

Let M = banana

0. a) Append sentinel character

M = banana\0

1) a) Append three zeros at end

M = banana\0\0\0  
0 1 2 3 4 5 6 7 8 9 10 ← indices

b) Find R<sub>12</sub>-idx

R<sub>12</sub>-idx → [1 2 4 5 7 8]  
→ [1 4 7 2 5 8]

c) Convert to string triplets and sort

R<sub>12</sub>-str → [ana ana\0\0\0 naa na\0\0\0]  
→ [0\0\0\0\0 ana ana\0\0\0 naa\0\0\0]

\* Since there are duplicates, we do find its recursive suffix array, which will be the R<sub>12</sub> order

d) Group into buckets

R<sub>12</sub>-str → [0\0\0\0\0 ana ana\0\0\0 naa]  
start at → 1 2 3 4 ← bucket index

e) Form recursive message M<sup>(2)</sup> from R<sub>12</sub>-idx

R<sub>12</sub>-idx → [1 4 7 2 5 8]  
M<sup>(2)</sup> → [2 2 1 4 3 1] ← convert these into their corresponding bucket indices

2) a) Append three zeros at end

M<sup>(2)</sup> → 2 2 1 4 3 1 0 0 0  
0 1 2 3 4 5 6 7 8

b) Find R<sub>12</sub>-idx

R<sub>12</sub>-idx<sup>(2)</sup> → [1 2 4 5]

c) Convert to string triplets and sort

R<sub>12</sub>-str<sup>(2)</sup> → [214 310 143 100]  
→ [100 143 214 310] ← R<sub>12</sub>(2)

\* Since there are no duplicates, we can directly get R<sub>12</sub>(2)

d) Find R<sub>12</sub>

R<sub>12</sub>(2) → [5 2 1 4]

e) Find R<sub>0</sub>(2)-idx

R<sub>0</sub>(2)-idx → [0 3 6]

f) Convert to 2-ary tuples and sort

R<sub>0</sub>-str<sup>(2)</sup> → [20 43 06]  
→ [06 20 43]

value in M<sup>(2)</sup> index in M<sup>(2)</sup>

g) Get second elements as R<sub>0</sub>(2)

R<sub>0</sub>(2) → [6 0 3]

h) Merge R<sub>0</sub>(2) and R<sub>12</sub>(2)

[6 0 3] [5 2 1 4]

5 i) [6 0 3] [5 2 1 4]  
M<sup>(2)</sup> → 2 M<sup>(2)</sup> → 1

2 ii) [6 0 3] [5 2 1 4]  
M<sup>(2)</sup> → 2 M<sup>(2)</sup> → 1

1 iii) [6 0 3] [5 2 1 4]  
M<sup>(2)</sup> → 2 M<sup>(2)</sup> → 2

0 iv) [6 0 3] [5 2 1 4]  
M<sup>(2)</sup> → 2 M<sup>(2)</sup> → 3

4 v) [6 0 3] [5 2 1 4]  
M<sup>(2)</sup> → 4 M<sup>(2)</sup> → 3

3 vi) [6 0 3] [5 2 1 4]

i) Collate suffix array

SA<sup>(2)</sup> → [5 2 1 0 4 3]

3) a) Convert SA<sup>(2)</sup> back to R<sub>12</sub> using R<sub>12</sub>-idx

SA<sup>(2)</sup> → [5 2 1 0 4 3] ← treat these as positions in R<sub>12</sub>-idx

R<sub>12</sub>-idx → [1 4 7 2 5 8]

R<sub>12</sub> → [8 7 4 1 5 2]

b) Find R<sub>0</sub>-idx

R<sub>0</sub>-idx → [0 3 6]

c) Convert to 2-ary tuples and sort

R<sub>0</sub>-str → [b0 a3 a6]  
→ [a3 a6 b0]

\* Since there are duplicate first elements, we have to omit these duplicates into their next elements

R<sub>0</sub>-str → [a3 a6 b0]

sort these

R<sub>0</sub>-str sub A → [a3 a6] → [a3 76]

replace with second element plus one

→ [43 76] → [23 16]

replace with position in R<sub>12</sub> plus one

→ [23 16]

sort and merge with original R<sub>12</sub>-str

→ [1 6 23]

R<sub>0</sub>-str → [16 23 b0]

d) Get second elements as R<sub>0</sub>

R<sub>0</sub> → [6 3 0]

\* M<sup>(2)</sup> in R<sub>0</sub> is smaller, so push this value to the suffix array  
\* M<sup>(2)</sup> for both indices are the same  
\* R<sub>12</sub> mod 3 is 1, so compare relative orders in R<sub>12</sub>  
\* pos of R<sub>0</sub> is smaller, so push this value to the suffix array

push remaining elements at end of suffix array

no more elements!

4) Merge R0 and R12

M = banana\0000  
 0 1 2 3 4 5 6 7 8 9 10

R0 [630] R12 [874152]  
 ^ ^  
 ignore this

7 [630] [874152]  
 ^ ^  
 M → a M → 0

6 [630] [874152]  
 ^ ^  
 M → a M → a  
 next idx → 7 next idx → 5  
 M → 0 M → n

· X. Min R12 is smaller, so push this value to the suffix array

· X. n for both values are the same!

· X. R12 mod 3 is 2, so compare M of next indices

· X. M of R0 is smaller, so push this value to the suffix array

3 [630] [874152]  
 ^ ^  
 M → a M → a  
 next idx → 4 next idx → 5  
 M → a M → n

· X. M for both values are the same!

· X. R12 mod 3 is 2, so compare M of next indices

· X. M of R0 is smaller, so push this value to the suffix array

1 [630] [874152]  
 ^ ^  
 M → b M → a

2 [630] [874]52]  
 ^ ^  
 M → b M → a

0 [630] [874152]  
 ^ ^  
 M → b M → n

5 [630] [874152]  
 ^ ^  
 no more elements! push remaining elements at end of suffix array

5) Collate suffix array

SA → [7 6 3 4 1 0 5 2]

6) Map suffix array against M to get the BWT

SA → [7 6 3 4 1 0 5 2]

6 5 2 3 0 1 4 2 ← these are indices in M

M → banana\0  
 0 1 2 3 4 5 6 7

BWT(M) → annab\0aa