Parikh Matrices

Mike Müller

Institut für Informatik
Christian-Albrechts-Universität zu Kiel
D-24098 Kiel, Germany

23. April 2013

Parikh Vectors

Let $\Sigma = \{a_1, a_2, \ldots, a_k\}$ be an ordered alphabet.

The **Parikh vector** of a word $w \in \Sigma^*$ is defined as

$$\Psi(w) = (|w|_{a_1}, |w|_{a_2}, \ldots, |w|_{a_n}).$$

Example: $\Sigma = \{a, b, c\}, w = abbcabac, \Psi(w) = (3, 3, 2)$

Mapping $w \mapsto \Psi(w)$ is obviously **not injective**.

Every $v \in \mathbb{N}^k$ is Parikh vector of a word $w \in \Sigma^*$.

Parikh Matrices

Let $\Sigma = \{a_1, a_2, \ldots, a_k\}$ be an ordered alphabet.
For simplicity, let $k = 3$ and $\Sigma = \{a, b, c\}$ here.

The **Parikh matrix** of a word $w \in \Sigma^*$ is a 4×4 upper triangular matrix, defined as

$$\Psi_M(w) = \begin{pmatrix} 1 & |w|_a & |w|_{ab} & |w|_{abc} \\ 0 & 1 & |w|_b & |w|_{bc} \\ 0 & 0 & 1 & |w|_c \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Then for a word $w = w_1 w_2 w_3 \cdots w_n$, we have:

$$\Psi_M(w) = \Psi_M(w_1)\Psi_M(w_2)\Psi_M(w_3)\cdots\Psi_M(w_n)$$

Properties of the Parikh Matrix Mapping

Mapping $w \mapsto \Psi_M(w)$ is also not injective:

$$\Psi_M(cab) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \Psi_M(acb)$$

But this mapping is not surjective:

$$M = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is **not** a parikh matrix!

Problem: Is a given Matrix M a Parikh Matrix?

Easy for binary Alphabet: M is parikh matrix iff $|w|_a \cdot |w|_b \geq |w|_{ab}$.

For larger alphabets: Only algorithm with runtime polynomial in $|w|$ known.

But input (matrix) is of size $\mathcal{O}(\log |w|)$!

$\Rightarrow$ Exponential in inputsize!

Can we do better?