

การเพิ่มประสิทธิภาพของสถานะและรางวัลในการเรียนรู้แบบเสริมกำลังเชิงลึก
ร่วมกับตรรกะฟัซซีเพื่อการจัดการสัญญาณไฟจราจร

STATE AND REWARD OPTIMIZATION IN DEEP REINFORCEMENT LEARNING
WITH FUZZY LOGIC FOR TRAFFIC SIGNAL MANAGEMENT



วิทยานิพนธ์นี้สำหรับการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรมหาบัณฑิต

สาขาวิชาวิศวกรรมไฟฟ้าและคอมพิวเตอร์

คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

พ.ศ.2568

KMITL-2025-EN-M-027-006

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

STATE AND REWARD OPTIMIZATION IN DEEP REINFORCEMENT LEARNING
WITH FUZZY LOGIC FOR TRAFFIC SIGNAL MANAGEMENT



A THESIS SUBMITTED IN FULFILLMENT
OF THE REQUIREMENT FOR THE DEGREE OF
MASTER OF ENGINEERING IN ELECTRICAL AND COMPUTER ENGINEERING
SCHOOL OF ENGINEERING
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
2025
KMITL-2025-EN-M-027-006

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



COPYRIGHT 2025

SCHOOL OF ENGINEERING

KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หัวข้อวิทยานิพนธ์	การเพิ่มประสิทธิภาพของสถานะและรางวัลในการเรียนรู้แบบเสริมกำลังเชิงลึกร่วมกับตรรกะฟัซซีเพื่อการจัดการสัญญาณไฟจราจร
นักศึกษา	นายธีรพันธ์ มีโภคกิจ
รหัสประจำตัว	66016056
ปริญญา	วิศวกรรมศาสตรมหาบัณฑิต
สาขาวิชา	วิศวกรรมไฟฟ้าและคอมพิวเตอร์
พ.ศ.	2568
อาจารย์ที่ปรึกษาวิทยานิพนธ์	ผศ.ดร.สุเมฆ วิศยทัตภิชิน

บทคัดย่อ

วิทยานิพนธ์นี้นำเสนอวิธีการจัดการสัญญาณไฟจราจรด้วยการเรียนรู้แบบเสริมกำลังเชิงลึก ร่วมกับตรรกะฟัซซีเพื่อลดระยะเวลาการรอคอยของรถยนต์ที่รอสัญญาณไฟจราจรบริเวณทางแยก พร้อมทั้งเพิ่มประสิทธิภาพการทำงานของวิธีการ โดยได้มีการปรับปรุงการเก็บค่าข้อมูลสถานะ ให้สามารถสะท้อนถึงตำแหน่งและระยะเวลาการรอคอยของรถยนต์ในแต่ละช่องทางได้อย่างละเอียดมากยิ่งขึ้น นอกจากนี้ ได้มีการออกแบบฟังก์ชันรางวัล ให้สะท้อนความสมดุลของการจราจรในทุกทิศทาง โดยใช้ระยะเวลาการรอคอยสะสมสูงสุดจากแต่ละกลุ่มช่องจราจรในแต่ละแยกเป็นเกณฑ์การประเมิน ระยะเวลาการรอคอยสะสมทั้งหมดของระบบ ซึ่งช่วยลดความเอนเอียงของระบบที่เคยให้ความสำคัญกับช่องทางตรงเป็นหลักในงานวิจัยก่อนหน้า ทำให้ระบบสามารถตอบสนองต่อสถานการณ์จราจรได้อย่างรอบด้านและมีประสิทธิภาพมากยิ่งขึ้น

สำหรับการทดลองในวิทยานิพนธ์นี้จะใช้ซอฟต์แวร์ที่ชื่อว่า Simulation of Urban Mobility (SUMO) ในการสร้างสภาพแวดล้อมจำลองการจราจรที่หลากหลายและประเภทของยานพาหนะที่แตกต่างกัน ผลการทดลองที่ได้เมื่อทดลองในสภาพแวดล้อมต้นแบบ พบว่าเวลาการรอคอยทั้งหมดโดยเฉลี่ยสามารถลดลงไปได้ 18.46% เมื่อเปรียบเทียบกับวิธีการควบคุมสัญญาณไฟจราจรโดยใช้การเรียนรู้แบบเสริมกำลังเชิงลึกร่วมกับตรรกะฟัซซีแบบดั้งเดิม นอกจากนี้ ระบบการให้ความสำคัญกับรถพยาบาลสามารถลดเวลาการรอคอยของรถพยาบาลได้อย่างมีนัยสำคัญ โดยสรุปแล้ว วิธีการที่นำเสนอขึ้นให้ผลลัพธ์ที่ดีที่สุดในทุกสภาพแวดล้อมที่ใช้ในการทดลอง

Thesis	State and Reward Optimization in Deep Reinforcement Learning with Fuzzy Logic for Traffic Signal Management
Student	Mr. Teerapun Meepokgit
Student ID.	66016056
Degree	Master of Engineering
Program	Electrical and Computer Engineering
Year	2025
Thesis Advisor	Asst. Prof. Sumek Wisayataksin, Ph.D

ABSTRACT

We propose a traffic signal management method using deep reinforcement learning combined with fuzzy logic to reduce vehicle waiting time at intersections and improve the efficiency of the method. The state representation has been improved to more accurately reflect the position and waiting time of vehicles in each lane. In addition, the reward function has been designed to reflect balanced traffic flow in all directions. The system-wide cumulative waiting time is evaluated based on the maximum cumulative waiting time from each group of lanes at every arm, which helps minimize the bias that leads the system to favor straight-through traffic, as observed in existing studies. This design enables the system to respond to various traffic conditions more comprehensively and efficiently.

For the experiment, Simulation of Urban Mobility (SUMO) software was used to simulate various environments and vehicle types. The results show that, when using the proposed method in a prototype environment, the average total waiting time could be reduced by 18.46% compared with the traffic light control method using a conventional Deep Q-network with fuzzy logic. Additionally, an ambulance prioritization system significantly reduced the ambulance waiting time. In summary, the proposed method yielded the best results in all environments used in the experiments.

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงได้ด้วยความกรุณาจากอาจารย์ ผศ.ดร.สุเมฆ วิตยทัตถิณ อาจารย์ที่ปรึกษาวิทยานิพนธ์ที่ได้ให้คำแนะนำ ปรึกษา ชี้แนะแนวทางการแก้ไขปัญหา ข้อบกพร่องต่างๆ และถ่ายทอดความรู้ที่เป็นประโยชน์อย่างยิ่งในกระบวนการทำวิจัย รวมถึงการช่วยพัฒนาความคิดและแนวทางการทำงานของผู้วิจัยให้มีความสมบูรณ์มากยิ่งขึ้น และยังคงเป็นกำลังใจสำคัญในช่วงเวลาที่ยากลำบาก คำพูดที่ให้ความหวังและการสนับสนุนอย่างจริงใจของอาจารย์เป็นสิ่งที่ช่วยให้ผู้วิจัยสามารถเดินหน้าต่อไปได้ จนกระทั่งวิทยานิพนธ์ฉบับนี้สำเร็จลุล่วงไปได้ด้วยดี ผู้วิจัยขอกราบขอบพระคุณเป็นอย่างสูง ไว้ ณ ที่นี้

ขอขอบพระคุณ คุณพ่อและคุณแม่ ที่เป็นกำลังใจสำคัญและคอยสนับสนุนในทุกด้านตลอดระยะเวลาที่ผ่านมา ความรักและความหวังใจจากครอบครัวช่วยเสริมสร้างพลังใจให้ผู้วิจัยสามารถก้าวผ่านอุปสรรคต่าง ๆ ได้อย่างมั่นคง

สุดท้ายนี้ ผู้วิจัยขอขอบคุณเพื่อน ๆ ทุกคนที่อยู่เคียงข้างในช่วงเวลาที่ยากลำบาก การสนับสนุนและกำลังใจจากทุกคนเป็นสิ่งสำคัญที่ช่วยให้ผู้วิจัยก้าวผ่านความท้อแท้และความเหนื่อยล้า การได้รับคำปลอบโยน คำแนะนำ และรอยยิ้มจากทุกคนในทุกช่วงเวลาของการทำวิจัย ช่วยสร้างพลังใจให้ผู้วิจัยมีแรงผลักดันในการทำวิจัยนี้จนสำเร็จ

ธีรพันธ์ มีโภคกิจ

สารบัญ

	หน้า
บทคัดย่อภาษาไทย.....	I
บทคัดย่อภาษาอังกฤษ.....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII
บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาและความสำคัญของปัญหา.....	1
1.2 ความมุ่งหมายและวัตถุประสงค์ของการวิจัย.....	4
1.3 ขอบเขตการวิจัย.....	4
1.4 ประโยชน์ที่คาดว่าจะได้รับ.....	5
บทที่ 2 ทฤษฎีที่เกี่ยวข้อง.....	6
2.1 แนวคิดของโมเดลต้นแบบ.....	6
2.2 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning).....	8
2.3 โครงข่ายประสาทเทียม (Neural Network).....	9
2.3.1 ฟังก์ชันกระตุ้น (Activation Function).....	11
2.4 การเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning).....	13
2.5 นโยบาย (Policy).....	14
2.5.1 การผสม Epsilon-Greedy Policy และ Deterministic Greedy Policy.....	14
2.6 ตรรกะแบบฟัซซี (Fuzzy Logic).....	15
2.6.1 ฟัซซีเซต (Fuzzy Set).....	15
2.6.2 การดำเนินการทางฟัซซีเซต.....	18
2.6.3 โครงสร้างการทำงานพื้นฐานของตรรกะแบบฟัซซี.....	19
2.6.4 การอนุมานฟัซซีแบบแมมดานี.....	20
2.6.5 ฟังก์ชันสมาชิก (Membership Function).....	21
บทที่ 3 การออกแบบระบบ.....	23
3.1 แนวทางการพัฒนาเพิ่มเติมจากโมเดลต้นแบบ.....	23
3.2 ซอฟต์แวร์ที่ใช้ในการทดลอง.....	24

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น อนุญาตให้นำไปใช้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.3 สภาพแวดล้อม (Environment)	25
3.4 สถานะ (State)	27
3.5 การกระทำ (Action)	31
3.6 รางวัล (Reward)	32
3.6.1 ตัวอย่างปัญหาจากการใช้เวลารอสะสมทั้งหมดในวิธีต้นแบบ	32
3.6.2 การคำนวณรางวัลจากเวลารอสะสมของระบบในวิธีที่นำเสนอ	33
3.7 ตัวแทน (Agent).....	34
3.8 ข้อมูลที่จำเป็น (Necessary Information).....	35
3.9 การออกแบบการจำลอง (Simulation Design).....	36
3.9.1 การตั้งค่าโครงข่าย (Network Setting)	36
3.9.2 การตั้งค่าสภาพแวดล้อม (Environment Setting).....	39
3.10 ตรรกะแบบฟัซซี (Fuzzy Logic).....	41
บทที่ 4 ผลการทดลอง.....	50
4.1 ผลการฝึกอบรมในสภาพแวดล้อมแบบสมดุลง (ALL-4).....	51
4.2 สภาพแวดล้อมแบบสมดุลง (ALL-4).....	52
4.3 สภาพแวดล้อมการจราจรแบบขับรถชิดซ้าย	53
4.4 สภาพแวดล้อม 3 ช่องจราจร (Lane) ใน 1 แยก	54
4.5 สภาพแวดล้อม 3 ช่องจราจร (Lane) ใน 2 แยก	56
4.6 สภาพแวดล้อม 3 ช่องจราจร (Lane) ใน 3 แยก	57
4.7 ระบบจัดลำดับความสำคัญรถพยาบาล	59
บทที่ 5 สรุป ข้อจำกัดและข้อเสนอแนะ.....	61
5.1 บทสรุป	61
5.2 ข้อจำกัด	63
5.3 ข้อเสนอแนะ	64
บรรณานุกรม.....	65
ประวัติผู้เขียน.....	70

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญตาราง

ตารางที่	หน้า
ตารางที่ 2.1 ระดับของการเป็นสมาชิกของเซตด้วยหนุ่มสาว	16
ตารางที่ 2.2 การดำเนินการแบบ AND.....	18
ตารางที่ 2.3 การดำเนินการแบบ OR.....	18
ตารางที่ 2.4 การดำเนินการแบบ NOT	19
ตารางที่ 3.1 การปรับปรุงรูปแบบการแทนสถานะและฟังก์ชันรางวัลจากโมเดลต้นแบบ.....	24
ตารางที่ 3.2 ขนาดของเซลล์ที่ใช้ในการแบ่งช่องจราจร	28
ตารางที่ 3.3 ระดับการแบ่งระยะเวลารอคอย	29
ตารางที่ 3.4 ข้อมูลที่จำเป็นสำหรับระบบควบคุมสัญญาณไฟจราจรด้วยการเรียนรู้แบบเสริมกำลังเชิงลึก	36
ตารางที่ 3.5 การตั้งค่าพารามิเตอร์ของโครงข่ายประสาทเทียม	37
ตารางที่ 3.6 การตั้งค่าพารามิเตอร์ของสภาพแวดล้อม	40
ตารางที่ 3.7 คำอธิบายฟังก์ชันสมาชิกของเฟสไฟเขียว (GP)	45
ตารางที่ 3.8 คำอธิบายฟังก์ชันสมาชิกของเฟสไฟแดง (RP).....	45
ตารางที่ 3.9 กฎการอนุมาน	45
ตารางที่ 3.10 คำอธิบายฟังก์ชันสมาชิกของเอาต์พุต	46
ตารางที่ 4.1 ผลการทดลองในสภาพแวดล้อมแบบสมดุล (ALL-4).....	53
ตารางที่ 4.2 ผลการทดลองในสภาพแวดล้อมการจราจรแบบขับรถชิดซ้าย	54
ตารางที่ 4.3 ผลการทดลองในสภาพแวดล้อมแบบ 3 ช่องจราจรใน 1 แยก.....	55
ตารางที่ 4.4 ผลการทดลองในสภาพแวดล้อมแบบ 3 ช่องจราจรใน 2 แยก.....	56
ตารางที่ 4.5 ผลการทดลองในสภาพแวดล้อมแบบ 3 ช่องจราจรใน 3 แยก.....	58
ตารางที่ 4.6 ผลการทดลองระบบจัดลำดับความสำคัญรถพยาบาลที่มีระยะเวลาความปลอดภัยและ ระยะเวลาตรวจจับแตกต่างกัน.....	59
ตารางที่ 4.7 ผลการทดลองเมื่อมีระบบจัดลำดับความสำคัญของรถพยาบาลและไม่มีระบบจัดลำดับ ความสำคัญของรถพยาบาล.....	60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สารบัญรูป

รูปที่	หน้า
รูปที่ 2.1 วงจรการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning Cycle) [30]	9
รูปที่ 2.2 โครงข่ายประสาทเทียม	10
รูปที่ 2.3 กระบวนการคำนวณภายในนิวรอน	10
รูปที่ 2.4 ฟังก์ชัน ReLU	11
รูปที่ 2.5 ฟังก์ชัน Sigmoid	12
รูปที่ 2.6 ฟังก์ชัน Tanh	12
รูปที่ 2.7 (ก) ตรรกะแบบดั้งเดิม (ข) ตรรกะแบบฟัซซี	15
รูปที่ 2.8 (ก) เซตวัยหนุ่มสาวแบบดั้งเดิม (ข) เซตวัยหนุ่มสาวแบบฟัซซี.....	16
รูปที่ 2.9 (ก) เซตช่วงอายุแบบดั้งเดิมที่มีหลายเซต (ข) เซตช่วงอายุแบบฟัซซีที่มีหลายเซต.....	17
รูปที่ 2.10 โครงสร้างการทำงานพื้นฐานของตรรกะแบบฟัซซี.....	19
รูปที่ 2.11 ฟังก์ชันความเป็นสมาชิกสามเหลี่ยม เมื่อ $a = -5$, $b = 0$, $c = 5$	21
รูปที่ 2.12 ฟังก์ชันความเป็นสมาชิกสี่เหลี่ยมคางหมู เมื่อ $a = -10$, $b = -5$, $c = 5$, $d = 10$	22
รูปที่ 2.13 ฟังก์ชันความเป็นสมาชิกเกาส์เซียน เมื่อ $c = 0$, $\sigma = 2$	22
รูปที่ 3.1 (ก) สภาพแวดล้อมทางแยกที่เป็นต้นแบบ (ข) สภาพแวดล้อมที่มี 3 ช่องจราจรใน 1 แยก เรียกว่า สภาพแวดล้อมแบบ N-3 (ค) สภาพแวดล้อมที่มี 3 ช่องจราจรใน 2 แยก เรียกว่าสภาพแวดล้อมแบบ E-3 S-3 (ง) สภาพแวดล้อมที่มี 3 ช่องจราจรใน 3 แยก เรียกว่าสภาพแวดล้อมแบบ S-3 N-3 W-3	26
รูปที่ 3.2 ระยะเวลาจับรถพยาบาลในระบบจัดลำดับความสำคัญรถพยาบาล.....	27
รูปที่ 3.3 (ก) ภาพรวมการจราจรบริเวณทางแยก (ข) ภาพรวมการจราจรทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์ (ค) เวลารอคอยของรถแต่ละคันทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์ (ง) ค่าสถานะทางด้านทิศใต้ของทางแยกที่จะถูกเก็บไว้.....	30
รูปที่ 3.4 ทิศทางของเฟสสัญญาณไฟจราจร	31
รูปที่ 3.5 ตัวอย่างสถานการณ์ที่ส่งผลต่อความคลาดเคลื่อนของเวลารอสะสมในวิธีต้นแบบ.....	33
รูปที่ 3.6 กลุ่มของช่องจราจร	34
รูปที่ 3.7 (ก) การสร้างสภาพการจราจรใน 1 ตอนโดยอิงตามประเภทของยานพาหนะ (ข) การสร้างสภาพการจราจรใน 1 ตอนโดยอิงตามทิศทางของยานพาหนะ (ค) การสร้างสภาพการจราจรใน 1 ตอนโดยอิงตามทิศทางของทางแยก	40

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

รูปที่ 3.8 (ก) ภาพรวมการจราจรบริเวณทางแยก
 (ข) ภาพรวมการจราจรทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์
 (ค) ค่าที่ระบุตำแหน่งของรถแต่ละคันทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์
 (ง) ค่าที่ระบุตำแหน่งของรถทางด้านทิศใต้ของทางแยกที่จะถูกเก็บไว้..... 42

รูปที่ 3.9 ฟังก์ชันเซตของเฟสไฟเขียว (GP)..... 43

รูปที่ 3.10 ฟังก์ชันเซตของเฟสไฟแดง (RP) 44

รูปที่ 3.11 ฟังก์ชันเซตของเอาต์พุต..... 47

รูปที่ 3.12 ตัวอย่างผลลัพธ์ที่ได้รับจากกระบวนการทำดีฟิซซิฟิเคชัน 49

รูปที่ 4.1 (ก) กราฟแสดงผลการฝึกรวมอิงตามรางวัลที่มีค่าติดลบ
 (ข) กราฟแสดงผลการฝึกรวมอิงตามเวลารอคอยรวมโดยเฉลี่ย
 (ค) กราฟแสดงผลการฝึกรวมอิงตามเวลารอคอยโดยเฉลี่ยของยานพาหนะที่จะเลี้ยวซ้าย
 (ง) กราฟแสดงผลการฝึกรวมอิงตามเวลารอคอยโดยเฉลี่ยต่อคัน 52



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
 ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1

บทนำ

1.1 ความเป็นมาและความสำคัญของปัญหา

ในปัจจุบัน หลายประเทศทั่วโลกต้องเผชิญหน้ากับปัญหาการจราจรติดขัด โดยการเพิ่มขึ้นของประชากรในเมืองส่งผลให้จำนวนของรถยนต์บนถนนเพิ่มขึ้นอย่างมีนัยสำคัญ ซึ่งเป็นต้นเหตุของปัญหานี้ ผลที่ตามมาคือต้องใช้ระยะเวลาในการเดินทางที่นานมากขึ้น โดยเฉพาะอย่างยิ่งในช่วงที่มีความต้องการในการสัญจรสูง ซึ่งในปี พ.ศ. 2567 ผู้คนที่อาศัยอยู่ในกรุงเทพฯ ต้องเผชิญกับการสูญเสียเวลาอย่างมากจากการจราจรที่หนาแน่น รวมทั้งหมด 74 ชั่วโมง โดยเพิ่มขึ้น 17% จากปีก่อนหน้า [1]

สัญญาณไฟจราจรเป็นเครื่องมือที่ใช้กันมากที่สุดในการบริหารจัดการระบบการจราจรภายในเมือง อย่างไรก็ตาม ปัญหาที่มักเกิดขึ้นเมื่อการควบคุมสัญญาณไฟจราจรที่ไม่สอดคล้องกับปริมาณรถยนต์บนท้องถนน นำไปสู่การจราจรที่ติดขัดและเวลารอคอยที่นานมากขึ้น ดังนั้น ในการศึกษาครั้งนี้ เรานำเสนอวิธีการที่ใช้ในการควบคุมสัญญาณไฟจราจรโดยใช้การเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning) ร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) เพื่อที่จะเพิ่มประสิทธิภาพของการควบคุมสัญญาณไฟจราจร วิธีนี้จะใช้ข้อมูลจากจำนวนรถและเวลารอคอย (Waiting Time) เพื่อปรับสัญญาณไฟจราจรตามสภาพแวดล้อมการจราจรในขณะนั้น โดยมีวัตถุประสงค์เพื่อที่จะลดเวลารอคอยและการจราจรที่ติดขัดให้น้อยที่สุด พร้อมทั้งเพิ่มความปลอดภัยบนท้องถนน

เมื่อการจราจรติดขัดลดลง รถยนต์สามารถเคลื่อนตัวได้คล่องขึ้น จะสามารถลดโอกาสเกิดอุบัติเหตุหรือสถานการณ์อันตรายได้ [2] การจราจรที่มีความคล่องตัวสามารถลดการหยุดกะทันหันและเพิ่มระยะห่างระหว่างรถยนต์แต่ละคัน ซึ่งทั้งสองสาเหตุนี้เป็นสาเหตุทั่วไปที่สามารถทำให้เกิดอุบัติเหตุได้ และในขณะเดียวกันยังสามารถเพิ่มประสิทธิภาพการตอบสนองของผู้ขับขี่ต่อสถานการณ์ต่าง ๆ และลดความเครียดได้อีกด้วย ดังนั้น เมื่อการจราจรติดขัดลดลง จึงมีส่วนช่วยให้สภาพแวดล้อมบนท้องถนนโดยรวมปลอดภัยยิ่งขึ้น

ในปี พ.ศ. 2520 Pappis และ Mamdani [3] ได้นำเสนอการใช้ตรรกะแบบฟัซซี (Fuzzy Logic) เป็นแนวทางในการควบคุมสัญญาณไฟจราจร ซึ่งนำมาใช้กับบริเวณทางแยกเดี่ยว (Single Intersection) และหลังจากนั้นได้มีนักวิจัยหลายท่านได้ศึกษาการประยุกต์ใช้ตรรกะแบบฟัซซี (Fuzzy Logic) ในการควบคุมสัญญาณไฟจราจรกันอย่างแพร่หลาย โดยรวบรวมข้อมูลจากสถานการณ์ต่าง ๆ บริเวณทางแยก เช่น จำนวนของรถยนต์ที่รอสัญญาณไฟแดง จำนวนของรถยนต์ที่จะได้รับสัญญาณไฟเขียว [4-6] เวลาเริ่มต้นของวงจร โอกาสที่จะเปลี่ยนแปลงสัญญาณไฟ [5] ความยาวของคิว [6,7] เวลารอคอย [8] และจำนวนของรถยนต์ทั้งหมด [9] ตัวแปรเหล่านี้ถูกใช้ในการ

วิเคราะห์และกำหนดเฟสของสัญญาณไฟจราจร [5,6] หรือระยะเวลาของสัญญาณไฟเขียว [4,7,10] เอกสารนี้เป็นเอกสารที่สงวนลิขสิทธิ์สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เครื่องอนุมานหรือการตีความ (Inference Engine) ใช้ฐานกฎ (Rule base) ซึ่งไม่จำเป็นต้องใช้โครงสร้างโมเดลทางคณิตศาสตร์ที่ซับซ้อน [7] โดยจะสร้างกฎที่มีความคล้ายคลึงกับการที่มนุษย์ใช้ในการควบคุมสัญญาณไฟจราจรด้วยการใช้เหตุผล [8,10,11] อย่างไรก็ตาม ตรรกะแบบฟัซซี (Fuzzy Logic) ยังคงมีข้อจำกัดเนื่องจากไม่สามารถเรียนรู้ข้อมูลด้วยตัวเองและไม่สามารถใช้กับสถานการณ์ที่มีความซับซ้อนมากได้

ปัญญาประดิษฐ์ (Artificial Intelligence) เข้ามามีบทบาทสำคัญในการแก้ปัญหาต่าง ๆ มากมายได้สำเร็จ โดยที่การเรียนรู้แบบมีผู้สอน (Supervised Learning) การเรียนรู้แบบไม่มีผู้สอน (Unsupervised Learning) และการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) [12,13] ถูกนำมาใช้แก้ปัญหาหลาย ๆ อย่าง วิธีการเรียนรู้เหล่านี้จัดอยู่ในประเภทของการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งแตกต่างจากตรรกะแบบฟัซซี (Fuzzy logic) ที่ไม่ได้ใช้อัลกอริทึมในการเรียนรู้

การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) มักจะรู้จักกันในชื่อของ Q-Learning เป็นวิธีการที่เกี่ยวข้องกับการเรียนรู้ซ้ำ ๆ และการใช้ค่าคิว (Q-Value) ที่คำนวณได้เพื่อเลือกการกระทำที่เหมาะสมที่สุด [14,15] โดยหลักการพื้นฐานสะท้อนถึงมนุษย์ที่มีการเรียนรู้ผ่านการลองผิดลองถูก รวบรวมจากข้อมูลเรียลไทม์มากำหนดความเหมาะสมของการดำเนินการตามสภาพแวดล้อมในขณะนั้น [16] มีการประเมินผลการกระทำซึ่งเป็นรางวัล (Reward) ตามที่กำหนดไว้ โดยรางวัล (Reward) ใช้ในการคำนวณค่าคิว (Q-Value) และจัดเก็บไว้ในตารางที่เรียกว่าตารางคิว (Q-Table) ตารางนี้จะใช้เก็บข้อมูลของทุกสถานการณ์และใช้เป็นข้อมูลอ้างอิงในการตัดสินใจ [17,18]

เนื่องจากความซับซ้อนของสถานการณ์การจราจรและสถานะ (State) จำนวนมาก ซึ่งเป็นข้อจำกัดของ Q-Learning ในการจัดเก็บข้อมูลในตารางคิว (Q-Table) ที่มีขนาดใหญ่เกินไป ทำให้เกิดการประยุกต์ใช้ระหว่างการเรียนรู้เชิงลึก (Deep Learning) และการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) เรียกวิธีการนี้ว่า Deep Q-Learning หรือ Deep Q-network (DQN) ซึ่งวิธีการนี้สามารถจัดการกับสถานะ (State) ขนาดใหญ่และซับซ้อนได้อย่างมีประสิทธิภาพ และถูกนำมาใช้กันอย่างแพร่หลายในหลายงานวิจัย เนื่องจากมีความเรียบง่ายและมีประสิทธิภาพ วิธีการนี้มีการปรับปรุงการเรียนรู้ให้ดีขึ้นโดยใช้การเล่นซ้ำประสบการณ์ (Experience Replay) โดยโครงข่ายประสาทเทียม (Neural Network) จะมาแทนที่ตารางคิว (Q-Table) ซึ่งช่วยลดความยุ่งยากในการกำหนดค่าและการปรับแต่ง โดยที่ไม่ต้องมีการออกแบบที่ซับซ้อนเกินไป โดยจะทำนายค่าคิว (Q-Value) และปรับน้ำหนักในโครงข่ายประสาทเทียม (Neural Network) เพื่อประมาณและปรับให้เข้ากับค่าคิว (Q-Value) ที่เหมาะสมที่สุดสำหรับสถานการณ์นั้น ๆ สิ่งนี้แตกต่างกับ Q-Learning แบบดั้งเดิม ซึ่งการเรียนรู้จะเกิดขึ้นตั้งแต่เริ่มต้นจึงทำให้ Deep Q-Learning มีความยืดหยุ่นมากกว่า Q-Learning [19] อย่างไรก็ตาม หลักการพื้นฐานยังคงมีความใกล้เคียงกัน ในวิธีการนี้สถานะ (State)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

หมายถึง การรวบรวมข้อมูลสถานการณ์ เช่น ตำแหน่งของรถยนต์ [20] ความเร็ว [21,22] และจำนวนรถยนต์ในแต่ละช่องจราจร (Lane) [23]

การกระทำ (Action) ของ Deep Q-network (DQN) คือ เฟสของสัญญาณไฟจราจรที่เหมาะสมที่สุดสำหรับสถานการณ์ที่กำหนดจะถูกเลือก อย่างไรก็ตาม ระยะเวลาของสัญญาณไฟเขียวจะถูกกำหนดไว้ล่วงหน้า และไม่ได้ปรับให้เข้ากับสถานการณ์ในขณะนั้น [23,24] โดยที่การคำนวณมูลค่ารางวัล (Reward) จะแตกต่างกันไปในแต่ละการศึกษา โดยปัจจัยหลักคือเวลารอคอย [24,25] และบางส่วนใช้ความยาวคิวและความเร็วของรถ [21]

การควบคุมการจราจรโดยใช้การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) นอกเหนือจาก Deep Q-network (DQN) ที่ได้รับความนิยมอย่างแพร่หลายแล้ว ยังมีวิธีการอื่นๆ เช่น Twin Delayed Deep Deterministic Policy Gradient (TD3) และ Deep Deterministic Policy Gradient (DDPG) วิธีการเหล่านี้มีประสิทธิภาพสูงในการจัดการกับปัญหาในพื้นที่การดำเนินการที่ต่อเนื่อง (Continuous Action Space) หรือการควบคุมที่มีความซับซ้อนมาก ๆ เช่น สำหรับยานยนต์อัตโนมัติหรือรถยนต์ไร้คนขับที่ต้องการรับค่าอย่างต่อเนื่อง ซึ่งช่วยให้สามารถปรับตัวเข้ากับสภาพแวดล้อมได้ตลอดเวลา หรือ การควบคุมอุณหภูมิที่ต้องปรับอินพุตอย่างต่อเนื่องเพื่อรักษาสถานะที่ต้องการ

สำหรับ DDPG [26] มีการใช้แนวทางนักแสดง-นักวิจารณ์ (Actor-Critic) โดยที่นักแสดง (Actor) เป็นผู้กำหนดนโยบายการควบคุมและผู้วิจารณ์ (Critic) จะประเมินนโยบายนี้ ทำให้มีประสิทธิภาพสำหรับสถานการณ์การควบคุมอย่างต่อเนื่อง (Continuous Control)

เพื่อแก้ไขปัญหาใน DDPG ได้มีการนำ TD3 [27] ที่ได้รับการพัฒนาโดยการใช้อัปเดตนโยบายล่าช้า (Delayed Policy Updates) และใช้เทคนิค Clipped Double Q-Learning ซึ่งช่วยเพิ่มความเสถียรของโมเดล

อย่างไรก็ตาม DQN มีความเหมาะสมสำหรับการนำไปใช้ในการตัดสินใจในพื้นที่การดำเนินการที่ไม่ต่อเนื่อง (Discrete Action Space) มากกว่า [28] เช่น การปรับเฟสสัญญาณไฟจราจรโดยจะเลือกเฟสใหม่หลังจากเฟสก่อนหน้าเสร็จสิ้นแล้วเท่านั้น วิธีการนี้จะช่วยจัดการปัญหาเหล่านี้ได้อย่างมีประสิทธิภาพมากขึ้น และตรงไปตรงมา อีกทั้งยังมีโครงสร้างและการนำไปใช้ที่ง่ายกว่าอีกด้วย

การศึกษาของ Tunc และ Soylemez [29] ใช้การรวมกันระหว่าง Deep Q-Learning และตรรกะแบบฟัซซี (Fuzzy Logic) ในการเปลี่ยนเฟสของสัญญาณไฟจราจรและกำหนดระยะเวลาสัญญาณไฟเขียวตามสภาพแวดล้อมการจราจร โดยที่ใช้ Deep Q-Learning เลือกเฟสของสัญญาณไฟจราจรที่เหมาะสม และใช้ตรรกะแบบฟัซซี (Fuzzy Logic) ในการคำนวณระยะเวลาของสัญญาณไฟเขียว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลลัพธ์จากการศึกษาของ Tunc และ Soylemez แสดงให้เห็นว่าวิธีการนี้ให้ผลลัพธ์ที่ดีกว่าวิธีการอื่น ๆ เช่น การควบคุมสัญญาณไฟจราจรแบบดั้งเดิมที่มีการกำหนดเวลาไว้ล่วงหน้า การใช้ Deep Q-Learning เพียงอย่างเดียว และการใช้ตรรกะแบบฟัซซี (Fuzzy Logic) เพียงอย่างเดียว

อย่างไรก็ตาม หนึ่งปัญหาที่พบในการศึกษานี้เกี่ยวข้องกับการจราจรแบบซับซ้อน โดยเฉพาะฟังก์ชันรางวัล (Reward Function) ที่ใช้ อาจคำนวณรางวัลที่ไม่เหมาะสม ซึ่งส่งผลให้รถยนต์ที่อยู่ในช่องจราจร (Lane) ที่ต้องการจะเลี้ยวซ้ายต้องรอสัญญาณไฟนานเกินไป เป็นผลมาจากการคำนวณค่าของรางวัล (Reward) ที่ทำให้ระบบเรียนรู้ว่าสามารถทำให้มีประสิทธิภาพมากยิ่งขึ้นได้ โดยการให้สัญญาณไฟตรงไป ดังนั้น ในการศึกษาจะใช้ Deep Q-Network ที่ได้รับการปรับปรุงและตรรกะแบบฟัซซี (Fuzzy Logic) เพื่อเพิ่มประสิทธิภาพในการจัดการสัญญาณไฟจราจรที่ทางแยกเดี่ยว (Single Intersection)

1.2 ความมุ่งหมายและวัตถุประสงค์ของการวิจัย

วิทยานิพนธ์นี้มุ่งเน้นไปที่การพัฒนาการควบคุมสัญญาณไฟจราจรให้มีประสิทธิภาพมากยิ่งขึ้น และเปรียบเทียบผลลัพธ์ที่ได้กับวิธีการแบบดั้งเดิมที่ยังคงใช้กันอยู่อย่างแพร่หลายในปัจจุบันและวิธีการของงานวิจัยที่มีอยู่ก่อนหน้านี้ โดย Tunc และ Soylemez [29] ที่ใช้การเรียนรู้แบบเสริมกำลังเชิงลึกร่วมกับตรรกะแบบฟัซซี ซึ่งวิธีการของเรามีคุณลักษณะของการเรียนรู้แบบเสริมกำลังเชิงลึกที่ได้รับการปรับปรุงดังต่อไปนี้

1. การปรับปรุงสถานะ (State Improvement) : สถานะ (State) ได้รับการปรับปรุงโดยแก้ไขให้สะท้อนกับมุมมองของบุคคลที่ใช้ในการตัดสินใจในการควบคุมสัญญาณไฟจราจร โดยที่สถานะ (State) จะเก็บข้อมูลของเวลารอคอย (Waiting Time) ที่ถูกแบ่งออกเป็น 10 ระดับและตำแหน่งของรถยนต์

2. การปรับปรุงรางวัล (Reward Improvement) : รางวัล (Reward) ได้รับการปรับปรุงโดยคำนวณจากผลรวมของเวลารอคอยของรถยนต์ทุกคันในแต่ละช่องจราจร (Lane) และนำผลรวมของเวลารอคอยของช่องจราจร (Lane) ที่ใช้สัญญาณไฟร่วมกันที่มีค่าสูงที่สุดจากแต่ละทางแยกมาใช้ในการคำนวณรางวัล (Reward)

1.3 ขอบเขตการวิจัย

การวิจัยนี้ดำเนินการโดยใช้โปรแกรม SUMO (Simulation of Urban Mobility) ในการจำลองสถานการณ์จราจรบริเวณทางแยกแบบสี่แยก โดยจำลองให้ใกล้เคียงกับสภาพการจราจรจริงมากที่สุด ทั้งในด้านโครงสร้างทางแยก จำนวนช่องจราจร และประเภทของยานพาหนะ ซึ่งประกอบด้วยรถยนต์ส่วนบุคคล รถบัส รถมินิบัส และรถพยาบาล

สภาพแวดล้อมจำลองครอบคลุมกรณีทางแยกที่มีจำนวนช่องจราจรไม่เท่ากันในแต่ละทิศทาง การจำลองดำเนินการในช่วงเวลา 5,400 วินาทีต่อรอบ โดยมียานพาหนะจำนวน 1,000 คันต่อรอบ เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่ออนุญาตให้นำไปใช้ประโยชน์ในการศึกษาไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เพื่อทดสอบประสิทธิภาพของระบบควบคุมสัญญาณไฟจราจรภายใต้สถานการณ์ที่แตกต่างกัน เกณฑ์ที่ใช้ประเมินผลได้แก่ เวลาเฉลี่ยในการรอของยานพาหนะโดยรวม เวลาเฉลี่ยของรถที่เลี้ยวซ้าย และเวลาเฉลี่ยที่รถแต่ละคันต้องรอ นอกจากนี้ ยังมีการทดสอบระบบการให้ความสำคัญกับรถพยาบาล ภายใต้เงื่อนไขของการจำลองจราจร เพื่อประเมินประสิทธิภาพในการลดระยะเวลาการรอของรถพยาบาล และเพิ่มความปลอดภัยในการผ่านทางแยกในกรณีฉุกเฉิน

1.4 ประโยชน์ที่คาดว่าจะได้รับ

การควบคุมสัญญาณไฟจราจรที่ใช้การเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning) และตรรกะแบบฟัซซี (Fuzzy Logic) คาดว่าจะช่วยลดปัญหาการจราจรติดขัดและลดระยะเวลาในการรอของยานพาหนะบริเวณทางแยกได้อย่างมีประสิทธิภาพ อันจะนำไปสู่การเพิ่มความคล่องตัวในการเดินทาง และสามารถประยุกต์ใช้เป็นแนวทางในการพัฒนาระบบควบคุมสัญญาณไฟจราจรอัจฉริยะในอนาคต



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 2

ทฤษฎีที่เกี่ยวข้อง

หลักการควบคุมสัญญาณไฟจราจรแบบปรับได้โดยใช้การเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning หรือ DRL) และตรรกะแบบฟัซซี (Fuzzy Logic) เกี่ยวข้องกับการใช้การเรียนรู้แบบเสริมกำลังเชิงลึก (DRL) ซึ่งใช้โครงข่ายประสาทเทียม (Neural Network) และการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ร่วมกันกับตรรกะฟัซซี (Fuzzy Logic) เพื่อจัดการสัญญาณไฟจราจรตามสภาพการจราจร โมเดลการเรียนรู้แบบเสริมกำลังเชิงลึก (DRL) เรียนรู้วิธีการควบคุมที่มีประสิทธิภาพสูงสุดโดยการโต้ตอบกับสภาพแวดล้อมการจราจรและรับรางวัล (Reward) ในรูปแบบของการลดเวลารอคอยเพื่อกำหนดเฟสของสัญญาณไฟจราจร ในขณะเดียวกัน ตรรกะแบบฟัซซี (Fuzzy Logic) จะกำหนดระยะเวลาของไฟเขียว ซึ่งขึ้นอยู่กับจำนวนรถบริเวณทางแยก พื้นฐานของวิธีการที่ใช้สำหรับการควบคุมสัญญาณไฟจราจรแบบปรับได้ในการศึกษาแบ่งเนื้อหาออกเป็น ส่วนย่อยดังนี้

2.1 แนวคิดของโมเดลต้นแบบ

งานวิจัยที่เป็นโมเดลต้นแบบ โดย Tunc และ Soylemez [29] ที่นำมาใช้เป็นแนวทางในการพัฒนาระบบควบคุมสัญญาณไฟจราจรในวิทยานิพนธ์ฉบับนี้ ได้เสนอแนวทางการควบคุมสัญญาณไฟจราจร โดยประยุกต์ใช้การเรียนรู้แบบเสริมกำลังเชิงลึกและตรรกะแบบฟัซซี เพื่อนำมาใช้ในการจัดการสัญญาณไฟจราจรบริเวณทางแยก ซึ่งถือเป็นจุดสำคัญที่ส่งผลต่อประสิทธิภาพของระบบจราจรโดยรวม การควบคุมที่มีประสิทธิภาพในบริเวณนี้สามารถช่วยลดเวลารอคอยของยานพาหนะ ลดปริมาณก๊าซคาร์บอนไดออกไซด์จากการเผาไหม้ที่ไม่จำเป็น และช่วยให้การจราจรมีความลื่นไหลมากยิ่งขึ้น โดยมีแนวคิดหลัก คือ การแบ่งหน้าที่การตัดสินใจออกเป็น 2 ส่วน ได้แก่

1. การกำหนดลำดับของเฟสสัญญาณไฟจราจร โดยใช้การเรียนรู้แบบเสริมกำลังเชิงลึก ซึ่งเรียนรู้จากประสบการณ์ในอดีตและเลือกการกระทำที่ให้ผลตอบแทนรวมสูงที่สุดในระยะยาว
2. การกำหนดระยะเวลาของสัญญาณไฟเขียว โดยใช้ตรรกะแบบฟัซซีซึ่งสามารถจัดการกับข้อมูลที่ไม่น่าเชื่อถือและเปลี่ยนแปลงตลอดเวลาได้อย่างยืดหยุ่น โดยไม่จำเป็นต้องตั้งค่าระยะเวลาแบบคงที่เหมือนระบบควบคุมแบบดั้งเดิม

ในด้านการออกแบบสถานะของระบบ ซึ่งเป็นองค์ประกอบสำคัญของการเรียนรู้แบบเสริมกำลัง ได้นำเสนอแนวทางการแบ่งช่องทางจราจรในแต่ละทิศทางออกเป็นเซลล์ย่อย จะมีเซลล์ทั้งหมด 10 เซลล์ต่อช่องทาง รวม 80 เซลล์ และเซลล์แต่ละเซลล์มีขนาดไม่เท่ากัน โดยเซลล์ที่อยู่ใกล้เส้นหยุดจะมีขนาดเล็กกว่าเซลล์ที่อยู่ห่างออกไปจากทางแยก ซึ่งแนวคิดนี้สะท้อนถึงหลักการที่ว่า ยานพาหนะที่อยู่ใกล้ทางแยกมากจะมีผลต่อประสิทธิภาพของระบบมากกว่า โดยการแทนค่าสถานะของแต่ละเอกสารนี้เป็นเอกสารที่สแกนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เซลล์จะใช้ค่าบูลีน คือ หากตรวจพบว่ามียานพาหนะอยู่ในเซลล์นั้น จะกำหนดค่าเป็น 1 หากไม่มีจะกำหนดเป็น 0 ซึ่งทำให้ข้อมูลสถานะมีลักษณะเป็นเวกเตอร์ของค่าไบนารีที่สะท้อนถึงตำแหน่งของยานพาหนะในแต่ละช่วงระยะทาง

การกระทำที่ระบบสามารถเลือกกระทำได้ ประกอบด้วยเฟสสัญญาณไฟจราจรต่าง ๆ เช่น การเปิดสัญญาณไฟเขียวให้รถที่วิ่งจากทิศตะวันออกไปทิศตะวันตกหรือจากทิศเหนือไปทิศใต้ ทั้งในกรณีทางตรงและการเลี้ยว โดยในช่วงเวลาใดเวลาหนึ่ง ระบบจะสามารถเลือกเฟสได้เพียงเฟสเดียว การตัดสินใจจะดำเนินไปตามนโยบายที่ได้เรียนรู้ผ่านประสบการณ์ในสภาพการจราจรที่หลากหลาย

ค่ารางวัลที่ใช้ในการเรียนรู้ ถูกคำนวณตามเวลารอสะสมของยานพาหนะทั้งหมดในระบบ โดยเปรียบเทียบค่าระหว่างสองช่วงเวลา หากระบบสามารถลดเวลารอลงได้ จะได้รับค่ารางวัลที่เป็นบวก แต่หากทำให้เวลารอเพิ่มขึ้น จะได้รับค่ารางวัลที่เป็นลบ แนวทางนี้สะท้อนถึงจุดมุ่งหมายของระบบในการลดเวลารอของรถทุกคันให้น้อยที่สุดเท่าที่เป็นไปได้ โดยมีสมการที่ใช้ในการคำนวณค่ารางวัล ดังแสดงในสมการที่ 2.1 และ 2.2

$$Twt_t = \sum_{i=1}^n wt_{i,t} \quad (2.1)$$

โดยที่ $wt_{i,t}$ คือ เวลารอสะสมของยานพาหนะในเวลาขั้นตอนที่ t
 n คือ จำนวนยานพาหนะทั้งหมดในเวลาขั้นตอนที่ t

$$r_t = Twt_{t-1} - Twt_t \quad (2.2)$$

โดยที่ Twt_t คือ เวลารอสะสมของยานพาหนะทั้งหมดในช่วงเวลาปัจจุบัน
 Twt_{t-1} คือ เวลารอสะสมของยานพาหนะทั้งหมดในช่วงเวลาก่อนหน้า

ในส่วนของตรรกะแบบฟัซซี ซึ่งใช้กำหนดระยะเวลาของสัญญาณไฟเขียว จะรับข้อมูลจากจำนวนยานพาหนะที่อยู่ในช่องทางที่กำลังจะได้รับสัญญาณไฟเขียว และช่องทางที่กำลังรอสัญญาณไฟแดง แล้วนำค่าดังกล่าวไปประเมินผ่านกฎฟัซซีที่ออกแบบไว้ล่วงหน้า เพื่อกำหนดช่วงเวลาเปิดสัญญาณไฟเขียวที่เหมาะสม โดยช่วงเวลาที่ได้จะไม่ตายตัว แต่สามารถปรับเปลี่ยนได้ในแต่ละสถานการณ์ ช่วยให้ระบบมีความยืดหยุ่นสูงและลดผลกระทบจากการตั้งค่าแบบตายตัวที่อาจไม่เหมาะสมกับสภาพจราจรจริง

สำหรับการจำลองได้ใช้โปรแกรม Simulation of Urban Mobility (SUMO) ในการสร้างสถานการณ์จำลองของการจราจรที่บริเวณสี่แยก โดยกำหนดให้มีการสร้างยานพาหนะจำนวน 3000 คัน ต่อหนึ่งตอนการจำลอง (Episode) แต่ละตอนใช้เวลาจำลอง 5400 วินาที หรือ เทียบเท่ากับ 90 นาที ทั้งนี้ยานพาหนะทั้งหมดเป็นประเภทเดียวกัน มีความยาว 5 เมตร ความเร็วเริ่มต้น 36 กม./ชม. ความเร็ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สูงสุด 90 กม./ชม. อัตราเร่ง 1 ม./วินาที² และอัตราชะลอ 4.5 ม./วินาที² โดยยานพาหนะจะผ่านทางแยกเพียงครั้งเดียวในแต่ละตอน และมีการกระจายทิศทางเคลื่อนที่อย่างสม่ำเสมอในทุกทิศทาง ความน่าจะเป็น โดยประมาณ 75% ของรถทั้งหมดจะวิ่งตรงไป ส่วนที่เหลืออีก 25% จะแยกไปทางซ้ายหรือขวา

และจากผลการทดลอง พบว่า แนวทางที่ผสมผสานการเรียนรู้แบบเสริมกำลังเชิงลึกเข้ากับตรรกะแบบฟัซซี สามารถลดเวลารอของยานพาหนะบริเวณแยกสี่ทางได้อย่างมีนัยสำคัญ เมื่อเปรียบเทียบกับวิธีควบคุมแบบดั้งเดิม รวมถึงวิธีการอื่นที่ทำการทดลองเปรียบเทียบภายในงานวิจัย ระบบสามารถลดจำนวนรถที่รออยู่ ลดระยะเวลาการเฉลี่ย และยังรักษาเสถียรภาพของการจราจรได้ดี โดยผลการทดลองสะท้อนให้เห็นว่าแนวทางดังกล่าวเป็นวิธีที่มีประสิทธิภาพมากที่สุดในกลุ่มวิธีที่ได้ศึกษาไว้ในงานวิจัย จึงเป็นพื้นฐานสำคัญในการคัดเลือกแนวทางดังกล่าวมาใช้เป็นต้นแบบ อันนำไปสู่การพัฒนาและต่อยอดในงานวิจัยที่นำเสนอในวิทยานิพนธ์ฉบับนี้

2.2 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)

การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) [30,31] เป็นอัลกอริทึมการเรียนรู้ของเครื่อง (Machine Learning) ประเภทหนึ่ง ซึ่งเป็นระบบที่เรียนรู้ผ่านการลองผิดลองถูกด้วยตนเอง ทำงานผ่านการกระทำ (Action) เพื่อเพิ่มประสิทธิภาพผลตอบแทน (Reward) และรับความรู้ผ่านประสบการณ์จริง เพื่อให้ได้มาซึ่งผลตอบแทนสูงสุด ซึ่งประกอบด้วยปัจจัยที่ใช้ในการเรียนรู้พื้นฐาน ดังแสดงในรูปที่ 2.1

สิ่งแวดล้อม (Environment) : สภาพแวดล้อมเฉพาะที่ต้องการดำเนินการ

ตัวแทน (Agent) : เปรียบเสมือนคนงานที่ทำหน้าที่ควบคุมภายในสภาพแวดล้อมเฉพาะ

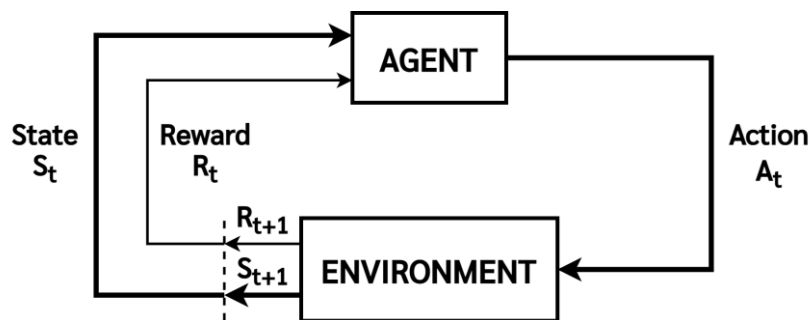
การกระทำ (Action) : การกระทำที่เกี่ยวข้องซึ่งสามารถดำเนินการได้ในสภาพแวดล้อมเฉพาะนั้น

สถานะ (State) : กระบวนการตัดสินใจใช้สถานะปัจจุบันหรือสถานการณ์ของสิ่งแวดล้อมเพื่อตัดสินใจ

รางวัล (Reward) : มูลค่ารางวัลคือการวัดประสิทธิภาพของการดำเนินการ

นโยบาย (Policy) : กฎหรือวิธีการในการพิจารณาทางเลือกที่ดีที่สุดสำหรับการกระทำตามสถานะปัจจุบัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 2.1 วงจรการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning Cycle) [30]

Markov Decision Process (MDP) [32] ใช้เพื่อกำหนดการเลือกการกระทำ (Action) กระบวนการเลือกเกี่ยวข้องกับการเลือกการกระทำ (Action) ที่เชื่อว่าจะให้ผลตอบแทนหรือผลตอบแทนสะสมสูงสุด โดยที่ผลตอบแทนสะสม (Cumulative Reward) หมายถึง ผลรวมของรางวัลทั้งหมดที่ได้รับจากช่วงเวลาปัจจุบันถึงอนาคต อย่างไรก็ตาม เนื่องจากความไม่แน่นอนตามธรรมชาติของอนาคตเมื่อคำนวณผลตอบแทนสะสม (Cumulative Reward) มีความจำเป็นต้องลดความสำคัญของผลตอบแทนที่คาดหวังลงด้วย Discount Factor (γ) ที่มีค่าตั้งแต่ 0 ถึง 1 ซึ่งถูกนำไปใช้ในสมการที่ 2.3

$$G_t = R_{t+1} + \gamma R_{t+2} + \dots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \quad (2.3)$$

โดยที่ G_t คือ ค่าผลตอบแทนสะสม
 R_{t+1} คือ รางวัลที่ได้รับจากการเปลี่ยนจากสถานะ S_t เป็น S_{t+1}
 γ คือ อัตราการลดค่าผลตอบแทนในอนาคต
 S_t คือ สถานะ ณ เวลาที่ t
 S_{t+1} คือ สถานะใหม่ หลังจากกระทำการในสถานะ S_t

2.3 โครงข่ายประสาทเทียม (Neural Network)

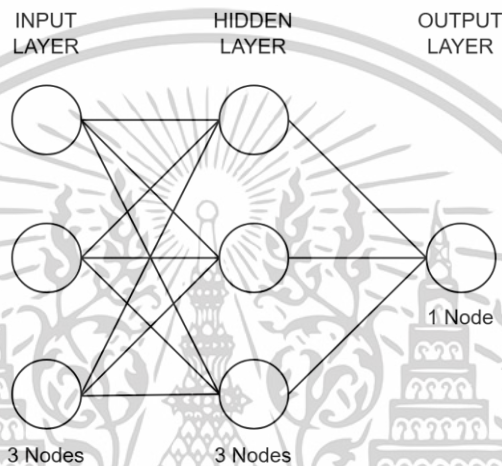
โครงข่ายประสาทเทียม (Neural Network) เป็นระบบการคำนวณเชิงคณิตศาสตร์ที่ได้รับแรงบันดาลใจจากการทำงานของระบบประสาทในสมองของมนุษย์ โดยถูกออกแบบมาเพื่อเลียนแบบกระบวนการเรียนรู้และการประมวลผลข้อมูลของเซลล์ประสาท (Neurons) ในสมอง ซึ่งโครงสร้างของ Neural Network ถูกนำมาใช้เพื่อแก้ปัญหาที่ซับซ้อน เช่น การจดจำรูปแบบ (Pattern Recognition) การทำนายข้อมูล (Prediction) และการตัดสินใจ (Decision Making)

โครงข่ายประสาทเทียมประกอบด้วยชั้น (Layers) หลายชั้น ซึ่งแต่ละชั้นประกอบด้วยโหนด (Nodes) จำนวนหนึ่งที่เชื่อมต่อกัน โดยชั้นแรกแรกสุด คือ ชั้นรับเข้า (Input Layer) ที่รับข้อมูล

นำเข้า เช่น ข้อมูลตัวเลข ข้อมูลรูปภาพ หรือข้อมูลเสียง โดยที่จำนวนนิวรอนในชั้นนี้ขึ้นอยู่กับขนาดของข้อมูลนำเข้า นอกจากนี้ยังมีชั้นซ่อน (Hidden Layer) ซึ่งทำหน้าที่ซ่อนข้อมูลนำเข้าก่อนจะส่งไปยังชั้นสุดท้าย (Output Layer) ซึ่งทำหน้าที่ทำนายผลลัพธ์ที่ต้องการ นอกจากนี้ยังมีชั้นรับเข้า (Input Layer) ที่รับข้อมูลนำเข้า เช่น ข้อมูลตัวเลข ข้อมูลรูปภาพ หรือข้อมูลเสียง โดยที่จำนวนนิวรอนในชั้นนี้ขึ้นอยู่กับขนาดของข้อมูลนำเข้า นอกจากนี้ยังมีชั้นซ่อน (Hidden Layer) ซึ่งทำหน้าที่ซ่อนข้อมูลนำเข้าก่อนจะส่งไปยังชั้นสุดท้าย (Output Layer) ซึ่งทำหน้าที่ทำนายผลลัพธ์ที่ต้องการ นอกจากนี้ยังมีชั้นรับเข้า (Input Layer) ที่รับข้อมูลนำเข้า เช่น ข้อมูลตัวเลข ข้อมูลรูปภาพ หรือข้อมูลเสียง โดยที่จำนวนนิวรอนในชั้นนี้ขึ้นอยู่กับขนาดของข้อมูลนำเข้า นอกจากนี้ยังมีชั้นซ่อน (Hidden Layer) ซึ่งทำหน้าที่ซ่อนข้อมูลนำเข้าก่อนจะส่งไปยังชั้นสุดท้าย (Output Layer) ซึ่งทำหน้าที่ทำนายผลลัพธ์ที่ต้องการ

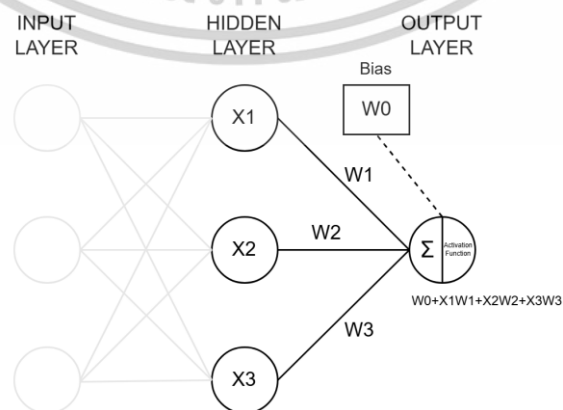
ของข้อมูลที่น่าเข้ามา ส่วนชั้นที่อยู่บริเวณตรงกลางของโครงข่ายประสาทเทียม เรียกว่า ชั้นที่ซ่อนอยู่ (Hidden Layers) โดยที่จำนวนชั้นที่ซ่อนอยู่และจำนวนนิวรอนในแต่ละชั้นส่งผลต่อความสามารถของโมเดลในการเรียนรู้ฟังก์ชันที่ซับซ้อน และชั้นสุดท้าย คือ ชั้นส่งออก (Output Layer) เป็นชั้นที่ให้ผลลัพธ์สุดท้าย เช่น การจำแนกประเภท (Classification) หรือค่าที่คาดการณ์ (Regression) โดยที่จำนวนนิวรอนในชั้นส่งออกจะขึ้นอยู่กับลักษณะของปัญหา

ในรูปที่ 2.2 เป็นการยกตัวอย่างของโครงข่ายประสาทเทียมอย่างง่าย โดยจะประกอบไปด้วย 3 ชั้น โดยที่โหนด (Nodes) ในแต่ละชั้น คือ นิวรอน (Neurons)



รูปที่ 2.2 โครงข่ายประสาทเทียม

นิวรอนในโครงข่ายประสาทเทียมทำหน้าที่คล้ายกับเซลล์ประสาทในสมองของมนุษย์ โดยรับข้อมูลเข้า (Input) จากแหล่งต่าง ๆ ทำการประมวลผลข้อมูลเหล่านั้น และส่งค่าผลลัพธ์ (Output) ออกไปยังนิวรอน (Neuron) อื่น ๆ ในเครือข่าย โดยที่นิวรอนแต่ละตัวจะเชื่อมต่อกันผ่านค่าถ่วงน้ำหนัก (Weights) และค่าคงที่ (Bias) ดังแสดงตัวอย่างในรูปที่ 2.3



รูปที่ 2.3 กระบวนการคำนวณภายในนิวรอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับข้าราชการเพื่อการศึกษาเท่านั้น ไม่นิยมนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ซึ่งแสดงถึงความสำคัญของข้อมูลนำเข้าที่แต่ละนิวรอนได้รับ โดยมีกระบวนการคำนวณภายในนิวรอนดังนี้

1. การคูณค่าข้อมูลนำเข้า (Input) กับค่าถ่วงน้ำหนัก (Weights)
2. การรวมผลถ่วงน้ำหนัก (Weighted Sum)
3. การผ่านค่าที่ได้ผ่านฟังก์ชันกระตุ้น (Activation Function) เพื่อกำหนดค่าผลลัพธ์

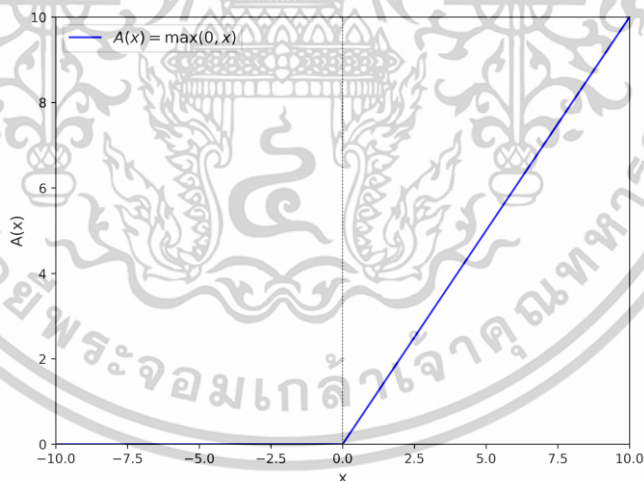
2.3.1 ฟังก์ชันกระตุ้น (Activation Function)

ฟังก์ชันกระตุ้น (Activation Function) จะเป็นการทำให้ผลลัพธ์ให้อยู่ในช่วงที่ต้องการ ซึ่งมีบทบาทสำคัญในการทำให้โครงข่ายสามารถเรียนรู้ความซับซ้อนและสร้างรูปแบบที่ไม่เป็นเส้นตรง (Non-linear patterns) ได้

โดยฟังก์ชันกระตุ้นที่นิยมใช้งานโดยทั่วไป ได้แก่

1. ฟังก์ชัน ReLU (Rectified Linear Unit Function) เป็นฟังก์ชันกระตุ้นที่ได้รับความนิยมมากในโครงข่ายประสาทเทียม (Neural Networks) ฟังก์ชันนี้จะคืนค่าเป็นศูนย์ หากอินพุตมีค่าต่ำกว่าศูนย์ และคืนค่าเท่ากับอินพุตที่นำเข้ามา หากอินพุตมีค่ามากกว่าหรือเท่ากับศูนย์ ดังแสดงในรูปที่ 2.4 โดยมีนิยามดังนี้

$$A(x) = \max(0, x)$$

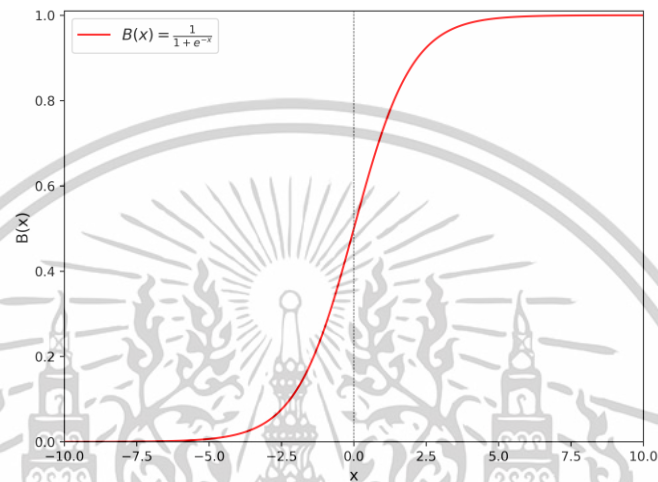


รูปที่ 2.4 ฟังก์ชัน ReLU

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ฟังก์ชัน Sigmoid เป็นฟังก์ชันกระตุ้นที่เปลี่ยนค่าของอินพุตให้กลายเป็นช่วงระหว่าง 0 ถึง 1 เหมาะสำหรับการใช้งานที่ต้องการแสดงความน่าจะเป็น (Probability) ดังแสดงในรูปที่ 2.5 โดยมีนิยามดังนี้

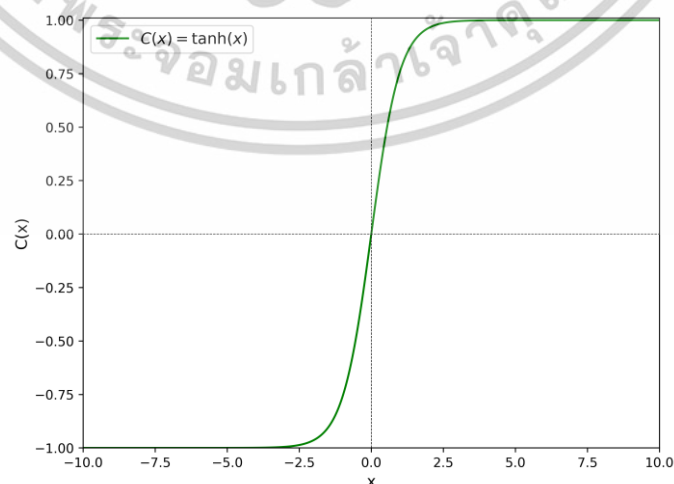
$$B(x) = \frac{1}{1 + e^{-x}}$$



รูปที่ 2.5 ฟังก์ชัน Sigmoid

3. ฟังก์ชัน Tanh เป็นฟังก์ชันกระตุ้นที่เปลี่ยนค่าของอินพุตให้อยู่ในช่วงระหว่าง -1 ถึง 1 ดังแสดงในรูปที่ 2.6 โดยมีนิยามดังนี้

$$C(x) = \tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



รูปที่ 2.6 ฟังก์ชัน Tanh

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.4 การเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning)

การเรียนรู้แบบเสริมกำลังเชิงลึก (DRL) [34] เป็นการนำการเรียนรู้เชิงลึก (Deep Learning) ร่วมกับการเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) ซึ่งใช้โครงข่ายประสาทเทียม (Neural Network) เพื่อจัดการกับโครงสร้างและชุดข้อมูลที่ซับซ้อน โดยทั่วไปแล้ว DRL จะใช้ในงานที่ต้องเรียนรู้เพื่อตัดสินใจในสภาพแวดล้อมที่มีความแปรปรวนและความไม่แน่นอน ซึ่งมีการนำไปประยุกต์ใช้ที่หลากหลายในสาขาต่าง ๆ เช่น การเล่นเกม [35] หุ่นยนต์ [36] การซื้อขายหุ้น [37] การแพทย์ [38] การจัดการพลังงาน [39] และการบริหารจัดการจราจร [40]

สำหรับ DRL [41] มีการใช้ Q-Function หรือที่เรียกว่า Deep Q-network (DQN) เป็นวิธีการที่ใช้กันอย่างแพร่หลาย โดยมีโครงข่ายประสาทเทียมที่สามารถประมาณค่าคิว (Q-Value) ผ่านการเล่นข้อมูลซ้ำและเรียนรู้จากประสบการณ์ (Experience Replay) ซึ่งเป็นการวัดประสิทธิภาพสำหรับการกระทำ (Action) ที่กำหนด โดยพิจารณาจากสถานะ (State) และการกระทำ (Action) จากนั้นจึงนำไปใช้เพื่อกำหนดแนวทางปฏิบัติที่ดีที่สุด สถานการณ์ปัจจุบัน

สมการหลักของ DQN คือ Q-Function และ Loss Function

$$Q(s, a) = r(s, a) + \gamma \max_{a'} Q(s', a') \quad (2.4)$$

สมการ 2.4 เรียกว่า สมการเบลล์แมน (Bellman Equation) [42] สมการนี้ใช้เพื่ออัปเดตค่าคิว (Q-Value) ของการกระทำปัจจุบันในสถานะ s , รวมถึงรางวัล (Reward) และค่าคิว (Q-Value) ที่ถูกลดโดยทันที [29]

โดยที่ $Q(s, a)$ คือ ค่าคิว (Q-Value) ซึ่งประเมินมูลค่าคาดหวังของรางวัลสะสมในสถานะ s และเลือกการกระทำ a

γ คือ อัตราการลดค่าของรางวัลในอนาคต

$r(s, a)$ คือ รางวัลที่ได้รับจากการกระทำ a ในสถานะ s

s' คือ สถานะถัดไป หลังจากการกระทำ a

a' คือ การกระทำในสถานะถัดไป s'

$\max_{a'} Q(s', a')$ คือ ค่าคิว (Q-Value) สูงสุด ในสถานะ s' เมื่อเลือกการกระทำที่ดีที่สุด a'

$$MSE = \frac{1}{n} \sum_{i=1}^n (Q(s, a) - q(s, a))^2 \quad (2.5)$$

สำหรับ DQN มี Loss Function ที่นำไปใช้คือ Mean Squared Error (MSE) [42] ตามสมการ 2.5 คำนวณโดยการเปรียบเทียบค่าคิวที่คาดการณ์ไว้ (Predicted Q-Value) กับค่าคิวจริง (Actual Q-Value)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ m คือ จำนวนข้อมูลทั้งหมด

$q(s, a)$ คือ ค่าคิว (Q-Value) ที่คาดการณ์จากโครงข่ายประสาทเทียม

$Q(s, a)$ คือ ค่าคิว (Q-Value) ที่เป็นเป้าหมาย ตามสมการ 2.4

2.5 นโยบาย (Policy)

นโยบาย (Policy: π) หมายถึง แนวทางที่ตัวแทนใช้ในการเลือกการกระทำในแต่ละสถานะ เพื่อให้ได้ผลตอบแทนที่มากที่สุด โดยประเภทของนโยบายจะแบ่งออกเป็น 2 ประเภทหลัก คือ

1. Deterministic Policy เป็นการกำหนดให้ตัวแทนเลือกการกระทำ (a) แบบตายตัวเสมอเมื่ออยู่ในสถานะ (s) โดยมีนิยามดังนี้

$$\pi(s) = a$$

2. Stochastic Policy เป็นการกำหนดให้ตัวแทนเลือกการกระทำ (a) ในสถานะ (s) โดยขึ้นอยู่กับความน่าจะเป็น โดยมีนิยามดังนี้

$$\pi(a|s) = P(A_t = a | S_t = s)$$

สำหรับนโยบายที่ใช้ใน Q-Learning และ DQN ส่วนใหญ่แล้วจะทำให้มีความสามารถทั้งในการใช้ความรู้ที่มีอยู่ในปัจจุบัน (Exploitation) เป็นการเลือกการกระทำที่ให้ค่าคิว (Q-value) สูงสุด และการสำรวจ (Exploration) เป็นการสุ่มสำรวจการกระทำอื่นเพื่อค้นหาวีธีที่อาจจะดีกว่า เรียกว่า Epsilon-Greedy Policy

2.5.1 การผสม Epsilon-Greedy Policy และ Deterministic Greedy Policy

การใช้ Epsilon-Greedy Policy [43] เพื่อให้เกิดการสำรวจในช่วงเริ่มต้น ช่วยให้โมเดลพิจารณาสถานะ (State) ที่ไม่รู้จัก ซึ่งค่าคิว (Q-Value) ที่ได้จาก Deep Q-Network จะเป็นตัวกำหนดเพื่อช่วยในการเลือกการกระทำ (Action) ที่เหมาะสมที่สุด อย่างไรก็ตาม ค่าคิว (Q-Value) ที่ได้อาจไม่เหมาะสมที่สุด ในบางครั้งการกระทำ (Action) บางอย่างอาจให้รางวัลที่มากกว่าได้ ดังนั้นจึงต้องการใช้ความน่าจะเป็นเล็กน้อยในการใช้การกระทำ (Action) ที่ไม่เคยเกิดขึ้นมาก่อน โดยใช้ค่าเอปซิลอน (Epsilon: ϵ) เป็นตัวกำหนดระหว่างการสำรวจและการใช้ความรู้ ซึ่งมักจะเริ่มต้นด้วยค่าที่สูงเพื่อนำการสำรวจ และลดค่าลงเรื่อย ๆ ระหว่างการฝึกเพื่อให้เลือกการกระทำตามค่าคิว (Q-Value) เป็นหลัก เมื่อตัวแทนมีความรู้ในการเลือกการกระทำที่มากเพียงพอแล้ว นโยบายที่ใช้มักอยู่ในรูปแบบ Deterministic Greedy Policy ซึ่งเป็นเลือกการกระทำที่ให้ค่าคิว (Q-value) สูงสุดเท่านั้น โดยมีนิยามดังนี้

$$\pi(s) = \begin{cases} \textit{Exploration}, & \epsilon \\ \textit{argmax}_a Q(s, a), & 1 - \epsilon \end{cases}$$

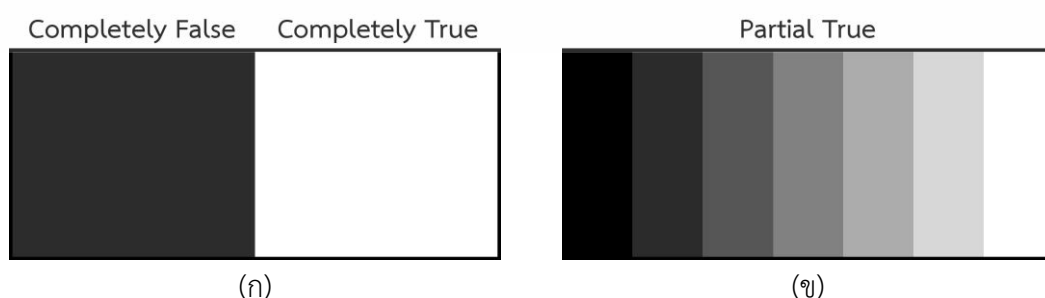
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6 ตรรกะแบบฟัซซี (Fuzzy Logic)

ตรรกะแบบฟัซซี (Fuzzy Logic) [44,45] ได้รับการพัฒนาขึ้นในปี พ.ศ. 2508 โดย L. A. Zadeh ซึ่งเป็นอีกวิธีหนึ่งในการคิดนอกเหนือจากตรรกะแบบดั้งเดิม (Classical Logic) โดยที่ปกติแล้วจะพิจารณาเพียงแค่ถูกหรือผิด ใช่หรือไม่ใช่ จริงหรือไม่จริง คือ การคิดแบบดิจิทัลหรือตรรกะบูลีน (Boolean logic) เป็นค่าไบนารีที่มีการใช้เพียงค่า 0 และ 1 แต่ตรรกะแบบฟัซซี (Fuzzy Logic) เป็นวิธีการวิเคราะห์สถานการณ์ที่อาจเกิดความสับสน คลุมเครือหรือไม่ชัดเจน ซึ่งจะมีค่าได้อยู่ระหว่างตั้งแต่ 0 ถึง 1 เช่น หากพิจารณาถึงเจตสึว่าสีชมพูอ่อนหรือสีชมพูเข้มว่าเป็นสีแดงหรือไม่ การใช้วิธีคิดตรรกะแบบฟัซซี (Fuzzy Logic) จะพิจารณาจากเจตสนั้นว่ามีความเป็นสีแดงมากแค่ไหน โดยที่สีแดงสดอาจมีค่าเป็น 1 (หมายถึงเป็นสีแดงเต็มที่) ในขณะที่สีชมพูอ่อนอาจมีค่าประมาณ 0.4 หรือ 0.5 (เป็นสีแดงในระดับต่ำ) และสีชมพูเข้มอาจมีค่าประมาณ 0.7 หรือ 0.8 (เป็นสีแดงในระดับที่มากกว่า) ในทางตรงกันข้าม หากพิจารณาถึงเจตสึว่าสีชมพูอ่อนว่าเป็นสีขาวหรือไม่ สีชมพูอ่อนอาจมีค่าประมาณ 0.6 ถึง 0.7 เนื่องจากมีความสว่างมาก ส่วนสีชมพูเข้มอาจมีค่าประมาณ 0.2 ถึง 0.4 เนื่องจากมีความเข้มกว่า อย่างไรก็ตาม ในการพิจารณาของแต่ละบุคคลอาจมีมุมมองที่แตกต่างกัน ซึ่งตรรกะแบบฟัซซี (Fuzzy Logic) ถูกนำมาใช้ในหลายสาขาเพื่อช่วยในการตัดสินใจในสถานการณ์ที่สับสนหรือไม่ชัดเจนและเป็นประโยชน์ในสถานการณ์ที่มีความไม่แน่นอนอย่างมาก

2.6.1 ฟัซซีเซต (Fuzzy Set)

ฟัซซีเซต (Fuzzy Set) เป็นกลุ่มหรือเซตของข้อมูลที่ใช้ในตรรกะแบบฟัซซี (Fuzzy Logic) เพื่อแสดงข้อมูลที่ไม่สามารถระบุได้อย่างชัดเจนว่าอยู่ในกลุ่มหรือชุดใดชุดหนึ่ง ซึ่งสื่อถึงความสับสนคลุมเครือและไม่แน่นอน โดยข้อมูลในเซตนั้นมีระดับของการเป็นสมาชิก (Degree of Membership) ที่ไม่จำเป็นต้องเป็น 1 หรือ 0 เสมอไป หมายความว่าสมาชิกของฟัซซีเซตสามารถอยู่ในระดับบางส่วน (Partial Membership) หรือบางส่วนของความเป็นจริง (Partial True) ได้ นั่นคือระดับของการเป็นสมาชิก (Degree of Membership) ของข้อมูลสามารถมีค่าอยู่ในช่วงตั้งแต่ 0 ถึง 1 ซึ่งแตกต่างจากเซตแบบดั้งเดิม (Crisp Set) ที่กำหนดให้ระดับของการเป็นสมาชิก (Degree of Membership) ในเซตมีเพียงสองสถานะคือ ใช่ (1) หรือ ไม่ใช่ (0) เท่านั้น ดังแสดงในรูปที่ 2.7



รูปที่ 2.7 (ก) ตรรกะแบบดั้งเดิม (ข) ตรรกะแบบฟัซซี

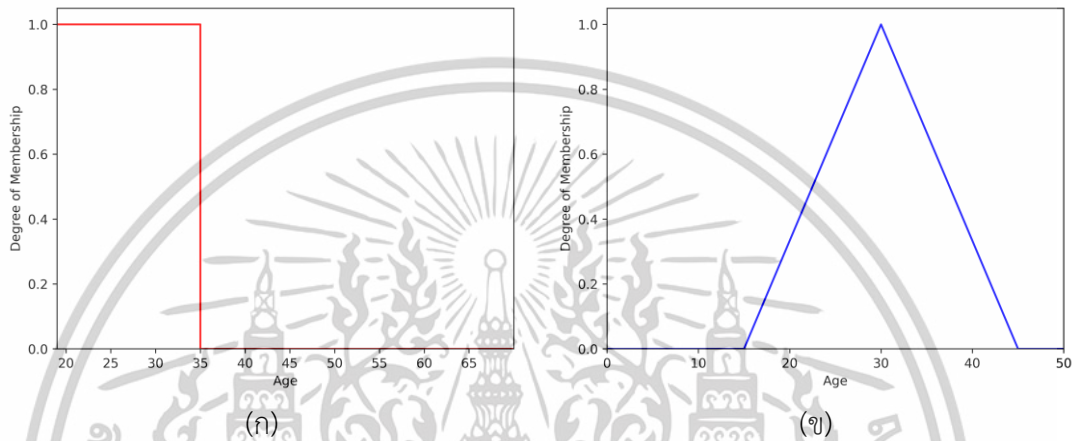
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมาชิกแต่ละตัวในฟัซซีเซต (Fuzzy Set) จะมีค่าที่เรียกว่าระดับของการเป็นสมาชิก (Degree of Membership) ซึ่งมีค่าตั้งแต่ 0 ถึง 1 โดยมีความหมายดังนี้

ค่า 0 หมายถึง ไม่เป็นสมาชิกของเซตนั้นเลย

ค่า 1 หมายถึง เป็นสมาชิกของเซตนั้นอย่างแน่นอน

ค่าระหว่าง 0 ถึง 1 หมายถึง เป็นสมาชิกของเซตในระดับบางส่วน (Partial Membership) ซึ่งบอกได้ว่าสมาชิคนั้นมีความใกล้เคียงหรือเป็นสมาชิกของเซตนั้นมากน้อยเพียงใด



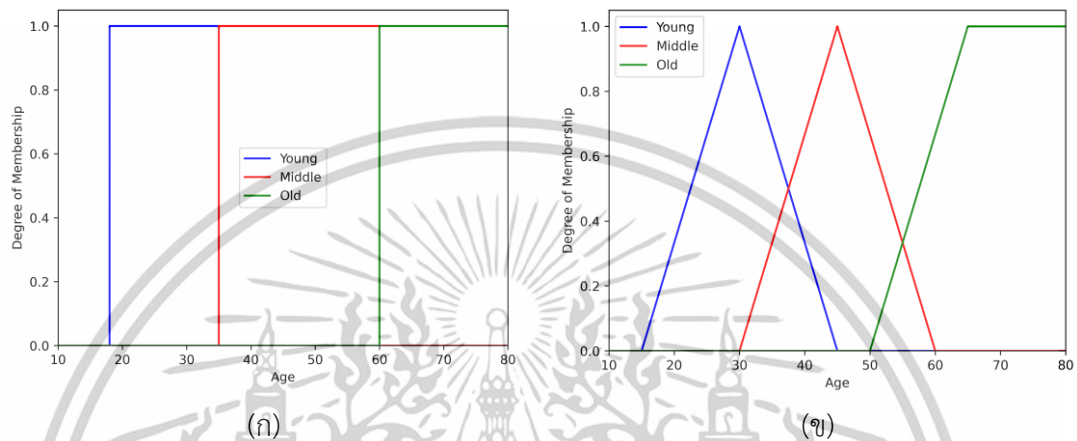
รูปที่ 2.8 (ก) เซตวัยหนุ่มสาวแบบดั้งเดิม (ข) เซตวัยหนุ่มสาวแบบฟัซซี

ตารางที่ 2.1 ระดับของการเป็นสมาชิกของเซตวัยหนุ่มสาว

อายุ (ปี)	ระดับความเป็นสมาชิก	
	เซตแบบดั้งเดิม	ฟัซซีเซต
20	1	0.33
25	1	0.66
28	1	0.87
30	1	1.00
32	1	0.87
39	0	0.40

ตัวอย่าง การเปรียบเทียบระหว่างเซตแบบดั้งเดิม (Crisp Set) กับ ฟัซซีเซต (Fuzzy Set) เกี่ยวกับการแบ่งช่วงอายุของคน โดยในเซตแบบดั้งเดิม (Crisp Set) เช่น วัยหนุ่มสาวจะกำหนดเกณฑ์อายุอย่างชัดเจนคือคนที่มีอายุอยู่ในช่วงอายุ 18 ปี ถึง 35 ปี จะมีระดับของการเป็นสมาชิก (Degree of Membership) ที่มีค่าเป็น 1 เช่น ในตารางที่ 2.1 คนที่มีอายุ 28 ปี จะถือว่าเป็นวัยหนุ่มสาว โดยมีเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ระดับของการเป็นสมาชิกวัยหนุ่มสาวเท่ากับ 1 ส่วนคนที่มีอายุ 39 ปี จะไม่ถือว่าเป็นวัยหนุ่มสาว โดยมีระดับของการเป็นสมาชิกวัยหนุ่มสาวเท่ากับ 0 แต่ในฟัซซีเซต (Fuzzy Set) ระดับของการเป็นสมาชิกจะค่อย ๆ มีค่าเพิ่มขึ้นและลดลงตามอายุ ดังแสดงในรูปที่ 2.8 (ข) คนที่มีอายุ 28 ปี จะอยู่ในวัยหนุ่มสาวด้วยระดับการเป็นสมาชิกที่มีค่าประมาณ 0.33 ส่วนคนที่มีอายุ 39 ปี จะอยู่ในวัยหนุ่มสาวด้วยระดับการเป็นสมาชิกที่มีค่าประมาณ 0.40 ซึ่งแสดงถึงความเป็นจริงบางส่วน (Partial True)



รูปที่ 2.9 (ก) เซตช่วงอายุแบบดั้งเดิมที่มีหลายเซต (ข) เซตช่วงอายุแบบฟัซซีที่มีหลายเซต

ในกรณีที่เซตแบบดั้งเดิม (Crisp Set) และฟัซซีเซต (Fuzzy Set) มีหลายเซต โดยจะแสดงตัวอย่างการแบ่งช่วงอายุออกเป็น 3 ช่วงอายุของเซตแบบดั้งเดิม (Crisp Set) ดังแสดงในรูปที่ 2.9 (ก) คือ 1.วัยหนุ่มสาวคือคนที่มีอายุอยู่ในช่วงอายุ 18 ปี ถึง 35 ปี 2.วัยกลางคนคือคนที่มีอายุอยู่ในช่วงอายุ 36 ปี ถึง 60 ปี และ 3.วัยสูงอายุคือคนที่มีอายุตั้งแต่ 61 ปีขึ้นไป โดยจะมีระดับของการเป็นสมาชิก (Degree of Membership) ที่มีค่าเป็น 1 ก็ต่อเมื่อมีช่วงอายุอยู่ในเกณฑ์ที่กำหนดเท่านั้น แต่ในส่วนของฟัซซีเซต (Fuzzy Set) ดังแสดงในรูปที่ 2.9 (ข) จะกำหนดช่วงอายุดังนี้ 1.วัยหนุ่มสาวคือคนที่มีอายุอยู่ในช่วงอายุ 15 ปี ถึง 45 ปี 2.วัยกลางคนคือคนที่มีอายุอยู่ในช่วงอายุ 30 ปี ถึง 60 ปี และ 3.วัยสูงอายุคือคนที่มีอายุตั้งแต่ 50 ปีขึ้นไป โดยจะมีระดับของการเป็นสมาชิก (Degree of Membership) ของเซตที่มีค่าอยู่ระหว่าง 0 ถึง 1 เมื่อพิจารณาจากรูปที่ 2.9 (ข) จะพบว่าในส่วนของฟัซซีเซต (Fuzzy Set) จะมีการซ้อนทับกันอยู่ในระหว่าง 3 ช่วงอายุนี้ แสดงถึงความเป็นไปได้ที่สมาชิกหรือข้อมูลหนึ่งสามารถอยู่ในหลายฟัซซีเซต (Fuzzy Set) พร้อมกันได้ในระดับที่ต่างกัน ซึ่งเป็นความคลุมเครือหรือความไม่แน่นอนของข้อมูล โดยสะท้อนให้เห็นถึงธรรมชาติของข้อมูลที่ไม่ได้มีการแบ่งแยกอย่างชัดเจนและมีความยืดหยุ่นในการจัดกลุ่ม

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2.6.2 การดำเนินการทางพีชคณิต

การดำเนินการทางพีชคณิตโดยทั่วไป จะมีการดำเนินการด้วยกัน 3 แบบ คือ Union, Intersection และ Complement

2.6.2.1 ยูเนียน (Union)

ยูเนียนของพีชคณิตจะเป็นการดำเนินการแบบ AND (AND Operation) ในตารางที่ 2.2

ตารางที่ 2.2 การดำเนินการแบบ AND

AND		
A	B	Min (A,B)
0	0	0
0	1	0
1	0	0
1	1	1

2.6.2.2 อินเตอร์เซกชัน (Intersection)

อินเตอร์เซกชันของพีชคณิตจะเป็นการดำเนินการแบบ OR (OR Operation) ในตารางที่ 2.3

ตารางที่ 2.3 การดำเนินการแบบ OR

OR		
A	B	Max (A,B)
0	0	0
0	1	1
1	0	1
1	1	1

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

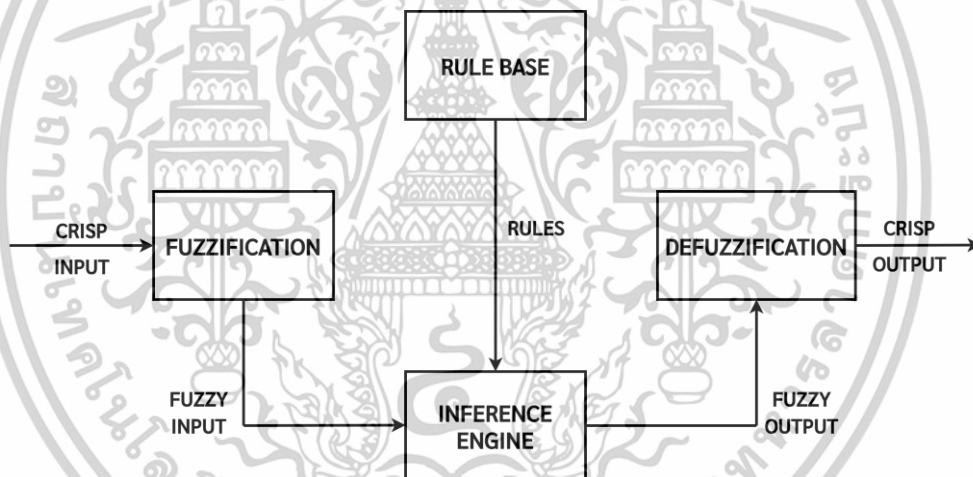
2.6.2.3 คอมพลิเมนต์ (Complement)

คอมพลิเมนต์ ของฟัซซีเซตจะเป็นการดำเนินการแบบ NOT (NOT Operation) ในตารางที่ 2.4

ตารางที่ 2.4 การดำเนินการแบบ NOT

OR	
A	1-A
0	1
1	0

2.6.3 โครงสร้างการทำงานพื้นฐานของตรรกะแบบฟัซซี



รูปที่ 2.10 โครงสร้างการทำงานพื้นฐานของตรรกะแบบฟัซซี

โครงสร้างการทำงานพื้นฐานของตรรกะแบบฟัซซีดังแสดงในรูปที่ 2.10 มีขั้นตอนการทำงานดังต่อไปนี้

1. การทำฟัซซีฟิเคชัน (Fuzzification) เป็นการคำนวณหาค่าระดับของการเป็นสมาชิก (Degree of Membership) จากฟังก์ชันสมาชิก (Membership Function) โดยจะทำการแปลงข้อมูลอินพุตที่เป็นค่าตัวเลขชัดเจน (Crisp Input) หรือเซตแบบดั้งเดิม (Crisp Set) ให้เป็นอินพุตที่เป็นฟัซซี (Fuzzy Input) หรือฟัซซีเซต (Fuzzy Set)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ฐานกฎ (Rule Base) ส่วนนี้เก็บกฎทั้งหมดของระบบ ซึ่งจะอยู่ในรูปแบบของคำสั่ง IF-THEN โดยแสดงถึงความสัมพันธ์ระหว่างอินพุตและเอาต์พุตในสถานการณ์ต่าง ๆ ที่สร้างขึ้นโดยผู้เชี่ยวชาญที่ควบคุมกระบวนการตัดสินใจโดยใช้กฎทางภาษา (Linguistic Rule)

3. กลไกการอนุมาน (Inference Engine) เป็นส่วนที่ทำหน้าที่ในการประมวลผลข้อมูล โดยใช้กฎที่อยู่ในฐานกฎ (Rule Base) เพื่อคำนวณผลลัพธ์ที่เป็นฟัซซี (Fuzzy Output) จากข้อมูลอินพุตที่ผ่านกระบวนการทำฟัซซีฟิเคชัน (Fuzzification) มาแล้ว

4. การทำดีฟัซซีฟิเคชัน (Defuzzification) เป็นการแปลงค่าผลลัพธ์ที่เป็นฟัซซี (Fuzzy Output) ให้กลับมาเป็นค่าตัวเลขชัดเจน (Crisp Output) ที่สามารถนำไปใช้งานได้จริง

2.6.4 การอนุมานฟัซซีแบบแมมดานี

การอนุมานฟัซซีแบบแมมดานี (Mamdani Fuzzy Inference) เป็นวิธีการให้เหตุผลแบบคลุมเครือในระบบตรรกะแบบฟัซซี (Fuzzy Logic) ที่นิยมใช้ในการควบคุมระบบ โดยใช้กฎ (Rule) ที่เป็นกฎทางภาษา (Linguistic Rule) และมีโครงสร้างแบบ IF-THEN เพื่อสร้างความสัมพันธ์ระหว่างอินพุตและเอาต์พุตในสถานการณ์ที่ข้อมูลไม่ชัดเจนหรือกำกวม โดยการอนุมานฟัซซีแบบแมมดานีประกอบไปด้วย 4 ขั้นตอนคือการทำฟัซซีฟิเคชัน การประเมินกฎ การรวมกฎและการทำดีฟัซซีฟิเคชัน โดยมีรายละเอียดดังต่อไปนี้

1. การทำฟัซซีฟิเคชัน (Fuzzification) โดยแปลงค่าอินพุตที่เป็นตัวเลขชัดเจน (Crisp Input) เป็นค่าอินพุตฟัซซี (Fuzzy Input) โดยใช้ฟังก์ชันสมาชิก (Membership Function)

2. การประเมินกฎ (Rule Evaluation) ขั้นตอนนี้ทำการประมวลผลของอินพุตฟัซซีจากกฎที่มีอยู่ในฐานกฎ (Rule Base) โดยใช้การอินเตอร์เซกชัน (AND) เพื่อหาค่าผลลัพธ์ของแต่ละกฎจากอินพุตฟัซซีที่เข้าเงื่อนไข ซึ่งอาจเข้าเงื่อนไขมากกว่าหนึ่งกฎพร้อม ๆ กัน

3. การรวมกฎ (Aggregation) นำค่าผลลัพธ์ที่ได้จากกฎทั้งหมดมารวมกัน โดยใช้การยูเนียน (OR)

4. การทำดีฟัซซีฟิเคชัน (Defuzzification) โดยแปลงค่าผลลัพธ์ที่เป็นฟัซซี (Fuzzy Output) ที่ได้จากกระบวนการการประเมินกฎ (Rule Evaluation) กลับไปเป็นค่าที่ชัดเจน (Crisp Output) เพื่อหาค่าผลลัพธ์ที่แน่นอนในการนำไปใช้งานจริง วิธีที่นิยมใช้มากที่สุดคือการใช้วิธีเซนทรอยด์ (Centroid Method) หรือ วิธีการหาจุดศูนย์กลางถ่วง (Center of Gravity หรือ COG) ซึ่งเป็นการหาค่ากลางของพื้นที่ใต้กราฟของฟังก์ชันสมาชิก (Membership Function) ของเอาต์พุตที่เป็นฟัซซี (Fuzzy Output) โดยมีสูตรในการคำนวณดังสมการที่ 2.6

$$\text{Crisp Output } (\bar{x}) = \frac{\sum_{i=1}^n A_i x_i}{A_{\text{Total}}} \quad (2.6)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ A_i คือ พื้นที่ของรูปที่ i

X_i คือ ตำแหน่งจุด Centroid ของรูปที่ i

A_{Total} คือ พื้นที่รวมของรูปทั้งหมด

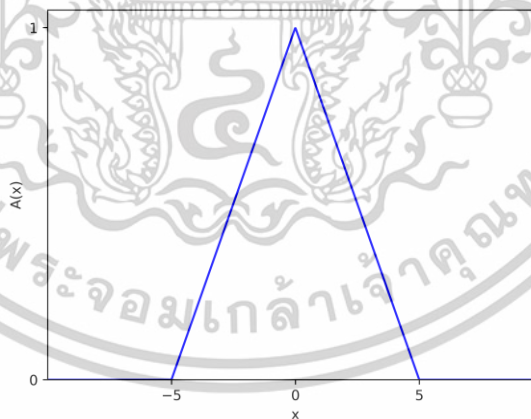
2.6.5 ฟังก์ชันสมาชิก (Membership Function)

ฟuzzy เซต (Fuzzy Set) แต่ละเซตจะมีฟังก์ชันสมาชิก (Membership Function) เป็นฟังก์ชันที่ใช้คำนวณค่าระดับของการเป็นสมาชิก (Degree of Membership) ของข้อมูลหรือสมาชิกในฟuzzy เซต (Fuzzy Set) ว่าค่าของข้อมูลหรือสมาชิกนั้นอยู่ในเซตในระดับใด ซึ่งค่านี้จะบ่งบอกระดับความเป็นจริงหรือระดับความเกี่ยวข้องของข้อมูลกับเซตนั้น โดยค่าระดับของการเป็นสมาชิกจะอยู่ในช่วงตั้งแต่ 0 ถึง 1

ตัวอย่างของฟังก์ชันสมาชิกที่นิยมใช้งานโดยทั่วไป เช่น

1. ฟังก์ชันความเป็นสมาชิกสามเหลี่ยม (Triangular Membership Function) มีลักษณะเป็นกราฟรูปสามเหลี่ยม ใช้ระบุค่าระดับของการเป็นสมาชิกตามช่วงที่ชัดเจน ดังแสดงในรูปที่ 2.11 โดยมีนิยามดังนี้

$$A(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ \frac{c-x}{c-b}, & b \leq x \leq c \\ 0, & x > c \end{cases}$$

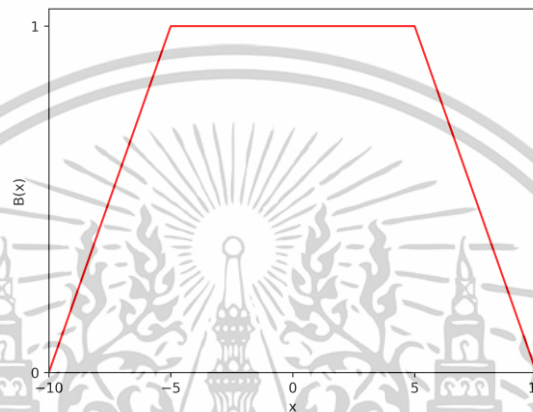


รูปที่ 2.11 ฟังก์ชันความเป็นสมาชิกสามเหลี่ยม เมื่อ $a = -5$, $b = 0$, $c = 5$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

2. ฟังก์ชันความเป็นสมาชิกสี่เหลี่ยมคางหมู (Trapezoidal Membership Function) มีลักษณะเป็นกราฟสี่เหลี่ยมคางหมู ใช้สำหรับช่วงค่าที่กว้างขึ้น ดังแสดงในรูปที่ 2.12 โดยมีนิยามดังนี้

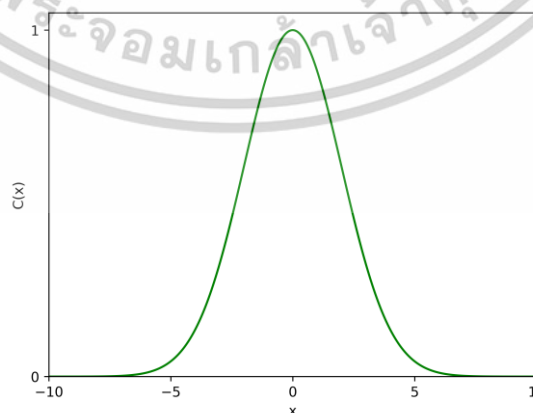
$$B(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x < b \\ 1, & b \leq x < c \\ \frac{d-x}{d-c}, & c \leq x < d \\ 0, & x \geq d \end{cases}$$



รูปที่ 2.12 ฟังก์ชันความเป็นสมาชิกสี่เหลี่ยมคางหมู เมื่อ $a = -10$, $b = -5$, $c = 5$, $d = 10$

3. ฟังก์ชันความเป็นสมาชิกเกาส์เซียน (Gaussian Membership Function) มีลักษณะเป็นกราฟโค้งเรียบแบบปกติ ใช้เมื่อค่าระดับของการเป็นสมาชิกค่อย ๆ เพิ่มขึ้นและลดลงแบบไม่ทันทีทันใด ดังแสดงในรูปที่ 2.13 โดยมีนิยามดังนี้

$$C(x) = \exp\left(-\frac{(x-c)^2}{2\sigma^2}\right)$$



รูปที่ 2.13 ฟังก์ชันความเป็นสมาชิกเกาส์เซียน เมื่อ $c = 0$, $\sigma = 2$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 3

การออกแบบระบบ

ในการบริหารจัดการระบบการจราจรในเมือง หนึ่งในความท้าทายหลักคือการเพิ่มประสิทธิภาพการควบคุมสัญญาณไฟจราจรที่ทางแยกเพื่อลดการจราจรติดขัดและลดเวลารอคอย อย่างไรก็ตาม ระบบควบคุมสัญญาณไฟจราจรแบบเดิมมักอาศัยแผนการกำหนดเวลาของสัญญาณไฟไว้ล่วงหน้า ซึ่งอาจไม่มีประสิทธิภาพภายใต้สภาพการจราจรที่เปลี่ยนแปลงตลอดเวลา ทำให้การเดินทางเกิดความล่าช้าและการจราจรติดขัดเพิ่มมากขึ้น

เพื่อแก้ไขข้อจำกัดนี้ เราจึงนำเสนอการควบคุมสัญญาณไฟจราจรโดยใช้ Deep Q-Network ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) การตอบสนองของระบบจะแตกต่างกันไปตามระดับความแออัดของการจราจรบริเวณทางแยก ช่วยให้สัญญาณไฟจราจรเปลี่ยนเฟสได้โดยอัตโนมัติผ่านการเรียนรู้ด้วย Deep Q-Network เพื่ออำนวยความสะดวกในการสัญจรที่ทางแยก นอกจากนี้ระยะเวลาของสัญญาณไฟเขียวยังถูกปรับให้เหมาะสมโดยใช้ตรรกะแบบฟัซซี (Fuzzy Logic) โดยเราใช้โมเดลต้นแบบที่พัฒนาขึ้นจากการศึกษาที่มีอยู่ โดย Tunc และ Soylemez [29] และนำมาเป็นเป้าหมายสำหรับการเปรียบเทียบผลลัพธ์ เนื้อหาของการออกแบบระบบแบ่งออกเป็นส่วนย่อยดังต่อไปนี้

3.1 แนวทางการพัฒนาเพิ่มเติมจากโมเดลต้นแบบ

จากการศึกษางานวิจัยต้นแบบ โดย Tunc และ Soylemez [29] พบว่า แม้โมเดลที่นำเสนอจะมีประสิทธิภาพที่ดีในการควบคุมสัญญาณไฟจราจรเพื่อลดเวลารอคอยสะสมของยานพาหนะ อย่างไรก็ตาม ยังมีองค์ประกอบบางส่วนที่สามารถพัฒนาเพิ่มเติมได้ ดังนี้

1. รูปแบบการแทนสถานะของระบบ (State Representation) ในโมเดลต้นแบบใช้เพียงข้อมูลแบบไบนารีว่ามียานพาหนะอยู่ในตำแหน่งนั้นหรือไม่ ซึ่งไม่สามารถสะท้อนถึงระยะเวลาการรอคอยสะสมของรถแต่ละช่องทางได้อย่างชัดเจน
2. ฟังก์ชันรางวัล (Reward Function) ที่ใช้ในการเรียนรู้แบบเสริมกำลัง มุ่งเน้นที่ผลรวมเวลารอสะสมทั้งหมด ส่งผลให้บางเฟสที่เป็นเฟสของช่องทางที่มีปริมาณยานพาหนะเป็นจำนวนมาก ถูกเลือกบ่อยเกินไป ขณะที่บางช่องทางที่มีปริมาณยานพาหนะน้อยกว่ากลับต้องรอเป็นระยะเวลานาน
3. ไม่มีการออกแบบระบบสำหรับรองรับสถานการณ์พิเศษ เช่น การให้ความสำคัญแก่ยานพาหนะฉุกเฉิน
4. ประเภทของยานพาหนะในการจำลอง กำหนดให้รถทุกคันเป็นประเภทเดียวกัน ขนาดเท่ากัน ความเร็ว และพฤติกรรมเหมือนกันทั้งหมด ซึ่งไม่สะท้อนสภาพการจราจรจริงที่มีรถหลากหลายประเภท เช่น รถยนต์ส่วนบุคคล รถโดยสาร และรถบรรทุกขนาดใหญ่ ที่มีพฤติกรรมการเคลื่อนที่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และการใช้พื้นที่บนถนนแตกต่างกัน ที่ส่งผลต่อรูปแบบการแทนสถานะของระบบ โดยเฉพาะในบริเวณเซลล์ที่อยู่ใกล้กับเส้นหยุดบริเวณทางแยก ซึ่งเป็นจุดสำคัญ ดังนั้น เพื่อพัฒนาโมเดลการควบคุมสัญญาณไฟจราจรให้มีประสิทธิภาพยิ่งขึ้น วิทยานิพนธ์นี้จึงเสนอแนวทางการพัฒนาเพิ่มเติมจากโมเดลต้นแบบใน 4 ด้าน ได้แก่

1. การปรับปรุงรูปแบบการแทนสถานะของระบบ ให้สามารถสะท้อนระยะเวลารอสะสมของยานพาหนะแต่ละช่องทางได้ในหลายระดับ
2. การออกแบบฟังก์ชันรางวัลใหม่ที่คำนึงถึงความสมดุลในแต่ละทิศทาง
3. การออกแบบระบบเพื่อเพิ่มความสามารถในการรองรับยานพาหนะฉุกเฉิน เช่น รถพยาบาล
4. การเพิ่มประเภทของยานพาหนะในการจำลอง

โดยรายละเอียดของการปรับปรุงในข้อ 1 และ 2 สามารถสรุปได้ดังตารางที่ 3.1

ตารางที่ 3.1 การปรับปรุงรูปแบบการแทนสถานะและฟังก์ชันรางวัลจากโมเดลต้นแบบ

องค์ประกอบ	วิธีต้นแบบ	วิธีที่นำเสนอ	จุดที่ปรับปรุง
สถานะ	ใช้เพียงค่าบูลีน (0 หรือ 1) เพื่อระบุว่ามียานพาหนะหรือไม่ในแต่ละเซลล์	ใช้ระดับเวลารอ (10 ระดับ) ร่วมกับตำแหน่งของรถ	แสดงข้อมูลที่ละเอียดขึ้น และสะท้อนสภาพการจราจรได้แม่นยำกว่า
ฟังก์ชันรางวัล	ใช้ผลรวมเวลารอสะสมของรถทุกคัน	ใช้ค่ารอสะสมที่มากที่สุดของแต่ละช่องจราจรในกลุ่มที่ใช้สัญญาณเดียวกัน	ลดความเอนเอียงต่อรถทางตรงและเพิ่มความสำคัญให้กับรถเลี้ยวมากขึ้น

3.2 ซอฟต์แวร์ที่ใช้ในการทดลอง

การจำลองสภาพแวดล้อมถูกสร้างขึ้นโดยใช้ SUMO (Simulation of Urban Mobility) เวอร์ชัน 1.18.0 และ Python 3.9.17 [46] โดย SUMO ใช้ Microscopic Models ในการจำลองจราจรเป็นหลัก ซึ่งรองรับการจำลองทั้งระดับเล็กและใหญ่ สามารถใช้ได้ทั้งในกรณีการทดลองขนาดเล็ก (เช่น การจำลองแยกจราจรเดียว) และการจำลองเครือข่ายถนนในระดับเมือง โมเดลนี้มุ่งเน้นไปที่พฤติกรรมและปฏิสัมพันธ์ของยานพาหนะแต่ละคันภายในระบบจราจร การขับขี่ของรถคันหน้า เช่น การเร่งความเร็ว ลดความเร็ว และการหยุดรถ จะส่งผลต่อพฤติกรรมการขับขี่ของยานพาหนะดังกล่าว การเคลื่อนตัวของยานพาหนะแต่ละคันในกระแสการจราจร (Traffic Flow) จะขึ้นอยู่กับปัจจัยต่อไปนี้

Car-Following Models : โมเดลจะกำหนดระยะทางและความเร็วของยานพาหนะตามสภาพการจราจรที่เปลี่ยนแปลง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

Lane-Changing Models : โมเดลจะกำหนดพฤติกรรมกรรมการเปลี่ยนช่องจราจร (Lane) ของยานพาหนะ ตามกฎเกณฑ์ต่าง ๆ เช่น ช่องว่างและการพิจารณาด้านความปลอดภัย

นอกจากนี้ โปรแกรม SUMO ยังรองรับการควบคุมสัญญาณไฟจราจรแบบไดนามิก โดยสามารถทำงานร่วมกับ Python ผ่าน TraCI (Traffic Control Interface) ซึ่งเป็น API สำหรับควบคุมการจำลองแบบเรียลไทม์ ทำให้สามารถปรับค่าหรือเก็บข้อมูลระหว่างการทำการทดลองได้อย่างยืดหยุ่น เช่น การเปลี่ยนสถานะสัญญาณไฟจราจร การตรวจสอบตำแหน่งของรถ ความเร็วของรถ เวลาที่ใช้ในการเดินทาง ความหนาแน่นของการจราจร ปริมาณรถที่ผ่านแต่ละแยก เป็นต้น อีกทั้งยังมีเครื่องมือประกอบที่หลากหลาย ที่ช่วยให้การสร้างและปรับแต่งสภาพแวดล้อมจำลองทำได้ง่ายขึ้น เช่น การสร้างโครงข่ายถนน การสร้างเส้นทางของรถ หรือการจัดการการจราจรในรูปแบบต่าง ๆ

3.3 สภาพแวดล้อม (Environment)

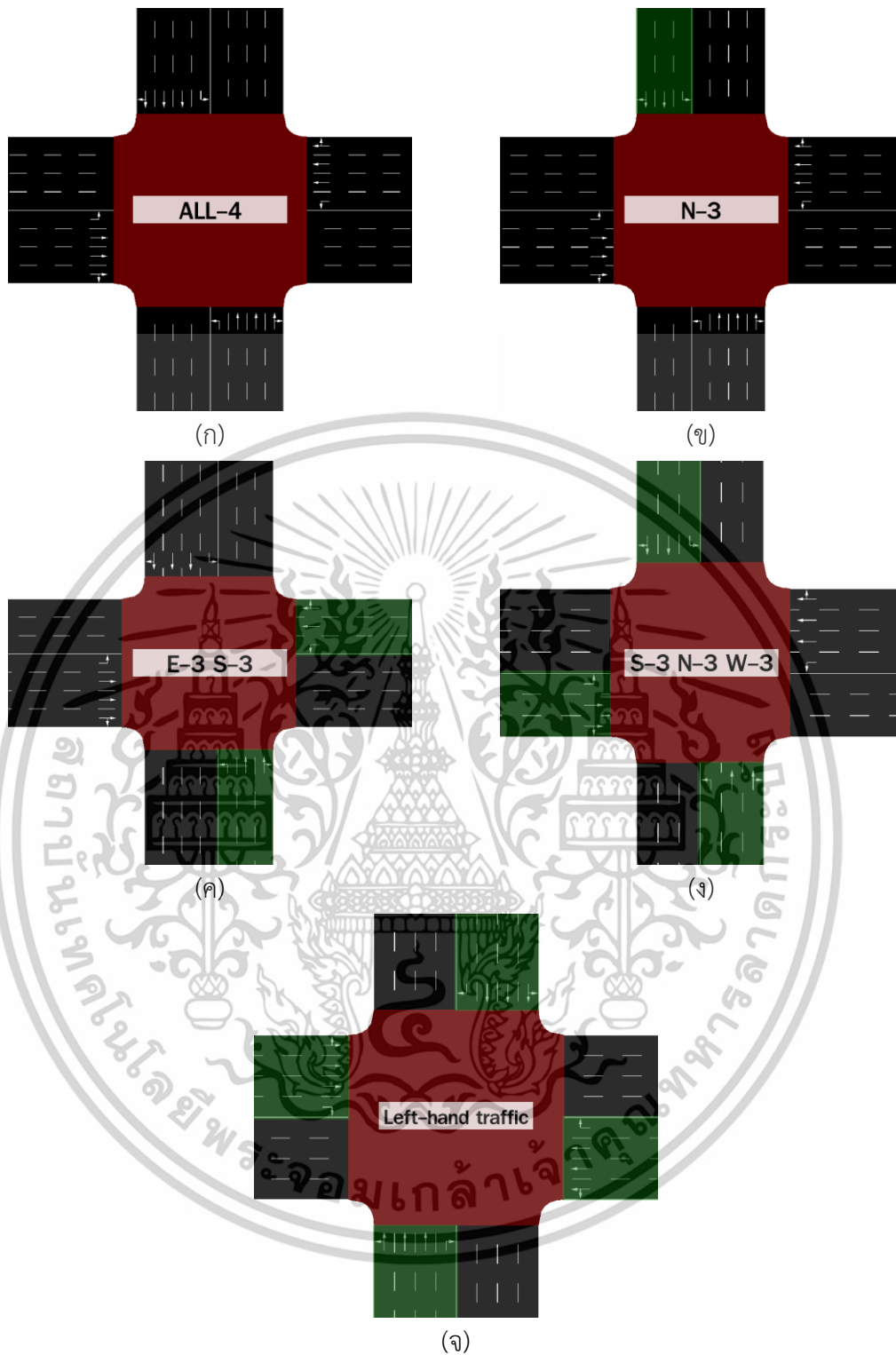
ในการศึกษาครั้งนี้ เราใช้ต้นแบบสภาพแวดล้อมของทางแยกโดยอิงตามวิธีการต้นแบบ [29] ซึ่งทางแยกประกอบด้วยสี่แยก โดยที่แต่ละแยกมี 4 ช่องจราจร (Lane) และมีความยาว 750 เมตร

อย่างไรก็ตาม เพื่อให้มั่นใจถึงทิศทางการจราจรที่ถูกต้องและมีประสิทธิภาพ ยานพาหนะจะต้องใช้ช่องจราจรที่เหมาะสมเมื่อเลี้ยวซ้ายหรือขวา ดังนั้น หากยานพาหนะต้องการเลี้ยวขวา จำเป็นต้องใช้ช่องจราจรทางขวาสุด และเมื่อต้องการที่จะเลี้ยวซ้าย ยานพาหนะจะต้องใช้ช่องจราจรซ้ายสุด ส่วนยานพาหนะที่ต้องการขับตรงไปสามารถใช้ช่องจราจรทางด้านขวาสุดได้ ดังแสดงในรูปที่ 3.1 (ก)

นอกจากนี้ เรายังมีการออกแบบสภาพแวดล้อมการจราจรอีกหลายแบบ โดยในช่วงแรกแต่ละแยกจะมีโครงสร้างที่สมดุลโดยมี 4 ช่องจราจร (Lane) เท่ากันทั้งหมด และจำนวนช่องจราจร (Lane) จะค่อย ๆ ลดลงเหลือ 3 ช่องจราจร (Lane) ทีละแยก โดยจะลดจำนวนช่องจราจร (Lane) ในแต่ละแยกจากทุกทิศทางให้ครอบคลุมแต่ละการออกแบบที่เป็นไปได้

การออกแบบสภาพแวดล้อมการจราจรประเภทสุดท้ายคือการขับตรงทางซ้ายโดยมี 4 ช่องจราจร (Lane) เท่ากันในทุกแยก ดังแสดงในรูปที่ 3.1 (ข)-(จ)

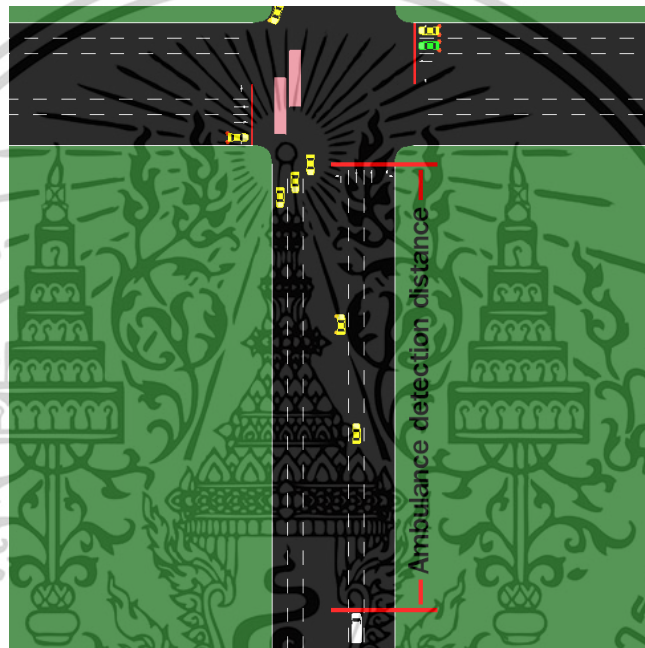
ในรูปที่ 3.1 (ข)-(จ) มีข้อแตกต่างเพียงอย่างเดียวคือโครงสร้างของช่องจราจร (Lane) ซึ่งจะลดช่องจราจร (Lane) ทางตรงลง และในรูปที่ 3.1 (จ) มีการเปลี่ยนแปลงของสภาพการจราจรที่จะขับชี้ทางชิดด้านซ้ายของถนน



รูปที่ 3.1 (ก) สภาพแวดล้อมทางแยกที่เป็นต้นแบบ (ข) สภาพแวดล้อมที่มี 3 ช่องจราจรใน 1 แยก เรียกว่า สภาพแวดล้อมแบบ N-3 (ค) สภาพแวดล้อมที่มี 3 ช่องจราจรใน 2 แยก เรียกว่าสภาพแวดล้อมแบบ E-3 S-3 (ง) สภาพแวดล้อมที่มี 3 ช่องจราจรใน 3 แยก เรียกว่าสภาพแวดล้อมแบบ S-3 N-3 W-3 (จ) สภาพแวดล้อมที่เป็นแบบขับชิดซ้าย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

นอกจากนี้ยังมีระบบจัดลำดับความสำคัญของรถพยาบาลที่จะให้ความสำคัญกับรถพยาบาล โดยให้ผ่านทางแยกก่อน อย่างไรก็ตาม ต้องคำนึงถึงความปลอดภัยด้วย ดังนั้น จึงจำเป็นต้องคำนึงถึงระยะเวลาความปลอดภัย ซึ่งหมายถึงระยะเวลาก่อนจะเปลี่ยนเป็นสัญญาณไฟเขียวเพื่อให้รถพยาบาล ผ่านไปได้ พร้อมทั้งระยะเวลาหลังจากรถพยาบาลผ่านทางแยกไปแล้วและการเลือกเฟสของสัญญาณไฟจึงจะกลับสู่สภาวะปกติ ซึ่งระยะเวลาความปลอดภัยยังส่งผลต่อระยะเวลาการตรวจจับอีกด้วย หากระยะเวลาความปลอดภัยเพิ่มขึ้น จำเป็นต้องจัดเตรียมเวลาให้เพียงพอเพื่อให้สัญญาณไฟเขียว เปิดใช้งานก่อนที่รถพยาบาลจะมาถึงทางแยก จึงมีความจำเป็นที่ต้องพิจารณาระยะการตรวจจับ ดังแสดงในรูปที่ 3.2



รูปที่ 3.2 ระยะตรวจจับรถพยาบาลในระบบจัดลำดับความสำคัญของรถพยาบาล

3.4 สถานะ (State)

สถานะ (State) เป็นหนึ่งในองค์ประกอบหลักของวิธีการที่นำเสนอ สถานะปัจจุบัน (Current State) ของสภาพแวดล้อม (Environment) คือ ความรู้ที่ใช้ในการปรับปรุงกระบวนการตัดสินใจของ DQN

เพื่อให้ระบบมีความรู้เพื่อใช้ในการควบคุมสัญญาณไฟจราจรให้มีประสิทธิภาพมากขึ้น จะต้องรวบรวมข้อมูลรถยนต์บริเวณทางแยกให้เพียงพอ ข้อมูลสถานะ (State) เป็นข้อมูลที่จำเป็นที่สามารถส่งให้แก่ผู้ควบคุม (Agent) ในการควบคุมสัญญาณไฟจราจรได้ทันที โดยพิจารณาตำแหน่งของยานพาหนะในสภาพแวดล้อม (Environment)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวแปรสถานะ (State) จะเก็บข้อมูลเกี่ยวกับเวลารอคอย โดยแบ่งออกเป็น 10 ระดับ ข้อมูลเหล่านี้ใช้เพื่อระบุตำแหน่งของรถและเวลารอคอย ตัวแปร S_t แสดงถึงสถานะ (State) ของทางแยก ณ เวลาที่กำหนดในขั้นตอนที่ t

มีการแบ่งช่องจราจร (Lane) บริเวณทางแยก โดยแต่ละช่องจราจร (Lane) จะถูกแบ่งออกเป็นส่วนที่เรียกว่าเซลล์ ขนาดของเซลล์ถูกออกแบบให้ไม่เท่ากัน ดังแสดงในตารางที่ 3.2 โดยเซลล์ที่ใกล้ทางแยกที่สุดจะแคบ และกว้างขึ้นเมื่อเคลื่อนห่างออกจากทางแยก [20] เนื่องจากมีประสิทธิภาพที่ดีกว่า เมื่อเทียบกับการใช้เซลล์ที่มีขนาดเท่ากันในทุกระยะทาง [29]

ตารางที่ 3.2 ขนาดของเซลล์ที่ใช้ในการแบ่งช่องจราจร

เซลล์ที่	ความกว้างของเซลล์ (เมตร)	ระยะห่างจากทางแยก (เมตร)
1	7	7
2	7	14
3	7	21
4	7	28
5	12	40
6	20	60
7	40	100
8	60	160
9	240	400
10	350	750

การกำหนดค่าของเซลล์ ถ้าภายในเซลล์มียานพาหนะและไม่มีเวลารอคอย ค่าจะเป็น 1 อย่างไรก็ตาม ถ้าเวลารอคอยไม่เป็นศูนย์ ค่าที่กำหนดจะถูกแบ่งออกเป็น 10 ระดับตามเวลารอคอย ดังแสดงในตารางที่ 3.3 และเซลล์มีค่าเป็น 0 ในกรณีที่ภายในเซลล์ไม่มียานพาหนะ

การใช้การออกแบบสถานะ (State) ของวิธีการที่เรานำเสนอ ทำให้สามารถสังเกตทั้งการมีอยู่ของยานพาหนะและเวลารอคอยได้ สิ่งนี้แตกต่างจากวิธีการที่ใช้ Deep Q-Network ร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) ที่เป็นต้นแบบ [29] ซึ่งเซลล์มีค่าเพียง 0 และ 1 แสดงถึงการมีหรือไม่มียานพาหนะอยู่ภายในเซลล์

ในแต่ละกลุ่มช่องจราจร (Lane) จะมีทั้งหมด 10 เซลล์ โดยที่ 3 ช่องจราจร (Lane) ทางด้านขวาจะถูกรวมเข้าด้วยกันให้เป็นกลุ่มเดียวกันและใช้สัญญาณไฟจราจรร่วมกัน ดังนั้นค่าที่บันทึกไว้ในเซลล์จะถูกเลือกจากค่าสูงสุดในแต่ละเซลล์ของทั้ง 3 ช่องจราจร (Lane)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.3 ระดับการแบ่งระยะเวลารอคอย

ระดับ	เวลารอคอย (วินาที)
1	0
2	1-6
3	7-12
4	13-18
5	19-24
6	25-30
7	31-36
8	37-42
9	43-48
10	มากกว่า 48

นอกจากนี้ วิธีการที่เรานำเสนอ ความยาวของรถแต่ละประเภทจะแตกต่างกันไป และเนื่องจากขนาดของเซลล์ที่อยู่ใกล้ทางแยกจะแคบ ดังนั้น หมายความว่ายานพาหนะที่มีความยาวมากกว่าขนาดของเซลล์สามารถอยู่ในเซลล์ถัดไปได้ ซึ่งค่าสูงสุดในแต่ละเซลล์จะถูกบันทึกลงในที่เก็บข้อมูลของสถานะ (State) ดังแสดงในรูปที่ 3.3 สิ่งนี้แตกต่างจากวิธีการที่ใช้ Deep Q-Network ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) ที่เป็นต้นแบบ [29] ซึ่งมียานพาหนะเพียงประเภทเดียวและความยาวของยานพาหนะน้อยกว่าขนาดของเซลล์

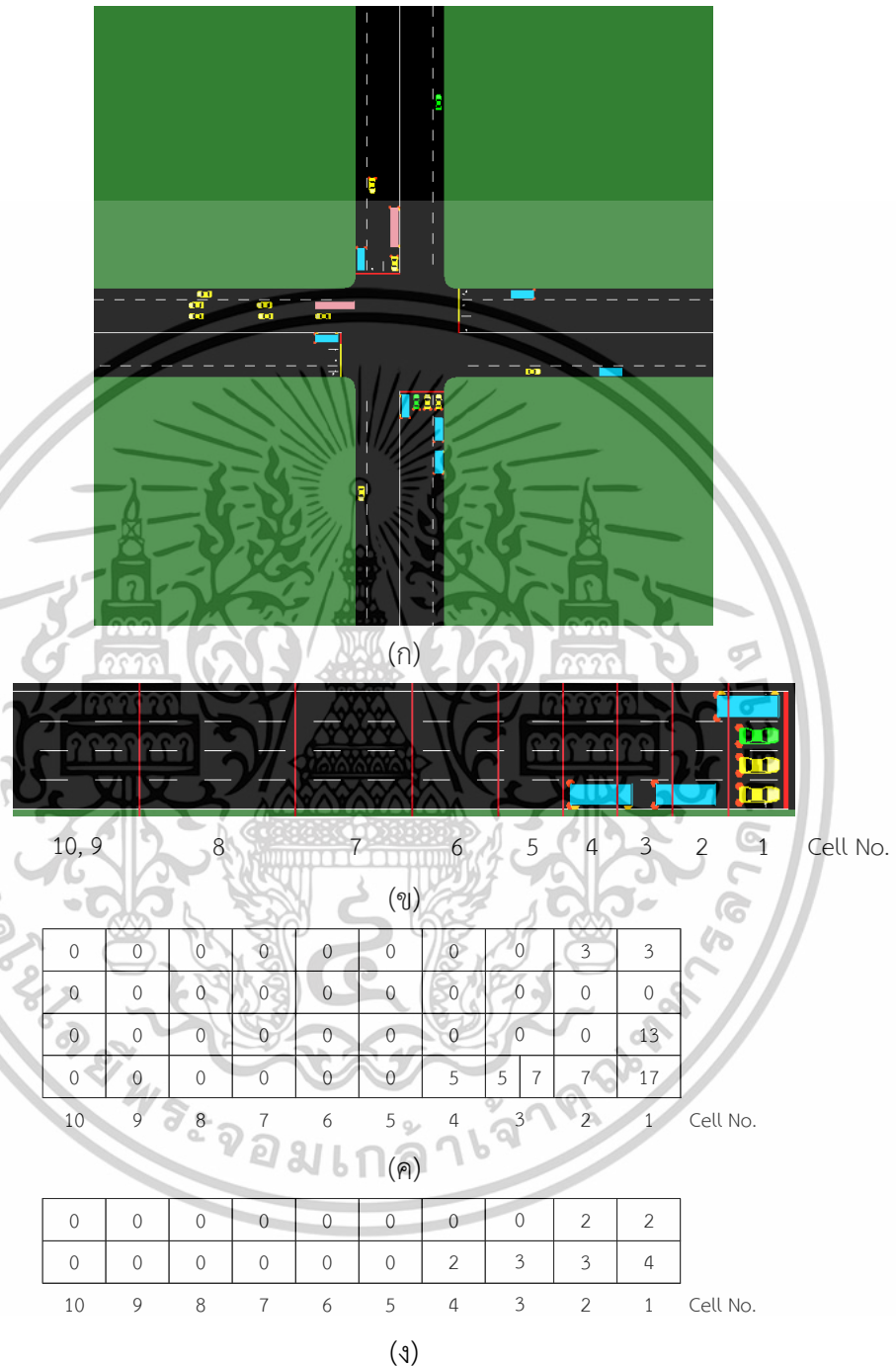
ในแต่ละสภาพแวดล้อมที่ใช้ในการทดลอง แม้รูปแบบของทางแยกหรือจำนวนช่องจราจรในแต่ละทิศจะแตกต่างกันไป แต่รูปแบบของข้อมูลสถานะที่ป้อนเข้าสู่ระบบยังคงใช้โครงสร้างเดียวกันทั้งหมด เพื่อให้สามารถใช้โมเดลเดียวกันได้ในทุกสภาพแวดล้อม ไม่ว่าจะ เป็นแบบที่ทุกทิศทางมีจำนวนช่องจราจรเท่ากัน หรือบางทิศทางมีเพียง 3 ช่องจราจร

การจัดกลุ่มช่องจราจรสำหรับเก็บค่าข้อมูลจะอิงตามกลุ่มช่องจราจรที่ใช้สัญญาณไฟเดียวกัน โดยจะเลือกใช้ค่าสูงสุดจากช่องจราจรในกลุ่มเดียวกันในแต่ละเซลล์ ตัวอย่างเช่น กลุ่มช่องจราจรของรถทางตรงที่มี 2 หรือ 3 ช่องจราจร จะถูกรวมและแทนด้วยค่าเพียงหนึ่งค่าที่มากที่สุดในแต่ละเซลล์เดียวกัน

ดังนั้น โครงสร้างของข้อมูลสถานะจึงยังคงมีจำนวนคงที่เท่ากับ 80 ค่าในทุกกรณี เพื่อให้สามารถทำงานได้โดยใช้โมเดลเดียวกันในทุกสภาพแวดล้อม

โดยรูปที่ 3.3 แสดงตัวอย่างการเก็บสถานะ (State) ทางด้านทิศใต้ของทางแยก ซึ่งรูปที่ 3.3 (ค) แสดงเวลารอคอยของรถแต่ละคันที่ถูกแบ่งออกเป็นเซลล์ตามรูปที่ 3.3 (ข) หากภายในเซลล์มียานพาหนะ

มากกว่าหนึ่งคัน ค่าสูงสุดภายในเซลล์จะถูกจัดเก็บเป็นสถานะ (State) ดังแสดงไว้ในเซลล์ที่ 3 ของรูปที่ 3.3 (ค) และ (ง)



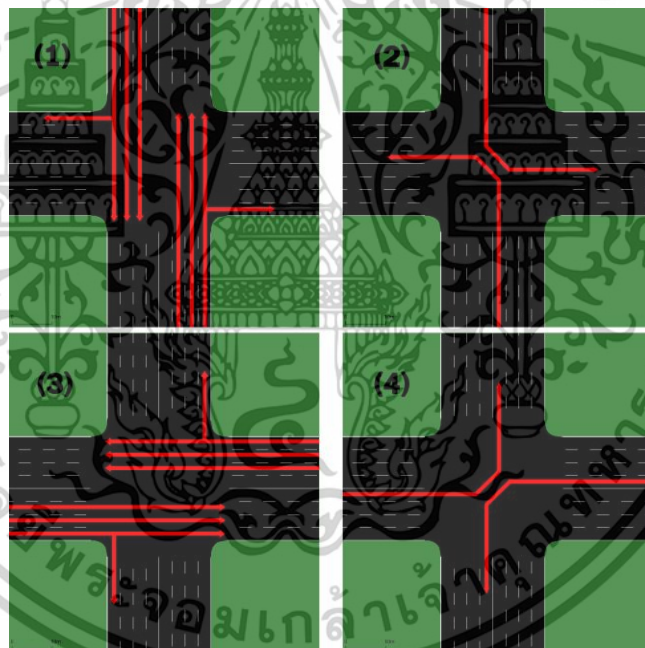
รูปที่ 3.3 (ก) ภาพรวมการจราจรบริเวณทางแยก (ข) ภาพรวมการจราจรทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์ (ค) เวลาารอคอยของรถแต่ละคันทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์ (ง) ค่าสถานะทางด้านทิศใต้ของทางแยกที่จะถูกเก็บไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.5 การกระทำ (Action)

การกระทำ (Action) เกี่ยวข้องกับการเลือกเฟสสัญญาณไฟจราจรของตัวแทน (Agent) ที่ให้รางวัลสูงสุดตามข้อมูลสถานะปัจจุบัน (Current State) โดย Deep Q-Network เลือกการกระทำ (Action) จากเอาต์พุตโดยตรงหรือเลือกแบบสุ่มโดยใช้ Epsilon-Greedy ซึ่งช่วยเพิ่มการสำรวจสถานะ (State) ที่ไม่รู้จัก

ในส่วนของทิศทางสัญญาณไฟจราจร ดังแสดงในรูปที่ 3.4 สัญญาณไฟจราจรจะสลับไปที่เฟสไฟเขียวตามระยะเวลาที่ตั้งไว้ซึ่งกำหนดโดยการคำนวณจากตรรกะแบบฟัซซี (Fuzzy Logic) โดยที่เฟสถัดไปจะถูกเลือกหลังจากระบบได้กำหนดระยะเวลาไฟเขียวแล้ว หากเฟสเดิมถูกเลือกจะดำเนินต่อไปโดยไม่มีสัญญาณไฟเหลือง อย่างไรก็ตาม หากตัวแทน (Agent) เลือกเฟสอื่น สัญญาณไฟจะเปลี่ยนเป็นไฟเหลืองเป็นเวลา 3 วินาที ก่อนที่จะเปลี่ยนเป็นไฟเขียวสำหรับเฟสที่ถูกเลือกอยู่ในปัจจุบัน โดยที่ทุกสภาพแวดล้อมจะใช้เฟสสัญญาณไฟจราจรในลักษณะเดียวกัน ดังแสดงในรูปที่ 3.4



รูปที่ 3.4 ทิศทางของเฟสสัญญาณไฟจราจร

ระบบจะเลือกเฟสที่เหมาะสมของสัญญาณไฟเขียวจาก 1 ใน 4 เฟสที่กำหนดไว้แล้ว ดังนี้

1. ไฟเขียวสำหรับทิศเหนือ-ใต้
2. ไฟเขียวสำหรับการเลี้ยวซ้ายทิศเหนือ-ใต้
3. ไฟเขียวสำหรับทิศตะวันออก-ตะวันตก
4. ไฟเขียวสำหรับการเลี้ยวซ้ายทิศตะวันออก-ตะวันตก

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.6 รางวัล (Reward)

รางวัล (Reward) คือการวัดผลที่สามารถใช้เพื่อประเมินผลลัพธ์จากการกระทำ (Action) ของตัวแทน (Agent) ได้ โดยการคำนวณเริ่มต้นเมื่อตัวแทนได้เลือกเฟสของสัญญาณไฟเขียวเสร็จสิ้นในแต่ละครั้ง ซึ่งเป็นเวลาที่ตัวแทน (Agent) เปลี่ยนจากสถานะ S_t เป็นสถานะ S_{t+1}

การเลือกพารามิเตอร์ที่ใช้ในการคำนวณผลตอบแทน (Reward) เป็นสิ่งสำคัญอย่างยิ่ง เพราะมีความจำเป็นที่ต้องใช้เป็นข้อมูลในการคำนวณค่าคิว (Q-Value) เพื่อให้ตัวแทน (Agent) พิจารณาในเลือกการกระทำ (Action) และนำไปใช้ในการเรียนรู้เพิ่มเติม

เราใช้เวลารอคอยในการคำนวณรางวัล เนื่องจากวัตถุประสงค์หลักของการศึกษาค้นคว้าครั้งนี้ คือการลดเวลารอคอยที่ทางแยก เพื่อลดการจราจรที่ติดขัด และปรับปรุงประสิทธิภาพการควบคุมสัญญาณไฟจราจร อย่างไรก็ตาม ปัญหาที่พบในวิธีการที่ใช้ Deep Q-Network ร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) ที่เป็นต้นแบบ [29] ซึ่งเกี่ยวข้องกับสภาพแวดล้อมต้นแบบที่การจราจรซับซ้อน โดยเฉพาะฟังก์ชันการให้รางวัลที่จะคำนวณเวลารอคอยรวมของรถยนต์ทุกคัน

ดังนั้นปัญหาที่เกิดจากฟังก์ชันการให้รางวัลนี้ จึงส่งผลให้รถที่จะเลี้ยวซ้ายต้องรอสัญญาณไฟจราจรเป็นเวลานาน เนื่องจากกระบวนการให้รางวัล ทำให้ตัวแทน (Agent) เลือกเฟสสัญญาณไฟที่ให้มีความสำคัญกับยานพาหนะที่จะตรงไปมากกว่า เพื่อให้ได้ผลตอบแทนที่ดีกว่า

3.6.1 ตัวอย่างปัญหาจากการใช้เวลารอสะสมทั้งหมดในวิธีต้นแบบ

ในระบบควบคุมสัญญาณไฟจราจรตามวิธีต้นแบบ การคำนวณรางวัลจะอ้างอิงจากผลรวมของเวลารอสะสมของรถทุกคันในระบบ ซึ่งอาจก่อให้เกิดความคลาดเคลื่อนในเชิงตรรกะ โดยเฉพาะในกรณีที่มีรถหลายคันในกลุ่มของช่องจราจรที่ใช้สัญญาณไฟจราจรร่วมกัน

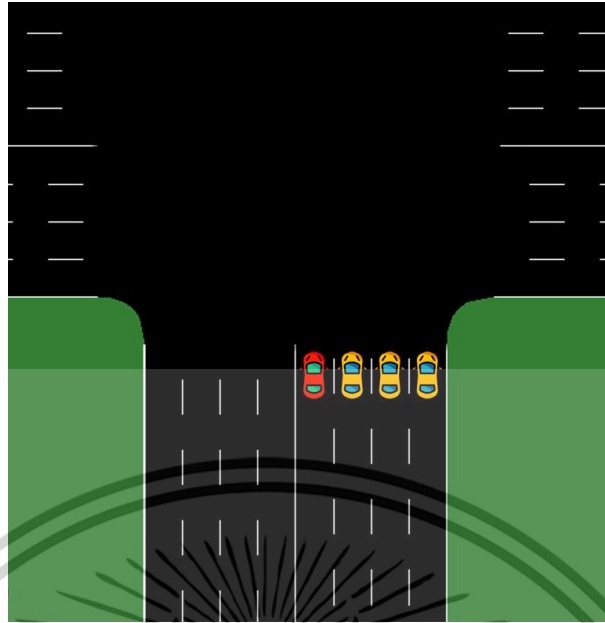
ตัวอย่าง ณ ช่วงเวลาหนึ่ง ดังแสดงในรูปที่ 3.5 มีรถทั้งหมด 4 คันอยู่ในทางแยกเดียวกัน ได้แก่ รถที่รอเลี้ยวซ้ายจำนวน 1 คัน (รถสีแดง) ซึ่งกำหนดให้มีเวลารอสะสม 10 วินาที และรถที่อยู่ในกลุ่มช่องจราจรสำหรับตรงไป 3 คัน (รถสีเหลือง) โดยแต่ละคันมีเวลารอสะสม 10 วินาที เช่นกัน

ระบบจะประเมินว่าเวลารอสะสมรวมของทั้งระบบเท่ากับ 40 วินาที ซึ่งได้จากผลรวมเวลารอสะสมของรถทั้ง 4 คัน

อย่างไรก็ตาม ในทางปฏิบัติ รถทั้ง 3 คันที่อยู่ในกลุ่มช่องจราจรสำหรับตรงไปซึ่งใช้สัญญาณไฟจราจรร่วมกัน ควรถูกพิจารณาเป็นกลุ่มเดียว ที่มีเวลารอสะสมเพียง 10 วินาที เท่านั้น เนื่องจากเมื่อได้รับสัญญาณไฟเขียว รถเหล่านี้จะเคลื่อนที่ไปพร้อมกัน

การคำนวณในลักษณะดังกล่าวจึงทำให้ระบบเข้าใจคลาดเคลื่อนว่า การปล่อยสัญญาณไฟเขียวให้กับรถในกลุ่มช่องจราจรสำหรับตรงไป จะสามารถลดเวลารอสะสมของระบบลงได้มากถึง 30 วินาที จึงถือว่าไม่สมเหตุผล

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.5 ตัวอย่างสถานการณ์ที่ส่งผลต่อความคลาดเคลื่อนของเวลารอสะสมในวิธีต้นแบบ

3.6.2 การคำนวณรางวัลจากเวลารอสะสมของระบบในวิธีที่นำเสนอ

ในรูปที่ 3.6 แสดงรูปการแบ่งกลุ่มของช่องจราจร โดยที่กลุ่ม 1, 3, 5 และ 7 เป็นกลุ่มช่องจราจร (Lane) สำหรับเลี้ยวซ้าย และกลุ่ม 2, 4, 6 และ 8 เป็นกลุ่มช่องจราจร (Lane) สำหรับตรงไปและเลี้ยวขวา การคำนวณเวลารอคอยสะสมในกลุ่มช่องจราจร (Lane) สำหรับตรงไปและเลี้ยวขวาของรถทุกคันจะสูงกว่ากลุ่มช่องจราจร (Lane) สำหรับเลี้ยวซ้าย เนื่องจากมี 3 ช่องจราจร (Lane) ดังนั้น เราจึงนำเวลารอสะสมในแต่ละช่องจราจร (Lane) มาใช้เพื่อแก้ไขปัญหานี้ โดยจะเลือกค่าสูงสุดจากช่องจราจร (Lane) ในแต่ละแยกที่ใช้เฟสสัญญาณไฟจราจรร่วมกัน ฟังก์ชันการให้รางวัลนี้แตกต่างจากวิธีการที่ใช้ Deep Q-Network ร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) ที่เป็นต้นแบบ [29] ซึ่งจะคำนวณเวลารอคอยสะสมของรถยนต์ทุกคัน

การคำนวณเวลารอคอยในแต่ละช่องจราจร (Lane) จะแสดงในสมการ 3.1 โดยเวลารอคอยสะสมจะเพิ่มขึ้นเมื่อรถเคลื่อนที่ด้วยความเร็วต่ำกว่า 0.1 เมตรต่อวินาที และฟังก์ชันการให้รางวัลแสดงอยู่ในสมการ 3.2

$$Twt_t = \sum_{i=1}^4 G_{2i-1,t} + \sum_{i=1}^4 G_{2i,t} \quad (3.1)$$

โดยที่ $G_{2i-1,t}$ คือ เวลารอคอยทั้งหมดในกลุ่มช่องจราจร (Lane) สำหรับเลี้ยวซ้ายในเวลาขั้นตอนที่ t ดังแสดงในรูปที่ 3.6

$G_{2i,t}$ คือ เวลารอคอยรวมสูงสุดในกลุ่มช่องจราจร (Lane) สำหรับตรงไปและเลี้ยว

ขวาในเวลาขั้นตอนที่ t ดังแสดงในรูปที่ 3.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$r_t = Twt_{t-1} - Twt_t \quad (3.2)$$

โดยที่ Twt_t คือ เวลารอคอยสะสมของยานพาหนะทั้งหมดในเวลาชั้นตอนที่ t ซึ่งเป็นช่วงเวลาปัจจุบัน

Twt_{t-1} คือ เวลารอคอยสะสมของยานพาหนะทั้งหมดในเวลาชั้นตอนที่ $t-1$ ซึ่งเป็นช่วงเวลาก่อนหน้า



รูปที่ 3.6 กลุ่มของช่องจราจร

3.7 ตัวแทน (Agent)

ตัวแทน (Agent) ใช้นโยบาย Epsilon-Greedy ในการสำรวจสภาพแวดล้อม และช่วยตัวแทน (Agent) ทำความรู้จักกับสภาพแวดล้อมในระหว่างที่เริ่มต้นการเรียนรู้ สำหรับนโยบายความน่าจะเป็นนี้ มีความน่าจะเป็น ϵ ที่จะเลือกการกระทำ (Action) แบบสุ่ม และมีความน่าจะเป็น $1-\epsilon$ ที่จะเลือกการกระทำ (Action) ที่เหมาะสมที่สุด โดยมีการระบุ Epsilon Value สำหรับแต่ละตอน ตามสมการ 3.3 [20]

$$\epsilon = 1 - \frac{n}{N} \quad (3.3)$$

โดยที่ n คือ ตอนปัจจุบัน

N คือ ตอนทั้งหมด

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

3.8 ข้อมูลที่จำเป็น (Necessary Information)

ตารางที่ 3.4 จะเป็นการสรุปหมวดหมู่ ข้อมูลที่จำเป็น เงื่อนไข และค่าใช้จ่ายที่เกี่ยวข้องในการใช้ระบบควบคุมสัญญาณไฟจราจรด้วยการเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning) และตรรกะแบบฟัซซี (Fuzzy Logic) การทำความเข้าใจข้อมูลเหล่านี้เป็นสิ่งสำคัญเนื่องจากช่วยให้เข้าใจข้อกำหนดพื้นฐานและข้อจำกัดในทางปฏิบัติที่เกี่ยวข้องกับการพัฒนาและการทำงานของระบบ ซึ่งการนำระบบไปใช้งานจริง จำเป็นต้องดำเนินการและเตรียมข้อมูล ต่อไปนี้

1. การเก็บข้อมูลจราจร เพื่อให้ระบบสามารถรับรู้สถานะของสี่แยกได้แบบเรียลไทม์ จำเป็นต้องติดตั้งอุปกรณ์ตรวจจับรถยนต์และสภาพการจราจร โดยใช้ เซนเซอร์ตรวจจับจำนวนและตำแหน่งรถ เช่น ใ้กล้อง CCTV พร้อมระบบประมวลผลภาพ (Computer Vision) หรือ ใช้ LiDAR ซึ่งข้อมูลที่ได้จะถูกนำมาใช้เป็นอินพุต ได้แก่ ตำแหน่งรถในแต่ละช่องทาง (แบ่งเป็นเซลล์) เวลาที่รถแต่ละคันรออยู่ ความเร็วรถ (สำหรับกรองรถที่ยังไม่หยุดสนิท) และประเภทของรถ โดยที่ข้อมูลเหล่านี้จะถูกใช้ในการสร้างสถานะของระบบ สำหรับนำเข้าสู่โมเดล Deep Q-Network และการประมวลผลด้วยตรรกะแบบฟัซซี

2. ระบบประมวลผลและเซิร์ฟเวอร์ จำเป็นต้องมีหน่วยประมวลผลที่สามารถรันโมเดลได้อย่างต่อเนื่อง เช่น คอมพิวเตอร์ตั้งโต๊ะ หรืออุปกรณ์ประมวลผลขนาดเล็ก (Edge Device) เช่น NVIDIA Jetson หรือ Raspberry Pi ซึ่งรองรับการประมวลผลด้าน AI และมีความเหมาะสมสำหรับการควบคุมแยกเดี่ยวโดยเฉพาะ

การใช้ Edge Device จะช่วยลดความซับซ้อนและประหยัดค่าใช้จ่าย เพราะสามารถติดตั้งและทำงานอยู่ที่จุดแยกโดยตรง โดยไม่ต้องส่งข้อมูลไปประมวลผลที่ศูนย์กลาง

3. ระบบการควบคุมสัญญาณไฟจราจร ในการใช้งานจริง จำเป็นต้องพัฒนาระบบควบคุมที่สามารถรับคำสั่งจากโมเดลได้โดยตรง อย่างไรก็ตาม หากระบบควบคุมสัญญาณไฟเดิมในพื้นที่รองรับการเชื่อมต่อจากภายนอก เช่น ผ่าน API สำหรับสื่อสารผ่านเครือข่าย หรือสามารถเชื่อมต่อโดยตรงผ่าน GPIO หรือ I/O Board ก็สามารถใช้งานร่วมกับระบบเดิมได้ โดยไม่จำเป็นต้องเปลี่ยนอุปกรณ์ทั้งหมด

เพื่อให้ระบบมีความปลอดภัย ควรมีระบบสลับกลับเป็นโหมดการทำงานแบบกำหนดเวลาคงที่ (Fixed-Time) ได้ หากระบบอัตโนมัติไม่ทำงาน หรือให้เจ้าหน้าที่ที่สามารถสั่งหยุดระบบชั่วคราวได้ในกรณีจำเป็น

ตารางที่ 3.4 ข้อมูลที่จำเป็นสำหรับระบบควบคุมสัญญาณไฟจราจรด้วยการเรียนรู้แบบเสริมกำลังเชิงลึก

หมวดหมู่	ข้อมูลที่จำเป็น	เงื่อนไข	ค่าใช้จ่าย
อุปกรณ์เก็บข้อมูล	จำนวนและตำแหน่งของยานพาหนะ	การติดตั้งเซนเซอร์ กล้อง หรืออุปกรณ์ตรวจจับ	ค่าใช้จ่ายในการติดตั้ง การบำรุงรักษา และเวลาที่ใช้ในการประมวลผลข้อมูล
ระบบประมวลผลและเซิร์ฟเวอร์	การจัดเก็บข้อมูล การประมวลผล และการฝึกโมเดล	การจัดการข้อมูล การประมวลผลข้อมูล การตั้งค่าเซิร์ฟเวอร์ และความปลอดภัย	ค่าใช้จ่ายสำหรับฮาร์ดแวร์ ซอฟต์แวร์ และเวลาที่ใช้สำหรับพัฒนา และจัดการระบบ
ระบบควบคุมสัญญาณไฟจราจร	สัญญาณไฟปัจจุบัน ระยะเวลา และการควบคุม	การนำไปใช้กับระบบ การควบคุมจราจรที่มีอยู่	ค่าใช้จ่ายในการติดตั้ง การปรับปรุง และอัปเดตระบบ

3.9 การออกแบบการจำลอง (Simulation Design)

ส่วนนี้จะเน้นที่การอธิบายโครงสร้างของ Deep Q-Network และการออกแบบสภาพแวดล้อมการจราจรจำลอง

3.9.1 การตั้งค่าโครงข่าย (Network Setting)

ส่วนของโครงสร้างโครงข่ายประสาทเทียม (Neural Network Structure) โครงข่ายของ Deep Q-Network ถูกสร้างขึ้นโดยใช้ Keras และ TensorFlow โดยที่เลเยอร์อินพุต (Input Layer) ของโครงข่ายประสาทเทียมรับสถานะ (State) ที่ได้จากสภาพแวดล้อม (Environment) ซึ่งอินพุตประกอบด้วย 80 โหนด โดยมี 2 ชั้นที่ซ่อนอยู่ (Hidden Layer) แต่ละชั้นประกอบด้วย 100 โหนด โดยมี Activation Function ที่ใช้ในแต่ละเลเยอร์คือ Rectified Linear Unit (ReLU) แต่ละค่าที่ได้รับจากแต่ละโหนดในเลเยอร์ก่อนหน้าจะถูกถ่วงน้ำหนัก (Weighted) ก่อนนำไปประมวลผลต่อไป โดยเลเยอร์แรกที่ซ่อนอยู่ (First Hidden Layer) จะได้รับค่าจากเลเยอร์อินพุต (Input Layer) ส่วนต่อมาผลลัพธ์ของเลเยอร์ที่ซ่อนอยู่แรก (First Hidden Layer) จะถูกส่งผ่านไปยังเลเยอร์ที่ซ่อนอยู่ที่สอง (Second Hidden Layer) และผลลัพธ์ของเลเยอร์ที่ซ่อนอยู่ที่สอง (Second Hidden Layer) จะถูกส่งไปยังเลเยอร์เอาต์พุต (Output Layer) ซึ่งคำนวณค่าคิว (Q-Value) สำหรับการกระทำ (Action) ทั้ง 4 เฟส โดยพารามิเตอร์การเรียนรู้ที่ใช้ในการทดลองแสดงอยู่ในตารางที่ 3.5

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.5 การตั้งค่าพารามิเตอร์ของโครงข่ายประสาทเทียม

พารามิเตอร์	ค่าที่ใช้
จำนวนเลเยอร์ซ่อนอยู่	2
จำนวนโนดในเลเยอร์ซ่อนอยู่	100
ขนาดแบตช์	64
อัตราการเรียนรู้	0.001
จำนวนรอบการฝึกฝน	500
ขนาดหน่วยความจำ	50,000
ปัจจัยส่วนลด	0.75
จำนวนสถานะ (อินพุต)	80
จำนวนการกระทำ (เอาต์พุต)	4

เราอธิบายกระบวนการเรียนรู้ของระบบเป็นรหัสเทียม (Pseudo code) ในอัลกอริทึม 1 โดยตอนต้นของแต่ละตอน ค่าเอปไซลอน (Epsilon Value) จะถูกคำนวณเพื่อพิจารณาว่าจะสำรวจ (Explore) หรือใช้ประโยชน์ (Exploit) หรือไม่ จากนั้นระบบจะควบคุมสัญญาณไฟจราจรตามค่าเอปไซลอน (Epsilon Value) และจัดเก็บสถานะ (State) การกระทำ (Action) รางวัล (Reward) และสถานะต่อไป (Next State) ในหน่วยความจำ (Memory) โดยใช้โครงสร้างข้อมูล Deque จาก Python's Collections Module เพื่อจัดการหน่วยความจำอย่างมีประสิทธิภาพ กระบวนการจะดำเนินต่อไปจนกระทั่งถึงขั้นตอนสูงสุด (Maximum Step) หลังจากนั้น สถิติของตอนจะถูกบันทึก และเริ่มกระบวนการฝึกอบรมที่เรียกว่าการเล่นซ้ำประสบการณ์ (Experience Replay) ขั้นตอนการฝึกอบรมที่สำคัญนี้เกี่ยวข้องกับการใช้ข้อมูลเกี่ยวกับสถานะ (State) การกระทำ (Action) รางวัล (Reward) และสถานะถัดไป (Next State) เพื่อปรับปรุงการเรียนรู้จากประสบการณ์ สุดท้ายจะวนซ้ำจนกว่าจะครบจำนวนตอนทั้งหมด (Total Episodes) ซึ่งเป็นจุดที่โมเดลได้รับการบันทึกและเตรียมพร้อมสำหรับการทดสอบ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

อัลกอริทึม 1 การฝึกอบรมโมเดล

อินพุต: จำนวนชั้นที่ซ่อนอยู่, จำนวนโหนดของชั้นที่ซ่อนอยู่, ขนาดของแบตช์, อัตราการเรียนรู้, จำนวนรอบการฝึก, ปัจจัยส่วนลด, จำนวนสถานะ, จำนวนการกระทำ, จำนวนขั้นตอนสูงสุดจำนวนรถที่ถูกสร้างทั้งหมด, จำนวนตอนทั้งหมด

เอาต์พุต: โมเดลการควบคุมสัญญาณไฟจราจร และ สถิติของแต่ละตอน

1: กำหนด: episode = 0

2: ในขณะที่ current episode < total episodes ให้ทำ

3: คำนวณค่า epsilon

4: กำหนด: step = 0

5: ในขณะที่ step < max steps ให้ทำ

6: รับสถานะปัจจุบัน

7: คำนวณรางวัล

8: เพิ่มข้อมูลตัวอย่างลงในหน่วยความจำ

9: ถ้า ตัวเลขสุ่มระหว่าง 0.0 to 1.0 < ค่า epsilon แล้ว

10: เลือกการกระทำแบบสุ่ม

11: มีฉะนั้น

12: เลือกการกระทำโดยใช้การคาดการณืจากโมเดล

13: จบเงื่อนไข

14: ถ้า เฟสของสัญญาณไฟเปลี่ยน แล้ว

15: กำหนดเฟสเป็นไฟเหลือง

16: กำหนดระยะเวลาไฟเหลือง

17: ในขณะที่ yellow duration > 0 ให้ทำ

18: step = step + 1

19: yellow duration = yellow duration - 1

20: จบลูป

21: จบเงื่อนไข

22: กำหนดเฟสไฟเขียวตามการกระทำ

23: รับระยะเวลาไฟเขียวตามการคำนวณจากตรรกะแบบฟัชซี

24: กำหนดระยะเวลาไฟเขียว

อัลกอริทึม 1 (ต่อ)

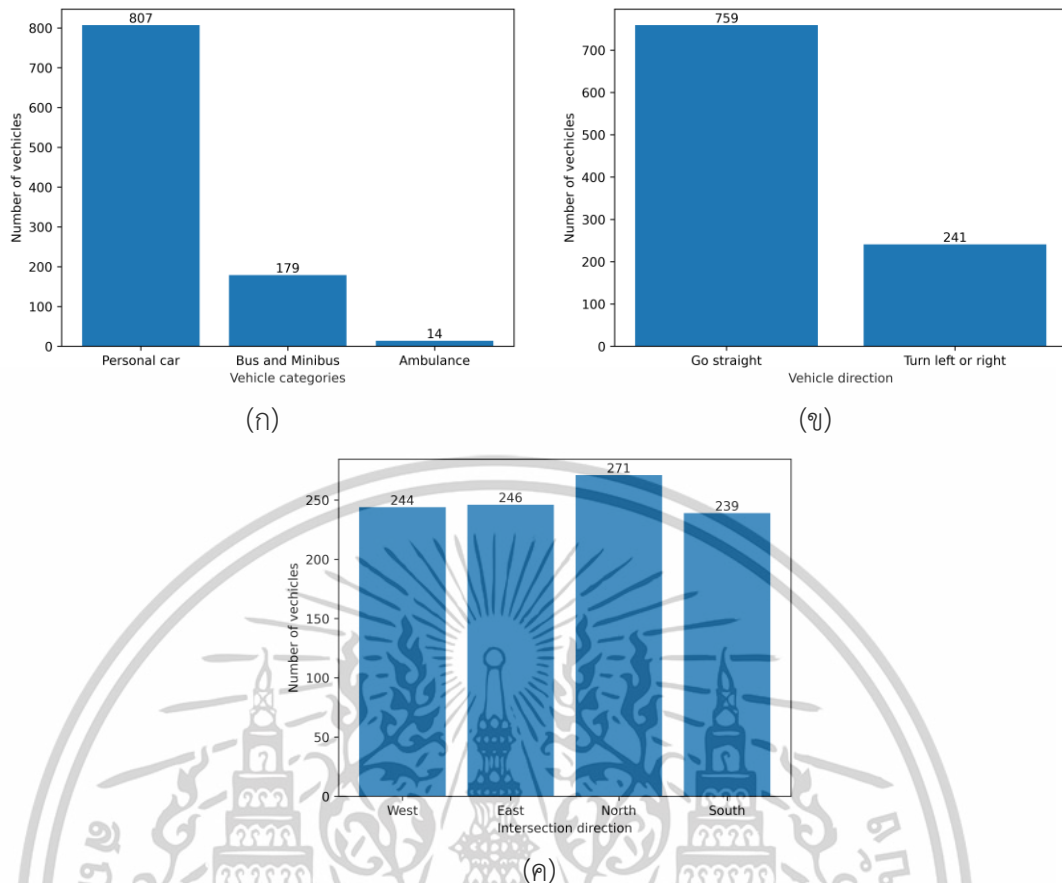
- 25: **ในขณะที่** green duration > 0 **ให้ทำ**
- 26: step = step + 1
- 27: green duration = green duration – 1
- 28: **จบloop**
- 29: **จบloop**
- 30: บันทึกสถิติของตอน
- 31: **สำหรับ** epoch ในช่วง training epochs **ให้ทำ**
- 32: รับตัวอย่างข้อมูลจากแบตช์ = batch size
- 33: ฝึกฝนโมเดล
- 34: **จบloop**
- 35: episode = episode + 1
- 36: **จบloop**
- 37: บันทึกโมเดล
-

3.9.2 การตั้งค่าสภาพแวดล้อม (Environment Setting)

จากคุณสมบัติของซอฟต์แวร์ SUMO ที่สามารถรองรับการควบคุมสัญญาณไฟจราจรแบบเรียลไทม์และสามารถสร้างสถานการณ์จำลองได้อย่างหลากหลายและยืดหยุ่น งานวิจัยนี้จึงได้นำ SUMO มาสร้างสภาพแวดล้อมจำลองสำหรับฝึกสอนและทดสอบระบบควบคุมสัญญาณไฟจราจร โดยสภาพแวดล้อมจำลองถูกออกแบบให้มีการปล่อยยานพาหนะเข้ามาแบบสุ่มตามช่วงเวลา โดยใช้ Random Seed ใน Python สำหรับสร้างยานพาหนะแบบสุ่มในแต่ละทิศทาง เพื่อให้แน่ใจว่าแต่ละตอนจะมีสถานการณ์ของกระแสการจราจร (Traffic Flow) ที่เท่ากันเสมอ เมื่อทำซ้ำเพื่อเปรียบเทียบกับวิธีอื่นหรือเพื่อตรวจสอบผลลัพธ์ โดยแต่ละตอนมีสถานการณ์ของกระแสการจราจร (Traffic Flow) ที่แตกต่างกันในการเรียนรู้ 500 ตอนและทดสอบ 25 ตอน ยานพาหนะที่สร้างขึ้นในสถานการณ์ที่กำหนดจะผ่านทางแยกเพียงครั้งเดียว และจำนวนรถที่มาจากแต่ละแยกจะมีปริมาณใกล้เคียงกัน

เราให้ความน่าจะเป็นที่รถจะเลี้ยวซ้ายหรือขวาแทนที่จะตรงไป เป็นอัตราส่วนประมาณ 1 ต่อ 3 ตามลำดับ โดยมีการสร้างรถยนต์ 1000 คันต่อเนื่องเป็นเวลา 5400 วินาทีใน 1 ตอน ด้วยความน่าจะเป็นที่จะปล่อยรถยนต์หนึ่งคันต่อวินาทีถูกกำหนดโดยใช้การแจกแจงแบบไวบูลล์ (Weibull Distribution) ที่มีพารามิเตอร์รูปร่างเท่ากับ 2 โดยรูปที่ 3.7 แสดงจำนวนยานพาหนะที่สร้างขึ้นตามการตั้งค่าใน 1 ตอน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 3.7 (ก) การสร้างสภาพการจราจรใน 1 ตอนโดยอิงตามประเภทของยานพาหนะ (ข) การสร้างสภาพการจราจรใน 1 ตอนโดยอิงตามทิศทางของยานพาหนะ (ค) การสร้างสภาพการจราจรใน 1 ตอนโดยอิงตามทิศทางของทางแยก

เราสร้างสถานการณ์สำหรับการจำลองด้วยแนวคิดที่ว่ารถยนต์ 1,000 คันจะผ่านทางแยกในเวลา 5,400 วินาที หรือ 90 นาที การใช้รถจำนวนดังกล่าวถือว่าเหมาะสมและครอบคลุม เนื่องจากทำให้เกิดสถานการณ์การจราจรที่ซับซ้อนพอสมควร ในขณะเดียวกันก็รักษาความซับซ้อนของระบบและความต้องการทรัพยากรให้อยู่ในขีดจำกัดที่จัดการได้ ความสมดุลนี้ช่วยให้การเรียนรู้และการควบคุมการจำลองมีประสิทธิภาพมากขึ้น โดยการตั้งค่าพารามิเตอร์สำหรับสภาพแวดล้อมแสดงอยู่ในตารางที่ 3.6

ตารางที่ 3.6 การตั้งค่าพารามิเตอร์ของสภาพแวดล้อม

พารามิเตอร์	ค่าที่ใช้
จำนวนขั้นตอนสูงสุด (วินาที)	5400
จำนวนรถที่สร้างทั้งหมด	1000
จำนวนตอนทั้งหมด	500

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้拿去ใช้ประโยชน์ด้านการค้า

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

การจำลองใช้เวลาประมาณ 150-180 วินาทีต่อหนึ่งตอนสำหรับขั้นตอนการฝึกฝน โดยรวมแล้วใช้เวลาประมาณ 24 ชั่วโมงสำหรับการฝึกทั้งหมด 500 ตอน ส่วนในขั้นตอนการทดสอบ การจำลองใช้เวลาประมาณ 130-140 วินาทีต่อหนึ่งตอน โดยรวมแล้วใช้เวลาประมาณ 1 ชั่วโมงสำหรับการทดสอบทั้งหมด 25 ตอน ซึ่งเวลาในการทดสอบกับสภาพแวดล้อมที่แตกต่างกัน เช่น การเปลี่ยนจำนวนช่องจราจรของแต่ละทิศทาง ใช้เวลาในการทดสอบที่ใกล้เคียงกันทั้งหมด โดยที่การทดลองทั้งหมดดำเนินการบนเครื่องคอมพิวเตอร์ที่มีหน่วยประมวลผล Intel Core i5-12400F หน่วยความจำ (RAM) 16 GB และการ์ดจอ NVIDIA GeForce RTX 3050

สำหรับการทดลองนี้ ยานพาหนะมีทั้งหมด 4 ประเภท ได้แก่ รถยนต์ส่วนบุคคล รถบัส รถมินิบัส และรถพยาบาล ซึ่งแต่ละประเภทมีคุณลักษณะดังนี้

รถยนต์ส่วนบุคคลมีความยาว 4.5 เมตร ความเร็วสูงสุด 80 กม./ชม. และอัตราเร่ง 1 ม./วินาที²

รถบัสมีความยาว 12 เมตร และรถมินิบัสมีความยาว 7 เมตร ทั้งสองประเภทนี้มีความเร็วสูงสุด 60 กม./ชม. และอัตราเร่ง 1 ม./วินาที²

รถพยาบาลมีความยาว 6.5 เมตร ความเร็วสูงสุด 120 กม./ชม. และอัตราเร่ง 2 ม./วินาที²

โดยยานพาหนะทั้งหมดมีอัตราการชะลอตัว 4.5 ม./วินาที² และความเร็วออกเดินทาง 36 กม./ชม. ระยะห่างขั้นต่ำระหว่างยานพาหนะสองคันคือ 2.5 เมตร และค่าความไม่สมบูรณ์แบบของผู้ขับขี่ (Driver Imperfection) เท่ากับ 0.5

เรากำหนดขีดจำกัดความเร็วสำหรับยานพาหนะทั่วไปตามกฎหมายการขับขี่ของท้องถิ่นในเขตเมือง อย่างไรก็ตาม รถพยาบาลอาจต้องขับด้วยความเร็วสูงในกรณีฉุกเฉิน ดังนั้น มีการตั้งขีดจำกัดความเร็วของรถพยาบาลให้สูงขึ้น

3.10 ตรรกะแบบฟัซซี (Fuzzy Logic)

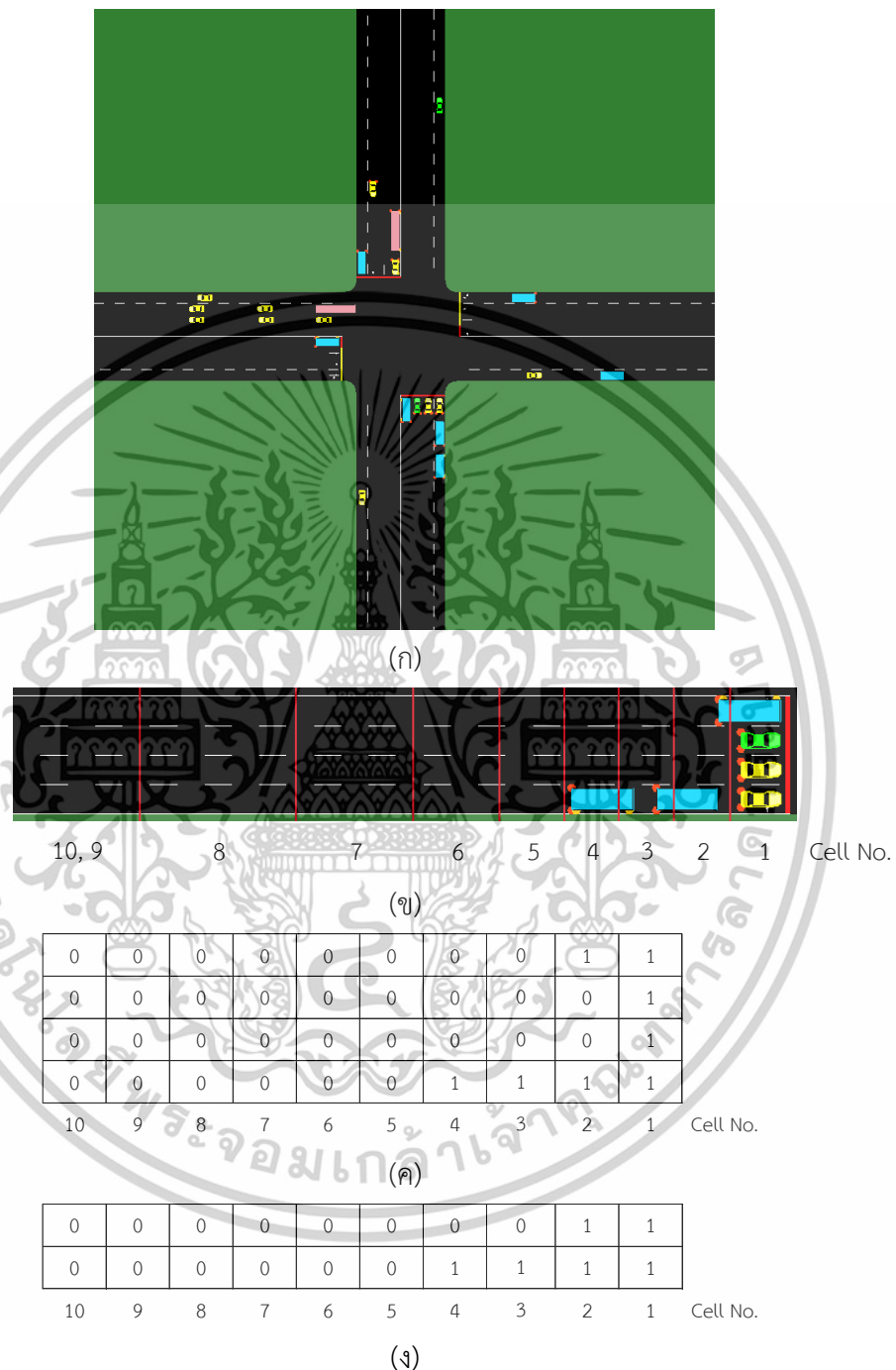
Scikit-Fuzzy Library หรือ Skfuzzy [44] ถูกใช้สำหรับการคำนวณตรรกะแบบฟัซซี (Fuzzy Logic) โดยพารามิเตอร์ที่เลือกใช้เป็นอินพุต ได้แก่ เฟสไฟเขียว (GP) และ เฟสไฟแดง (RP) [29] โดยที่ GP (Green Phase) คือผลรวมของค่าที่อยู่ในเซลล์ของกลุ่มช่องจราจรที่จะได้รับสัญญาณไฟเขียวในเฟสนั้น และ RP (Red Phase) คือผลรวมของค่าที่อยู่ในเซลล์ของกลุ่มช่องจราจรที่เป็นสัญญาณไฟแดงอยู่ในเฟสนั้น

โดยอินพุตที่ใช้ในการประมวลผลถูกออกแบบให้คล้ายกับข้อมูลสถานะของ DQN โดยอิงจากตำแหน่งของยานพาหนะในช่องจราจร แต่ละความละเอียดลง เหลือเพียงข้อมูลการมีอยู่ของรถในแต่ละเซลล์ เนื่องจากในส่วนนี้เน้นการประเมินภาพรวมของสถานการณ์จราจรในแต่ละช่วงเวลา จึงไม่จำเป็นต้องใช้ข้อมูลเวลาที่ละเอียดเหมือนใน DQN ซึ่งช่วยลดความซับซ้อนในการประมวลผล

ทั้งนี้ ค่าในแต่ละเซลล์จะมีค่าเป็น 1 หากมีรถอยู่ในตำแหน่งนั้น และเป็น 0 หากไม่มีรถอยู่ ดังแสดงในรูปที่ 3.8 ส่งผลให้ค่า GP มีค่าอยู่ในช่วงระหว่าง 0 ถึง 20 (เนื่องจากมี 2 ทิศทางที่จะได้รับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณไฟเขียว และมีจำนวนเซลล์รวมกันสูงสุด 20 เซลล์) ส่วนค่า RP จะมีค่าอยู่ช่วงระหว่าง 0 ถึง 60 (ซึ่งได้จาก 6 ทิศทางที่เหลื้รวมกัน)



รูปที่ 3.8 (ก) ภาพรวมการจราจรบริเวณทางแยก (ข) ภาพรวมการจราจรทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์ (ค) ค่าที่ระบุตำแหน่งของรถแต่ละคันทางด้านทิศใต้ของทางแยกที่แบ่งออกเป็นเซลล์ (ง) ค่าที่ระบุตำแหน่งของรถทางด้านทิศใต้ของทางแยกที่จะถูกเก็บไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และระยะเวลาของสัญญาณไฟเขียวคือค่าเอาต์พุต โดยจะมีค่าอยู่ในช่วง 0 ถึง 30 วินาที ในการทดลองนี้ ได้สร้างชุดข้อมูลที่รวมชุดค่าผสมที่เป็นไปได้ทั้งหมดของ GP และ RP ไว้ก่อนแล้ว เนื่องมาจากการประมวลผลของตรรกะแบบฟัซซี (Fuzzy Logic) ซึ่งการเลือกเฟสใหม่มีความหน่วง 0.5 ถึง 1 วินาที ส่งผลให้ระบบเกิดความล่าช้าในกระบวนการที่กำหนดเวลาสำหรับไฟเขียว ค่าอินพุต GP และ RP จะถูกใช้เพื่อรับค่าจากชุดข้อมูลผลลัพธ์ที่ถูกสร้างไว้ก่อนแล้วแทนการประมวลผลตรรกะแบบฟัซซี (Fuzzy Logic) ซ้ำในทุก ๆ เฟส

ฟัซซีเซตของเฟสไฟเขียว (GP) มีทั้งหมด 5 เซต ดังแสดงในรูปที่ 3.9 โดยมีนิยามดังนี้

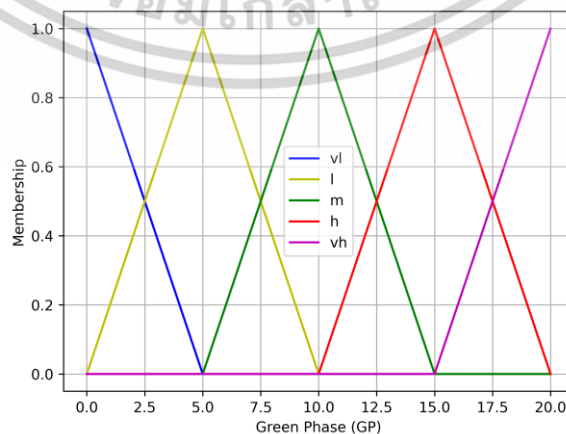
$$GP, \text{Very Low}(x) = \begin{cases} -0.2x + 1, & 0 \leq x \leq 5 \\ 0, & x > 5 \end{cases} \quad (3.4)$$

$$GP, \text{Low}(x) = \begin{cases} 0, & x < 0 \\ 0.2x, & 0 \leq x \leq 5 \\ -0.2x + 2, & 5 < x \leq 10 \\ 0, & x > 10 \end{cases} \quad (3.5)$$

$$GP, \text{Medium}(x) = \begin{cases} 0, & x < 5 \\ 0.2x - 1, & 5 \leq x \leq 10 \\ -0.2x + 3, & 10 < x \leq 15 \\ 0, & x > 15 \end{cases} \quad (3.6)$$

$$GP, \text{High}(x) = \begin{cases} 0, & x < 10 \\ 0.2x - 2, & 10 \leq x \leq 15 \\ -0.2x + 4, & 15 < x \leq 20 \\ 0, & x > 20 \end{cases} \quad (3.7)$$

$$GP, \text{Very High}(x) = \begin{cases} 0, & x < 15 \\ 0.2x - 3, & 15 \leq x \leq 20 \end{cases} \quad (3.8)$$



รูปที่ 3.9 ฟัซซีเซตของเฟสไฟเขียว (GP)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับงานวิจัยเท่านั้นเพื่อการศึกษาเท่านั้น เมื่อผู้ใดเห็นนำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ฟังก์ชันเซตของเฟสไฟแดง (RP) มีทั้งหมด 5 เซต ดังแสดงในรูปที่ 3.10 โดยมีนิยามดังนี้

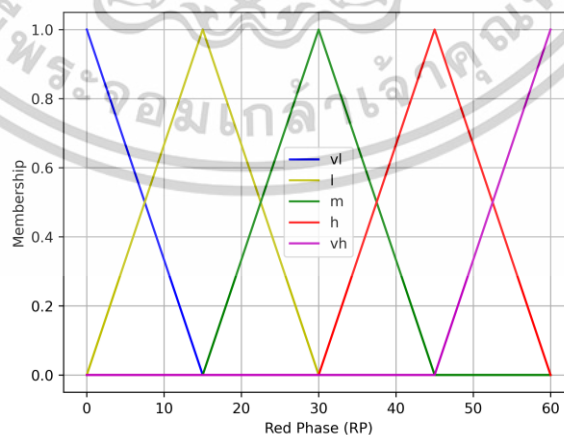
$$RP, \text{Very Low}(x) = \begin{cases} -\frac{2}{30}x + 1, & 0 \leq x \leq 15 \\ 0, & x > 15 \end{cases} \quad (3.9)$$

$$RP, \text{Low}(x) = \begin{cases} 0, & x < 0 \\ \frac{2}{30}x, & 0 \leq x \leq 15 \\ -\frac{2}{30}x + 2, & 15 < x \leq 30 \\ 0, & x > 30 \end{cases} \quad (3.10)$$

$$RP, \text{Medium}(x) = \begin{cases} 0, & x < 15 \\ \frac{2}{30}x - 1, & 15 \leq x \leq 30 \\ -\frac{2}{30}x + 3, & 30 < x \leq 45 \\ 0, & x > 45 \end{cases} \quad (3.12)$$

$$RP, \text{High}(x) = \begin{cases} 0, & x < 30 \\ \frac{2}{30}x - 2, & 30 \leq x \leq 45 \\ -\frac{2}{30}x + 4, & 45 < x \leq 60 \\ 0, & x > 60 \end{cases} \quad (3.13)$$

$$RP, \text{Very High}(x) = \begin{cases} 0, & x < 45 \\ \frac{2}{30}x - 3, & 45 \leq x \leq 60 \end{cases} \quad (3.14)$$



รูปที่ 3.10 ฟังก์ชันเซตของเฟสไฟแดง (RP)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.7 คำอธิบายฟังก์ชันสมาชิกของเฟสไฟเขียว (GP)

GP			
ลำดับ	ระดับของการเป็นสมาชิก	คำอธิบาย	ช่วงค่า
1	vl	ต่ำมาก	0-5
2	l	ต่ำ	0-10
3	m	ปานกลาง	5-15
4	h	สูง	10-20
5	vh	สูงมาก	15 เป็นต้นไป

ตารางที่ 3.8 คำอธิบายฟังก์ชันสมาชิกของเฟสไฟแดง (RP)

RP			
ลำดับ	ระดับของการเป็นสมาชิก	คำอธิบาย	ช่วงค่า
1	vl	ต่ำมาก	0-10
2	l	ต่ำ	0-30
3	m	ปานกลาง	15-45
4	h	สูง	30-60
5	vh	สูงมาก	45 เป็นต้นไป

ตารางที่ 3.9 กฎการอนุมาน

	เฟสไฟแดง (RP)					
		vl	l	m	h	vh
เฟสไฟเขียว (GP)	vl	vvs	vvs	vvs	vvs	vvs
	l	s	s	vs	vs	vs
	m	m	m	m	s	s
	h	l	l	m	m	s
	vh	vvl	vl	l	m	m

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 3.10 คำอธิบายฟังก์ชันสมาชิกของเอาต์พุต

ระยะเวลาไฟเขียว (Green Duration)			
ลำดับ	ระดับของการเป็นสมาชิก	คำอธิบาย	ช่วงค่า
1	vvs	สั้นมากที่สุด	0-5
2	vs	สั้นมาก	0-10
3	s	สั้น	5-15
4	m	ปานกลาง	10-20
5	l	ยาว	15-25
6	vl	ยาวมาก	20-30
7	vvl	ยาวมากที่สุด	25 เป็นต้นไป

ฟังก์ชันเซตของเอาต์พุต (Fuzzy Output) หรือ ระยะเวลาสัญญาณไฟเขียว มีทั้งหมด 7 เซต ดังแสดงในรูปที่ 3.11 โดยมีนิยามดังนี้

$$GreenDuration, Very\ very\ short(x) = \begin{cases} -0.2x + 1, & 0 \leq x \leq 5 \\ 0, & x > 5 \end{cases} \quad (3.15)$$

$$GreenDuration, Very\ short(x) = \begin{cases} 0, & x < 0 \\ 0.2x, & 0 \leq x \leq 5 \\ -0.2x + 2, & 5 < x \leq 10 \\ 0, & x > 10 \end{cases} \quad (3.16)$$

$$GreenDuration, Short(x) = \begin{cases} 0, & x < 5 \\ 0.2x - 1, & 5 \leq x \leq 10 \\ -0.2x + 3, & 10 < x \leq 15 \\ 0, & x > 15 \end{cases} \quad (3.17)$$

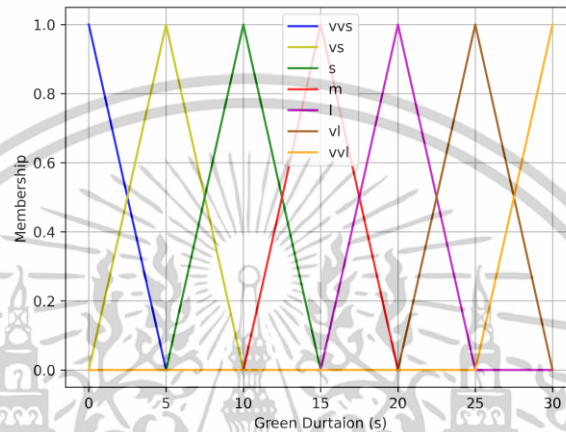
$$GreenDuration, Medium(x) = \begin{cases} 0, & x < 10 \\ 0.2x - 2, & 10 \leq x \leq 15 \\ -0.2x + 4, & 15 < x \leq 20 \\ 0, & x > 20 \end{cases} \quad (3.18)$$

$$GreenDuration, Long(x) = \begin{cases} 0, & x < 15 \\ 0.2x - 3, & 15 \leq x \leq 20 \\ -0.2x + 5, & 20 < x \leq 25 \\ 0, & x > 25 \end{cases} \quad (3.19)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$GreenDuration, Very\ long(x) = \begin{cases} 0, & x < 20 \\ 0.2x - 4, & 20 \leq x \leq 25 \\ -0.2x + 6, & 25 < x \leq 30 \\ 0, & x > 30 \end{cases} \quad (3.20)$$

$$GreenDuration, Very\ very\ long(x) = \begin{cases} 0, & x < 25 \\ 0.2x - 5, & 25 \leq x \leq 30 \end{cases} \quad (3.21)$$



รูปที่ 3.11 ฟังก์ชันเซตของเอาต์พุต

การประมวลผลตรรกะแบบฟัซซี (Fuzzy Logic) ใช้กลไกการอนุมานด้วยวิธีแบบแมมดานิ (Mamdani Fuzzy Inference) [44] ขั้นตอนแรกคือ การทำฟัซซีฟิเคชัน (Fuzzification) ซึ่งรับค่าอินพุตที่เป็นตัวเลขที่มีค่าแน่นอน (Crisp Input) และใช้ฟังก์ชันสมาชิก (Membership Function) แปลงให้เข้ากับระดับของการเป็นสมาชิก (Degree of Membership) ดังแสดงในรูปที่ 3.9 และ 3.10 ซึ่งจะได้ข้อมูลที่เป็นฟัซซี (Fuzzy Input) โดยคำอธิบายของฟังก์ชันการเป็นสมาชิกจะแสดงอยู่ในตารางที่ 3.7 และ 3.8 จากนั้นจะถูกส่งไปยังกระบวนการของกลไกการอนุมาน (Inference Engine) โดยใช้แนวคิด IF-THEN โดยอิงตามกฎที่กำหนดไว้ ดังที่แสดงในตารางที่ 3.9 อย่างไรก็ตาม อินพุตสามารถเข้ากับกฎหลายข้อได้ จากนั้นผลลัพธ์ที่ได้จากกระบวนการของกลไกการอนุมาน (Inference Engine) โดยใช้กฎที่กำหนดไว้ทั้งหมด ซึ่งจะถูกรวมเข้าด้วยกันโดยใช้ตัวดำเนินการยูเนียน (Union Operation) และใช้ฟังก์ชันสมาชิก (Membership Function) ดังแสดงในรูปที่ 3.11 คำอธิบายของฟังก์ชันสมาชิกแสดงอยู่ในตารางที่ 3.10 จากนั้นนำผลลัพธ์ที่ได้เข้าสู่กระบวนการทำให้ฟัซซีฟิเคชัน (Defuzzification) ซึ่งจะทำให้การแปลงผลลัพธ์ที่เป็นฟัซซี (Fuzzy Output) ให้เป็นค่าที่ชัดเจน (Crisp Output) ด้วยวิธีการคำนวณหาจุดศูนย์กลางมวลหรือจุดศูนย์กลางถ่วง (COG) [44] บนแกน X

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตัวอย่างการประมวลผลตรรกะแบบฟัซซีในการคำนวณหาระยะเวลาของสัญญาณไฟเขียว โดยใช้กลไกการอนุมานวิธีแบบแมมดานี (Mamdani Fuzzy Inference) โดยกำหนดให้เฟสไฟเขียว (GP) มีค่าเท่ากับ 7 และเฟสไฟแดง (RP) มีค่าเท่ากับ 3

1. การทำฟัซซีฟิเคชัน (Fuzzification) โดยเริ่มจากการหาระดับของการเป็นสมาชิกเฟสไฟเขียว (GP) ซึ่ง GP ที่มีค่าเท่ากับ 7 จะเข้ากับการเป็นสมาชิกของฟัซซีเซตของเฟสไฟเขียวมีรตต่ำ (Low) และเฟสไฟเขียวมีรตปานกลาง (Medium) ซึ่งคำนวณค่าระดับของการเป็นสมาชิกได้ดังนี้

$$GP, Low(7) = 0.2(7) + 2 = 0.6 \text{ และ } GP, Medium(7) = 0.2(7) - 1 = 0.4$$

และหาระดับของการเป็นสมาชิกเฟสไฟแดง (RP) ซึ่ง RP ที่มีค่าเท่ากับ 3 จะเข้ากับการเป็นสมาชิกของฟัซซีเซตของเฟสไฟแดงมีรตต่ำมาก (Very Low) และเฟสไฟแดงมีรตต่ำ (Low) ซึ่งคำนวณค่าระดับของการเป็นสมาชิกได้ดังนี้

$$RP, Very Low(3) = -\frac{2}{30}(3) + 1 = 0.8 \text{ และ } RP, Low(3) = \frac{2}{30}(3) = 0.2$$

2. การประเมินกฎ (Rule Evaluation) ทำการประมวลผลของอินพุตฟัซซีจากกฎที่มีอยู่ในฐานกฎ (Rule Base) ในตารางที่ 3.9 โดยใช้การอินเตอร์เซกชัน (AND) ซึ่งจะเข้ากับกฎดังนี้

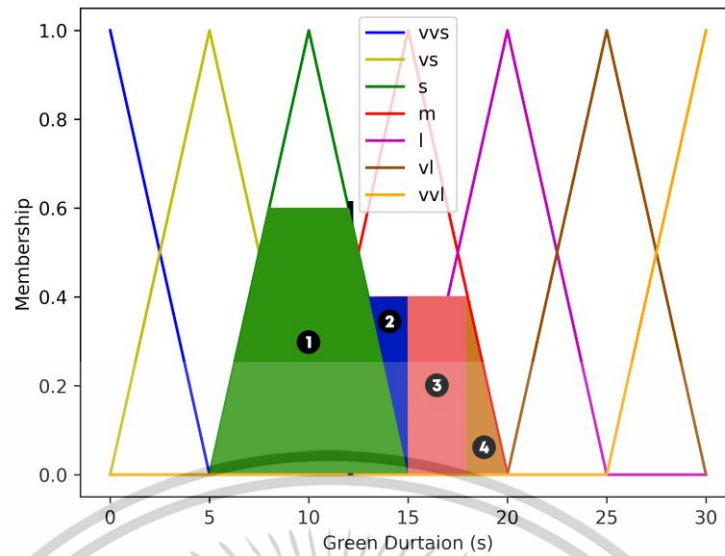
ถ้า เฟสไฟเขียวมีรตต่ำ (Low) และ เฟสไฟแดงมีรตต่ำมาก (Very Low) แล้ว ให้ระยะสัญญาณไฟเขียวสั้น (Short) จะได้ GP, Low(7) AND RP, Very low(3) เท่ากับ $\min(0.6, 0.8) = 0.6$

ถ้า เฟสไฟเขียวมีรตต่ำ (Low) และ เฟสไฟแดงมีรตต่ำ (Low) แล้ว ให้ระยะสัญญาณไฟเขียวสั้น (Short) จะได้ GP, Low(7) AND RP, Low(3) เท่ากับ $\min(0.6, 0.2) = 0.2$

ถ้า เฟสไฟเขียวมีรตปานกลาง (Medium) และ เฟสไฟแดงมีรตต่ำมาก (Very Low) แล้ว ให้ระยะสัญญาณไฟเขียวปานกลาง (Medium) จะได้ GP, Medium(7) AND RP, Very Low(3) เท่ากับ $\min(0.4, 0.2) = 0.2$

ถ้า เฟสไฟเขียวมีรตปานกลาง (Medium) และ เฟสไฟแดงมีรตต่ำ (Low) แล้ว ให้ระยะสัญญาณไฟเขียวปานกลาง (Medium) จะได้ GP, Medium(7) AND RP, Low(3) เท่ากับ $\min(0.4, 0.8) = 0.4$

3. การรวมกฎ (Aggregation) นำค่าผลลัพธ์ที่ได้จากการประเมินกฎ (Rule Evaluation) ทั้งหมดมารวมกัน โดยใช้การยูเนียน (OR) ซึ่งจะได้ผลลัพธ์ที่เป็นฟัซซี (Fuzzy Output) ดังแสดงในรูปที่ 3.12



รูปที่ 3.12 ตัวอย่างผลลัพธ์ที่ได้รับจากกระบวนการทำดีฟัซซิฟิเคชัน

4. การทำดีฟัซซิฟิเคชัน (Defuzzification) นำผลลัพธ์ที่เป็นฟัซซี (Fuzzy Output) ดังแสดงในรูปที่ 3.12 แปลงกลับไปเป็นค่าที่ชัดเจน (Crisp Output) โดยการใช้วิธีเซนทรอยด์ (Centroid Method) หรือ วิธีการหาจุดศูนย์กลาง (Center of Gravity : COG) ซึ่งคำนวณได้ดังนี้

$$\text{รูปหมายเลข 1 มีพื้นที่ } A_1 = \frac{1}{2}(14)(0.6) = 4.2 \text{ และมีจุดศูนย์กลาง } x_1 = 5 + \frac{10}{2} = 10$$

$$\text{รูปหมายเลข 2 มีพื้นที่ } A_2 = \frac{1}{2}(2)(0.4) = 0.4 \text{ และมีจุดศูนย์กลาง } x_2 = 15 - \frac{2}{3} = \frac{43}{3}$$

$$\text{รูปหมายเลข 3 มีพื้นที่ } A_3 = (3)(0.4) = 1.2 \text{ และมีจุดศูนย์กลาง } x_3 = 15 + \frac{3}{2} = 16.5$$

$$\text{รูปหมายเลข 4 มีพื้นที่ } A_4 = \frac{1}{2}(2)(0.4) = 0.4 \text{ และมีจุดศูนย์กลาง } x_4 = 18 + \frac{2}{3} = \frac{56}{3}$$

$$\text{จะได้ Crisp Output } (\bar{x}) = \frac{(4.2)(10) + (0.4)\left(\frac{43}{3}\right) + (1.2)(16.5) + (0.4)\left(\frac{56}{3}\right)}{4.2+0.4+1.2+0.4} = \frac{75}{6.2} \approx 12.097$$

ผลลัพธ์ที่ได้คือจุดศูนย์กลางจะอยู่ที่ 12.097 บนแกน X ดังนั้น ระยะเวลาของสัญญาณไฟเขียวในเฟสนี้คือ 13 วินาที

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 4

ผลการทดลอง

ผลการทดลองแบ่งออกเป็น 7 ส่วน ส่วนแรกเป็นผลของกระบวนการฝึกอบรมโมเดลและส่วนที่เหลือเป็นผลการทดลอง โดยใช้โมเดลที่ผ่านการฝึกอบรมเสร็จสิ้นแล้ว ผลลัพธ์จะแสดงค่าเฉลี่ยที่ได้รับจากการทดลองทั้งหมด 25 ตอน ซึ่งเป็นสถานการณ์ที่แตกต่างกัน ผลการทดลองสำหรับการควบคุมสัญญาณไฟจราจรจะมีทั้งหมด 3 แบบ

1. การควบคุมสัญญาณไฟจราจรแบบดั้งเดิมเป็นวิธีที่ส่วนใหญ่ยังใช้กันทั่วโลกในปัจจุบันซึ่งให้สัญญาณไฟสีเขียวเป็นเวลา 10 วินาทีตามลำดับและทำซ้ำสำหรับแต่ละแยกดังแสดงในรูปที่ 5 โดยในผลการทดลอง เรียกว่า วิธีแบบดั้งเดิม

2. การควบคุมสัญญาณไฟจราจรโดยใช้ Deep Q-Network ร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) ที่เป็นต้นแบบ โดยในผลการทดลองเรียกว่า วิธี DQN แบบดั้งเดิมร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]

3. การควบคุมสัญญาณไฟจราจรโดยใช้วิธีการที่น่าเสนอ

ในตารางผลการทดลอง ALL-4 หมายถึง ช่องจราจรขาออก (Outgoing Lane) มี 4 ช่องจราจร (Lane) เท่ากันทั้งหมด E-3 หมายถึงถนนที่มาจากทางทิศตะวันออกมี 3 ช่องจราจรขาออก (Outgoing Lane) N-3 หมายถึงถนนที่มาจากทางทิศเหนือมี 3 ช่องจราจรขาออก (Outgoing Lane) S-3 หมายถึงถนนที่มาจากทางทิศใต้มี 3 ช่องจราจรขาออก (Outgoing Lane) และ W-3 หมายถึงถนนที่มาจากทางทิศตะวันตกมี 3 ช่องจราจรขาออก (Outgoing Lane)

ผลการทดลองในส่วนที่ 4.1-4.7 จะถูกเปรียบเทียบกับการใช้วิธีแบบดั้งเดิมและวิธี DQN แบบดั้งเดิมร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) โดยค่าต่อไปนี้ถูกใช้เพื่อเปรียบเทียบผลลัพธ์

1. WT_{AVG} คือ เวลารอคอยสะสมเฉลี่ยของยานพาหนะทุกคันที่อยู่ในช่องจราจรขาออก (Outgoing Lane) ต่อหนึ่งขั้นตอน (Step)

$$WT_{Total} = \sum_{i=1}^N WT[i] \quad (4.1)$$

โดยที่ N คือ จำนวนรถที่ทางแยกในช่องจราจรขาออก ณ ขั้นตอนที่ t

WT คือ เวลารอสะสมของยานพาหนะ

$$WT_{AVG} = \frac{1}{T} \times \sum_{i=1}^T WT_{Total}[i] \quad (4.2)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

โดยที่ T คือ ขั้นตอนสูงสุด

WTL_{Total} คือ เวลารอทั้งหมด ณ ขั้นตอนที่ t ดังสมการ 4.1

2. WTL_{AVG} คือ เวลารอคอยสะสมเฉลี่ยของยานพาหนะทุกคันที่อยู่ในช่องจราจรขาออก (Outgoing Lane) สำหรับเลียวย้ายต่อหนึ่งขั้นตอน (Step)

$$WTL_{Total} = \sum_{i=1}^N WTL[i] \quad (4.3)$$

โดยที่ N คือ จำนวนรถที่ทางแยกของช่องจราจรที่ใช้สำหรับเลียวย้าย ณ ขั้นตอนที่ t

WTL คือ เวลารอสะสมของรถในช่องจราจรที่ใช้สำหรับเลียวย้าย

$$WTL_{AVG} = \frac{1}{T} \times \sum_{i=1}^T WTL_{Total}[i] \quad (4.4)$$

โดยที่ T คือ ขั้นตอนสูงสุด

WTL_{Total} คือ เวลารอรวมของรถในช่องจราจรที่ใช้สำหรับเลียวย้าย ณ ขั้นตอนที่ t ดังสมการ 4.3

3. WTC_{AVG} คือ เวลารอคอยสะสมเฉลี่ยต่อรถยนต์หนึ่งคันที่อยู่ในช่องจราจรขาออก (Outgoing Lane) ต่อหนึ่งขั้นตอน (Step)

$$WTC_{AVG} = \frac{1}{T} \times \sum_{i=1}^T \frac{1}{Q} WTL_{Total}[i] \quad (4.5)$$

โดยที่ T คือ ขั้นตอนสูงสุด

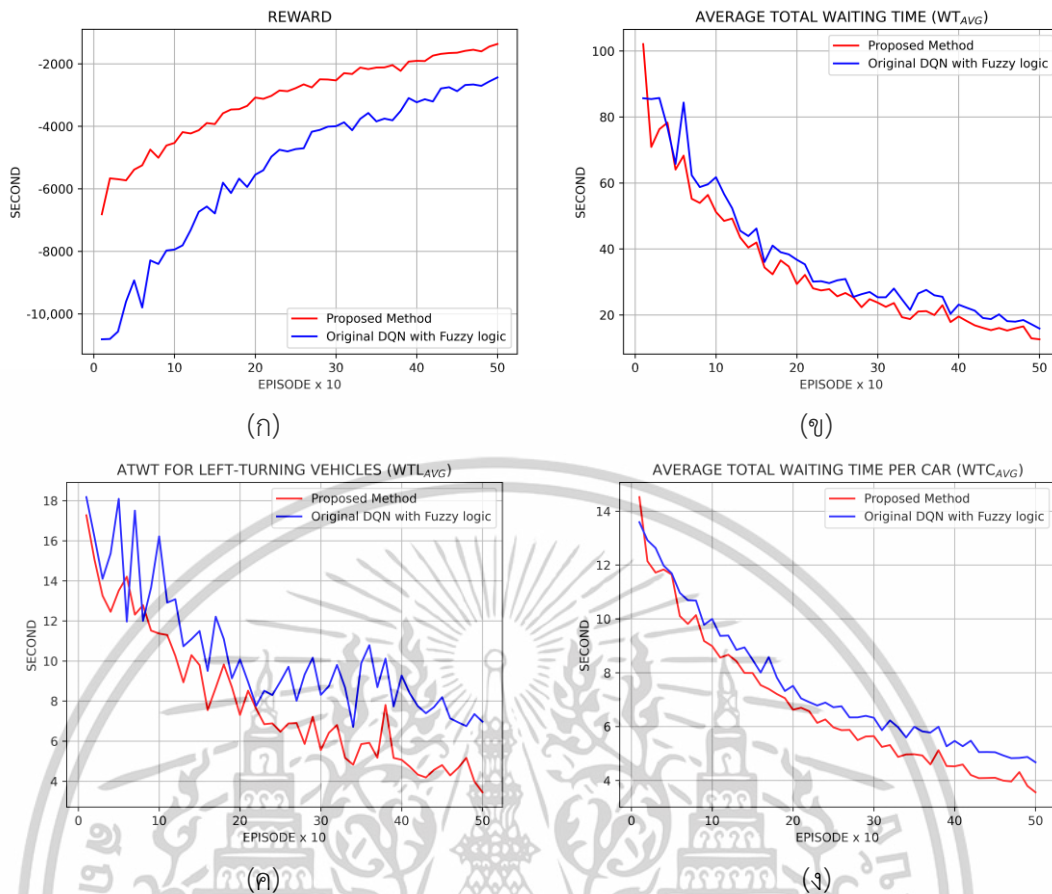
Q คือ จำนวนยานพาหนะที่ทางแยกที่มีความเร็วต่ำกว่า 0.1 เมตรต่อวินาที

WTL_{Total} คือ เวลารอคอยทั้งหมด ณ ขั้นตอนที่ t ดังสมการ 4.1

4.1 ผลการฝึกอบรมในสภาพแวดล้อมแบบสมมูล (ALL-4)

ระบบเสริมจลันกระบวนการเรียนรู้โดยใช้การเรียนรู้ทั้งหมด 500 ตอน ผลการทดลองที่แสดงในรูปที่ 4.1 (ก) ค่าของรางวัล (Reward) ที่มีค่าติดลบบ่งชี้ว่าเวลารอคอยเพิ่มขึ้นหลังจากเลือกการกระทำ (Action) ซึ่งกราฟแสดงแนวโน้มมูลค่ารวมของรางวัล (Reward) ที่มีค่าติดลบในแต่ละตอนมีแนวโน้มลดลง แสดงว่าระบบได้มีการเรียนรู้แล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูปที่ 4.1 (ก) กราฟแสดงผลการฝึกอบรมอิงตามรางวัลที่มีค่าติดลบ (ข) กราฟแสดงผลการฝึกอบรมอิงตามเวลารอคอยรวมโดยเฉลี่ย (ค) กราฟแสดงผลการฝึกอบรมอิงตามเวลารอคอยโดยเฉลี่ยของยานพาหนะที่จะเลี้ยวซ้าย (ง) กราฟแสดงผลการฝึกอบรมอิงตามเวลารอคอยโดยเฉลี่ยต่อคัน

ในรูปที่ 4.1 (ข)-(ง) กราฟแสดงให้เห็นว่าหลังจากเรียนรู้แล้ว วิธีการที่นำเสนอมีประสิทธิภาพเหนือกว่าการใช้วิธี DQN แบบดั้งเดิมร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic)

4.2 สภาพแวดล้อมแบบสมดุลง (ALL-4)

ผลการทดลองในสภาพแวดล้อมแบบเดิมที่มี 4 ช่องจราจรขาออก (Outgoing Lane) ในแต่ละแยก ในตารางที่ 4.1 สำหรับวิธีที่นำเสนอ ค่าของ WT_{AVG} , WTL_{AVG} และ WTC_{AVG} สามารถลดลงได้ 69.80%, 23.24% และ 60.78% เมื่อเปรียบเทียบกับวิธีแบบดั้งเดิม และสามารถลดลงได้ 18.46%, 40.36% และ 21.23% เมื่อเปรียบเทียบกับวิธี DQN แบบดั้งเดิมร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) ตามลำดับ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ตารางที่ 4.1 ผลการทดลองในสภาพแวดล้อมแบบสมดุลง (ALL-4)

ตัวชี้วัด	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม รวมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
$WT_{AVG}(s)$	43.57	16.14	13.16
$WTL_{AVG}(s)$	5.12	6.59	3.93
$WTC_{AVG}(s)$	9.46	4.71	3.71

มีข้อสังเกตว่าวิธี DQN แบบดั้งเดิมรวมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) ให้ผลลัพธ์ที่แย่งกว่าวิธีแบบดั้งเดิมเมื่อเทียบกับรถที่จะเลี้ยวซ้าย เนื่องจากการให้ฟังก์ชันการให้รางวัลที่ไม่เหมาะสม ซึ่งวิธีแบบดั้งเดิมให้สัญญาณไฟเขียวคงที่ 10 วินาทีตามลำดับในทุกเฟส

อย่างไรก็ตาม วิธีการที่นำเสนอซึ่งเราปรับเปลี่ยนการคำนวณรางวัล ให้ผลลัพธ์ที่ดีกว่าวิธีอื่น ๆ ทั้งหมดเมื่อเปรียบเทียบกับตัวชี้วัดทั้งหมด

4.3 สภาพแวดล้อมการจราจรแบบขับรถชิดซ้าย

ผลการทดลองที่แสดงไว้ในหัวข้อนี้ มีการใช้ค่าต่อไปนี้ เพื่อเปรียบเทียบผลลัพธ์แทนสมการ 4.4

1. WTR_{AVG} คือ เวลารอคอยสะสมเฉลี่ยของยานพาหนะทุกคันที่อยู่ในช่องจราจรขาออก (Outgoing Lane) สำหรับเลี้ยวขวาต่อหนึ่งขั้นตอน (Step)

$$WTR_{Total} = \sum_{i=1}^N WTR[i] \quad (4.6)$$

โดยที่ N คือ จำนวนรถที่ทางแยกของช่องจราจรที่ใช้สำหรับเลี้ยวขวา ณ ขั้นตอนที่ t

WTR คือ เวลารอสะสมของรถในช่องจราจรที่ใช้สำหรับเลี้ยวขวา

$$WTR_{AVG} = \frac{1}{T} \times \sum_{i=1}^T WTR_{Total}[i] \quad (4.7)$$

โดยที่ T คือ ขั้นตอนสูงสุด

WTR_{Total} คือ เวลารอรวมของรถในช่องจราจรที่ใช้สำหรับเลี้ยวขวา ณ ขั้นตอนที่ t

ดังสมการ 4.6

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ผลการทดลองแสดงไว้ในตารางที่ 4.2 สำหรับสภาพแวดล้อมการจราจรแบบซับซ้อนซ้ำโดยมี 4 ช่องจราจรขาออก (Outgoing Lane) ในแต่ละแยก

ตารางที่ 4.2 ผลการทดลองในสภาพแวดล้อมการจราจรแบบซับซ้อนซ้ำ

ตัวชี้วัด	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
WT_{AVG} (s)	43.45	16.22	12.51
WTR_{AVG} (s)	4.94	6.52	3.53
WTC_{AVG} (s)	9.44	4.64	3.66

เมื่อใช้วิธีการที่นำเสนอ ค่าของ WT_{AVG} , WTR_{AVG} และ WTC_{AVG} สามารถลดลงได้ 71.21%, 28.54% และ 61.23% เมื่อเปรียบเทียบกับวิธีแบบดั้งเดิม และสามารถลดลงได้ 22.87%, 45.86% และ 21.12% เมื่อเปรียบเทียบกับวิธี DQN แบบดั้งเดิมร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) ตามลำดับ

4.4 สภาพแวดล้อม 3 ช่องจราจร (Lane) ใน 1 แยก

ผลการทดลองแสดงไว้ในตารางที่ 4.3 สำหรับสภาพแวดล้อมที่ 1 แยก มี 3 ช่องจราจรขาออก (Outgoing Lane) และ 3 แยก มี 4 ช่องจราจรขาออก (Outgoing Lane) ซึ่งมีสภาพแวดล้อมที่เป็นไปได้ทั้งหมด 4 แบบ โดยผลลัพธ์จะแสดงค่าเฉลี่ยสำหรับสภาพแวดล้อมทั้ง 4 แบบ ซึ่งวิธีการที่นำเสนอ ค่าของ WT_{AVG} , WTL_{AVG} และ WTC_{AVG} สามารถลดลงได้ 78.01%, 14.73% และ 64.23% เมื่อเทียบกับวิธีแบบดั้งเดิม และสามารถลดลงได้ 19.12%, 39.92% และ 19.69% เมื่อเปรียบเทียบกับวิธี DQN แบบดั้งเดิมร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) ตามลำดับ

จากผลการทดลองในตารางที่ 4.3 พบว่าวิธีการแบบดั้งเดิมมีเวลารอคอยโดยรวมที่แย่ที่สุด อย่างไรก็ตาม วิธีการแบบดั้งเดิมให้ผลลัพธ์ที่ดีกว่า วิธี DQN แบบดั้งเดิมร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) เมื่อเทียบกับรถที่จะเลี้ยวซ้ายเพราะให้สัญญาณไฟเขียวเป็นเวลา 10 วินาที ตามลำดับในทุกเฟส อย่างไรก็ตาม วิธีการที่นำเสนอให้ผลลัพธ์ที่ดีกว่าวิธีอื่น ๆ ทั้งหมด เมื่อเปรียบเทียบกับตัวชี้วัดทั้งหมด

ตารางที่ 4.3 ผลการทดลองในสภาพแวดล้อมแบบ 3 ช่องจราจรใน 1 แยก

WT_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3	66.51	17.32	13.73
N-3	59.48	17.79	14.53
S-3	56.59	18.18	15.09
W-3	88.50	18.05	14.36
WTL_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3	5.05	6.93	4.07
N-3	5.02	6.96	4.09
S-3	5.00	7.23	4.70
W-3	5.01	7.37	4.26
WTC_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3	11.07	4.84	3.85
N-3	10.61	4.82	3.90
S-3	10.60	4.90	3.96
W-3	11.73	4.89	3.91

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.5 สภาพแวดล้อม 3 ช่องจราจร (Lane) ใน 2 แยก

ผลการทดลองแสดงไว้ในตารางที่ 4.4 สำหรับสภาพแวดล้อมที่ 2 แยก มี 3 ช่องจราจรขาออก (Outgoing Lane) และ 2 แยก มี 4 ช่องจราจรขาออก (Outgoing Lane) ซึ่งมีสภาพแวดล้อมที่เป็นไปได้ทั้งหมด 6 แบบ โดยผลลัพธ์จะแสดงค่าเฉลี่ยสำหรับสภาพแวดล้อมทั้ง 6 แบบ ซึ่งวิธีการที่นำเสนอ ค่าของ WT_{AVG} , WTL_{AVG} และ WTC_{AVG} สามารถลดลงได้ 81.66%, 9.22% และ 66.21% เมื่อเทียบกับวิธีแบบดั้งเดิม และสามารถลดลงได้ 20.23%, 45.14% และ 21.35% เมื่อเปรียบเทียบกับวิธี DQN แบบดั้งเดิมร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) ตามลำดับ

ตารางที่ 4.4 ผลการทดลองในสภาพแวดล้อมแบบ 3 ช่องจราจรใน 2 แยก

WT_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3 S-3	75.85	19.53	16.75
E-3 W-3	104.60	19.93	14.97
N-3 S-3	72.83	19.81	15.85
N-3 W-3	100.23	21.57	16.17
N-3 E-3	83.61	19.39	16.03
S-3 W-3	94.58	19.57	15.68
WTL_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3 S-3	5.10	8.29	4.88
E-3 W-3	5.11	8.38	4.42
N-3 S-3	5.14	8.30	4.54
N-3 W-3	5.06	9.21	4.76
N-3 E-3	5.03	8.22	4.43
S-3 W-3	5.02	8.07	4.62

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

WTC_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3 S-3	11.44	5.17	4.08
E-3 W-3	12.86	5.15	3.96
N-3 S-3	11.37	5.16	4.08
N-3 W-3	12.47	5.31	4.11
N-3 E-3	12.02	5.11	4.13
S-3 W-3	12.16	5.10	4.02

มีข้อสังเกตว่าผลลัพธ์ที่ได้สอดคล้องกับในส่วนที่ 4.4 เมื่อใช้วิธีการแบบดั้งเดิมซึ่งให้เวลารอคอยโดยรวมที่แย่ที่สุด แต่ให้ผลลัพธ์ที่ดีกว่าวิธี DQN แบบดั้งเดิมร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) เมื่อเทียบกับรถที่จะเลี้ยวซ้าย อย่างไรก็ตาม วิธีที่นำเสนอให้ผลลัพธ์ดีกว่าวิธีอื่น ๆ ทั้งหมด เมื่อเปรียบเทียบกับตัวชี้วัดทั้งหมด

4.6 สภาพแวดล้อม 3 ช่องจราจร (Lane) ใน 3 แยก

ผลการทดลองแสดงไว้ในตารางที่ 4.5 สำหรับสภาพแวดล้อมที่ 3 แยก มี 3 ช่องจราจรขาออก (Outgoing Lane) และ 1 แยก มี 4 ช่องจราจรขาออก (Outgoing Lane) ซึ่งมีสภาพแวดล้อมที่เป็นไปได้ทั้งหมด 4 แบบ โดยผลลัพธ์จะแสดงค่าเฉลี่ยสำหรับสภาพแวดล้อมทั้ง 4 แบบ ซึ่งวิธีการที่นำเสนอ ค่าของ WT_{AVG} , WTL_{AVG} และ WTC_{AVG} สามารถลดลงได้ 84.34%, 6.26% และ 67.54% เมื่อเทียบกับวิธีดั้งเดิม และสามารถลดลงได้ 19.98%, 48.46% และ 20.66% เมื่อเปรียบเทียบกับวิธี DQN แบบดั้งเดิมร่วมกันกับตรรกะแบบฟัซซี (Fuzzy Logic) ตามลำดับ

จากผลการทดลองในตารางที่ 4.5 พบว่าผลลัพธ์ที่ได้สอดคล้องกับในส่วนที่ 4.4 และ 4.5 โดยวิธีที่นำเสนอให้ผลลัพธ์ที่ดีกว่าทุกวิธีเมื่อเปรียบเทียบกับตัวชี้วัดทั้งหมด

ตารางที่ 4.5 ผลการทดลองในสภาพแวดล้อมแบบ 3 ช่องจราจรใน 3 แยก

WT_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3 S-3 W-3	122.02	22.48	17.37
E-3 S-3 N-3	95.74	21.09	17.86
E-3 N-3 W-3	118.49	22.38	16.90
S-3 N-3 W-3	113.16	21.20	17.51
WTL_{AVG} (s)			
สภาพแวดล้อม	Method		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3 S-3 W-3	5.06	9.63	4.80
E-3 S-3 N-3	5.05	8.70	4.69
E-3 N-3 W-3	5.02	9.11	4.71
S-3 N-3 W-3	5.00	9.21	4.67
WTC_{AVG} (s)			
สภาพแวดล้อม	วิธีการ		
	วิธีแบบดั้งเดิม	วิธี DQN แบบดั้งเดิม ร่วมกับกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
E-3 S-3 W-3	13.39	5.40	4.25
E-3 S-3 N-3	12.41	5.28	4.28
E-3 N-3 W-3	13.46	5.38	4.17
S-3 N-3 W-3	12.94	5.27	4.22

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

4.7 ระบบจัดลำดับความสำคัญรถพยาบาล

ผลการทดลองที่แสดงไว้ในหัวข้อนี้ มีค่าเพิ่มเติมต่อไปนี้จะถูกนำมาใช้เพื่อเปรียบเทียบผลลัพธ์

1. WTA_{Total} คือ ระยะเวลารอคอยสะสมของรถพยาบาลทั้งหมด

$$WTA_{Total} = \sum_{i=1}^A WTA[i] \quad (4.8)$$

โดยที่ A คือ จำนวนรถพยาบาล

WTA คือ เวลารอคอยสะสมของรถพยาบาล

ผลการทดลองในตารางที่ 4.6 แสดงข้อมูลระยะเวลาความปลอดภัยที่ 3 วินาทีและ 5 วินาที รวมถึงระยะการตรวจจับที่ 150 เมตร, 200 เมตร, 250 เมตร และ 300 เมตร

ตารางที่ 4.6 ผลการทดลองระบบจัดลำดับความสำคัญรถพยาบาลที่มีระยะเวลาความปลอดภัย และระยะการตรวจจับแตกต่างกัน

		WTA_{AVG} (s)			
ระยะเวลาความปลอดภัย (วินาที)		ระยะการตรวจจับ (เมตร)			
		150	200	250	300
3		12.87	12.98	13.14	13.53
5		14.17	13.23	13.48	13.92
		WTA_{Total} (s)			
ระยะเวลาความปลอดภัย (วินาที)		ระยะการตรวจจับ (เมตร)			
		150	200	250	300
3		0.24	0.24	0.00	0.00
5		2.96	1.20	0.20	0.16

ตัวชี้วัดสองตัวที่ถูกพิจารณาในการเลือกระยะเวลาความปลอดภัยและระยะตรวจจับที่เหมาะสมสำหรับระบบให้ความสำคัญของรถพยาบาล ได้แก่ เวลารอคอยทั้งหมดโดยเฉลี่ยและเวลารอคอยทั้งหมดของรถพยาบาล เมื่อวิเคราะห์เวลารอคอยทั้งหมดโดยเฉลี่ย พบว่าการใช้ระยะตรวจจับ 300 เมตร ให้ผลลัพธ์ที่มีค่าสูง ดังนั้นจึงถือว่าไม่เหมาะสมที่จะใช้ระยะตรวจจับ 300 เมตร และเมื่อวิเคราะห์เวลารอคอยทั้งหมดของรถพยาบาล ผลลัพธ์ที่มีระยะเวลาความปลอดภัย 3 วินาทีและระยะตรวจจับ 250 เมตร ให้ค่าที่เหมาะสมที่สุด ซึ่งบ่งชี้ว่าไม่มีเวลารอคอยสำหรับรถพยาบาล ดังนั้นเราจึงเลือกระยะเวลาความปลอดภัย 3 วินาทีและระยะตรวจจับ 250 เมตร สำหรับระบบให้

ความสำคัญของรถพยาบาล เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

จากผลการทดลองที่แสดงในตารางที่ 4.7 โดยใช้ระยะเวลาความปลอดภัย 3 วินาทีและระยะตรวจจับ 250 เมตร พบว่า เวลารอคอยรวมโดยเฉลี่ยนั้นมีค่าใกล้เคียงกันสำหรับวิธีที่นำเสนอ อย่างไรก็ตาม การใช้วิธี DQN แบบดั้งเดิมร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) มีเวลารอคอยรวมโดยเฉลี่ยสูงกว่าประมาณ 5.39% ซึ่งโดยรวมแล้ว วิธีที่นำเสนอให้ผลลัพธ์ที่ดีกว่า

ตารางที่ 4.7 ผลการทดลองเมื่อมีระบบจัดลำดับความสำคัญของรถพยาบาลและไม่มีระบบจัดลำดับความสำคัญของรถพยาบาล

มีระบบจัดลำดับความสำคัญของรถพยาบาล		
ตัวชี้วัด	วิธี DQN แบบดั้งเดิม ร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
WT_{AVG} (s)	17.01	13.14
WTA_{Total} (s)	0.04	0.00
ไม่มีระบบจัดลำดับความสำคัญของรถพยาบาล		
ตัวชี้วัด	วิธี DQN แบบดั้งเดิม ร่วมกับตรรกะแบบฟัซซี (Fuzzy Logic) [29]	วิธีที่นำเสนอ
WT_{AVG} (s)	16.14	13.16
WTA_{Total} (s)	65.24	67.60

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 5

สรุป ข้อจำกัดและข้อเสนอแนะ

5.1 บทสรุป

งานวิจัยนี้มีวัตถุประสงค์เพื่อพัฒนาและประเมินประสิทธิภาพของวิธีการควบคุมสัญญาณไฟจราจรที่ผสมผสานระหว่างการเรียนรู้แบบเสริมกำลังเชิงลึก (Deep Reinforcement Learning : DRL) กับตรรกะแบบฟัซซี (Fuzzy Logic) โดยมุ่งหวังเพื่อลดความแออัดของการจราจร ลดเวลาในการรอคอย และเพิ่มประสิทธิภาพของวิธีการควบคุมสัญญาณไฟจราจรจากงานวิจัยก่อนหน้านี้

ระบบที่นำเสนอในงานวิจัยนี้มีการปรับปรุงโครงสร้างของข้อมูลสถานะ (State) ให้สะท้อนข้อมูลเชิงเวลาการรอคอยของรถในแต่ละช่องจราจรได้อย่างละเอียด โดยจำแนกระยะเวลาออกเป็นระดับต่าง ๆ ทั้งหมด 10 ระดับ และพิจารณาความยาวของรถที่แตกต่างกันในแต่ละประเภท ซึ่งส่งผลต่อการเก็บค่าของสถานะ นอกจากนี้ ยังได้ออกแบบ ฟังก์ชันรางวัล (Reward Function) ให้สะท้อนความสมดุลระหว่างการไหลของจราจรในทุกทิศทาง โดยเลือกใช้เวลารอสะสมสูงสุดในแต่ละกลุ่มช่องจราจรเป็นเกณฑ์ในการคำนวณระยะเวลาการรอคอยสะสมทั้งหมดของระบบ ซึ่งช่วยปรับปรุงระบบจากงานวิจัยก่อนหน้านี้ ที่มักให้ความสำคัญกับช่องทางตรงมากกว่าช่องทางเลี้ยวให้มีความสมดุลมากยิ่งขึ้น

การทดลองได้มีการออกแบบและจำลองระบบควบคุมไฟจราจรโดยใช้ซอฟต์แวร์ Simulation of Urban Mobility (SUMO) ซึ่งเป็นซอฟต์แวร์โอเพ่นซอร์สที่ได้รับความนิยมสูงในการวิจัยด้านการจำลองการจราจร เนื่องจากสามารถรองรับการสร้างแบบจำลองระดับจุลภาค ที่คำนึงถึงพฤติกรรมของยานพาหนะแต่ละคันอย่างละเอียด โดยมีการจำลองพฤติกรรมการเล่น การเร่ง และการเบรกของยานพาหนะในแต่ละช่วงเวลา ผ่านการใช้โมเดล Car-Following และ Lane-Changing เพื่อสะท้อนการจราจรในสถานการณ์จริงให้มากที่สุด

การทดลองได้ทำการฝึกฝนโมเดลในสภาพแวดล้อมแบบที่มีลักษณะเป็นสี่แยกแบบแยกเดี่ยว (Isolated Intersection) ที่มีจำนวนช่องจราจรสมมาตรกันในทุกทิศทาง และนำโมเดลที่ได้ไปทดสอบในสถานการณ์ต่าง ๆ ที่จำลองผ่าน SUMO เช่น สี่แยกที่มีจำนวนช่องจราจรไม่สมมาตร (3 ช่องจราจรในบางทิศทาง) และการจำลองการจราจรแบบขับชิดซ้าย (Left-Hand Traffic) รวมถึงการจำลองสถานการณ์ฉุกเฉินที่มีการใช้ระบบกำหนดลำดับความสำคัญให้กับรถพยาบาล โดยในแต่ละการทดลอง ได้สร้างจำนวนรถที่หลากหลาย ได้แก่ รถยนต์ส่วนบุคคล รถมินิบัส รถบัส และรถพยาบาล ที่มีขนาดและความเร็วสูงสุดที่แตกต่างกัน พบว่าระบบควบคุมสัญญาณไฟจราจรที่นำเสนอแสดงให้เห็นถึงประสิทธิภาพที่ดีในทุกสถานการณ์ที่ทดสอบในขอบเขตของการจำลอง และมีความยืดหยุ่นในการประยุกต์ใช้งานกับสี่แยกที่มีโครงสร้างช่องจราจรไม่สมมาตร

โดยผลการทดลองแสดงให้เห็นว่า ระบบที่นำเสนอสามารถลดค่าเฉลี่ยของเวลารอสะสมทั้งหมด (WT_{AVG}), เวลารอของรถที่เลี้ยวซ้าย (WTL_{AVG}), และเวลารอเฉลี่ยต่อคัน (WTC_{AVG}) ได้

ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้คัดลอกเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

18.46%, 40.36%, และ 21.23% ตามลำดับ เมื่อเปรียบเทียบกับงานวิจัยก่อนหน้านี้ที่เป็นระบบต้นแบบเดิมในสภาพแวดล้อมต้นแบบ

นอกเหนือจากสภาพแวดล้อมต้นแบบที่มีจำนวนช่องจราจรเท่ากันในทุกทิศทางแล้ว ระบบที่นำเสนอยังได้มีการทดสอบในสภาพแวดล้อมที่มีจำนวนช่องจราจรไม่เท่ากันในบางทิศทาง และสภาพแวดล้อมจำลองที่มีการซับซ้อนแบบซิดซายซึ่งเป็นลักษณะจราจรในประเทศไทย

ผลการทดลองในสภาพแวดล้อมที่มีจำนวนช่องจราจรไม่เท่ากันในบางทิศทาง พบว่า ระบบสามารถลดค่าเฉลี่ยของเวลารอสะสมทั้งหมด (WT_{AVG}), เวลารอของรถที่เลี้ยวซ้าย (WTL_{AVG}), และเวลารอเฉลี่ยต่อคัน (WTC_{AVG}) ได้ดี เมื่อเปรียบเทียบกับวิธีดั้งเดิมและวิธีของงานวิจัยก่อนหน้านี้ที่เป็นระบบต้นแบบเดิม โดยสามารถลดค่าเฉลี่ยของเวลารอสะสมทั้งหมด (WT_{AVG}) ได้โดยเฉลี่ย 70-80% เมื่อเทียบกับวิธีดั้งเดิม และลดลงได้โดยเฉลี่ย 19-20% เมื่อเทียบกับวิธีของงานวิจัยก่อนหน้านี้ที่เป็นระบบต้นแบบเดิม และสามารถแก้ปัญหาของรถที่เลี้ยวซ้ายได้อย่างชัดเจน โดยสามารถลดเวลารอของรถที่เลี้ยวซ้าย ได้เฉลี่ยกว่า 39-40%

ผลการทดลองในสภาพแวดล้อมแบบซับซ้อนซิดซายซึ่งเป็นลักษณะจราจรในประเทศไทย พบว่า สามารถลดค่าเฉลี่ยของเวลารอสะสมทั้งหมด (WT_{AVG}) ได้ถึง 71.21% เมื่อเทียบกับวิธีดั้งเดิม และลดลงได้ 22.87% เมื่อเทียบกับวิธีของงานวิจัยก่อนหน้านี้ที่เป็นระบบต้นแบบเดิม และขณะเดียวกันยังสามารถลดเวลารอของรถที่เลี้ยวขวา (WTR_{AVG}) ได้อย่างชัดเจน

โดยภาพรวม จากผลการทดลองในแต่ละสภาพแวดล้อม พบว่าระบบที่นำเสนอสามารถลดค่าเฉลี่ยของเวลารอสะสมทั้งหมด ได้ประมาณ 18-22% และลดเวลารอของรถในช่องทางเลี้ยวได้มากถึง ประมาณ 40% เมื่อเปรียบเทียบกับวิธีต้นแบบ โดยผลลัพธ์ในแต่ละสภาพแวดล้อมมีแนวโน้มใกล้เคียงกัน สะท้อนถึงความสามารถของระบบในการทำงานได้อย่างสม่ำเสมอ แม้ในสภาพแวดล้อมที่มีความแตกต่างด้านโครงสร้าง

ทั้งนี้ โมเดลที่ใช้ในการทดลองได้รับการฝึกฝนจากสภาพแวดล้อมแบบสมมาตร (ที่มีจำนวนช่องจราจรเท่ากันทุกทิศทาง) เพียงครั้งเดียว ก่อนนำไปประยุกต์ใช้กับสภาพแวดล้อมอื่น ๆ ที่มีโครงสร้างไม่สมมาตร โดยไม่มีการฝึกฝนเพิ่มเติมในสภาพแวดล้อมที่ต่างออกไป ซึ่งผลลัพธ์ที่ได้ยังคงแสดงถึงประสิทธิภาพที่ดีในแต่ละสภาพแวดล้อม

นอกจากนี้ ในระบบกำหนดลำดับความสำคัญของรถพยาบาล หากกำหนดค่าระยะเวลาความปลอดภัย (Safety Duration) ที่ 3 วินาที และ ระยะตรวจจับ (Detection Distance) ที่ 250 เมตร ระบบสามารถลดเวลารอสะสมของรถพยาบาลลงเป็นศูนย์ ขณะที่เวลารอสะสมโดยเฉลี่ยของยานพาหนะทั่วไปอาจมีการเปลี่ยนแปลงเพียงเล็กน้อย ซึ่งอยู่ในระดับที่ไม่ส่งผลกระทบต่อประสิทธิภาพโดยรวมของระบบจราจรมากจนเกินไป แสดงให้เห็นถึงความสามารถของระบบในการบริหารจัดการลำดับความสำคัญ โดยคงไว้ซึ่งความสมดุลระหว่างความเร่งด่วนและความต่อเนื่องของการจราจร

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

และในส่วนของเวลาที่ใช้ในการจำลองภายใต้สภาพแวดล้อมที่แตกต่างกัน เช่น การเปลี่ยนจำนวนช่องจราจรของแต่ละทิศทาง ยังอยู่ในช่วงที่ใกล้เคียงกันทั้งหมด ซึ่งสามารถตีความได้ว่าการเปลี่ยนแปลงโครงสร้างของแยกจราจรที่ใช้ในการทดลอง ไม่ได้ส่งผลกระทบต่อระยะเวลาที่ใช้ในการประมวลผลของโมเดล อย่างไรก็ตาม ในกรณีที่จำนวนยานพาหนะในระบบเพิ่มขึ้น อาจทำให้เวลาที่ใช้ในการจำลองมีแนวโน้มเพิ่มขึ้นตามไปด้วย

สำหรับการนำระบบที่นำเสนอไปประยุกต์ใช้งานจริง ระบบนี้เหมาะสมกับการใช้งานในพื้นที่หรือเมืองขนาดกลางถึงขนาดเล็กที่มีเพียงสี่แยกสำคัญไม่กี่จุด โดยระบบสามารถตัดสินใจได้อย่างเป็นอิสระโดยไม่ต้องเชื่อมโยงกับทางแยกอื่น

5.2 ข้อจำกัด

แม้ว่าระบบควบคุมสัญญาณไฟจราจรที่นำเสนอจะแสดงให้เห็นถึงประสิทธิภาพที่ดีในสภาพแวดล้อมทั้งหมดที่ทดสอบในขอบเขตของการจำลอง อย่างไรก็ตาม การนำระบบไปใช้งานจริงยังคงต้องอยู่ภายใต้เงื่อนไข ดังนั้น จึงยังมีข้อจำกัดบางประการที่ควรได้รับการพิจารณาเพิ่มเติม ดังนี้

1. งานวิจัยนี้มุ่งเน้นไปการควบคุมสัญญาณไฟจราจรเฉพาะในสี่แยกเดี่ยวที่ไม่มีการเชื่อมต่อกับแยกอื่น ซึ่งยังไม่ได้ครอบคลุมถึงสภาพการจราจรในเมืองใหญ่ที่มีประกอบด้วยแยกจำนวนมากซึ่งมีผลกระทบซึ่งกันและกันอย่างซับซ้อน

2. แม้ว่าการใช้ซอฟต์แวร์ SUMO จะสามารถสร้างสภาพแวดล้อมจำลองที่ใกล้เคียงกับความเป็นจริงได้ในระดับหนึ่ง แต่ยังมีข้อจำกัดที่ผู้ใช้ในสถานการณ์จริงได้ไม่ครบถ้วน เช่น การตัดสินใจในเหตุการณ์ฉุกเฉิน การจอดรถซ้อนคัน หรือการเกิดอุบัติเหตุ ซึ่งล้วนเป็นปัจจัยที่อาจส่งผลกระทบต่อประสิทธิภาพของระบบควบคุมเมื่อใช้งานในสภาพแวดล้อมจริง

3. การฝึกโมเดลด้วยยานพาหนะ 1,000 คันใน 90 นาที อาจไม่รองรับสถานการณ์ที่มีความหนาแน่นของจราจรในระดับสูงมาก เช่น ช่วงเวลาเร่งด่วนในเขตเมืองขนาดใหญ่ที่มีปริมาณรถหลายหมื่นคันต่อชั่วโมง อาจจำเป็นต้องทำการฝึกฝนโมเดลใหม่ เพื่อให้สามารถตอบสนองต่อสถานการณ์ที่ไม่เคยพบมาก่อนได้อย่างเหมาะสม

4. แม้ว่าการใช้ตรรกะแบบฟัซซีจะช่วยให้การตัดสินใจของระบบมีความยืดหยุ่น และสามารถคำนวณระยะเวลาในการเปิดสัญญาณไฟเขียวได้อย่างเหมาะสมตามระดับความหนาแน่นของการจราจรในแต่ละทิศทาง แต่การออกแบบกฎและฟังก์ชันสมาชิก ยังคงอิงจากการกำหนดล่วงหน้าโดยมนุษย์ตามวิจารณ์ของผู้พัฒนา ซึ่งจำกัดความสามารถในการปรับตัว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่าจะกรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

5.3 ข้อเสนอแนะ

1. เพื่อให้ระบบสามารถจัดการการจราจรในระดับเมืองขนาดใหญ่ที่มีโครงสร้างซับซ้อนได้อย่างมีประสิทธิภาพ อาจพิจารณาพัฒนาระบบจากการควบคุมสี่แยกเดี่ยวไปสู่การควบคุมแบบเครือข่ายหลายแยก โดยใช้แนวทางระบบหลายเอเจนต์ที่สามารถแลกเปลี่ยนข้อมูลและตัดสินใจร่วมกันได้

2. แม้ระบบจำลองจะมีความสมจริงในระดับหนึ่ง แต่การนำไปใช้งานจริงยังต้องพิจารณาจากปัจจัยภายนอกเพิ่มเติม ดังนั้น จึงควรมีการทดสอบระบบร่วมกับข้อมูลจากพื้นที่จริง

3. เพื่อให้ระบบสามารถใช้งานได้อย่างมีประสิทธิภาพในสภาพแวดล้อมที่แตกต่างจากข้อมูลที่ใช้ฝึกฝนอาจพิจารณาออกแบบให้ระบบสามารถเรียนรู้และปรับปรุงตัวเองได้อย่างต่อเนื่องจากข้อมูลใหม่ที่ได้รับในระหว่างการใช้งานจริง โดยที่ไม่ต้องทำการฝึกฝนโมเดลใหม่ทั้งหมด ซึ่งจะช่วยให้โมเดลสามารถตอบสนองต่อสภาพการจราจรที่หลากหลายและเปลี่ยนแปลงอยู่ตลอดเวลาได้ดียิ่งขึ้น

4. เพื่อให้ตรรกะแบบฟัซซีสามารถเรียนรู้และปรับกฎได้อัตโนมัติ อาจพิจารณาการใช้ระบบอนุมานฟัซซีแบบปรับตัวได้ ซึ่งเป็นการผสมผสานระหว่างตรรกะแบบฟัซซีกับโมเดลประสาทเทียม อย่างไรก็ตาม แนวทางดังกล่าวอาจทำให้ระบบมีความซับซ้อนและต้องใช้ทรัพยากรในการประมวลผลมากขึ้น จึงควรพิจารณาความเหมาะสมตามบริบทของการใช้งานจริง

บรรณานุกรม

- [1] INRIX. "INRIX 2024 Global Traffic Scorecard." [Online]. Available : <https://inrix.com/press-releases/2024-global-traffic-scorecard-uk>. 2024.
- [2] Wang, C. "The Relationship between Traffic Congestion and Road Accidents: An Econometric Approach Using GIS." Ph.D. Thesis, Loughborough University, Leicestershire, UK, 2010.
- [3] C. P. Pappis and E. H. Mamdani, "A fuzzy logic controller for a traffic junction," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7, no. 10, 1977, pp. 707-717.
- [4] H. Taskin and R. Gumustas, "Simulation of traffic flow system and control using fuzzy logic," *Proc. 12th IEEE International Symposium on Intelligent Control*, 1997, pp. 325-330.
- [5] H.-H. Liu and P.-L. Hsu, "Design and simulation of adaptive fuzzy control on the traffic network," *Proc. 2006 SICE-ICASE International Joint Conference*, 2006, pp. 4961-4966.
- [6] G. H. Kulkarni and P. G. Waingankar, "Fuzzy logic based traffic light controller," *Proc. 2007 International Conference on Industrial and Information Systems*, 2007, pp. 107-110.
- [7] L. Wu, X. Zhang, and Z. Shi, "An intelligent fuzzy control for crossroads traffic light," *Proc. 2010 Second WRI Global Congress on Intelligent Systems*, 2010, vol. 3, pp. 28-32.
- [8] M. Firdous, F. U. D. Iqbal, N. Ghafoor, N. K. Qureshi, and N. Naseer, "Traffic light control system for four-way intersection and T-crossing using fuzzy logic," *Proc. 2019 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, 2019, pp. 178-182.
- [9] I. Tunc, A. Y. Yesilyurt, and M. T. Soylemez, "Intelligent traffic light control system simulation for different strategies with fuzzy logic controller," *Proc. 2019 11th International Conference on Electrical and Electronics Engineering (ELECO)*, 2019, pp. 830-834.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [10] S. Prontri, P. Wuttidittachotti, and S. Thajchayapong, "Traffic signal control using fuzzy logic," **Proc. 2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)**, 2015, pp. 1-6.
- [11] Y. Bi, J. Li, and X. Lu, "Single intersection signal control and simulation based on fuzzy logic," **Proc. 2011 Third International Conference on Intelligent Human-Machine Systems and Cybernetics**, 2011, vol. 2, pp. 87-89.
- [12] C. Sun, "Fundamental Q-learning algorithm in finding optimal policy," **Proc. 2017 International Conference on Smart Grid and Electrical Automation (ICSGEA)**, 2017, pp. 243-246.
- [13] D. Pandey and P. Pandey, "Approximate q-learning: An introduction," **Proc. 2010 second international conference on machine learning and computing**, 2010, pp. 317-320.
- [14] A. R. Rosyadi, T. A. B. Wirayuda, and S. Al-Faraby, "Intelligent traffic light control using collaborative Q-Learning algorithms," **Proc. 2016 4th International Conference on Information and Communication Technology (ICoICT)**, 2016, pp. 1-6.
- [15] Y. Liao and X. Cheng, "Study on traffic signal control based on q-learning," **Proc. 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery**, 2009, vol. 3, pp. 581-585.
- [16] Y. Liu, L. Liu, and W.-P. Chen, "Intelligent traffic light control using distributed multi-agent Q learning," **Proc. 2017 IEEE 20th international conference on intelligent transportation systems (ITSC)**, 2017, pp. 1-8.
- [17] B.-L. Ye, P. Wu, W. Wu, L. Li, Y. Zhu, and B. Chen, "Q-learning based traffic signal control method for an isolated intersection," **Proc. 2022 China Automation Congress (CAC)**, 2022, pp. 6063-6068.
- [18] M. A. M. Yusop, H. Mansor, T. S. Gunawan, and H. Nasir, "Intelligent Traffic Lights using Q-learning," **Proc. 2022 IEEE 8th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)**, 2022, pp. 200-204.
- [19] N. Sutisna *et al.*, "Deep Q-Network Model for Intelligent Traffic Light," **Proc. 2022 International Symposium on Electronics and Smart Devices (ISESD)**, 2022, pp. 1-5.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [20] A. Vidali, L. Crociani, G. Vizzari, and S. Bandini, "A Deep Reinforcement Learning Approach to Adaptive Traffic Lights Management," **Proc. Woa**, 2019, pp. 42-50.
- [21] T. Wu *et al.*, "Multi-agent deep reinforcement learning for urban traffic light control in vehicular networks," **IEEE Transactions on Vehicular Technology**, vol. 69, no. 8, 2020, pp. 8243-8256.
- [22] A. Abhishek, P. Nayak, K. P. Hegde, A. L. Prasad, and K. Nagegowda, "Smart traffic light controller using deep reinforcement learning," **Proc. 2022 3rd International Conference for Emerging Technology (INCET)**, 2022, pp. 1-5.
- [23] S. Zhancheng, "Research on application of deep reinforcement learning in traffic signal control," **Proc. 2021 6th International Conference on Frontiers of Signal Processing (ICFSP)**, 2021, pp. 17-21.
- [24] A. Tigga, L. Hota, S. Patel, and A. Kumar, "A deep Q-learning-based adaptive traffic light control system for urban safety," **Proc. 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)**, 2022, pp. 2430-2435.
- [25] N. Kodama, T. Harada, and K. Miyazaki, "Traffic signal control system using deep reinforcement learning with emphasis on reinforcing successful experiences," **IEEE Access**, vol. 10, 2022, pp. 128943-128950.
- [26] J. Yang, P. Wang, and Y. Ju, "Variable Speed Limit Intelligent Decision-Making Control Strategy Based on Deep Reinforcement Learning under Emergencies," **Sustainability**, vol. 16, no. 3, 2024, p. 965.
- [27] H. Jiang, H. Zhang, Z. Feng, J. Zhang, Y. Qian, and B. Wang, "A Multi-Objective Optimal Control Method for Navigating Connected and Automated Vehicles at Signalized Intersections Based on Reinforcement Learning," **Applied Sciences**, vol. 14, no. 7, 2024, p. 3124.
- [28] D. Tagesson, "a Comparison Between Deep Q-Learning and Deep Deterministic Policy Gradient for an Autonomous Drone in a Simulated Environment," ed, 2021.
- [29] I. Tunc and M. T. Soylemez, "Fuzzy logic and deep Q learning based control for traffic lights," **Alexandria Engineering Journal**, vol. 67, 2023, pp. 343-359.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [30] K.-L. A. Yau, Y.-W. Chong, X. Fan, C. Wu, Y. Saleem, and P.-C. Lim, "Reinforcement learning models and algorithms for diabetes management," **IEEE Access**, vol. 11, 2023, pp. 28391-28415.
- [31] L. Prashanth and S. Bhatnagar, "Reinforcement learning with function approximation for traffic signal control," **IEEE Transactions on Intelligent Transportation Systems**, vol. 12, no. 2, 2010, pp. 412-421.
- [32] R. Singh, A. Gupta, and N. B. Shroff, "Learning in constrained Markov decision processes," **IEEE Transactions on Control of Network Systems**, vol. 10, no. 1, 2022, pp. 441-453.
- [33] Z. Zhu, K. Lin, A. K. Jain, and J. Zhou, "Transfer learning in deep reinforcement learning: A survey," **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2023.
- [34] Y. Luo, T. Gang, and L. Chen, "Research on target defense strategy based on deep reinforcement learning," **IEEE Access**, vol. 10, 2022, pp. 82329-82335.
- [35] I. Oh, S. Rho, S. Moon, S. Son, H. Lee, and J. Chung, "Creating pro-level AI for a real-time fighting game using deep reinforcement learning," **IEEE Transactions on Games**, vol. 14, no. 2, 2021, pp. 212-220.
- [36] A. N. Kumar and S. Kochuvila, "Mobile service robot path planning using deep reinforcement learning," **IEEE Access**, 2023.
- [37] Y. Ansari *et al.*, "A deep reinforcement learning-based decision support system for automated stock market trading," **IEEE Access**, vol. 10, 2022, pp. 127469-127501.
- [38] S. Yin *et al.*, "Left Ventricle Contouring in Cardiac Images in the Internet of Medical Things via Deep Reinforcement Learning," **IEEE Internet of Things Journal**, vol. 10, no. 20, 2023, pp. 17705-17717.
- [39] R. Lu, Z. Jiang, H. Wu, Y. Ding, D. Wang, and H.-T. Zhang, "Reward shaping-based actor-critic deep reinforcement learning for residential energy management," **IEEE Transactions on Industrial Informatics**, vol. 19, no. 3, 2022, pp. 2662-2673.
- [40] X. Liang, X. Du, G. Wang, and Z. Han, "A deep reinforcement learning network for traffic light cycle control," **IEEE Transactions on Vehicular Technology**, vol. 68, no. 2, 2019, pp. 1243-1253.

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

- [41] A. Haydari and Y. Yilmaz, "Deep reinforcement learning for intelligent transportation systems: A survey," **IEEE Transactions on Intelligent Transportation Systems**, vol. 23, no. 1, 2020, pp. 11-32.
- [42] R. Sarikhani and F. Keynia, "Cooperative spectrum sensing meets machine learning: Deep reinforcement learning approach," **IEEE Communications Letters**, vol. 24, no. 7, 2020, pp. 1459-1462.
- [43] B. Mughal, Z. M. Fadlullah, M. M. Fouda, and S. Ikki, "Optimizing packet forwarding performance in multiband relay networks via customized reinforcement learning," **IEEE Open Journal of the Communications Society**, vol. 3, 2022, pp. 973-985.
- [44] M. E. M. Ali, A. Durdu, S. A. Çeltek, and A. Yilmaz, "An adaptive method for traffic signal control based on fuzzy logic with webster and modified webster formula using SUMO traffic simulator," **IEEE Access**, vol. 9, 2021, pp. 102985-102997.
- [45] N. Khan, D. A. Elizondo, L. Deka, and M. A. Molina-Cabello, "Fuzzy logic applied to system monitors," **IEEE Access**, vol. 9, 2021, pp. 56523-56538.
- [46] B. Lju and Z. Ding, "A distributed deep reinforcement learning method for traffic light control," **Neurocomputing**, vol. 490, 2022, pp. 390-399.

ประวัติผู้เขียน

ชื่อ-นามสกุล นายธีรพันธ์ มีโชคกิจ
วัน เดือน ปีเกิด 7 มีนาคม 2543
ที่อยู่ 25,27 ซอยเทียนทะเล 19 แยก 1 แขวงท่าข้าม เขตบางขุนเทียน
กรุงเทพมหานคร 10150

ประวัติการศึกษา:

2565

สำเร็จการศึกษา วิศวกรรมศาสตรบัณฑิต
สาขาวิชาวิศวกรรมอิเล็กทรอนิกส์ (เกียรตินิยมอันดับ 1 เหรียญทอง)
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า
ไม่ว่ากรณีใดๆทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้