

อัลกอริทึมไทยเรียงลำดับคำไทย

Sorting for TIS-620

Within TIS-620 character set, sorting is less complicated than that on ISO/IEC 10646. We can assign the four levels of weights to each class of character in this manner:

Level	Order	Blanks	Ignored
1	digits (language-insensitive), English alphabet (case-insensitive), Thai consonants, Ru (ร), Lu (ล) (as TIS-620), kikkhahit, Thai vowels (as TIS-620)	-	punctuation marks, Yanakkhan, Pintu, Thantakhat, Mai Tokhu, Thai Tone Marks
2	Thai digit script language, Yanakkhan, Pintu, Thantakhat, Mai Tokhu, Mai Ek, Mai Tho, Mai Tri, Mai Chotbwa	Arabic digits, English alphabet, Thai consonants, Ru (ร), Lu (ล), kikkhahit, Thai vowels	punctuation marks
3	space, non-breaking space, low line , , highan - , comma , , senkolan , , colon , ,	digits, English alphabet, Thai consonants, Ru (ร), Lu (ล), kikkhahit, Thai vowels, Yanakkhan, Pintu,	-

ลักษณะการดำเนินการ

ขั้นตอนการพัฒนาอัลกอริทึมไทยเรียงลำดับคำไทยจำแนกได้ดังนี้ คือ

- ศึกษาอัลกอริทึมไทยเรียงลำดับคำไทยที่มีอยู่
- พัฒนาอัลกอริทึมไทยเรียงลำดับคำไทยจากที่มีอยู่ให้สามารถเรียงลำดับคำไทยตามพจนานุกรมฉบับราชบัณฑิตยสถาน และจัดการกับเครื่องหมายวรรคตอนได้
- ขยายขอบเขตของอัลกอริทึมให้อยู่ในรูปทั่วไปที่สามารถใช้ร่วมกับ ภาษาอังกฤษและภาษาอื่นๆ ได้โดยอิงมาตรฐาน ISO/IEC 14651 International String Ordering รวมทั้งตกลงกำหนดลำดับของเครื่องหมายวรรคตอนตามความเหมาะสม

ประโยชน์ที่ได้รับ

- การเรียงลำดับค่าไทยเป็นขั้นตอนวิธีที่สำคัญอย่างหนึ่งในการทำดัชนีและเรียงลำดับระเบียบอื่นต่างๆ ในฐานข้อมูล การมีอัลกอริทึมที่ละเอียดถูกต้องจึงส่งผลให้การใช้งานในระบบเป็นธรรมชาติมากขึ้น
- การเปิดเผย Source Code ในลักษณะ Free Software ทำให้เกิดทรัพยากรสาธารณะที่ทุกคนนำไปใช้งานที่สูงขึ้นไปได้ทันที และทำให้เกิดมาตรฐานที่ตรงกันได้ง่ายยิ่งขึ้น
- สนับสนุนการพัฒนาซอฟต์แวร์ สนับสนุนภาษาไทย ทำให้เกิดความสะดวกในการใช้งาน

วิจัยและพัฒนาโดย: งานวิจัยซอฟต์แวร์พื้นฐานและทั่วไป

สอบถามข้อมูลทางเทคนิคเพิ่มเติมได้ที่หน้า 249

รหัสผลงาน: BI101-43

ติดต่อสอบถามได้ที่: ฝ่ายพัฒนาธุรกิจและอุตสาหกรรม ศูนย์เทคโนโลยีอิเล็กทรอนิกส์และคอมพิวเตอร์แห่งชาติ

โทร. (662) 644-8150..99 ต่อ 610, 656..9

โทรสาร (662) 644-8122, 644-8137..8