

ปัญหา ภาษาอะไร (UnicodeCharacterRanges)

เป็นที่ทราบว่า Unicode เป็นรหัสใช้แทนตัวอักษรที่ใช้ 2 ไบต์ เป็นมาตรฐานที่ภาษาต่างๆ ทั่วโลกสามารถใช้รหัสเฉพาะของภาษาเอง ทำให้เอกสารสามารถรวมหลายๆ ภาษาไว้ในหน้าเอกสารเดียวกันได้ โดยนำไปเปิดที่เครื่องใดๆ จะสามารถเห็นฟอนต์ของภาษาได้มากกว่า 1 ภาษา ค่ารหัสยูนิโค้ดมาแยกภาษาที่รู้จักดังตารางที่ 1

ตารางที่ 1 แสดงค่ารหัสยูนิโค้ดและภาษาที่สอดคล้องกัน

Ranges	Character	Ranges	Character	Ranges	Character
0020-007F	BASIC LATIN	0E00-0E7F	THAI	3040-309F	HIRAGANA
0600-06FF	ARABIC	0E80-0EFF	LAO	30A0-30FF	KATAKANA

หมายเหตุ รหัสอื่นให้เป็น UNKNOWN

ให้รับค่ารหัสยูนิโค้ดมาเป็นเลขฐาน 16 ขนาด 2 ไบต์ แล้วนำมาจำแนกว่าเป็นภาษาใดตามตารางที่ 1

รูปแบบข้อมูลเข้า

บรรทัดแรก	รหัสยูนิโค้ด มีได้มากกว่า 1 ค่า แต่ไม่เกิน 1000 ค่า แต่ละค่าคั่นด้วยลูกน้ำ “,”
-----------	--

รูปแบบผลลัพธ์

บอกภาษาของรหัสยูนิโค้ดแต่ละตัว เรียงลำดับตามข้อมูลเข้า โดยบอกภาษาละหนึ่งบรรทัด พิมพ์เป็นตัวพิมพ์ใหญ่เท่านั้น

ตัวอย่าง

ข้อมูลเข้า	ผลลัพธ์
0010,0E24,0E82,3024,3042	UNKNOWN THAI LAO UNKNOWN HIRAGANA

ข้อมูลเข้า	ผลลัพธ์
8973,3100,308F,0043,0079,06F4,3209,0E56	UNKNOWN UNKNOWN HIRAGANA BASIC LATIN BASIC LATIN ARABIC UNKNOWN THAI

เมธอดที่น่าสนใจในคลาสสตริง (String)

Method signature	คำอธิบาย
int compareTo(String)	เปรียบเทียบสตริงเจ้าของเมธอด กับ สตริงในวงเล็บ ให้สตริงเจ้าของเมธอดเป็น text1 และสตริงในวงเล็บเป็น text2 เช่น “ANT”.compareTo(“CAT”) จะได้ -2 เพราะตัว A ลบ ตัว C เมธอดนี้จะเปรียบเทียบรหัสของตัวอักษรตามตำแหน่งตัวต่อตัว
int compareToIgnoreCase(String)	เปรียบเทียบสตริงเจ้าของเมธอด กับ สตริงในวงเล็บ ให้สตริงเจ้าของเมธอดเป็น text1 และสตริงในวงเล็บเป็น text2 เช่น “ant”.compareTo(“CAT”) จะได้ -2 เพราะตัว a ลบ ตัว c เมธอดนี้จะเปรียบเทียบรหัสของตัวอักษรตามตำแหน่งตัวต่อตัวโดยไม่สนใจ ตัวเล็กตัวใหญ่
int indexOf(String)	คืนค่าตำแหน่งของ String ในวงเล็บที่ปรากฏในสตริงเจ้าของเมธอด ถ้ามี หลายตำแหน่งให้คืนตำแหน่งแรกที่ปรากฏ เช่น “MISSISSIPPI”.indexOf(“SI”) จะได้ค่า 3
int indexOf(String,int)	คืนค่าตำแหน่งของ String ในวงเล็บที่ปรากฏในสตริงเจ้าของเมธอด โดย ตำแหน่งเริ่มต้นที่ทำให้เริ่มจากค่า int ที่ปรากฏในวงเล็บที่เป็นอาร์กิวเมนต์ ตัวที่สอง ถ้ามีหลายตำแหน่งให้คืนตำแหน่งแรกที่ปรากฏ เช่น “MISSISSIPPI”.indexOf(“SI”, 4) จะได้ค่า 5
String substring(int i)	คืนค่าสตริงที่ตัดตั้งแต่ตำแหน่ง i ถึงจบสตริง เช่น “MISSISSIPPI”.substring(3) จะได้ค่า “SISIPPI”
String substring(int i0,int i1)	คืนค่าสตริงที่ตัดตั้งแต่ตำแหน่ง i0 ถึง i1-1 เช่น “MISSISSIPPI”.substring(3,7) จะได้ค่า “SISI”