The prefix normal form of a binary word

Ferdinando Cicalese ∗ Gabriele Fici † Zsuzsanna Lipták ‡

Let \( w = w_1w_2\cdots w_n \) be a word over the alphabet \( \Sigma = \{a, b\} \). For every factor \( u \) of \( w \) we define the Parikh vector of \( u \) as the vector \( p_u = (|u|_a, |u|_b) \in \mathbb{N}^2 \) counting the number of occurrences of the letters of \( \Sigma \) in \( u \).

For any \( 1 \leq i \leq n \) one can define
\[
f_i(w) = \max\{|u|_a : u \text{ a factor of } w \text{ of length } i\},
\]
that is the maximum number of \( a \)'s occurring in a factor of \( w \) of length \( i \).

\textbf{Definition 1.} A finite word \( w = w_1w_2\cdots w_n \) over the alphabet \( \Sigma = \{a, b\} \) is in Prefix Normal Form if for every \( 1 \leq i \leq n \) one has \( |w_1w_2\cdots w_i|_a = f_i(w) \).

That is, a word \( w \) is in PNF if every factor of \( w \) contains no more \( a \)'s than the prefix of \( w \) of the same length.

\textbf{Example 1.} The word \( aabbaab \) is in PNF, while the word \( w = aabbaaba \) is not.

We state the following problem.

\textbf{Problem 1.} Let \( w \) be a word over \( \Sigma \) of length \( n > 0 \). Test whether \( w \) is in PNF.

We present a geometrical construction allowing to solve Problem 1 in time \( O(n^2) \). First draw \( w \) over a 2-dimensional grid in which each \( a \) is represented by an increasing segment and each \( b \) is represented by a decreasing segment. Then draw in the same grid, starting from the same initial point, all the suffixes of \( w \). The word \( w \) is then in PNF if and only if all the suffixes of \( w \) stay below \( w \).

![Figure 1: A binary word represented on a grid.](image-url)