# Back-To-Front Online Lyndon Forest Construction

Golnaz Badkobeh, Maxime Crochemore, *Jonas Ellert*, Cyril Nicaud

Goldsmiths
UNIVERSITY OF LONDON

Université Gustave Eiffel
INSTITUT D'ÉLECTRONIQUE ET D'INFORMATIQUE GASPARD-MONGE

technische universität dortmund

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

b a n a n a

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

a
ana
anana
**banana**
na
nana

b a n a n a

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$\forall i \in [2, |w|] : w \prec w[i..|w|]$

a
ana
anana
**banana**
na
nana

b a n a n a

**not Lyndon** ✗

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

a
ana
anana
**banana**
na
nana

b a n a n a

**not Lyndon** ✗

a n a n a s

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

a
ana
anana
**banana**
na
nana

b a n a n a

**not Lyndon ✗**

a n a n a s

**ananas**
anas
as
nanas
nas
s

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

a
ana
anana
**banana**
na
nana

b a n a n a

**not Lyndon** ✗

a n a n a s

**Lyndon** ✓

**ananas**
anas
as
nanas
nas
s

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

b a n a n a o r a n a n a s

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$$\forall i \in [2, |w|] : w \prec w[i..|w|]$$

b a n a n a o r a n a n a s

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$\forall i \in [2, |w|] : w \prec w[i..|w|]$

b a n a n a o r a n a n a s

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$\forall i \in [2, |w|] : w \prec w[i..|w|]$

b a n a n a o r a n a n a s

5

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

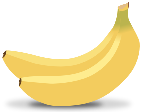
$\forall i \in [2, |w|] : w \prec w[i..|w|]$

| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Lyn$ = [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] |

$w$ is Lyndon $\iff$ $w$ is lex. smaller than all its suffixes

$\forall i \in [2, |w|] : w \prec w[i..|w|]$

| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Lyn =$ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |



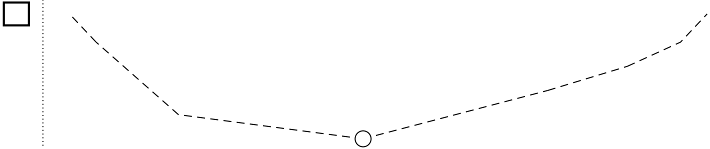
| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| *Lyn* | = [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] |

Lyndon forest: recursively find longest proper Lyndon suffix

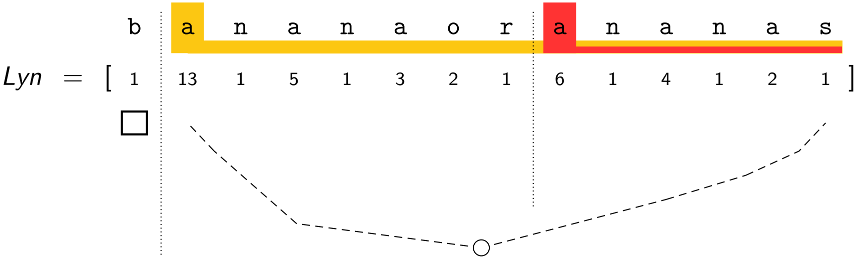
| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| *Lyn* = [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] |
| | □ | | | | | | | | | | | | | |

Lyndon forest: recursively find longest proper Lyndon suffix

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| *Lyn* = | [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] |

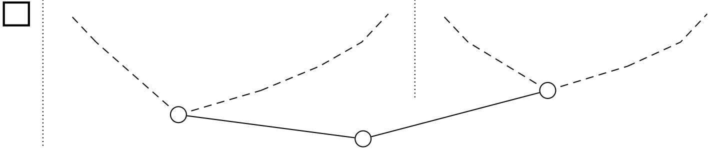


Lyndon forest: recursively find longest proper Lyndon suffix

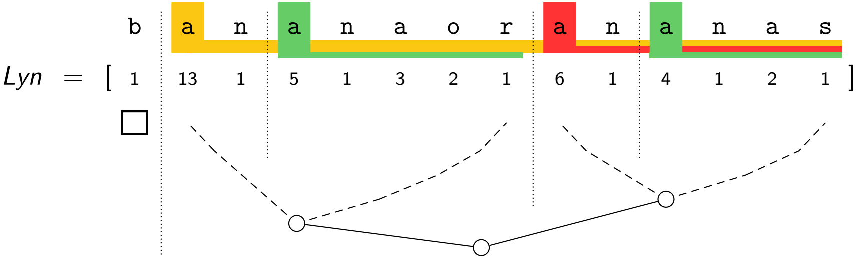



Lyndon forest: recursively find longest proper Lyndon suffix

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| *Lyn* = | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |



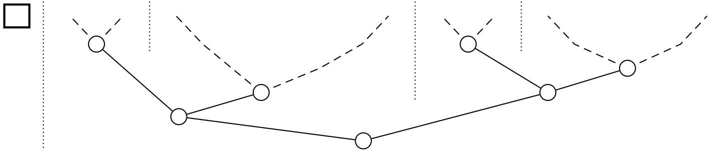
Lyndon forest: recursively find longest proper Lyndon suffix



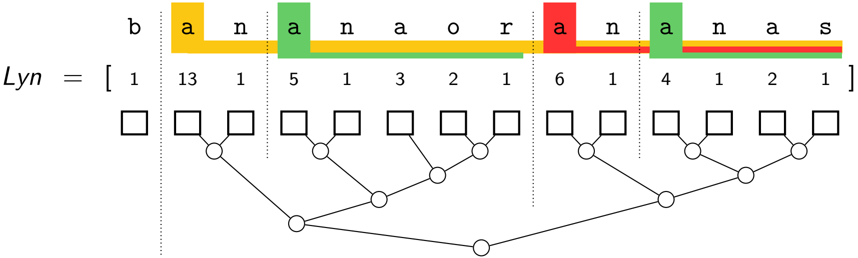

Lyndon forest: recursively find longest proper Lyndon suffix

| b | a | n | a | n | a | o | r | a | n | a | n | a | s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

$Lyn = [\ 1\ \ 13\ \ 1\ \ 5\ \ 1\ \ 3\ \ 2\ \ 1\ \ 6\ \ 1\ \ 4\ \ 1\ \ 2\ \ 1\ ]$



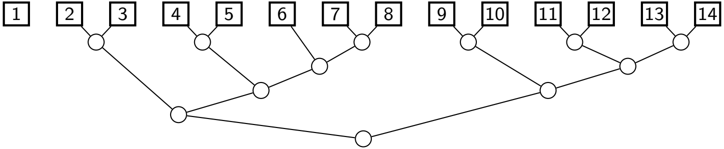
Lyndon forest: recursively find longest proper Lyndon suffix




Lyndon forest: recursively find longest proper Lyndon suffix

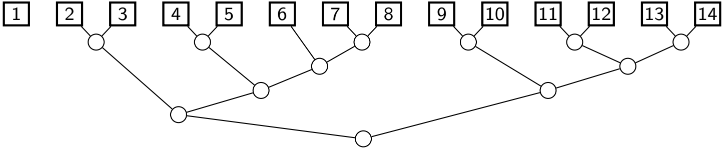
b a n a n a o r a n a n a s

$Lyn = [\ 1\ \ 13\ \ 1\ \ 5\ \ 1\ \ 3\ \ 2\ \ 1\ \ 6\ \ 1\ \ 4\ \ 1\ \ 2\ \ 1\ ]$



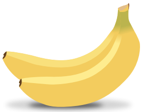

Lyndon forest: recursively find longest proper Lyndon suffix

| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Lyn$ = [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] |



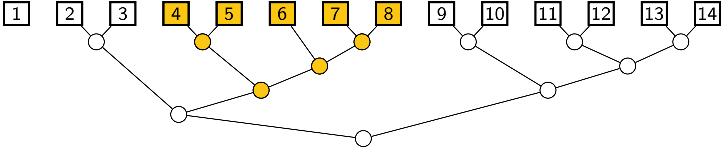

**For any $i$, leaves $\{i, \ldots, i + Lyn[i] - 1\}$ form a subtree.**

Lyndon forest: recursively find longest proper Lyndon suffix



b a n a n a o r a n a n a s

$Lyn = [\; 1 \;\; 13 \;\; 1 \;\; 5 \;\; 1 \;\; 3 \;\; 2 \;\; 1 \;\; 6 \;\; 1 \;\; 4 \;\; 1 \;\; 2 \;\; 1 \;]$




**For any $i$, leaves $\{i, \ldots, i + Lyn[i] - 1\}$ form a subtree.**

Lyndon forest: recursively find longest proper Lyndon suffix

| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Lyn =$ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| *Lyn* = | [ 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] |

**Goal:** Compute *Lyn* back-to-front online in $\mathcal{O}(n)$ time.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | `b` | `a` | `n` | `a` | `n` | `a` | `o` | `r` | `a` | `n` | `a` | `n` | `a` | `s` |
| *Lyn* = | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

**Goal:** Compute *Lyn* back-to-front online in $\mathcal{O}(n)$ time.

**Why?**

| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $Lyn =$ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

**Goal:** Compute *Lyn* back-to-front online in $\mathcal{O}(n)$ time.

**Why?**

- find all maximal repetitions (runs) [Bannai et al. 17, Ellert+Fischer 21]

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| *Lyn* = | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

**Goal:** Compute *Lyn* back-to-front online in $\mathcal{O}(n)$ time.

**Why?**
- find all maximal repetitions (runs) [Bannai et al. 17, Ellert+Fischer 21]
- full-text indexing (Lyndon SLPs) [Tsuruta et al. 20]

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| $Lyn =$ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

**Goal:** Compute *Lyn* back-to-front online in $\mathcal{O}(n)$ time.

**Why?**

- find all maximal repetitions (runs) [Bannai et al. 17, Ellert+Fischer 21]
- full-text indexing (Lyndon SLPs) [Tsuruta et al. 20]
- because it is combinatorially insightful and fun

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | b | a | n | a | n | a | o | r | a | n | a | n | a | s |
| $Lyn$ = | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 |

**Goal:** Compute $Lyn$ back-to-front online in $\mathcal{O}(n)$ time.

**Why?**

- find all maximal repetitions (runs) [Bannai et al. 17, Ellert+Fischer 21]
- full-text indexing (Lyndon SLPs) [Tsuruta et al. 20]
- because it is combinatorially insightful and fun

**Previously:** Compute $Lyn$ front-to-back in $\mathcal{O}(n)$ time, but not online [Bille et al. 20]

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| $Lyn =$ | [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] | |

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *Lyn* = | | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 | |
| *nss* = | | | | | | | | | | | | | | | | |

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\}$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *Lyn* = | [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 | ] |
| *nss* = | [ | | | | | | | | | | | | | | | ] |

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\}$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| *Lyn* = | [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 | ] |
| *nss* = | [ | | | | | | | | | | | | | | | ] |

[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| $Lyn$ = [ | | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 ] | |
| $nss$ = [ | | | | | 9 | | | | | | | | | | ] | |

[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

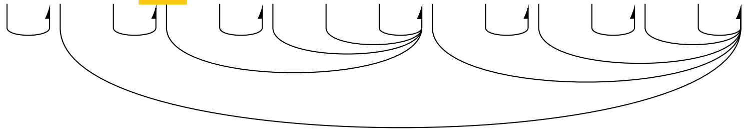
- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| *Lyn* = | [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 | ] |
| *nss* = | [ | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 | ] |

[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *Lyn* = | [ | 1 | 13 | 1 | 5 | 1 | 3 | 2 | 1 | 6 | 1 | 4 | 1 | 2 | 1 | ] |
| *nss* = | [ | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 | ] |

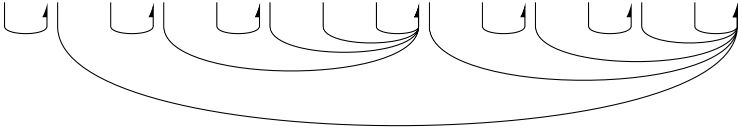


[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |

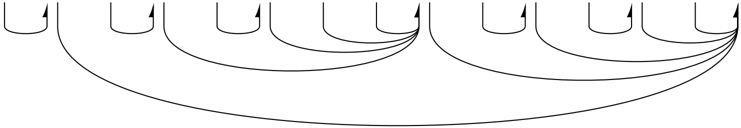
$nss = [\ 2\ \ 15\ \ 4\ \ 9\ \ 6\ \ 9\ \ 9\ \ 9\ \ 15\ \ 11\ \ 15\ \ 13\ \ 15\ \ 15\ ]$



[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \land w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ ] | | | | | | | | | | | | | | | | |
| *nss* = [ ] | | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 | |



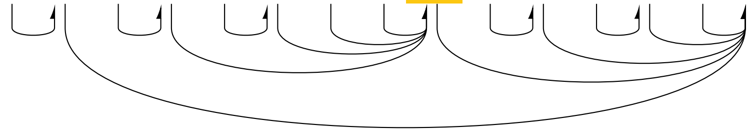
[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$
- previous smaller suffixes: $pss[i] = \min\{j \mid j < i \wedge w[j..|w|] \prec w[i..|w|]\}$

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |

$pss = [\quad]$

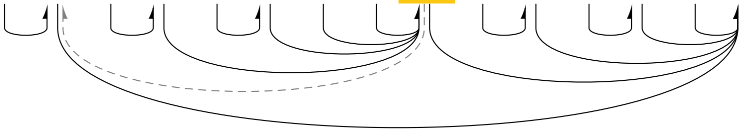
$nss = [\ 2\ \ 15\ \ 4\ \ 9\ \ 6\ \ 9\ \ 9\ \ 9\ \ 15\ \ 11\ \ 15\ \ 13\ \ 15\ \ 15\ ]$



[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$
- previous smaller suffixes: $pss[i] = \min\{j \mid j < i \wedge w[j..|w|] \prec w[i..|w|]\}$

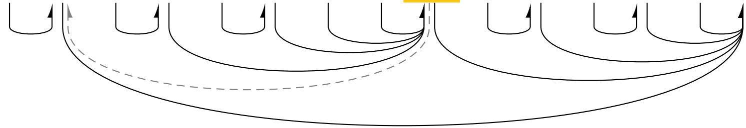
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = | [ | | | | | | | | | 2 | | | | | ] | |
| *nss* = | [ | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$
- previous smaller suffixes: $pss[i] = \min\{j \mid j < i \wedge w[j..|w|] \prec w[i..|w|]\}$

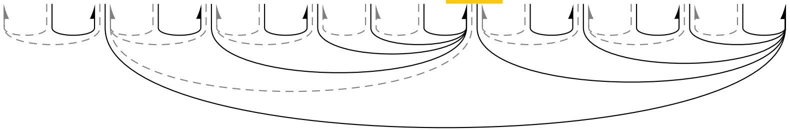
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = | [ | 0 | 0 | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 | ] |
| *nss* = | [ | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 | ] |



[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$
- previous smaller suffixes: $pss[i] = \min\{j \mid j < i \wedge w[j..|w|] \prec w[i..|w|]\}$

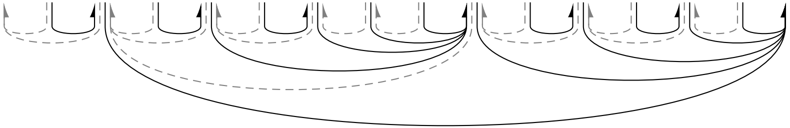
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| *pss* = [ | | 0 | 0 | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



[Hohlweg+Reutenauer 03]
[Franek et al. 16]
[Franek+Liut 19]

- next smaller suffixes: $nss[i] = \min\{j \mid j > i \wedge w[j..|w|] \prec w[i..|w|]\} = i + Lyn[i]$
- previous smaller suffixes: $pss[i] = \min\{j \mid j < i \wedge w[j..|w|] \prec w[i..|w|]\}$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | 0 | 0 | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | 2 | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



**for** $i = n$ **down to** 1 **do**

$\quad j \leftarrow i + 1$

$\quad$ **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**

$\qquad pss[j] \leftarrow i$

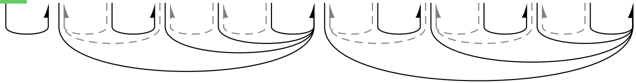
$\qquad j \leftarrow nss[j]$

$\quad nss[i] \leftarrow j$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | | | | | 4 | 4 | 6 | 7 | | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |

```
for i = n down to 1 do
    j ← i + 1

    while w[i..|w|+1] ≺ w[j..|w|+1] do
        pss[j] ← i
        j ← nss[j]

    nss[i] ← j
```

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | | | | | 4 | 4 | 6 | 7 | | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



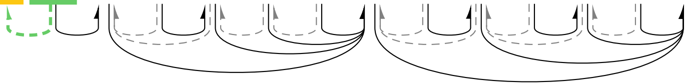
**for** $i = n$ **down to** 1 **do**
  $j \leftarrow i + 1$

  **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**
    $pss[j] \leftarrow i$
    $j \leftarrow nss[j]$

  $nss[i] \leftarrow j$

ananao.. ≺ nanaor..

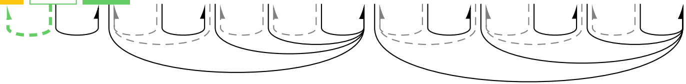
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* | = [ | | | | 2 | | 4 | 4 | 6 | 7 | | 9 | 9 | 11 | 11 | 13 | ] |
| *nss* | = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 | ] |



```
for i = n down to 1 do
    j ← i + 1

    while w[i..|w| + 1] ≺ w[j..|w| + 1] do
        pss[j] ← i
        j ← nss[j]

    nss[i] ← j
```

ananao.. ≺ nanaor..

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| *pss* = [ | | | | 2 | | 4 | 4 | 6 | 7 | | 9 | 9 | 11 | 11 | 13 | ] |
| *nss* = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 | ] |



**for** $i = n$ **down to** 1 **do**
  $j \leftarrow i + 1$

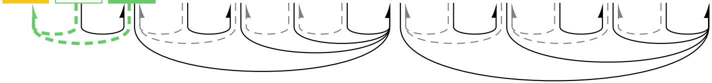
  **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**
    $pss[j] \leftarrow i$
    $j \leftarrow nss[j]$

  $nss[i] \leftarrow j$

ananao.. $\prec$ nanaor..

ananao.. $\prec$ anaora..

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| $pss$ = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | | 9 | 9 | 11 | 11 | 13 ] | |
| $nss$ = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



**for** $i = n$ **down to** 1 **do**
  $j \leftarrow i + 1$

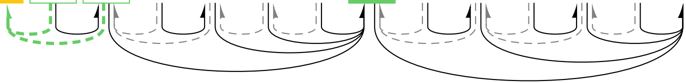
  **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**
    $pss[j] \leftarrow i$
    $j \leftarrow nss[j]$

  $nss[i] \leftarrow j$

ananao.. $\prec$ nanaor..

ananao.. $\prec$ anaora..

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



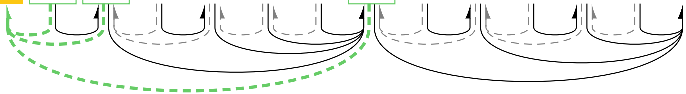
```
for i = n down to 1 do
    j ← i + 1

    while w[i..|w| + 1] ≺ w[j..|w| + 1] do
        pss[j] ← i
        j ← nss[j]

    nss[i] ← j
```

$\texttt{ananao..} \prec \texttt{nanaor..}$

$\texttt{ananao..} \prec \texttt{anaora..}$

$\texttt{ananao..} \prec \texttt{ananas\$}$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



**for** $i = n$ **down to** 1 **do**
  $j \leftarrow i + 1$

  **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**
    $pss[j] \leftarrow i$
    $j \leftarrow nss[j]$

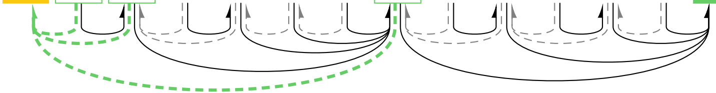
  $nss[i] \leftarrow j$

`ananao..` $\prec$ `nanaor..`

`ananao..` $\prec$ `anaora..`

`ananao..` $\prec$ `ananas$`

| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| *pss* | = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* | = [ | | | ? | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



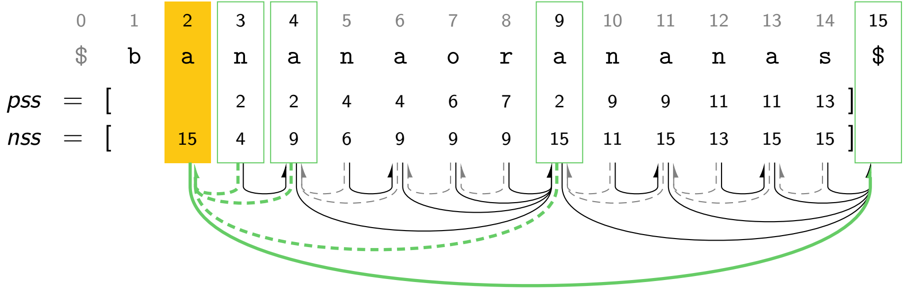
```
for i = n down to 1 do
    j ← i + 1

    while w[i..|w| + 1] ≺ w[j..|w| + 1] do
        pss[j] ← i
        j ← nss[j]

    nss[i] ← j
```

ananao.. ≺ nanaor..

ananao.. ≺ anaora..

ananao.. ≺ ananas\$

ananao.. ≺ \$




**for** $i = n$ **down to** 1 **do**

  $j \leftarrow i + 1$

  **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**

    $pss[j] \leftarrow i$

    $j \leftarrow nss[j]$

  $nss[i] \leftarrow j$

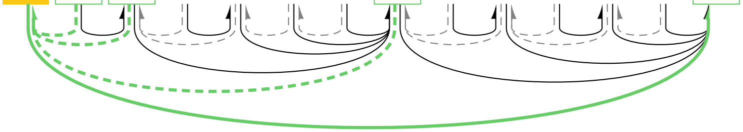
ananao.. $\prec$ nanaor..

ananao.. $\prec$ anaora..

ananao.. $\prec$ ananas$

ananao.. $\prec$ $

| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| $pss$ | = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| $nss$ | = [ | | | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



```
for i = n down to 1 do
    j ← i + 1

    while w[i..|w|+1] ≺ w[j..|w|+1] do
        pss[j] ← i
        j ← nss[j]

    nss[i] ← j
```

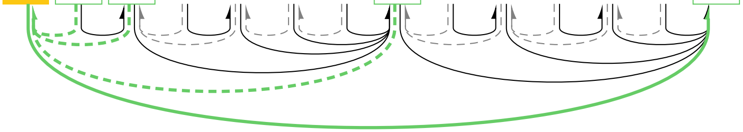
ananao.. $\prec$ nanaor.. $lce(2,3) = 0$

ananao.. $\prec$ anaora.. $lce(2,4) = 3$

ananao.. $\prec$ ananas$ $lce(2,9) = 5$

ananao.. $\prec$ $ $lce(2,15) = 0$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | \$ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | \$ |
| $pss =$ [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| $nss =$ [ | | | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



**for** $i = n$ **down to** 1 **do**
    $j \leftarrow i + 1$
    $\ell \leftarrow lce(i, j)$
    **while** $w[i..|w|+1] \prec w[j..|w|+1]$ **do**
        $pss[j] \leftarrow i$
        $j \leftarrow nss[j]$
        $\ell \leftarrow lce(i, j)$
    $nss[i] \leftarrow j$

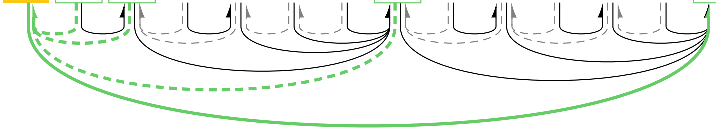
ananao.. $\prec$ nanaor..  $lce(2,3) = 0$

ananao.. $\prec$ anaora..  $lce(2,4) = 3$

ananao.. $\prec$ ananas\$  $lce(2,9) = 5$

ananao.. $\prec$ \$  $lce(2,15) = 0$

| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



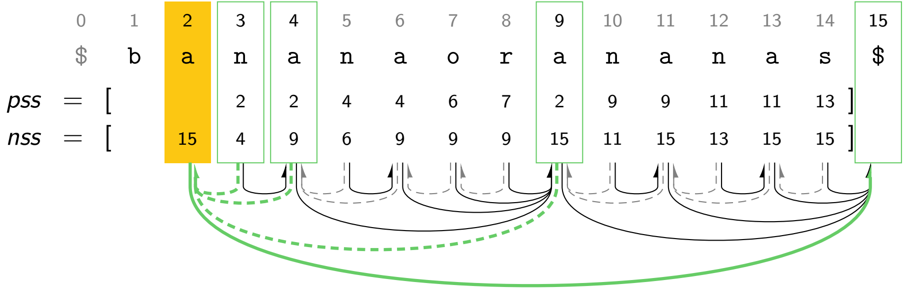
```
for i = n down to 1 do
    j ← i + 1
    ℓ ← lce(i, j)
    while w[i + ℓ] < w[j + ℓ] do
        pss[j] ← i
        j ← nss[j]
        ℓ ← lce(i, j)
    nss[i] ← j
```

ananao.. $\prec$ nanaor.. $lce(2,3) = 0$

ananao.. $\prec$ anaora.. $lce(2,4) = 3$

ananao.. $\prec$ ananas$ $lce(2,9) = 5$

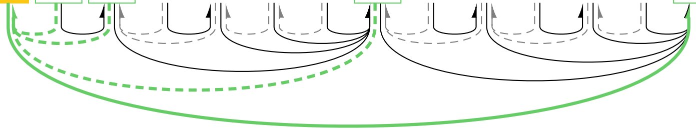
ananao.. $\prec$ $ $lce(2,15) = 0$




```
for i = n down to 1 do
    j ← i + 1
    ℓ ← lce(i, j)
    while w[i + ℓ] < w[j + ℓ] do
        pss[j] ← i
        j ← nss[j]
        ℓ ← lce(i, j)
    nss[i] ← j
```

Worst case $\mathcal{O}(n^2)$ time, $\mathcal{O}(n)$ iterations.

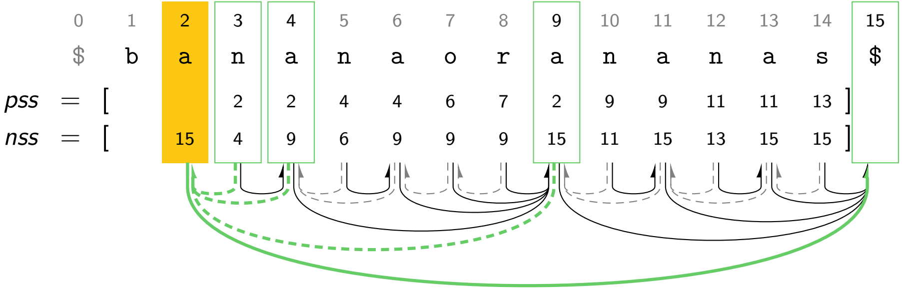
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $ | b | a | n | a | n | a | o | r | a | n | a | n | a | s | $ |
| *pss* = [ | | | | 2 | 2 | 4 | 4 | 6 | 7 | 2 | 9 | 9 | 11 | 11 | 13 ] | |
| *nss* = [ | | | 15 | 4 | 9 | 6 | 9 | 9 | 9 | 15 | 11 | 15 | 13 | 15 | 15 ] | |



```
for i = n down to 1 do
    j ← i + 1
    ℓ ← lce(i, j)
    while w[i + ℓ] < w[j + ℓ] do
        pss[j] ← i
        j ← nss[j]
        ℓ ← lce(i, j)
    nss[i] ← j
```

Worst case $\mathcal{O}(n^2)$ time, $\mathcal{O}(n)$ iterations.

Theorem 1: Expected $\mathcal{O}(n)$ time, if $w$ is drawn uniformly at random from $\Sigma^n$ (for any $\Sigma$ with $|\Sigma| \geq 2$).




```
for i = n down to 1 do
    j ← i + 1
    ℓ ← lce(i, j)
    while w[i + ℓ] < w[j + ℓ] do
        pss[j] ← i
        j ← nss[j]
        ℓ ← lce(i, j)
    nss[i] ← j
```

Worst case $\mathcal{O}(n^2)$ time, $\mathcal{O}(n)$ iterations.

Theorem 1: Expected $\mathcal{O}(n)$ time, if $w$ is drawn uniformly at random from $\Sigma^n$ (for any $\Sigma$ with $|\Sigma| \geq 2$).

Theorem 2: Extra tricks lead to $\mathcal{O}(n)$ worst-case time.

# Trick 1: Getting the first LCE for free

**for** $i = n$ **down to** $1$ **do**

  $j \leftarrow i + 1$

  $\ell \leftarrow lce(i, j)$

  **while** $w[i + \ell] < w[j + \ell]$ **do**

    $pss[j] \leftarrow i$

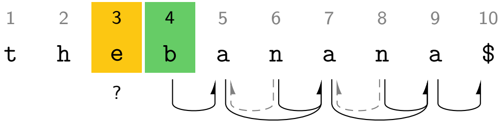
    $j \leftarrow nss[j]$

    $\ell \leftarrow lce(i, j)$

  $nss[i] \leftarrow j$

# Trick 1: Getting the first LCE for free

**for** $i = n$ **down to** $1$ **do**
    $j \leftarrow i + 1$
    $\ell \leftarrow lce(i, j)$
    **while** $w[i + \ell] < w[j + \ell]$ **do**
        $pss[j] \leftarrow i$
        $j \leftarrow nss[j]$
        $\ell \leftarrow lce(i, j)$
    $nss[i] \leftarrow j$



# Trick 1: Getting the first LCE for free

**for** $i = n$ **down to** $1$ **do**

  $j \leftarrow i + 1$

  $\ell \leftarrow lce(i, j)$

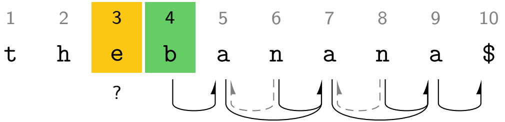
  **while** $w[i + \ell] < w[j + \ell]$ **do**

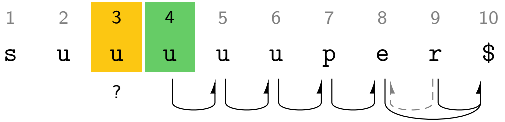
    $pss[j] \leftarrow i$

    $j \leftarrow nss[j]$

    $\ell \leftarrow lce(i, j)$

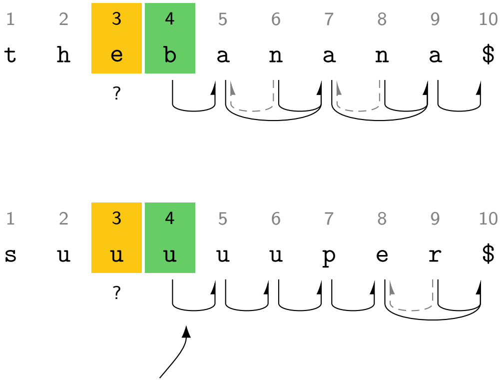
  $nss[i] \leftarrow j$



# Trick 1: Getting the first LCE for free

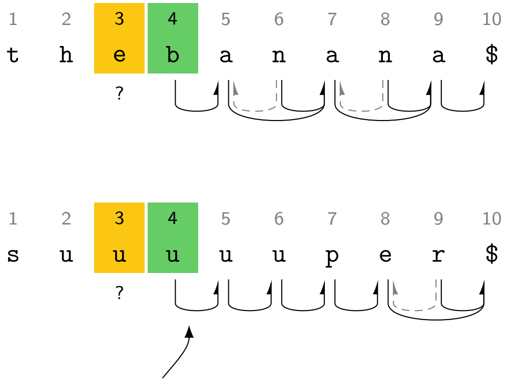
**for** $i = n$ **down to** 1 **do**
    $j \leftarrow i + 1$
    $\ell \leftarrow lce(i, j)$
    **while** $w[i + \ell] < w[j + \ell]$ **do**
        $pss[j] \leftarrow i$
        $j \leftarrow nss[j]$
        $\ell \leftarrow lce(i, j)$
    $nss[i] \leftarrow j$




$lce(4, 5)$ known from previous computation

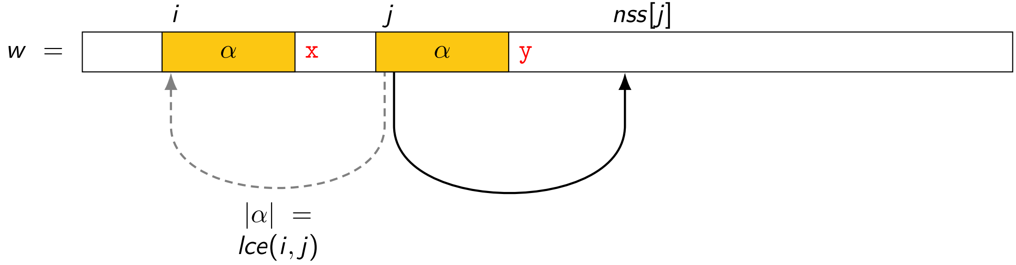
# Trick 1: Getting the first LCE for free

**for** $i = n$ **down to** 1 **do**
  $j \leftarrow i + 1$
  **if** $w[i] \neq w[j]$ **then** $\ell \leftarrow 0$
  **else** $\ell \leftarrow 1 + lce(j, j+1)$
  **while** $w[i+\ell] < w[j+\ell]$ **do**
    $pss[j] \leftarrow i$
    $j \leftarrow nss[j]$
    $\ell \leftarrow lce(i, j)$
  $nss[i] \leftarrow j$

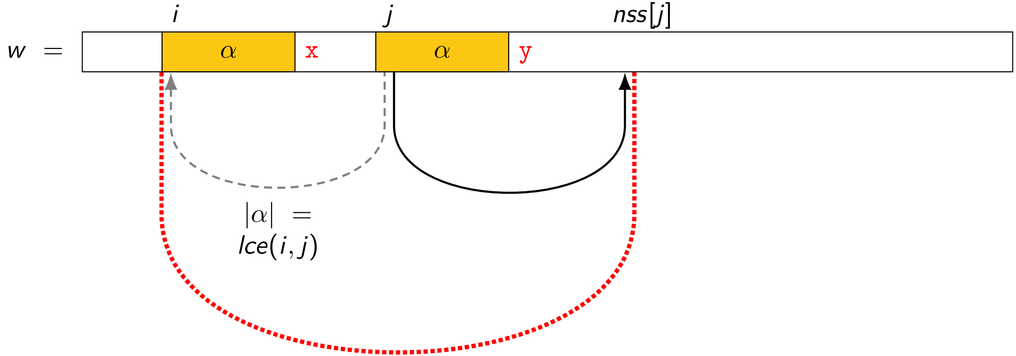



$lce(4,5)$ known from previous computation

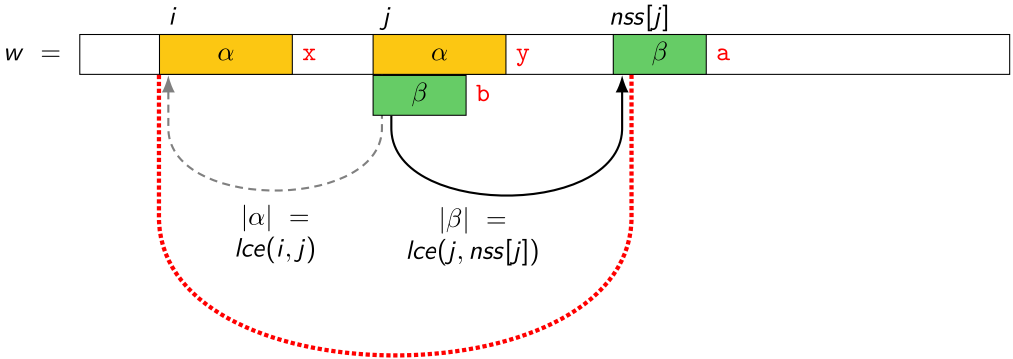
# Trick 2: Getting other LCEs more efficiently



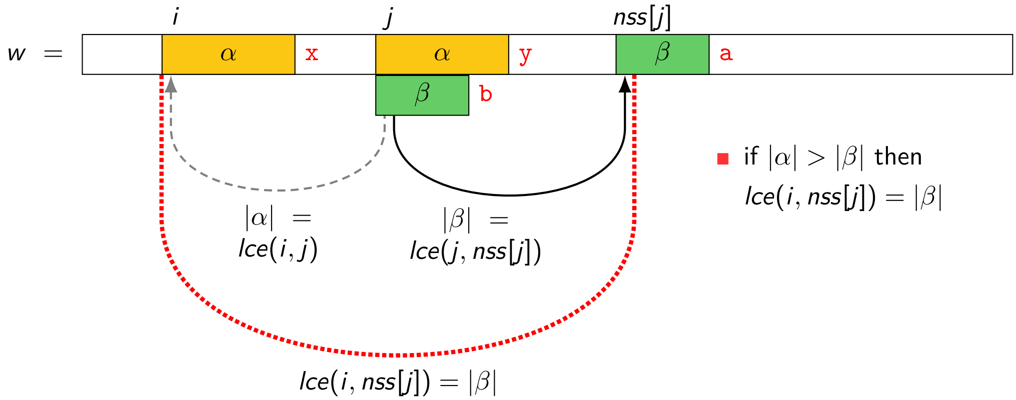
# Trick 2: Getting other LCEs more efficiently



# Trick 2: Getting other LCEs more efficiently

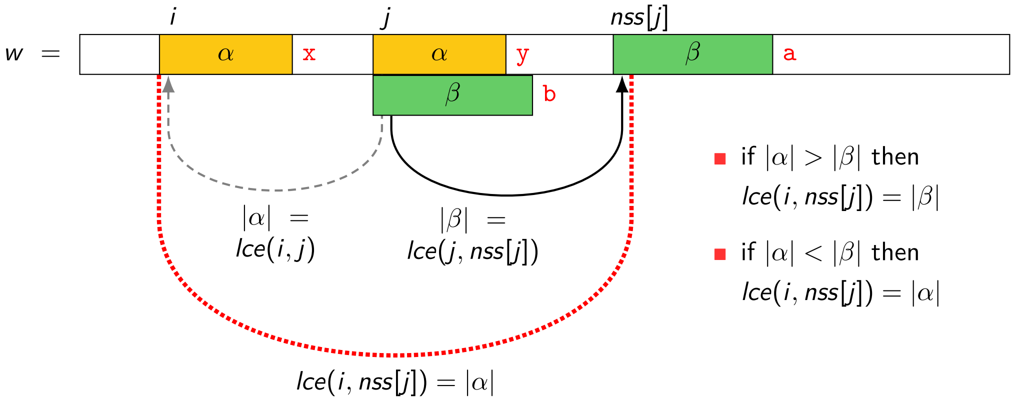


# Trick 2: Getting other LCEs more efficiently



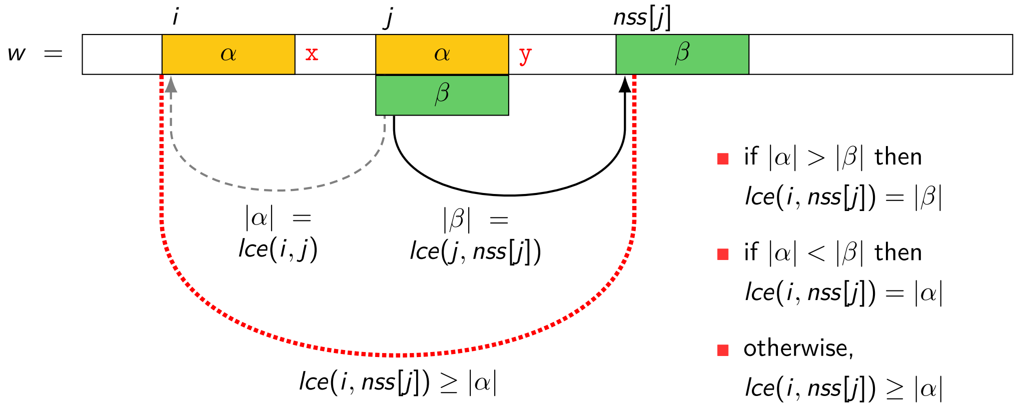

- if $|\alpha| > |\beta|$ then $lce(i, nss[j]) = |\beta|$

# Trick 2: Getting other LCEs more efficiently



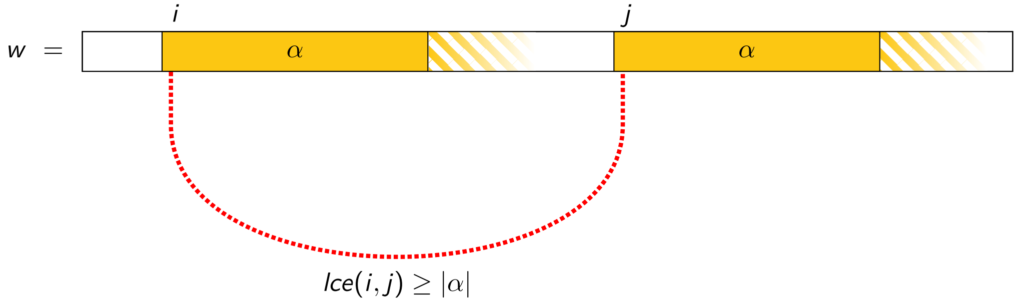

- if $|\alpha| > |\beta|$ then
  $lce(i, nss[j]) = |\beta|$
- if $|\alpha| < |\beta|$ then
  $lce(i, nss[j]) = |\alpha|$

# Trick 2: Getting other LCEs more efficiently

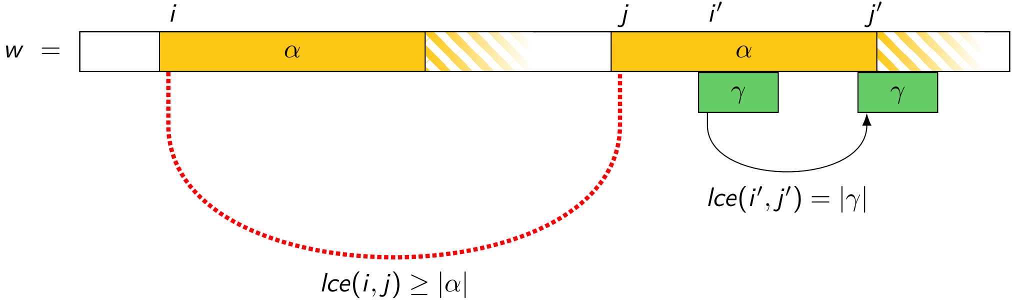



- if $|\alpha| > |\beta|$ then
  $lce(i, nss[j]) = |\beta|$
- if $|\alpha| < |\beta|$ then
  $lce(i, nss[j]) = |\alpha|$
- otherwise,
  $lce(i, nss[j]) \geq |\alpha|$

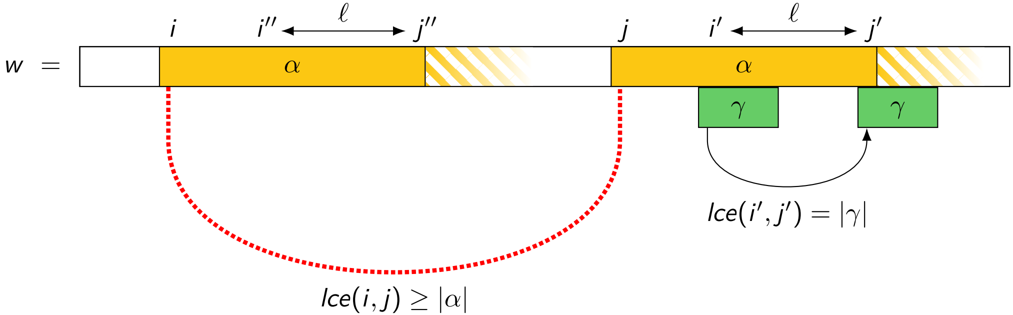
# Trick 3: Exploiting previously computed LCEs



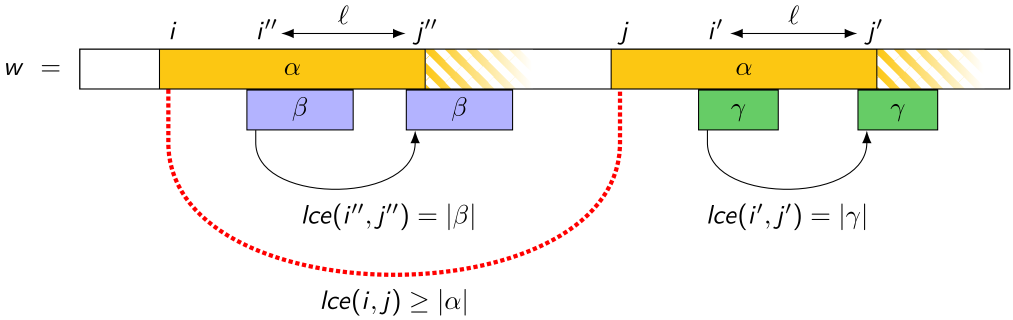
# Trick 3: Exploiting previously computed LCEs



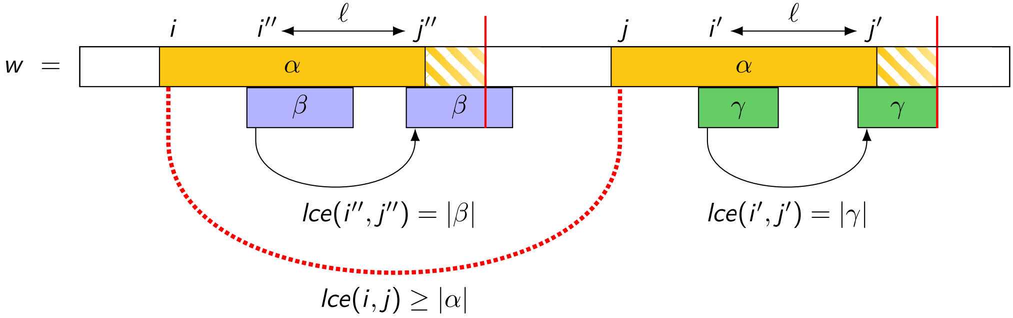
# Trick 3: Exploiting previously computed LCEs



# Trick 3: Exploiting previously computed LCEs

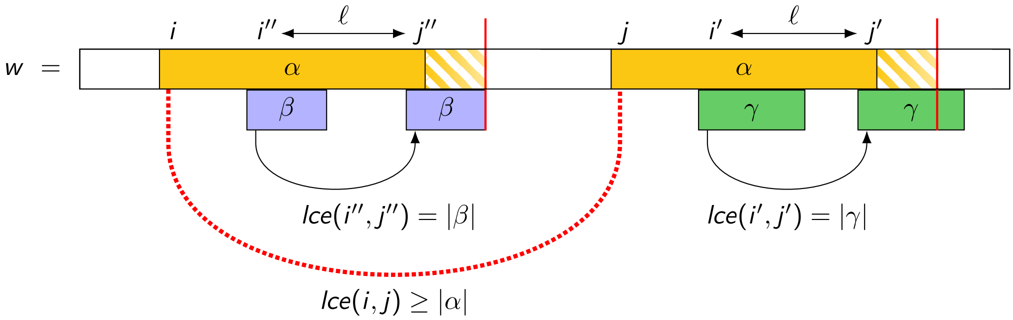


# Trick 3: Exploiting previously computed LCEs



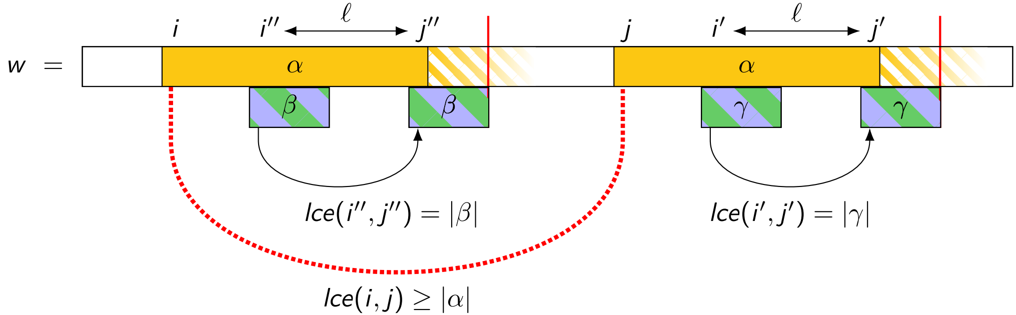

- if $|\beta| > |\gamma|$ then $lce(i, j) = |\alpha| + |\gamma| - 1$

# Trick 3: Exploiting previously computed LCEs



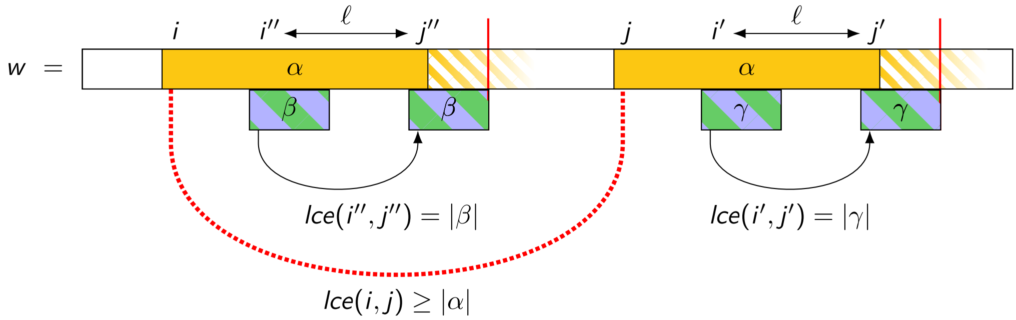

- if $|\beta| > |\gamma|$ then $lce(i, j) = |\alpha| + |\gamma| - 1$
- if $|\beta| < |\gamma|$ then $lce(i, j) = |\alpha| + |\beta| - 1$

# Trick 3: Exploiting previously computed LCEs




- if $|\beta| > |\gamma|$ then $lce(i, j) = |\alpha| + |\gamma| - 1$
- if $|\beta| < |\gamma|$ then $lce(i, j) = |\alpha| + |\beta| - 1$
- otherwise, $lce(i, j) \geq |\alpha| + |\beta| - 1$

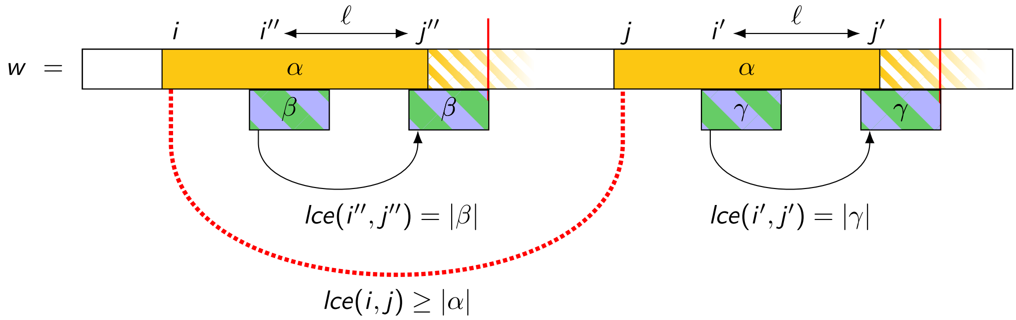
# Trick 3: Exploiting previously computed LCEs




**With a bit of extra care:**

- if $|\beta| > |\gamma|$ then $lce(i,j) = |\alpha| + |\gamma| - 1$
- if $|\beta| < |\gamma|$ then $lce(i,j) = |\alpha| + |\beta| - 1$
- otherwise, $lce(i,j) \geq |\alpha| + |\beta| - 1$

# Trick 3: Exploiting previously computed LCEs

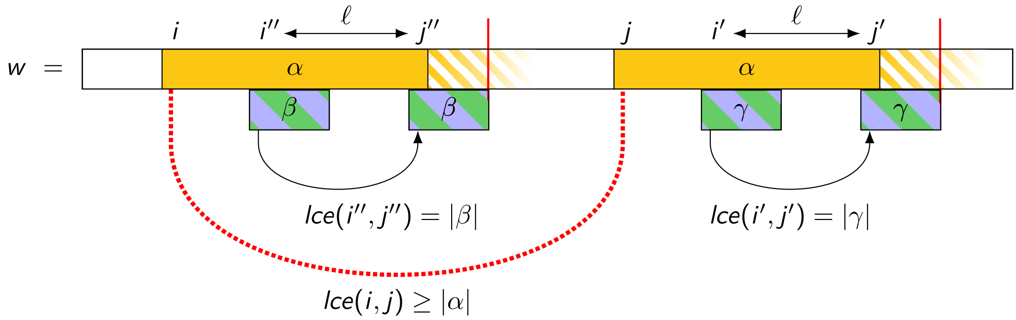



**With a bit of extra care:**

- there are always suitable $i'$ and $i''$

- if $|\beta| > |\gamma|$ then $lce(i, j) = |\alpha| + |\gamma| - 1$
- if $|\beta| < |\gamma|$ then $lce(i, j) = |\alpha| + |\beta| - 1$
- otherwise, $lce(i, j) \geq |\alpha| + |\beta| - 1$

# Trick 3: Exploiting previously computed LCEs




**With a bit of extra care:**

- there are always suitable $i'$ and $i''$
- the case $|\beta| = |\gamma|$ happens $\mathcal{O}(n)$ times

- if $|\beta| > |\gamma|$ then $lce(i, j) = |\alpha| + |\gamma| - 1$
- if $|\beta| < |\gamma|$ then $lce(i, j) = |\alpha| + |\beta| - 1$
- otherwise, $lce(i, j) \geq |\alpha| + |\beta| - 1$

# Summary

**Results for online Lyndon table computation:**

- naive algorithm: expected $\mathcal{O}(n)$ time, but worst case $\mathcal{O}(n^2)$
- can be improved to $\mathcal{O}(n)$ in the worst case

**Not in the presentation:**

- proof of expected time bound
- combining the tricks to achieve optimal time

**In the future:**

- use the online algorithm for detecting repetitions and for indexing

**Code on Github!**

# Summary

**Results for online Lyndon table computation:**

- naive algorithm: expected $\mathcal{O}(n)$ time, but worst case $\mathcal{O}(n^2)$
- can be improved to $\mathcal{O}(n)$ in the worst case

**Not in the presentation:**

- proof of expected time bound
- combining the tricks to achieve optimal time

**In the future:**

- use the online algorithm for detecting repetitions and for indexing

**Code on Github!**

**Thanks for listening!**