

วิทยานิพนธ์นี้นำเสนอวิธีแก้ไขคำผิดแบบไม่ตั้งใจโดยอัตโนมัติในภาษาไทย เพื่อนำไปใช้กับหุ่นยนต์สนทนา ซึ่งใช้คุณสมบัติอักขระข้างเคียงบนแป้นพิมพ์ช่วยในการแก้ไขคำผิด โดยได้มีการทดสอบเพื่อสังเกตพฤติกรรมในการพิมพ์ พบรูปแบบการพิมพ์ผิดในภาษาไทยว่า 93.54 เปอร์เซ็นต์จากคำผิดที่พบทั้งหมด เกิดจากความผิดพลาดทั้ง 4 กรณีประกอบกัน คือ แทนที่ เกิน ตก สลับ (เรียงตามลำดับปริมาณที่พบจากมากไปน้อย) และตำแหน่งอักขระที่ผิดนั้นเฉลี่ยอยู่ตำแหน่งที่ 58.36 เปอร์เซ็นต์ของความยาวคำ

ในงานวิจัยนี้นำรูปแบบการพิมพ์ผิดที่ได้มาออกแบบอัลกอริทึมแก้ไขคำผิด แล้วทำการทดสอบประสิทธิภาพด้วยบทสนทนาที่มีคำผิดแบบไม่ตั้งใจทั้งหมดจำนวน 120 ประโยค (ยกเว้นความผิดพลาดที่มาจากการพิมพ์ตก) พบว่าหุ่นยนต์สามารถตอบได้คิดเป็น 95 เปอร์เซ็นต์ ซึ่งแสดงให้เห็นว่าหุ่นยนต์สามารถทำงานได้ประสิทธิภาพมากขึ้นหากใช้อัลกอริทึมนี้

This thesis proposes an algorithm, which is an automatic correction algorithm for unintentional typing errors in Thai language, for communicating with chat robot. It uses the characteristic of adjacent alphabets on keyboard for error correction. Our investigation found that 93.54 percent of all misspelled words in Thai language contains four kinds of typing errors : substitution, insertion, deletion and transposition (arranged from maximum to minimum frequency). Average error position in misspelled words is at 58.36 percent of the word length.

This thesis uses the discovered typing error patterns to design an error correction algorithm. The algorithm is tested on 120 sentences with unintentional typing errors (except error from deletion typing). The chat robot is able to identify 95 percent of the errors. It shows that the chat robot can work more effectively if it uses this algorithm.