



Finding Strategy from Pattern of Winner by Q-Learning : Case Study of Downsized Splendor game



เรนศพล เดชาดวงสกุล ฐานิดา ก้อนทอง และชอฬาร์ เตชะโร

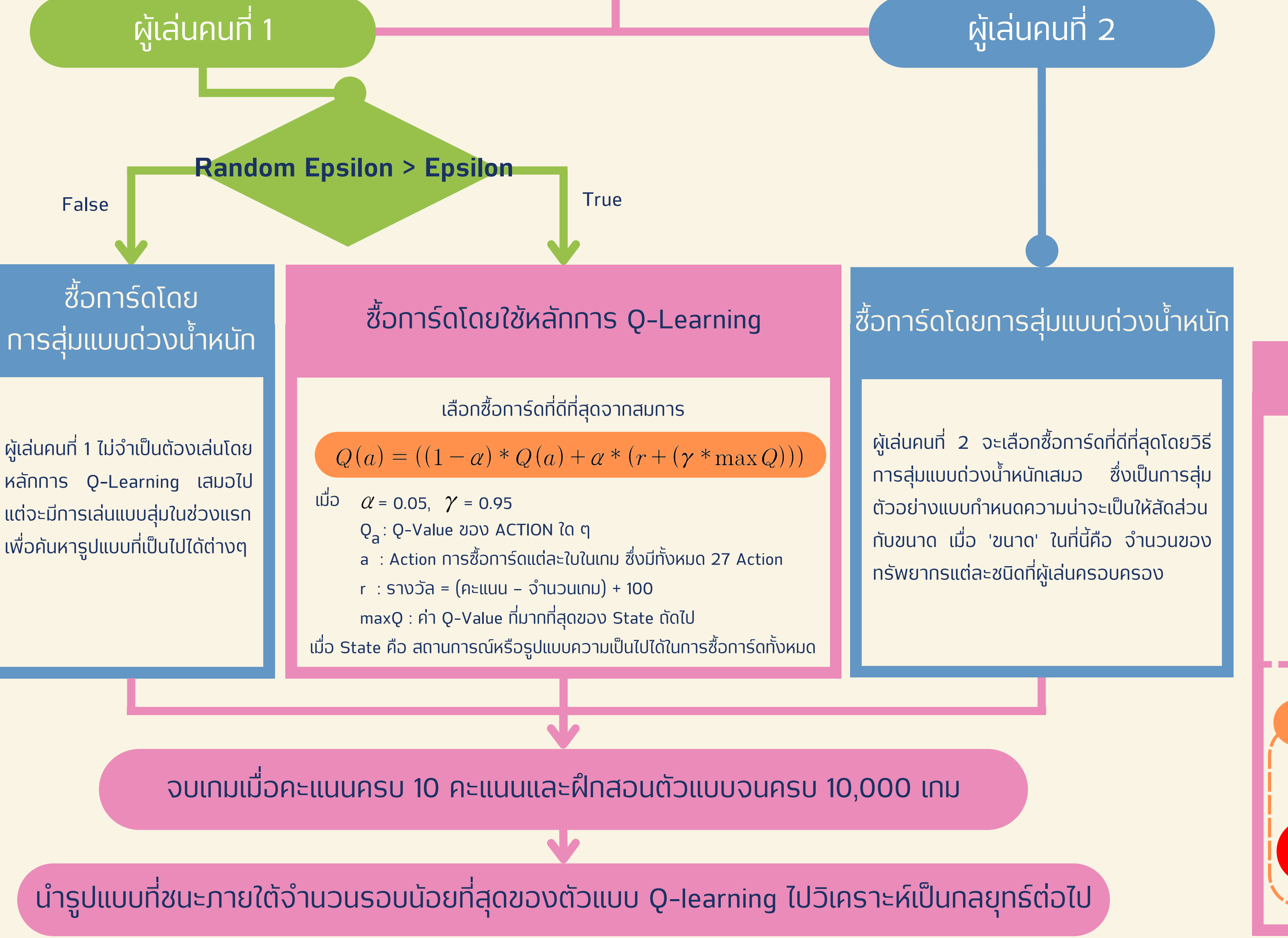
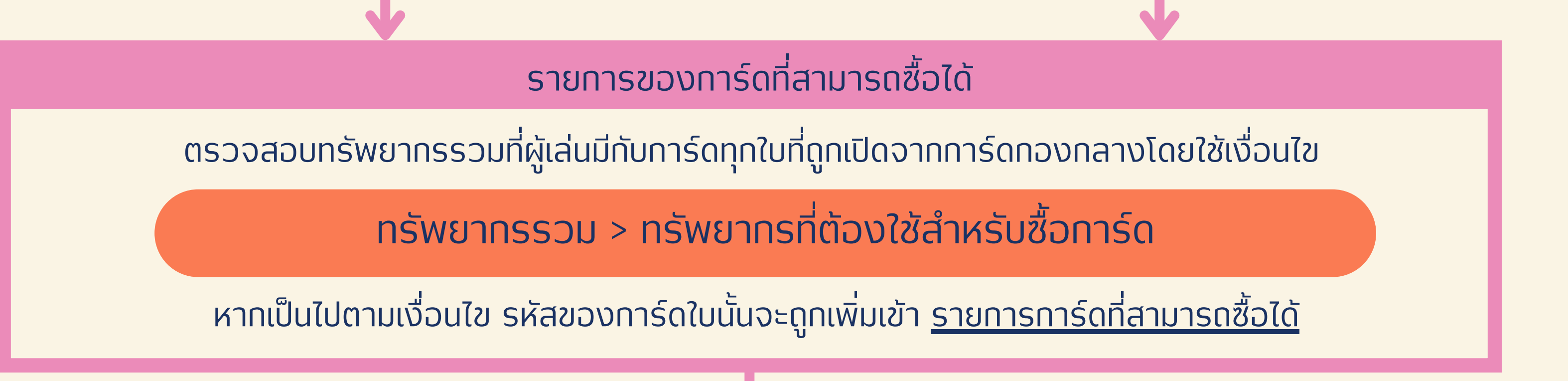
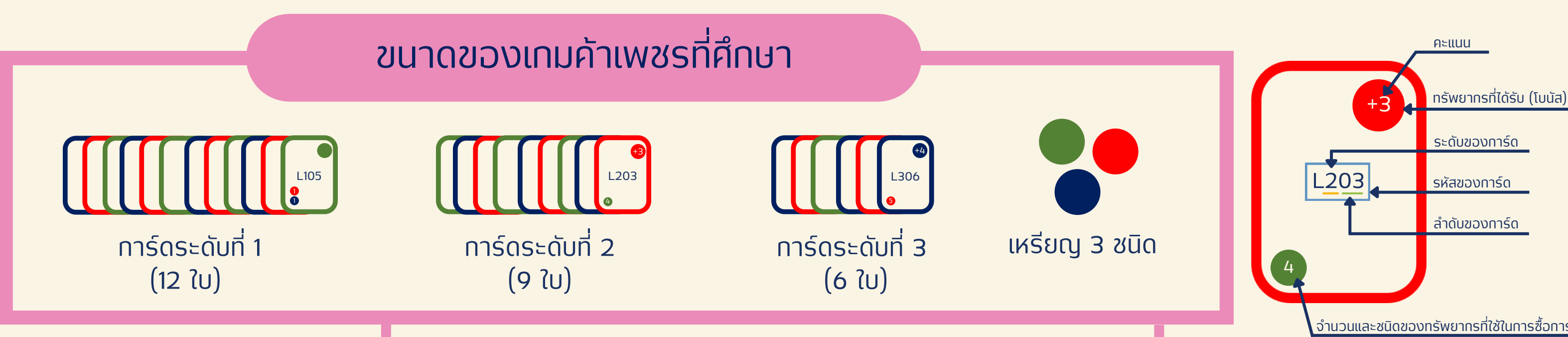
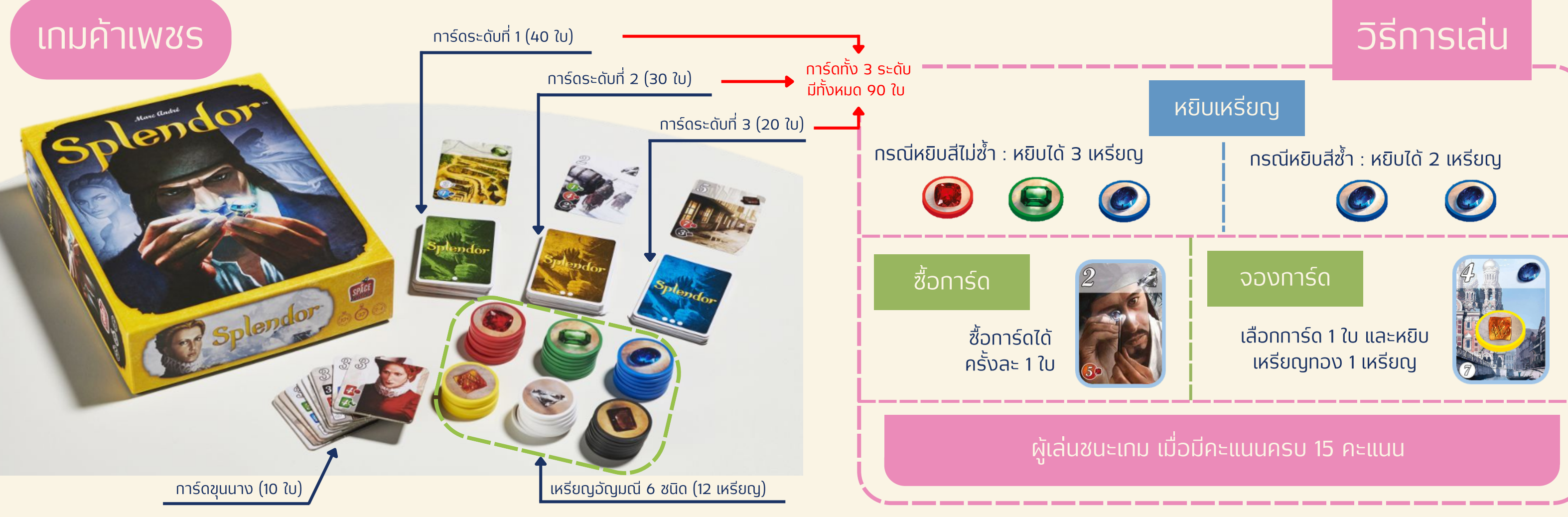
หลักสูตรคณิตศาสตร์ประยุกต์ สาขาวิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี มหาวิทยาลัยธรรมศาสตร์

บทคัดย่อ

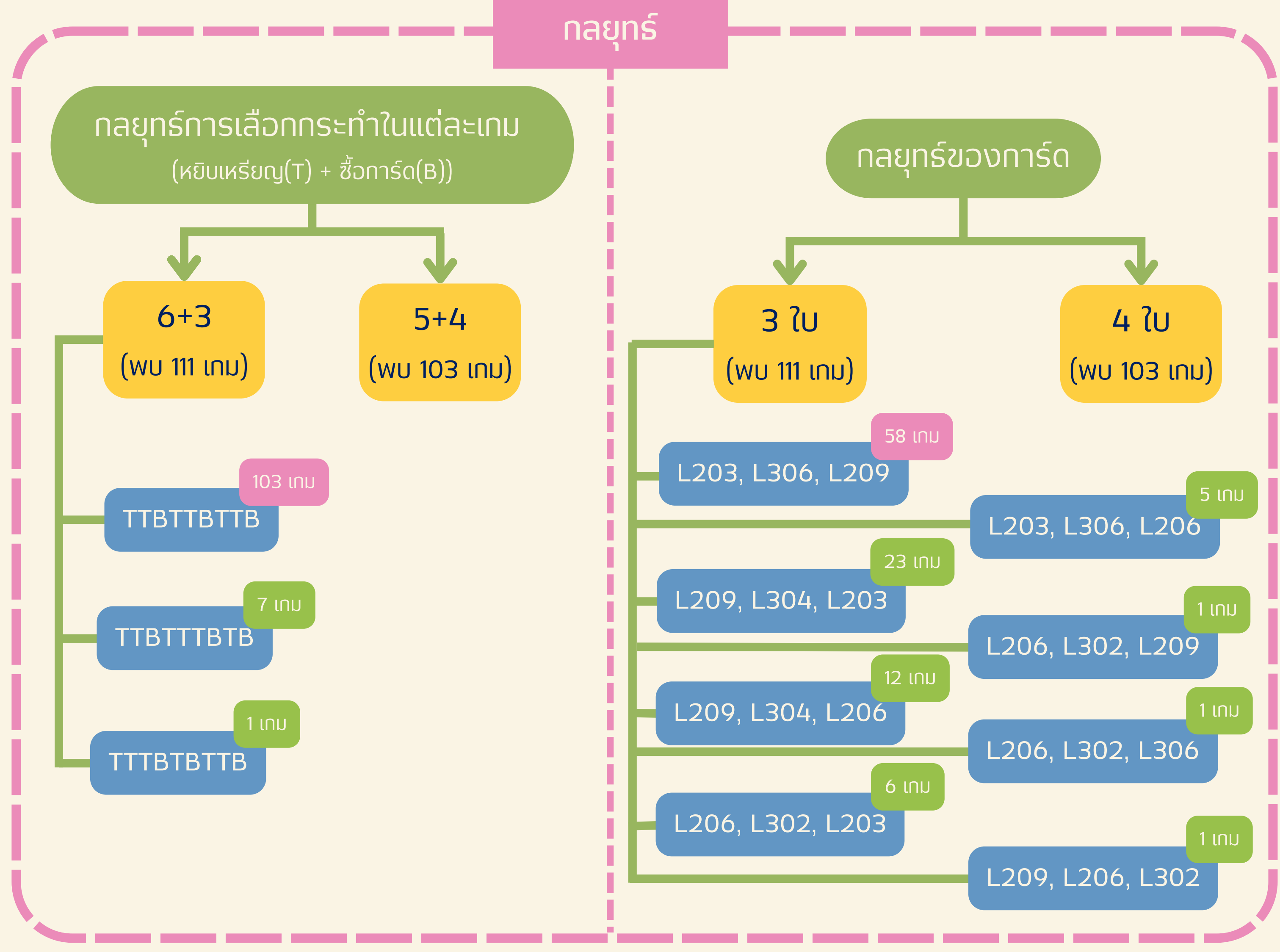
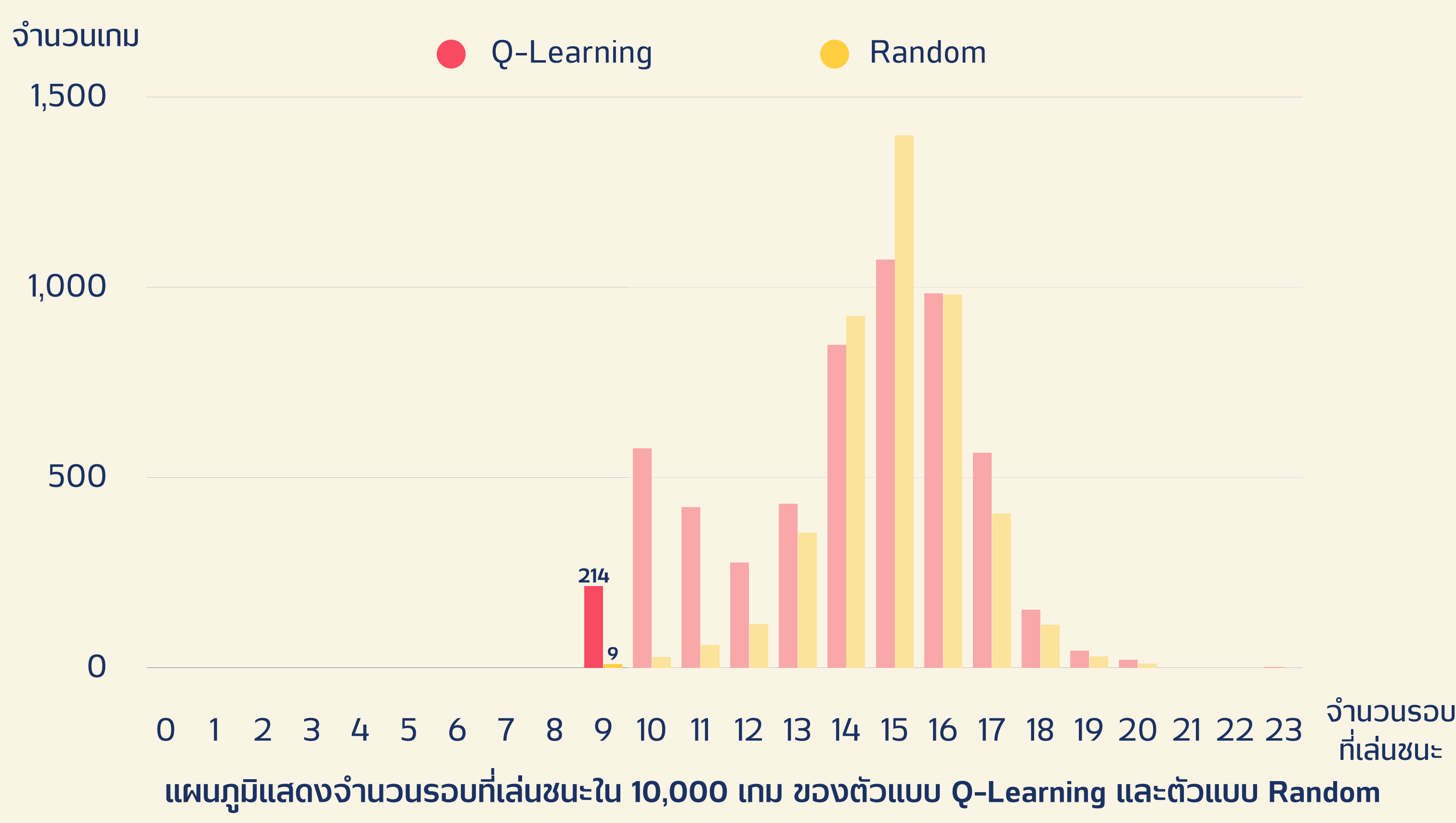
เกมค้าเพชร เป็นเกมประเภทวางแผนที่มีรูปแบบการเล่นที่หลากหลาย คณะผู้จัดทำจึงสนใจนำรูปแบบเหล่านั้นมาวิเคราะห์เพื่อจัดกลุ่มกลยุทธ์ที่สามารถใช้ได้จริง เพื่อให้ง่ายต่อการศึกษา คณะผู้จัดทำจึงนำเกมมาลดขนาด และใช้หลักการ Q-learning ในการหารูปแบบการเล่นที่ทำให้ชนะด้วยจำนวนรอบที่น้อยที่สุด จากการศึกษาพบว่า การใช้หลักการ Q-Learning เพื่อฝึกสอนตัวแบบผู้เล่นทั้งหมด 10,000 เกม ทำให้ทราบจำนวนรอบที่น้อยที่สุด คือ 9 รอบ ซึ่งมี 214 เกม นำมาจัดกลยุทธ์ได้ 2 กลุ่ม ได้แก่ กลยุทธ์การเลือกกระทำในแต่ละเกม และกลยุทธ์ของการ์ด โดยจำนวนรูปแบบที่พบนั้นเป็นเพียงส่วนหนึ่งที่ได้จากการฝึกสอนตัวแบบ ดังนั้นหากมีการฝึกสอนตัวแบบให้มากขึ้น จะทำให้เพิ่มโอกาสการหากลยุทธ์ที่สอดคล้องกับสถานการณ์ทั้งหมดที่คาดว่าจะเกิดขึ้นในเกมค้าเพชรฉบับเต็ม เพื่อที่จะสามารถนำผลลัพธ์ที่ได้จากการศึกษาไปใช้ในสถานการณ์จริงต่อไป

วิธีการดำเนินงาน

คณะผู้จัดทำได้ลดขนาดของเกมค้าเพชร และสร้างโปรแกรมสำหรับการเล่นเกมเพื่อนำไปฝึกสอนตัวแบบ โดยใช้ผู้เล่น 2 คน ผู้เล่นคนที่ 1 ใช้หลักการ Q-Learning และผู้เล่นคนที่ 2 ใช้หลักการสุ่มแบบถ่วงน้ำหนัก (Random) ซึ่งมีวัตถุประสงค์เพื่อหาจำนวนรอบที่น้อยที่สุดที่สามารถชนะด้วยคะแนน 10 คะแนน และนำรูปแบบการเล่นไปวิเคราะห์เป็นกลยุทธ์ต่อไป



ผลลัพธ์



จากผลลัพธ์ที่ได้จะเห็นว่า รูปแบบกลยุทธ์การเล่นที่ชนะภายใต้จำนวนรอบที่น้อยที่สุด 9 รอบ สามารถแบ่งออกเป็น 2 กลุ่ม คือ กลยุทธ์การเลือกกระทำในแต่ละเกม (หยิบเหรียญ(T)+ซื้อการ์ด(B)) และกลยุทธ์ของการ์ดที่ครอบครองหลังจบเกม รูปแบบที่ถูกพบมากที่สุด คือ 6+3 : TTBTBTBTB ซึ่งถูกพบทั้งหมด 103 เกม จาก 111 เกม และจากรูปแบบดังกล่าว พบว่า L203, L306, L209 เป็นกลุ่มการ์ดที่ถูกซื้อ มากที่สุด ซึ่งพบทั้งหมด 58 เกม จาก 103 เกม โดยสามารถแสดงตัวอย่างการเล่นเกม ดังต่อไปนี้

