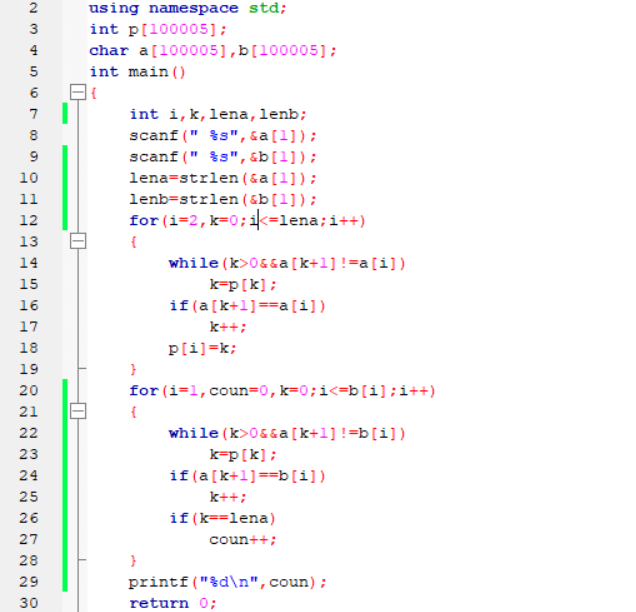
Knuth Morris Pratt(KMP)

Kmp เป็นอัลกอที่สามารถใช้หาว่าคำ a (key) มีอยู่ใน b กี่คำ โดยหลักการของ kmp จะมีอาเรย์ช่วยตัวนึง ทำหน้าที่ดังนิยามนี้ ให้ k แทนค่าใน p[i] แสดงว่าตัวอักษรตัวที่ i-k+1 ถึง i เหมือนกับตัวอักษร k ตัวเลข ก็คือช่องที่ 1 ถึง k นั่นเอง แต่ว่าค่า k นี้จะไม่เท่ากับ i แน่นอนนะ

โค้ด kmp เป็นดังนี้



จะค่อยๆ อธิบายรูปแบบสำคัญและก็แนวคิดของ kmp นะ

-อย่างแรกก็คือการที่เราเริ่ม index ที่ 1 เนื่องจากทำให้การเก็บค่าลงอาเรย์ p นั้นง่ายขึ้นนะ

-ต่อมาให้ดูที่บรรทัดที่ 12-19 ในลูปนี้จะเป็นการเจนค่าในอาเรย์ p เพื่อใช้ต่อไป ซึ่งจากนิยามที่พี่บอกไปข้างบนคือ ถ้า k เป็นค่าในอาเรย์ p[i] แล้ว k จะมีค่าที่มากที่สุดที่ไม่เท่ากับ i และทำให้ a[1 to k] เหมือนกับ a[i-k+1 to i] ทีนี้ให้สังเกตลูปข้างในตรงบรรทัดที่ 14 15 นะ ลุปนี้ใช้แนวคิดที่ว่า ถ้าเรารู้ว่าปัจจุบันมันเหมือนมา k ตัวละ แล้วพบกว่าตัวต่อไปไม่เหมือน แต่มันก็อาจจะเหมือนช่วงใด ซักช่วงข้างหน้าก็ได้ แล้วทางเดียวที่จะเหมือนก็คือ ตัวข้างหน้าตรงนั้นมีลักษณะที่เหมือนกันจุดเรา พุดเข้าใจยาก จะยกตัวอย่างอธิบายแทน ก็กรณี ababc ถ้าสมมุติเราหาคำนี้บน abababc เราจะพบว่าเวลาเราเช็คแล้วพบว่าผิด เช่นตอนวิ่งมาถึงตัวอักษรตัวที่ 5 ของสตริงยาวจะผิดจากสตริงแรก เราก็ไม่สามารถเริ่มหาใหม่ไม่ได้เพราะว่า 5 ตัวหลังของสตริงยาวเหมือนสตริงสั้นเปะ แต่ถ้าเราเริ่มใหม่มันจะเหลือแค่ 3 ตัว ทีนี้เราก็นึกดูว่าเราควรจะให้ค่าจำนวนตัวที่เหมือนกันแล้ว(k) ไปเป็นเท่าไหร่ดีตอันที่เช็คตัวที่ 5 คำตอบก็คือ 2 เพราะว่าเมื่อเราเช็คมาครบ 4 ตัวละจะเห็นว่า 2 ตัวหลังจะซ้ำกับ 2 ตัวแรก ก็คือ ab แล้วค่า 2 นี้ก็อยู่ใน p[4] พอดีเพราะจากนิยามของ p[i] จะบอกว่าเป็นค่าที่มากที่สุดที่ไม่เท่ากัน i และเป็นค่าที่ทำให้ a[1 to k] เหมือนกับ a[i-k+1 to i] จากนิยามนี้เราสรุปได้ว่าต่อให้ปัจจุบันเราจะเหมือนมา k ตัวแล้ว และตัวต่อไปของสตริงยาวไม่เหมือนกับสตริงสั้นตัวที่ k+1 แต่ก็ไม่ได้หมายความว่าจะไม่เหมือนกัน p[k]+1 หรือ p[p[k]]+1 หรือ p[p[p[k]]]+1 ... เพราะว่าตัวต่อไปไม่เหมือน แต่อาจเหมือนกับตัวก่อนหน้า ก็ได้ ก็เลยเป็นสาเหตุให้ใน while พอรู้ว่าตัวต่อไปจะไม่เหมือนก็เลยขยับมาเป็น p[k] แทน

-ต่อมาในบรรทัดที่ 16 17 ก็จากลูป while ข้างบน กรณีที่จะหลุดมามี 2 กรณีคือตรงกันกับไม่ตรงกันและ k=0 ซึ่งสำหรับกรณีตรงกันเนี่ย ถ้าลองคิดดูคือเราตรงกัน k ตัวละ แล้วตัวที่ k+1 ก็ตรงอีกแสดงว่าตอนนี้เราก็ควรจะตรงกัน k+1 แทน k ส่วนถ้าไม่ตรงก็เป็น 0 ไปแบบนั้นแหละ

-บรรทัดที่ 18 จะเป็นการเซ็ทค่าให้อาเรย์ p

\*สำหรับเรื่องนี้ อันนี้ก็เป็นคร่าวๆที่พี่พยายามอธิบายให้เข้าใจนะ แต่เรื่องนี้มันอธิบายยากมากจริงๆ ขอโทษถ้ายังทำให้เข้าใจไม่ได้ ก็จะพยายามพูดให้ดีกว่านี้เวลาไปสอนที่ ม.บู ละกัน