

String Structure

Resource

- <https://usaco.guide/adv/string-search?lang=cpp>
 - มีสรุป+โจทย์ให้ทำ
- <https://cp-algorithms.com/>
- Geekforgeeks
- Codeforces, etc.

Suffix Array

- $\mathcal{O}(N \log^2 N)$ version
- $\mathcal{O}(N \log N)$ version
- Longest Common Prefix (LCP) Array in $\mathcal{O}(N)$

Suffix Tree

- Building a suffix tree in $\mathcal{O}(N)$ with Ukkonen's Algorithm
 - ไม่น่าออกในปฏิบัติ ถึงออกก็ใช้อัลกอริทึมในข้อต่อไปเขียนง่ายกว่า
 - ใช้เวลาอ่านนานมาก เก็บไว้เรียนเป็นอย่างท้าย ๆ หรือไม่ต้องเรียนเลยก็ได้
- Building a suffix tree from a suffix array in $\mathcal{O}(N)$

Algorithms

- Aho-Corasick Algorithm
 - หลักการเดียวกับ KMP แต่เราสามารถรับมือกับ string หลายอันได้
 - $\mathcal{O}(\text{ผลรวมความยาว})$
 - Example: <https://open.kattis.com/contests/dnjvd8/problems/taboo>
- Manacher's Algorithm
 - Find number of palindrome substrings
 - Find longest palindrome substrings
 - Time complexity $\mathcal{O}(N)$
- Palindromic Tree / Eertree
 - **Problems involving palindromes of a string and its substrings.**
 - It can solve problems like 'longest palindrome in a string', 'count of plaindromic substrings' etc.
 - Example : https://oj.uz/problem/view/APIO14_palindrome
- Lyndon Factorization
 - Find maximum/minimum cyclic shift of a string

- $\mathcal{O}(N)$
- https://cp-algorithms.com/string/lyndon_factorization.html
- Extra : มีส่ร่อยลูกบ้ดอยู่หลายอัน ลูกบ้ดแต่ละอันมีตัวอักษรภาษาอังกฤษเขียนเอาไว้ เราสามารถ represent ส่ร่อยบ้ดเหล่านี้ด้วย string ที่มาจากตัวอักษรของเหล่าลูกบ้ดตามเข้มนาฬิกา(โดยเริ่มจากลูกไหนก็ได้ แล้วไล่ตามเข้มนาฬิกาจนครบทุกลูก แล้วหยุด)
 - เช่นมีส่ร่อยที่มีลูกบ้ดเรียงตามเข้มนาฬิกาเป็น 'a','b','c','d' (ลูกบ้ด 'd' อยู่ติดกับลูกบ้ด 'a')
 - เราสามารถ represent เป็น "abcd","bcda","cdab" หรือ "dabc" ก็ได้ (สังเกตว่าส่ร่อยอันเดียวสามารถ represent ได้หลายแบบ)
 - โจทย์: ให้ string representation มาทั้งหมด N string ($1 \leq N \leq 10^6$) โดย string representation บางอันอาจจะมาจากส่ร่อยลูกบ้ดอันเดียวกันกับอันอื่นก็ได้ จงหาว่า string representation ทั้งหมดนี้มาจากส่ร่อยลูกบ้ดที่แตกต่างกันทั้งหมดกี่อัน