

ข้อกำหนดและวิธีการส่ง :

1. ให้นักศึกษาเขียนโปรแกรมด้วยภาษา C++ โดยใช้ Dev-C++ เป็นเครื่องมือในการพัฒนาโปรแกรม
2. ให้นักศึกษาใส่เลขทะเบียน ชื่อ นามสกุล เป็น comment ที่ในบรรทัดแรกของไฟล์
3. ให้นักศึกษาส่งไฟล์โปรแกรมแค่ไฟล์เดียวและตั้งชื่อไฟล์โดยใช้เลขทะเบียนนักศึกษา เช่น 650961xxxx.cpp (ไม่ต้องส่งไฟล์ input)
4. ให้ส่งงานในกล่องส่งการบ้านที่อยู่ใน course web ของวิชาเท่านั้น ไม่รับงานช่องทางอื่น
5. นักศึกษาสามารถส่งการบ้านได้หลายครั้งหากยังไม่เลยกำหนดส่งโดยการบ้านเดิมจะถูกเขียนทับด้วยงานที่ส่งใหม่และผู้ตรวจจะตรวจและให้คะแนนเฉพาะฉบับที่ส่งเป็นครั้งสุดท้าย
6. การคัดลอกงานของผู้อื่นหรือการจ้างวานให้ผู้อื่นทำงานให้ถือเป็นการกระทำความผิดวินัยนักศึกษาอย่างร้ายแรง นักศึกษาจะไม่ได้รับคะแนนในการบ้านชิ้นนั้น และผู้สอนมีสิทธิดำเนินการลงโทษตามระเบียบข้อบังคับเกี่ยวกับการทำผิดวินัยนักศึกษา
7. การปกป้องชิ้นงานของตนเองไม่ให้ผู้อื่นนำไปคัดลอกถือเป็นหน้าที่ของนักศึกษา ถ้าพบว่ามี การคัดลอกกัน งานทุกฉบับที่เหมือนกันจะได้คะแนน 0 ทั้งกลุ่ม

โจทย์

ให้นักศึกษาสร้างดัชนีของคำในไฟล์เอกสารที่กำหนด โดยทำตามขั้นตอน ดังนี้

- โปรแกรมจะอ่านไฟล์เอกสารชื่อ document.txt ที่มีขนาดไม่เกิน 100 บรรทัด โดยอ่านทีละบรรทัด แล้วตัดคำ
- ตัดตัวเลขเครื่องหมายต่างๆ ออกจากคำ ได้แก่ dot (.), space (' '), tab ('\t'), newline('\n'), carriage-return ('\r'), form-feed ('\f'), comma (','), semicolon (';'), single quote (''), double quote (''), underscore ('_') และ minus ('-') ออก. โดยนักศึกษาอาจเลือกใช้ฟังก์ชัน

```
token = strtok( str, " .,();\"'-_1234567890\t\n\r"); //defined in string.h
```

token เป็นค่าที่ได้จากการตัดสตริง **str** และเป็นค่าที่ไม่มีตัวเลขและเครื่องหมายเหล่านั้น ดูตัวอย่างการใช้งานจากโปรแกรมตัวอย่างที่ให้ไว้

- ให้แปลงตัวอักษรของคำเป็นตัวพิมพ์เล็กทั้งหมด ในการเปรียบเทียบคำให้ถือว่าคำที่สะกดเหมือนกัน แต่ตัวอักษรตัวเล็กและใหญ่ต่างกันเป็นคำเดียวกัน (Hint : ฟังก์ชัน tolower(), toupper() ใน ctype.h สามารถใช้แปลงตัวอักษรเป็นตัวพิมพ์เล็กและตัวพิมพ์ใหญ่ได้)
- สร้างดัชนีของคำโดยใช้ Binary Search Tree เก็บคำทั้งหมดที่ตัดได้ยกเว้นคำที่เป็น **“a”, “an”, “the”** โดยที่แต่ละโหนดเก็บค่าของคำและบรรทัดที่คำนั้นปรากฏในเอกสาร (คำที่ซ้ำกันจะถูกเก็บที่โหนดเดียวกัน แต่เพิ่มหมายเลขบรรทัดที่พบ)
- พิมพ์ดัชนีของคำทั้งหมดเป็นตัวพิมพ์เล็กเรียงตามตัวอักษร พร้อมแสดงหมายเลขบรรทัดที่พบคำนั้นในเอกสาร ออกทางหน้าจอ ตัวอย่างดังภาพ

ตัวอย่าง output ที่ได้จากไฟล์ document.txt

=====INDEX.TXT=====		
A	abdominal	48
	abdominoscopy	47
	ability	6
	access	22
	acid	32
	acids	60
	additional	34
	all	36
	also	2 48
	and	2 4 8 20 30 34 34 36 36 36 38 46 54 54 60 62 64
	antacids	59
	aphasia	5
	are	36
	around	38 46
	artery	26
B	balance	60
	basal	36
	basic	50
	below	36
	blood	26 44 46 56
	blow	46
	bones	18
	brain	6 36
	breaks	24
	breast	38
	bridges	34
	broken	34
	bruise	46
	by	6 46
C	can	2
	cause	2
	caused	6 46 46
	causes	54
	cavity	48
	cerebrum	35 36 36
	certification	34

คำสั่ง

โปรแกรมของนักศึกษาจะต้องประกอบด้วย class ของ Binary Search Tree และมีอย่างน้อย 2 methods คือ InsertNode และ PrintTree โดยสามารถเลือก implement โครงสร้างเป็นแบบ array หรือ pointer ก็ได้