

193110

การอนุมานไวยากรณ์ไม่พึ่งบริบท เป็นส่วนสำคัญในการวิเคราะห์แก้ปัญหการรู้จำรูปแบบ และงานวิจัยต่างๆที่เกี่ยวกับการประมวลผลภาษาธรรมชาติได้เน้นไปที่วิธีการพัฒนาการอนุมาน ซึ่งปัญหาหลักของการอนุมานไวยากรณ์ไม่พึ่งบริบท คือ ใช้ค่าความซับซ้อนเชิงเวลาของการอนุมานไวยากรณ์ที่สูง ซึ่งมีผลงานวิจัยทางทฤษฎีระบุว่าไม่สามารถหาอัลกอริทึมอนุมานไวยากรณ์ไม่พึ่งบริบทที่ใช้ความซับซ้อนเชิงเวลาไม่เกินฟังก์ชันพหุนามได้

ดังนั้นงานวิจัยนี้ได้นำเสนออัลกอริทึมการอนุมานไวยากรณ์ไม่พึ่งบริบทสำหรับบางภาษาไม่พึ่งบริบทรวมทั้งภาษาสม่าเสมอ ที่มีความซับซ้อนเชิงเวลาไม่เกินฟังก์ชันพหุนาม ซึ่งหลักการทำงานของอัลกอริทึมจะพิจารณาสร้างกฎวนซ้ำจากข้อมูลตัวอย่างที่อยู่ในภาษาเรียกว่าข้อมูลตัวอย่างบวก และเพื่อไม่ให้ไวยากรณ์ไม่พึ่งบริบทมีความกว้างมากเกินไป งานวิจัยนี้ได้นำข้อมูลตัวอย่างที่ไม่อยู่ในภาษาเรียกว่าตัวอย่างลบมาร่วมพิจารณาด้วย

193110

Context-free grammar inference plays an important role in pattern recognition. Many researches in natural language processing are focused on how to improve an inference technique. In fact, the major problem is that it requires a high degree of computational complexity. Some theoretical results stated that the problem cannot be solved by any polynomial time algorithms.

This research is aimed to introduce an inference algorithm for some context-free languages including regular languages. The proposed algorithm needs a polynomial time complexity. Our concept is to create some recursive rules using positive samples. In order to avoid an overgeneralization problem, some negative samples are considered in the training process.