

Description

n-gram เป็นข้อมูลชนิดหนึ่งซึ่งเป็นการวิเคราะห์รูปแบบที่ติดกันของตัวอักษรในแฟ้มเอกสาร โดย n จะแทนความยาวของรูปแบบที่สนใจตัวอย่างเช่น สมมุติว่าเอกสารชิ้นหนึ่งมีข้อความที่ปรากฏคือ 'WOWWOWO'

ถ้าพิจารณาว่า $n = 1$ หรือ 1-gram จะได้ว่าข้อความนี้มี 1-gram (ตัวอักษรตัวเดียว) อยู่ 2 แบบ คือ W และ O โดยความถี่การเกิด $W = 4$ และ $O = 3$

ถ้าพิจารณาว่า $n = 2$ หรือ 2-gram จะได้ว่าข้อความนี้มี 2-gram (ตัวอักษร 2 ตัวที่อยู่ติดกัน) อยู่ 3 แบบ คือ WO และ OW และ WW โดยความถี่การเกิด $WO = 3$ และ $OW = 2$ และ $WW = 1$

ถ้าพิจารณาว่า $n = 3$ หรือ 3-gram จะได้ว่าข้อความนี้มี 3-gram (ตัวอักษร 3 ตัวที่อยู่ติดกัน) อยู่ 4 แบบ คือ WOW และ OWW และ WWO และ OWO โดยความถี่การเกิด $WOW = 2$ และ $OWW = 1$ และ $WWO = 1$ และ $OWO = 1$

เมื่อกำหนดข้อความที่เข้ามา ให้คำนวณหา 2-gram ที่มีความถี่มากที่สุดของข้อความนั้น และความถี่ที่เกิดขึ้น (ถ้ามีความถี่สูงสุดเท่ากันให้แสดง 2-gram ที่เจอตัวแรกในข้อความ)

by นายธนทรัพย์ เพิ่มพูล (<https://ejudge.it.kmitl.ac.th/account/3>)

14 November 2017, 22:25

Specification

➡ Input Specification

มี 1 บรรทัดเป็นข้อความที่ต้องการคำนวณ

** ข้อความดังกล่าวไม่ได้มีเพียงแค่ตัว O และ W และรับประกันว่า เป็นตัวอักษรภาษาอังกฤษ A-Z ตัวพิมพ์ใหญ่เท่านั้น

➡ Output Specification

มี 2 บรรทัด

บรรทัดแรกเป็น 2-gram ที่มีความถี่มากที่สุด
บรรทัดที่ 2 เป็นความถี่ที่เกิดขึ้นของ 2-gram ที่มีความถี่มากที่สุด

Sample Case

➡ Sample Input

WOWWOWO

➡ Sample Output

WO
3

⌚ Time Remaining

0
Day

1
Hour

26
Minutes

5
Seconds

966

i Information

Time Limit	1 Second
Memory Limit	32 MB
Language	py
Deadline	5 December 2017, 23:59
Submission Limit	<div>Unlimit</div>
Restrict Word	<div>No Restrict</div>
Required Word	<div>No Restrict</div>
Testcase	10 case
Rejudge Testcase	0 case
Full Score	100 Point
Bonus Score	900
Your Score	N/A

📁 Submit File

เลือกไฟล์

ไม่ได้เลือกไฟล์ใด

Submit