Assignment 2.
ZHANG Chongzhi 1155077072.

1. (a). Suffix trie.   Location.

S = GTAACTGTAGTG $ 1

Suffix: TAACTGTAGTG $ 2

AACTGTAGTG $ 3

ACTGTAGTG $.   4

CTGTAGTG $ 5

TGTAGTG $ 6

GTAGTG $ 7

TAGTG $ 8

AGTG $ 9

GTG $ 10

TG $ 11

G $ 12

$ 13.
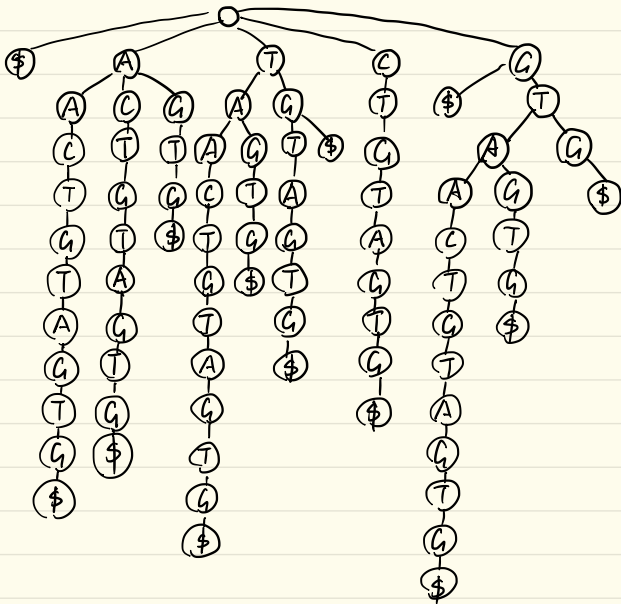
(b). suffix tree.
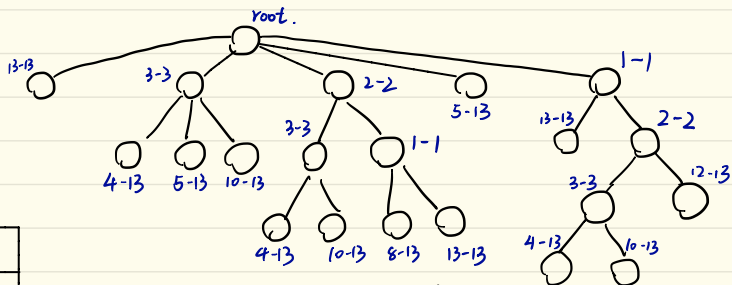
(c). suffix array

1	13	8	7
2	3	9	10
3	4	10	2
4	9	11	8
5	5	12	11
6	12	13	6
7	1		

Step: ①. Sort each level of the tree in lexicographic order.
② Find the left most Leaf each time, calculate
the Location of the suffix by using (the start location of
the leaf) − (The depth of the leaf), and fill
the Location to the array one by one.

(d). ① First gives all the suffixes.

```
  1 2 3 4 5 6 7 8 9 0 1 2 3
  G T A A C T G T A G T G $     1
    T A A C T G T A G T G $     2
      A A C T G T A G T G $     3
        A C T G T A G T G $.    4
          C T G T A G T G $     5
            T G T A G T G $     6
              G T A G T G $     7
                T A G T G $     8
                  A G T G $     9
                    G T G $     10
                      T G $     11
                        G $     12
                          $     13.
```

② and then sort it in lexicographic order.

$	13
AA CTGTAGTG $	3
ACTG TAGTG $	4
AGTG $	9
CTGTAGTG $	5
G $	12
GTAA CTGTAGTG $	1
GTAGTG $	7
GTG $	10
TAA CTGTAGTG $	2
TAGTG $	8
TG $	11
TG TAGTG $	6

(e). Because BTW is the last column of the sorted matrix.
Thus each character of BTW is the one before [starting position of suffix array].
Thus for each position saved in Suffix array, Find that position in S and return the character before that position.

For instance, the first position is 13, and we put 13 - 1 = 12th character, which is G to the first position of b.
The following do the same procedure.

we have   b = G T A T A T $ T A G G G C

(f). ①. Find all the rotation of $, and put them in a matrix.

②. Sort the rows of matrix by Lexicographic order.

$GTAACTGTAGTG

AACTGTAGTG$GT

ACTGTAGTG$GTA

AGTG$GTAACTGT

CTGTAGTG$GTAA

G$GTAACTGTAGT

GTAACTGTAGTG$

GTAGTG$GTAACT

GTG$GTAACTGTA

TAACTGTAGTG$G

TAGTG$GTAACTG

TG$GTAACTGTAG

TGTAGTG$GTAAC

③. Find the last column.
  we have b= GTATAT$TAGGGC

(g)   f   b   ①. Get the first column of Matrix by sorting b is Lexicographic order.
      $ G   ②. Mark the order of each character of both b and f.
③.   A₁   T₁   In order to Find GTA,
      A₂   A₁
      A₃   T₂   ③. Find As in f, before A, it is a T. Thus T₁ and T₂ has the
      C   A₂   chance.
⑤.   G₁   T₃   ④. Find T₁ and T₂ in f. before T, it is a G. Thus G₁ and G₂
      G₂ $ meet the requirment.
      G₃   T₄
      G₄   A₃   ⑤. Find the position of G₁ and G₂ in f. return the corrisponding position in suffix
④.   T₁   G₁   carray, which is 1 and 7.
      T₂   G₂
      T₃   G₃
      T₄   C

(h). No.

Consider the first row of the sorted Matrix, the first character of first row must be a $.

$ ←
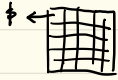
Thus, for all n > 0, the last character of first row have no chance to be a $ because there is only one $ in a row.

Sence there is a position in b that $ can never appear, there's no doubt that the $ does not have a equal probability of all n+1 positions in b.