN-C)

[References]

Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.

1. T. Harju, O. Ibarra, J. Karhumäki, A. Salomaa, Some decision problems concerning semilinearity and commutation, J. Comput. System Sci. 65 (2002) 278–294. MR1939772 (2003i:68070)
2. J. Honkala, On Parikh slender languages and power series, J. Comput. System Sci. 52 (1996) 185–190. MR1375813 (97b:68108)
3. J. Karhumäki, Generalized Parikh mappings and homomorphisms, Inform. Control 47 (1980) 155–164. MR0617486 (83c:68095)

4. W. Kuich, A. Salomaa, Semirings, Automata, Languages, Springer, Berlin, Heidelberg, New York, 1986. MR0817983 (87h:68093)
5. A. Mateescu, A. Salomaa, K. Salomaa, S. Yu, A sharpening of the Parikh mapping, Theoret. Inform. Appl. 35 (2001) 551–564. MR1922295 (2003c:68130)
6. J.J. Pansiot, A decidable property of iterated morphisms, Springer Lecture Notes in Computer Science 104 (1981) 152–158.
7. R.J. Parikh, On context-free languages, J. Assoc. Comput. Mach. 13 (1966) 570–581. MR0209093 (34 #8901)
8. G. Rozenberg, A. Salomaa (Eds.), Handbook of Formal Languages 1–3, Springer, Berlin, Heidelberg, New York, 1997. MR1469992 (98d:68001)
9. J. Sakarovitch, I. Simon, Subwords, in: M. Lothaire (Ed.), Combinatorics on Words, Addison-Wesley, Reading, MA, 1983, pp. 105–142. MR0675953 (84g:05002)
10. A. Salomaa, Formal Languages, Academic Press, New York, 1973. MR0438755 (55 #11661)