

Estruturas de Dados

Cristina Gomes Fernandes

Implementação de árvores de sufixo

Árvore para o texto $s[1..n]$.

Cada nó x tem os seguintes campos:

- $filho(x, c)$: filho correspondente ao caracter c
- $pai(x)$: pai de x
- $ini(x)$: índice inicial do rótulo da aresta $(x, pai(x))$
- $fim(x)$: índice final do rótulo da aresta $(x, pai(x))$
- $suf(x)$: índice de um sufixo (apenas para as folhas)

Busca em árvore de sufixo

BUSQUE (r, p, i, n, s):

Recebe um $i \leq n$ e busca $p[i..n]$ na árvore $r \neq \text{NIL}$.

Devolve uma tripla (b, x, i', j) onde
se $b = \text{VERDADE}$,

conseguiu soletrar $p[i..n]$ na árvore,
tendo terminado de soletrar no nó x ,
na letra de índice j de s ; neste caso, $i' = n$;
senão conseguiu soletrar $p[i..i' - 1]$ na árvore,
mas ocorreu um *mismatch* no nó x entre $s[j]$ e $p[i']$.

Busca em árvore de sufixo

BUSQUE (r, p, i, n, s) \triangleright busca $p[i..n]$ na árvore $r \neq \text{NIL}$, com $i \leq n$

- 1 **se** $\text{filho}(r, p[i]) = \text{NIL}$
- 2 **então devolva** (**FALSO**, $r, i, 0$)
- 3 $x \leftarrow \text{filho}(r, p[i])$
- 4 $j \leftarrow \text{ini}(x)$
- 5 **enquanto** $j \leq \text{fim}(x)$ **e** $i \leq n$ **faça**
- 6 **se** $s[j] = p[i]$
- 7 **então** $i \leftarrow i + 1$ $j \leftarrow j + 1$
- 8 **senão devolva** (**FALSO**, x, i, j)
- 9 **se** $i > n$
- 10 **então devolva** (**VERDADE**, $x, i - 1, j - 1$)
- 11 **senão devolva** **BUSQUE**(x, p, i, n, s)

Complexidade: $O(n)$

(n é o comprimento do padrão buscado)

Construção de árvore de sufixo

CONSTRUA (s, m): constrói a árvore de sufixo para $s[1..m]$.

Acrescente o símbolo \$ no final de s ajustando m .

Comece com uma única aresta rotulada por $s[1..m]$.

Para $i \leftarrow 2$ até m faça

 soletre $s[i..m]$ na árvore corrente

 se terminar no meio de uma aresta

 subdivida-a criando um novo nó e reparta o rótulo

 da aresta original onde terminou de soletrar

 crie um novo filho para o nó corrente e

 ponha como rótulo a parte não soletrada de $s[i..m]$

Construção de árvore de sufixo

CONSTRUA (s, m)

```
1   $r \leftarrow \text{INICIALIZA}(s, m)$ 
2   $i \leftarrow 2$ 
3  enquanto  $i \leq m$  faça
4     $(b, x, i', j) \leftarrow \text{BUSQUE}(r, s, i, n, s)$ 
5    se  $j \neq 0$   $\triangleright$  quebra a aresta no meio
6      então  $y \leftarrow \text{NOVACÉLULA}()$   $\triangleright$  apontadores NIL
7         $c \leftarrow s[\text{ini}(x)]$ 
8         $\text{filho}(\text{pai}(x), c) \leftarrow y$ 
9         $\text{pai}(y) \leftarrow \text{pai}(x)$ 
10        $\text{ini}(y) \leftarrow \text{ini}(x)$ 
11        $\text{fim}(y) \leftarrow j - 1$ 
12        $\text{pai}(x) \leftarrow y$ 
13        $\text{ini}(x) \leftarrow j$ 
14        $\text{filho}(y, s[j]) \leftarrow x$ 
```

Construção de árvore de sufixo

CONSTRUA (s, m)

1 ...

3 **enquanto** $i \leq m$ **faça**

4 $(b, x, j) \leftarrow \text{BUSQUE}(r, p, i, n, s)$

5 **se** $j \neq 0$ ▷ quebra a aresta no meio

6 **então** ...

15 **senão** $y \leftarrow x$

16 $z \leftarrow \text{NOVACÉLULA}()$

17 $\text{filho}(y, s[i]) \leftarrow z$

18 $\text{pai}(z) \leftarrow y$

19 $\text{ini}(z) \leftarrow i'$

20 $\text{fim}(z) \leftarrow m$

21 $\text{suf}(z) \leftarrow i$

22 $i \leftarrow i + 1$

23 **devolva** r

Complexidade: $O(m^2)$