การพัฒนาขั้นตอนวิธีการยิงประตูระยะไกล โดยวิธีการเรียนรู้แบบคิว สำหรับการแข่งขันฟุตบอลแบบ 2 มิติ

วิทยานิพนธ์

ของ

วรัญญู เกียรติจินดารัตน์

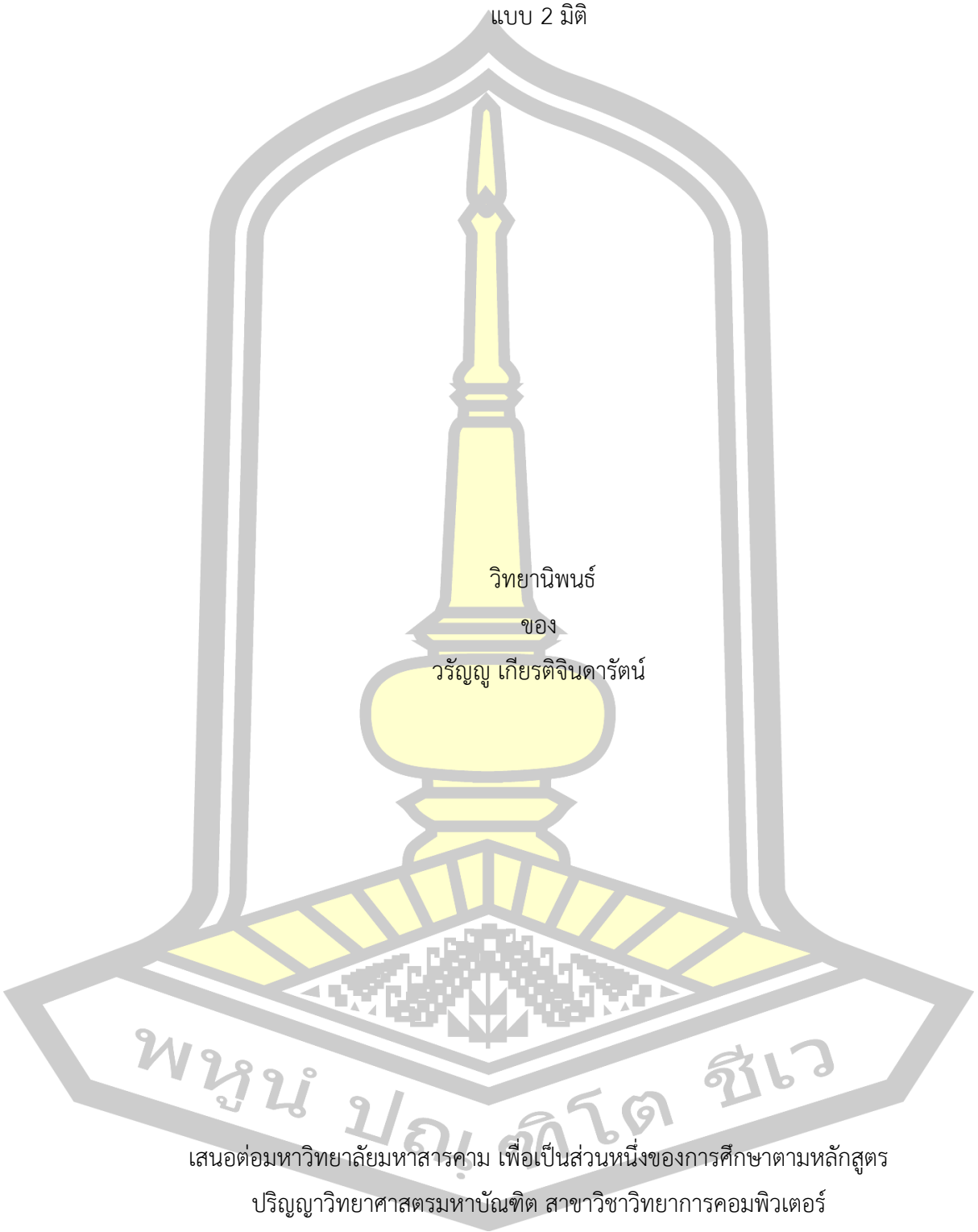
เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

มีนาคม 2562

สงวนลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

การพัฒนาขั้นตอนวิธีการยิงประตูระยะไกล โดยวิธีการเรียนรู้แบบคิว สำหรับการแข่งขันฟุตบอลแบบ 2 มิติ



วิทยานิพนธ์

ของ

วรัญญู เกียรติจินดารัตน์

เสนอต่อมหาวิทยาลัยมหาสารคาม เพื่อเป็นส่วนหนึ่งของการศึกษาตามหลักสูตร

ปริญญาวิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์

มีนาคม 2562

สงวนลิขสิทธิ์เป็นของมหาวิทยาลัยมหาสารคาม

The Development of Long Shooting Algorithm using Q-learning for 2D Soccer Games

Waranyoo Kiatjindarat

A Thesis Submitted in Partial Fulfillment of Requirements

for Master of Science (Computer Science)

March 2019

Copyright of Mahasarakham University

คณะกรรมการสอบวิทยานิพนธ์ ได้พิจารณาวิทยานิพนธ์ของนายวรัญญู เกียรติจินดารัตน์ แล้วเห็นสมควรรับเป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ของมหาวิทยาลัยมหาสารคาม

คณะกรรมการสอบวิทยานิพนธ์

.....ประธานกรรมการ
(ผศ. ดร. วรารัตน์ สงฆ์แป้น)

.....อาจารย์ที่ปรึกษาวิทยานิพนธ์หลัก
(ผศ. ดร. ฉัตรเกล้า เจริญผล)

.....กรรมการ
(ผศ. ดร. พัฒนพงษ์ ชมภูวิเศษ)

.....กรรมการ
(ผศ. ดร. พนิดา ทรงรัมย์)

มหาวิทยาลัยอนุมัติให้รับวิทยานิพนธ์ฉบับนี้ เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญา วิทยาศาสตรมหาบัณฑิต สาขาวิชาวิทยาการคอมพิวเตอร์ ของมหาวิทยาลัยมหาสารคาม

.....
(ผศ. ศศิธร แก้วมั่น)
คณบดีคณะวิทยาการสารสนเทศ

.....
(ผศ. ดร. กริสน์ ชัยมูล)
คณบดีบัณฑิตวิทยาลัย

พหุน ปณฺฑิโต ธีโร

ชื่อเรื่อง การพัฒนาขั้นตอนวิธีการยิงประตูระยะไกล โดยวิธีการเรียนรู้แบบคิว สำหรับการแข่งขันฟุตบอล แบบ 2 มิติ

ผู้วิจัย วรัญญู เกียรติจินดารัตน์

อาจารย์ที่ปรึกษา ผู้ช่วยศาสตราจารย์ ดร. ฉัตรเกล้า เจริญผล

ปริญญา วิทยาศาสตรมหาบัณฑิต **สาขาวิชา** วิทยาการคอมพิวเตอร์

มหาวิทยาลัย มหาวิทยาลัยมหาสารคาม **ปีที่พิมพ์** 2562

บทคัดย่อ

การแข่งขันหุ่นยนต์ระดับนานาชาติ หรือ RoboCup เป็นรายการแข่งขันเพื่อส่งเสริมงานวิจัยที่เกี่ยวข้องกับหุ่นยนต์และปัญญาประดิษฐ์ ซึ่งฟุตบอลเป็นหนึ่งในหัวข้อสำคัญในรายการนี้ การแข่งขันรายการ RoboCup 2D soccer simulation league จะตัดสินชนะหรือแพ้จากการยิงประตูของแต่ละฝ่าย ซึ่งในปัจจุบันการยิงประตูของหุ่นยนต์ต้องเลี้ยงลูกฟุตบอลเข้าไปยิงในกรอบเขตโทษเพียงเท่านั้น ซึ่งทำให้ไม่มีความหลากหลายในการยิงประตู จากปัญหาดังกล่าวงานวิจัยนี้จึงปรับปรุงขั้นตอนวิธีการยิงประตูจากระยะไกล เพื่อเพิ่มโอกาสในการยิงประตู โดยใช้อัลกอริธึมการเรียนรู้แบบเสริมกำลัง แบบ Q-learning การดำเนินการวิจัยจะแบ่งออกเป็น 4 ขั้นตอน ได้แก่ 1) การเก็บรวบรวมข้อมูล 2) การเลือกข้อมูล 3) การเรียนรู้แบบเสริมกำลัง แบบ Q-learning และ 4) การวัดประสิทธิภาพ โดยการทดลองจะแบ่งออกเป็น 2 การทดลอง ได้แก่ 1) การออกแบบการกระทำ (เฉพาะการกระทำ) และ 2) การออกแบบการกระทำและทิศทาง (การกระทำและทิศทาง) ซึ่งผลการทดลองของการทดลองที่ 1 พบว่ามีการเพิ่มขึ้นของการตัดสินใจยิงประตู ซึ่งคิดเป็น 80% เป็นการตัดสินใจยิงประตูที่เข้ากรอบประตูเพียง 20% เนื่องจากเป็นการกระทำที่ตัดสินใจทำโดยไม่รู้ทิศทาง จึงเป็นการตัดสินใจยิงประตูที่ไม่แน่นอน จึงต้องนำทิศทางของการกระทำมาตัดสินใจด้วย และผลการทดลองของการทดลองที่ 2 พบว่ามีการเพิ่มขึ้นของการตัดสินใจยิงประตู ซึ่งคิดเป็น 28% นอกจากนั้นการตัดสินใจยิงประตูทั้ง 28% ล้วนเป็นการตัดสินใจยิงประตูที่เข้ากรอบประตูทั้งหมด

คำสำคัญ : การเรียนรู้แบบเสริมกำลัง, การเรียนรู้แบบคิว, การจำลองการแข่งขันหุ่นยนต์โรโบคัพแบบ 2 มิติ, การยิงประตูจากระยะไกล

TITLE The Development of Long Shooting Algorithm using Q-learning for 2D Soccer Games

AUTHOR Waranyoo Kiatjindarat

ADVISORS Assistant Professor Chatklaw Jareanpon , Ph.D.

DEGREE Master of Science **MAJOR** Computer Science

UNIVERSITY Mahasarakham University **YEAR** 2019

ABSTRACT

RoboCup is the tournament for developing the robotics and artificial intelligence. Soccer is the most interested topic in RoboCup. RoboCup 2D soccer simulation league, the number of scoring decide the winner. Nowadays, the scoring of each agent must dribble into the penalty area, which do not the variety of shooting. This research improved long shooting for increasing the chance of the score by the Q-learning algorithm. The methodology of this research is divided into 4 processes: 1) Data collection 2) Feature selection 3) Q-learning algorithm and 4) Evaluation; moreover, the experimental has 2 experiments: 1) Design only the action and 2) Design the action and direction. The experimental result of the design only action increases the chance of score as 80% and 20%, that shooting directs into the goal because the decision of each agent does not know a direction, that is an unstable shooting. The experimental result of the design action and direction increases the chance of score as 28% and all shooting direct into the goal.

Keyword : Reinforcement learning, Q-learning, RoboCup 2D soccer simulation, Long shooting

กิตติกรรมประกาศ

วิทยานิพนธ์ฉบับนี้สำเร็จสมบูรณ์ได้ด้วยความกรุณาและความช่วยเหลืออย่างสูงยิ่งจาก ผู้ช่วยศาสตราจารย์ ดร.ฉัตรเกล้า เจริญผล ประธานกรรมการควบคุมวิทยานิพนธ์ ผู้ช่วยศาสตราจารย์ ดร.วรารัตน์ สงฆ์แป้น ประธานกรรมการสอบ ผู้ช่วยศาสตราจารย์ ดร.พัฒนพงษ์ ชมภูวิเศษ และผู้ช่วยศาสตราจารย์ ดร.พนิดา ทรงรัมย์ กรรมการสอบ

ขอขอบพระคุณบิดา มารดา และผู้เกี่ยวข้องที่เป็นกำลังใจ และให้ทุนสนับสนุนในการเรียน ซึ่งเป็นผลให้วิทยานิพนธ์นี้สำเร็จลุล่วงไปได้ด้วยดี

วรัญญู เกียรติจินดารัตน์

พหูน ปณฺฑิโต ชีเว

สารบัญ

หน้า

บทคัดย่อภาษาไทย ..... ง

บทคัดย่อภาษาอังกฤษ ..... จ

กิตติกรรมประกาศ ..... ฉ

สารบัญ ..... ช

สารบัญตาราง ..... ญ

สารบัญรูปภาพ ..... ฏ

บทที่ 1 บทนำ 1

1.1 หลักการและเหตุผล 1

1.2 วัตถุประสงค์ของการวิจัย 2

1.3 ความสำคัญของการวิจัย 2

1.4 ขอบเขตของการวิจัย 2

1.5 นิยามศัพท์เฉพาะ 2

บทที่ 2 ทฤษฎีและงานวิจัยที่เกี่ยวข้อง 3

2.1 ฟุตบอล 3

2.2 การแข่งขันหุ่นยนต์โรโบคัพ (RoboCup) และซอฟต์แวร์สำหรับการแข่งขันหุ่นยนต์โรโบคัพ 4

2.2.1 การแข่งขันหุ่นยนต์โรโบคัพ (RoboCup) 4

2.2.2 การติดตั้งซอฟต์แวร์สำหรับการแข่งขันหุ่นยนต์โรโบคัพ 4

2.2.2.1 การติดตั้งซอฟต์แวร์พื้นฐานสำหรับซอฟต์แวร์ rcssserver และ soccerwindow2 4

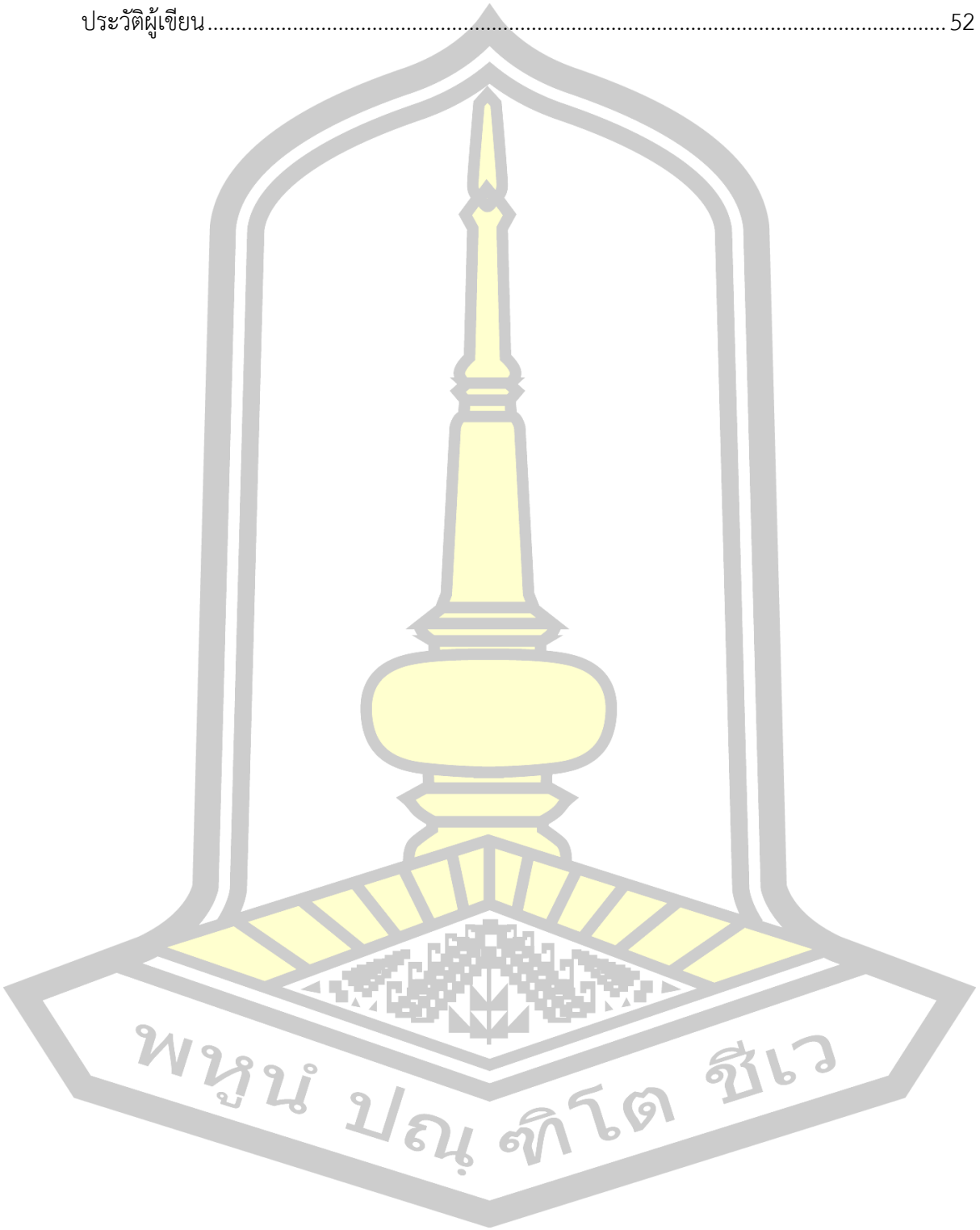
2.2.2.2 การติดตั้งซอฟต์แวร์ rcssserver และ soccerwindow2 5

2.2.3 เครื่องแม่ข่ายการแข่งขันหุ่นยนต์โรโบคัพ (RoboCup soccer server) 6

2.2.3.1 ระยะการมองเห็นของเอเจนต์ 6
2.2.3.2 การคำนวณระยะทางและทิศทางระหว่างเอเจนต์กับตำแหน่งเป้าหมาย 7
2.2.4 soccerwindow2 8
2.3 บันทึกข้อมูลการแข่งขัน (RoboCup competition log file หรือ Log file) 8
2.4 การเรียนรู้แบบเสริมกำลัง (Reinforcement learning) 12
2.4.1 Q-learning 12
2.5 การวัดประสิทธิภาพ (Performance Measurement) 13
2.6 งานวิจัยที่เกี่ยวข้อง 13
บทที่ 3 วิธีดำเนินการวิจัย 20
3.1 การเก็บรวบรวมข้อมูล (Data collection) 20
3.2 การเลือกข้อมูล (Feature selection) 22
3.3 การเรียนรู้แบบเสริมกำลัง (Reinforcement learning) 22
3.3.1 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning 24
3.3.2 การออกแบบเมทริกซ์ Q และ R 26
3.4 การวัดประสิทธิภาพ (Performance measurement) 30
บทที่ 4 ผลการวิจัยและการอภิปราย 31
4.1 เครื่องมือและข้อมูลที่ใช้ในการทดลอง 31
4.1.1 เครื่องมือที่ใช้ในการทดลอง 31
4.1.2 ข้อมูลที่ใช้ในการทดลอง 31
4.2 การออกแบบ Q-learning สำหรับการทดลองที่ 1 (เฉพาะการกระทำ) 34
4.3 การออกแบบ Q-learning สำหรับการทดลองที่ 2 (การกระทำและทิศทาง) 38
บทที่ 5 บทสรุป 48
5.1 สรุปผล 48
5.2 แนวทางการพัฒนา 49

บรรณานุกรม..... 50

ประวัติผู้เขียน..... 52



สารบัญตาราง

หน้า

ตาราง 2.1 การติดตั้งซอฟต์แวร์พื้นฐานสำหรับซอฟต์แวร์ rcssserver และ soccerwindow2 4

ตาราง 2.2 การติดตั้งซอฟต์แวร์ rcssserver 5

ตาราง 2.3 การติดตั้งซอฟต์แวร์ soccerwindow2 5

ตาราง 2.4 คำอธิบายของลูกฟุตบอลในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ 10

ตาราง 2.5 คำอธิบายของเอเจนต์แต่ละตัวในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ 10

ตาราง 2.6 คุณลักษณะข้อมูลที่ใช้ในการสร้างแบบจำลอง (1) 14

ตาราง 2.7 คุณลักษณะข้อมูลที่ใช้ในการสร้างแบบจำลอง (12) 16

ตาราง 2.8 ผลการทดลองระหว่างทีม Helios_Base กับทีม GPR-2D 18

ตาราง 3.1 คุณลักษณะที่ถูกนำมาใช้ในงานวิจัย 22

ตาราง 3.2 การทำงานของอัลกอริทึม Q-learning 23

ตาราง 3.3 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 1 24

ตาราง 3.4 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 2 25

ตาราง 3.5 เงื่อนไขการกำหนดรางวัลของการทดลองที่ 1 27

ตาราง 3.6 เงื่อนไขการกำหนดรางวัลของการทดลองที่ 2 27

ตาราง 4.1 เหตุการณ์ของการแข่งขันของการทดลองที่ 1 32

ตาราง 4.2 คุณลักษณะที่ถูกนำมาใช้ในการทดลองที่ 1 32

ตาราง 4.3 เหตุการณ์ของการแข่งขันของการทดลองที่ 2 32

ตาราง 4.4 คุณลักษณะที่ถูกนำมาใช้ในการทดลองที่ 2 33

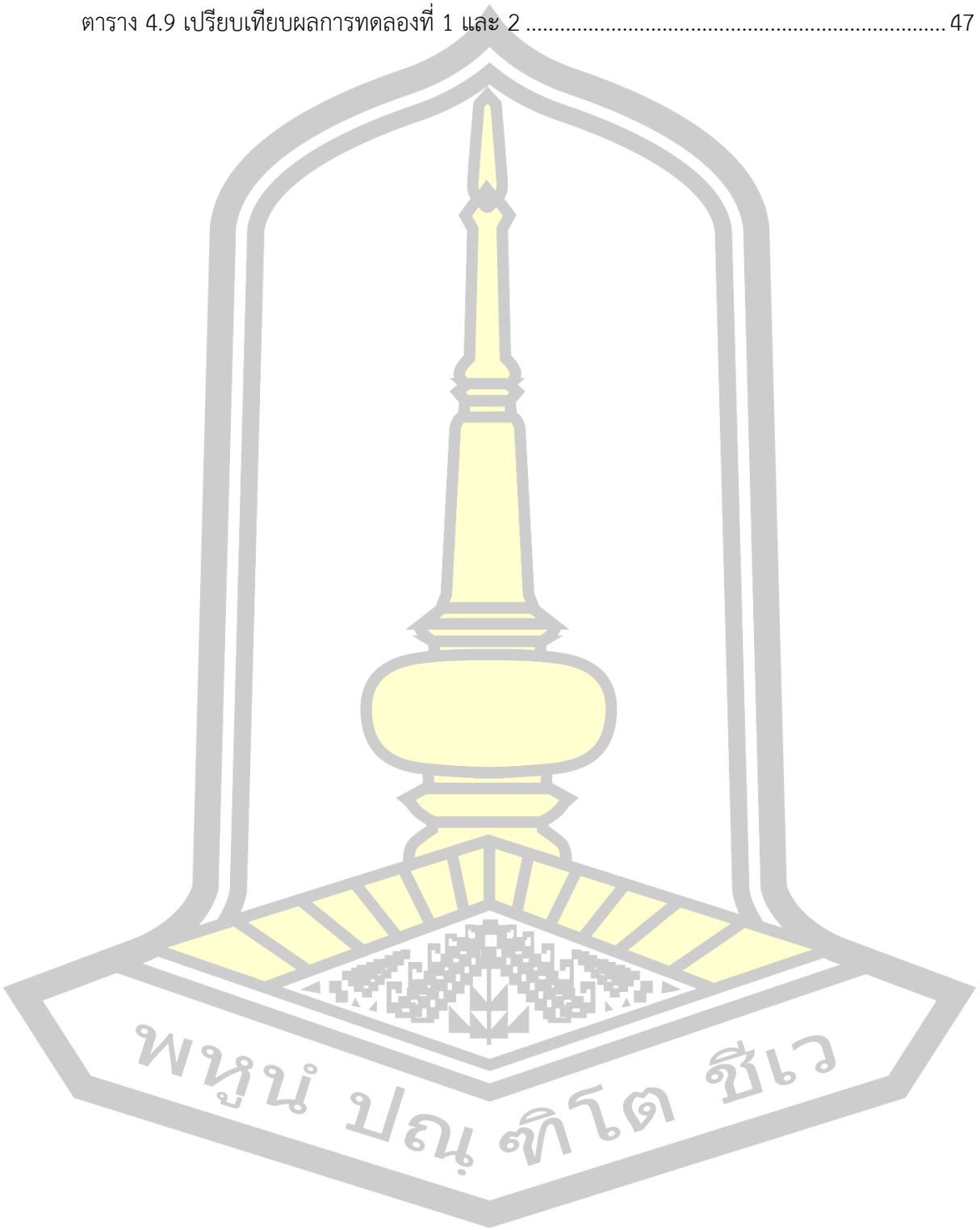
ตาราง 4.5 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 1 35

ตาราง 4.6 ตัวอย่างของผลการทดลองที่ 1 36

ตาราง 4.7 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 2 41

ตาราง 4.8 ตัวอย่างของผลการทดลองที่ 2 44

ตาราง 4.9 เปรียบเทียบผลการทดลองที่ 1 และ 2 47



สารบัญรูปภาพ

หน้า

รูปที่ 2.1 โครงสร้างของ RoboCup soccer server (10) 6

รูปที่ 2.2 ลักษณะการมองเห็นของเอเจนต์ (10) 7

รูปที่ 2.3 ซอฟต์แวร์ soccerwindow2 8

รูปที่ 2.4 ตัวอย่างข้อมูลภายในบันทึกข้อมูลการแข่งขันนามสกุล rcg 9

รูปที่ 2.5 คุณลักษณะของเอเจนต์แต่ละตัวในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ 9

รูปที่ 2.6 การทำงานของการเรียนรู้แบบเสริมกำลัง 12

รูปที่ 2.7 ลักษณะข้อมูลของต้นไม้ตัดสินใจทั้ง 4 ต้น (1) 14

รูปที่ 2.8 ขั้นตอนการปรับปรุงแผนการเล่นของทีม YuShan (11) 15

รูปที่ 2.9 พื้นที่ที่สนใจในการยิงประตู (พื้นที่หมายเลข 1 และ 2) (7) 18

รูปที่ 2.10 ระยะนอกกรอบเขตโทษที่ไกลที่สุดที่สามารถยิงประตูได้ (พื้นที่สีทึบ) 19

รูปที่ 3.1 กระบวนการพัฒนาขั้นตอนวิธีการเคลื่อนที่ของเกมรุก 20

รูปที่ 3.2 ระยะการประกบตัวของเอเจนต์ (สี่เหลี่ยมประ) 21

รูปที่ 3.3 ระยะการมองเห็นของเอเจนต์ (เส้นประ) 21

รูปที่ 3.4 ผังการทำงานของ Q-learning 24

รูปที่ 3.5 สร้างเมทริกซ์ Q 28

รูปที่ 3.6 สร้างเมทริกซ์ R 28

รูปที่ 3.7 การกระทำที่ถูกสุ่มเลือกที่มีโอกาสเป็นไปได้ (ในกรอบสี่เหลี่ยม) 28

รูปที่ 3.8 เลือกค่ามากที่สุดของสถานะถัดไปจากเมทริกซ์ Q 29

รูปที่ 3.9 แก้ไขเมทริกซ์ Q จากผลลัพธ์ที่คำนวณได้ (ในกรอบสี่เหลี่ยม) 29

รูปที่ 4.1 เมทริกซ์ Q ของการทดลองที่มีการกำหนดการกระทำ (Action) 34

รูปที่ 4.2 เมทริกซ์ Q ที่ถูกปรับรางวัลของการทดลองที่ 1 36

รูปที่ 4.3 การเลี้ยงลูกฟุตบอลและทิศทางของการเลี้ยงลูกฟุตบอล 39
รูปที่ 4.4 การส่งลูกฟุตบอลและทิศทางของการส่งลูกฟุตบอล 39
รูปที่ 4.5 การยิงประตูและทิศทางของการยิงประตู 40
รูปที่ 4.6 เมทริกซ์ Q ของการทดลองที่มีการกำหนดการกระทำและทิศทาง (Action และ 40
รูปที่ 4.7 เมทริกซ์ Q ที่ถูกปรับรางวัลของการทดลองที่ 2 42
รูปที่ 4.8 ผลลัพธ์การทดสอบของการทดลองที่ 2 43

พหุน ปณ ฑิโต ชีเว

บทที่ 1

บทนำ

1.1 หลักการและเหตุผล

RoboCup เป็นรายการการแข่งขันหุ่นยนต์ระดับนานาชาติ เพื่อส่งเสริมการพัฒนาด้านหุ่นยนต์และด้านปัญญาประดิษฐ์ ซึ่งการแข่งขันฟุตบอลเป็นหนึ่งในหัวข้อสำคัญของรายการนี้ (1-3) เนื่องจากฟุตบอลมีกติกาที่ซับซ้อน เช่น การล้ำหน้า การทำฟาวล์ ลูกตั้งเตะ เป็นต้น (2, 4) นอกจากนั้นในปี ค.ศ. 2050 จะมีการจัดแข่งขันฟุตบอลชิงแชมป์โลกระหว่างมนุษย์กับหุ่นยนต์ (1, 5) RoboCup 2D soccer simulation league เป็นรายการหนึ่งในการแข่งขันของ RoboCup ซึ่งเน้นเฉพาะการพัฒนาด้านปัญญาประดิษฐ์ เนื่องจากมีจำนวนผู้เล่น โค้ช สภาพแวดล้อม ที่เหมือนกับการแข่งขันฟุตบอลจริง นอกจากนั้นกติกามีความคล้ายกันกับการแข่งขันของมนุษย์ เช่น การล้ำหน้า การทำฟาวล์ เป็นต้น (1) โดยหัวข้อในงานวิจัยที่เกี่ยวข้องกับ RoboCup 2D soccer simulation league มีหลายด้าน เช่น การพัฒนาด้านปัญญาประดิษฐ์ โดยมีงานวิจัยของ Mota และคณะ (6) ได้พัฒนาระบบการตัดสินใจของเอเจนต์ โดยใช้อัลกอริทึม Bayesian SOM neural network การพัฒนาด้านเหมืองข้อมูล โดยมีงานวิจัยของ Karimi และ Ahmazadeh (1) ทำการปรับปรุงแผนการเล่น และคาดการณ์พฤติกรรมของคู่แข่ง โดยใช้เทคนิคทางเหมืองข้อมูล โดยใช้อัลกอริทึมต้นไม้ตัดสินใจ C4.5 เป็นต้น การพัฒนารูปแบบการเล่นและการตัดสินใจของเอเจนต์ยังเป็นส่วนสำคัญของการแข่งขันรายการนี้ เนื่องจากการส่งบอล การยิงประตู และด้านอื่นๆ ยังมีข้อจำกัด เช่น งานวิจัยของ Karimi และ Ahmazadeh (1) ปรับปรุงแผนการเล่น โดยจะเน้นการส่งบอลระยะสั้น เนื่องจากลดขนาดของปริภูมิสถานะ นอกจากนี้ยังมีงานวิจัยของ Neri และคณะ (7) ทำการปรับปรุงเกมรุกเพื่อเพิ่มโอกาสในการยิงประตู ซึ่งการยิงประตูจะพิจารณาเฉพาะในกรอบเขตโทษเท่านั้น แต่ในความเป็นจริงแล้วการยิงประตูไม่จำเป็นต้องเกิดเฉพาะในกรอบเขตโทษ เนื่องจากถ้าคู่ต่อสู้ที่มีแผนเกมรับที่เหนียวแน่นจะมีโอกาสที่จะนำลูกฟุตบอลไปอยู่ในกรอบเขตโทษได้น้อยลง จึงทำให้เสียโอกาสในการยิงประตูจากระยะอื่นๆ

จากปัญหาข้างต้นผู้วิจัยจึงเกิดแนวคิดในการปรับปรุงขั้นตอนวิธีการยิงประตูจากระยะไกลเพื่อเพิ่มโอกาสในการยิงประตู

1.2 วัตถุประสงค์ของการวิจัย

เพื่อพัฒนาอัลกอริทึมสำหรับการเคลื่อนที่ของเกมรุก โดยปรับปรุงการเพิ่มโอกาสการยิงประตูจากระยะไกล สำหรับ RoboCup 2D soccer simulation แบบอัตโนมัติ

1.3 ความสำคัญของการวิจัย

ได้อัลกอริทึมและการออกแบบเมทริกซ์ Q สำหรับการยิงประตูจากระยะไกล เพื่อประยุกต์ใช้งานอื่นๆ ของหุ่นยนต์ เช่น การยิงประตูของ RoboCup small size หรือ Humanoid เป็นต้น

1.4 ขอบเขตของการวิจัย

1.3.1 ข้อมูลที่ใช้ในงานวิจัยเป็นบันทึกข้อมูลการแข่งขันของการแข่งขันรายการ RoboCup ของปี ค.ศ. 2016 จำนวน 20 Log file โดยข้อมูลทั้งหมดไม่มีการยิงประตูจากระยะไกล

1.3.2 ใช้อัลกอริทึมการเรียนรู้แบบเสริมกำลัง แบบ Q-learning เพื่อปรับปรุงการยิงประตูของเอเจนต์จากระยะไกล สำหรับ RoboCup 2D soccer simulation

1.3.3 การวัดประสิทธิภาพจะวัดจากจำนวนการตัดสินใจยิงประตูจากระยะไกล

1.5 นิยามศัพท์เฉพาะ

1.4.1 การยิงประตูจากระยะไกล (Long shot goal) คือ การยิงประตูจากระยะนอกกรอบเขตโทษ ซึ่งจะเป็นระยะจากนอกกรอบเขตโทษทั้งหมดจนถึงระยะ f [t|b] [l|r] [20] (ธงบอกตำแหน่งแสดงดังรูปที่ 2.10)

1.4.2 การยิงประตูจากระยะใกล้ (Short shot goal) คือ การยิงประตูจากระยะในกรอบเขตโทษ

1.4.3 RoboCup 2D soccer simulation คือ การแข่งขันของ RoboCup ในประเภทการจำลองการแข่งขันบนซอฟต์แวร์ rcssserver เป็นเซิร์ฟเวอร์ และซอฟต์แวร์ soccerwindow2 เพื่อแสดงเหตุการณ์ที่เกิดขึ้นระหว่างการแข่งขัน

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยนี้จะพัฒนาอัลกอริทึมสำหรับการเคลื่อนที่ของเกมรุก โดยการเพิ่มโอกาสการยิงประตูจากระยะไกล สำหรับ RoboCup 2D soccer simulation แบบอัตโนมัติ จึงจำเป็นต้องศึกษา RoboCup 2D soccer simulation และการเรียนรู้แบบเสริมกำลัง โดยในบทนี้จะกล่าวถึงทฤษฎีและงานวิจัยที่เกี่ยวข้องสำหรับการพัฒนาขั้นตอนวิธีเพื่อปรับปรุงรูปแบบของเกมรุก สำหรับ RoboCup 2D soccer simulation ดังต่อไปนี้

2.1 ฟุตบอล

ฟุตบอล (Soccer) (8) เป็นกีฬาประเภททีม มีการแบ่งผู้เล่นออกเป็น 2 ทีม โดยแต่ละทีมจะมีผู้เล่น 11 คน ซึ่งจะใช้ลูกฟุตบอลในการเล่น โดยจะเป็นบนสนามสี่เหลี่ยมผืนผ้า มีความยาวประหว่าง 90-120 เมตร และมีความกว้างระหว่าง 70-90 เมตร พื้นสนามจะเป็นสนามหญ้าหรือสนามหญ้าเทียม โดยมีประตูอยู่กึ่งกลางของปลายสนามทั้ง 2 ฝั่ง เป้าหมายในการเล่นคือการทำคะแนนจากการยิงประตูคู่แข่ง

ตำแหน่งของผู้เล่น (8) จะแบ่งได้เป็น 4 ตำแหน่งหลักๆ ได้แก่ ผู้รักษาประตู กองหลัง กองกลาง และกองหน้า ซึ่งแต่ละตำแหน่งจะมีหน้าที่แตกต่างกันออกไป โดยทั่วไปแต่ละทีมจะมีผู้รักษาประตูเพียง 1 คน และที่เหลืออีก 10 คน จะขึ้นอยู่กับโค้ชที่จะเลือกแผนการเล่นว่าจะให้มีกองหลังกี่คนและทำหน้าที่อะไร กองกลางกี่คนและทำหน้าที่อะไร และกองหน้ากี่คนและทำหน้าที่อะไร

แผนการเล่น (9) ของแต่ละทีมจะขึ้นอยู่กับโค้ชผู้คุมทีม ซึ่งโค้ชแต่ละคนจะเลือกแผนการเล่นตามที่ตนชอบและจะพิจารณาจากผู้เล่นที่มีอยู่ว่าเหมาะสมกับแผนการเล่นแบบใด ตัวอย่างแผนการเล่น ระบบ 4-4-2 เป็นแผนการเล่นมาตรฐานสากล คือ มีกองหลัง 4 คน กองกลาง 4 คน และกองหน้า 2 คน ระบบ 4-3-3 เป็นแผนการเล่นที่เน้นเกมรุก คือ มีกองหลัง 4 คน กองกลาง 3 คน และกองหน้า 3 คน เป็นต้น โดยแต่ละแผนการเล่นจะมีคุณสมบัติที่แตกต่างกันออกไป เช่น เน้นการตั้งรับ เน้นการครองบอล เน้นการรุก เป็นต้น

2.2 การแข่งขันหุ่นยนต์โรโบคัพ (RoboCup) และซอฟต์แวร์สำหรับการแข่งขันหุ่นยนต์โรโบคัพ

2.2.1 การแข่งขันหุ่นยนต์โรโบคัพ (RoboCup)

RoboCup (5) เป็นรายการการแข่งขันหุ่นยนต์ระดับนานาชาติ มีจุดมุ่งหมายเพื่อส่งเสริงการวิจัย และพัฒนาทางด้านปัญญาประดิษฐ์ เริ่มจัดครั้งแรกเมื่อปี ค.ศ. 1997 มีการแข่งขันหลายประเภทดังนี้

1) RoboCup Soccer เป็นการแข่งขันที่มุ่งเน้นไปยังการเล่นฟุตบอล ซึ่งมีจุดมุ่งหมายเพื่อการพัฒนาระบบ Multi-agent โดยจะแบ่งการแข่งขันตามขนาดของหุ่นยนต์ ได้แก่ 1. Humanoid league 2. Middle size league 3. Small size league 4. Simulation league และ 5. Standard platform league

2) RoboCup Rescue เป็นการแข่งขันหุ่นยนต์สำหรับกู้ภัย ซึ่งมีจุดมุ่งหมายเพื่อการพัฒนาระบบ Multi-agent และโครงสร้างของข้อมูล โดยมีการแข่งขัน 2 ประเภท ได้แก่ 1. Robot league และ 2. Simulation league

3) RoboCup@Home เป็นการแข่งขันหุ่นยนต์สำหรับเป็นผู้ช่วยและการบริการ ซึ่งมีจุดมุ่งหมายเพื่อการพัฒนาด้านปฏิสัมพันธ์ของมนุษย์กับหุ่นยนต์ คอมพิวเตอร์วิทัศน์ การจดจำวัตถุ เป็นต้น

4) RoboCupJunior เป็นการแข่งขันสำหรับเด็กอายุไม่เกิน 18 ปี ซึ่งมีจุดมุ่งหมายเพื่อเป็นการปูพื้นฐานด้านหุ่นยนต์สำหรับเด็ก โดยมีการแข่งขัน 3 ประเภท ได้แก่ 1. Soccer 2. Dance และ 3. Rescue

2.2.2 การติดตั้งซอฟต์แวร์สำหรับการแข่งขันหุ่นยนต์โรโบคัพ

2.2.2.1 การติดตั้งซอฟต์แวร์พื้นฐานสำหรับซอฟต์แวร์ rcssserver และ soccerwindow2

ซอฟต์แวร์พื้นฐานที่สำคัญสำหรับการทำงานและแสดงผลของซอฟต์แวร์ rcssserver และ soccerwindow2 โดยมีขั้นตอนการติดตั้งดังตาราง 2.1

ตาราง 2.1 การติดตั้งซอฟต์แวร์พื้นฐานสำหรับซอฟต์แวร์ rcssserver และ soccerwindow2

ขั้นตอนที่ 1	เปิดซอฟต์แวร์ Terminal
ขั้นตอนที่ 2	$sudo apt-get install g++ build-essential libboost-all-dev qt4-dev-tools libaudio-dev libgtk-3-dev libxt-dev bison flex

2.2.2.2 การติดตั้งซอฟต์แวร์ rcssserver และ soccerwindow2

rcssserver และ soccerwindow2 สามารถใช้งานได้เฉพาะระบบปฏิบัติการ Linux เท่านั้น ซึ่งในงานวิจัยนี้ได้เลือกใช้ระบบปฏิบัติการ Ubuntu desktop 18.04 LTS และ rcssserver เป็นรุ่นที่ 15.5.0 และ soccerwindow2 เป็นรุ่นที่ 5.1.1

1. rcssserver เป็นซอฟต์แวร์ที่อยู่บนระบบปฏิบัติการ Ubuntu desktop 18.04 LTS การติดตั้ง rcssserver โดยสามารถดาวน์โหลดจาก https://github.com/rcsoccersim/rcssserver มีขั้นตอนการติดตั้งดังตาราง 2.2

ตาราง 2.2 การติดตั้งซอฟต์แวร์ rcssserver

ขั้นตอนที่ 1	เปิดซอฟต์แวร์ Terminal
ขั้นตอนที่ 2	ไปยังโฟลเดอร์ที่เก็บไฟล์ดาวน์โหลดของ rcssserver
ขั้นตอนที่ 3	$tar -zxpf rcssserver-15.5.0.tar.gz เพื่อทำการแตกไฟล์
ขั้นตอนที่ 4	$cd rcssserver-15.5.0 เพื่อเข้าไปในโฟลเดอร์ของ rcssserver
ขั้นตอนที่ 5	$./configure && make เพื่อตรวจสอบสถานะของซอฟต์แวร์พื้นฐานที่สำคัญสำหรับ rcssserver
ขั้นตอนที่ 6	$sudo make install เพื่อทำการติดตั้ง rcssserver

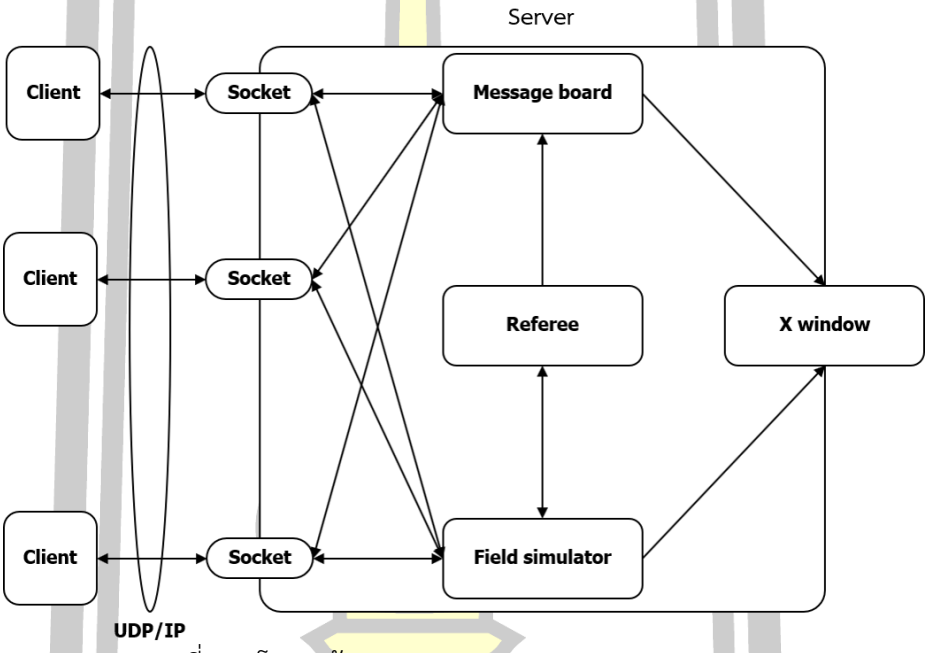
2. soccerwindow2 เป็นซอฟต์แวร์ที่อยู่บนระบบปฏิบัติการ Ubuntu desktop 18.04 LTS การติดตั้ง soccerwindow2 โดยสามารถดาวน์โหลดจาก https://osdn.net/projects/rctools/releases/68532 มีขั้นตอนการติดตั้งดังตาราง 2.3

ตาราง 2.3 การติดตั้งซอฟต์แวร์ soccerwindow2

ขั้นตอนที่ 1	เปิดซอฟต์แวร์ Terminal
ขั้นตอนที่ 2	ไปยังโฟลเดอร์ที่เก็บไฟล์ดาวน์โหลดของ soccerwindow2
ขั้นตอนที่ 3	$tar -zxpf soccerwindow2-5.1.1.tar.gz เพื่อทำการแตกไฟล์
ขั้นตอนที่ 4	$cd soccerwindow2-5.1.1 เพื่อเข้าไปในโฟลเดอร์ของ soccerwindow2
ขั้นตอนที่ 5	$./configure && make เพื่อตรวจสอบสถานะของซอฟต์แวร์พื้นฐานที่สำคัญสำหรับ soccerwindow2
ขั้นตอนที่ 6	$sudo make install เพื่อทำการติดตั้ง soccerwindow2

2.2.3 เครื่องแม่ข่ายการแข่งขันหุ่นยนต์โรโบคัพ (RoboCup soccer server)

RoboCup soccer server (10) เป็นระบบที่ใช้สำหรับจำลองการแข่งขันฟุตบอล ในรายการ RoboCup soccer 2D simulation league และ RoboCup soccer 3D simulation league โดยจะแบ่งเป็นฝั่งเซิร์ฟเวอร์กับฝั่งไคลเอนท์ ซึ่งฝั่งเซิร์ฟเวอร์จะทำหน้าที่จำลองขอบเขตของสนาม และการเคลื่อนที่ทั้งของเอเจนต์และลูกฟุตบอล ส่วนฝั่งไคลเอนท์จะควบคุมการเคลื่อนที่ของเอเจนต์หนึ่งตัว แต่ละทีมจะมีเอเจนต์ 11 ตัว การติดต่อกันระหว่างฝั่งเซิร์ฟเวอร์กับฝั่งไคลเอนท์จะติดต่อผ่านช่องทาง UDP/IP โดยมีโครงสร้างดังรูปที่ 2.1

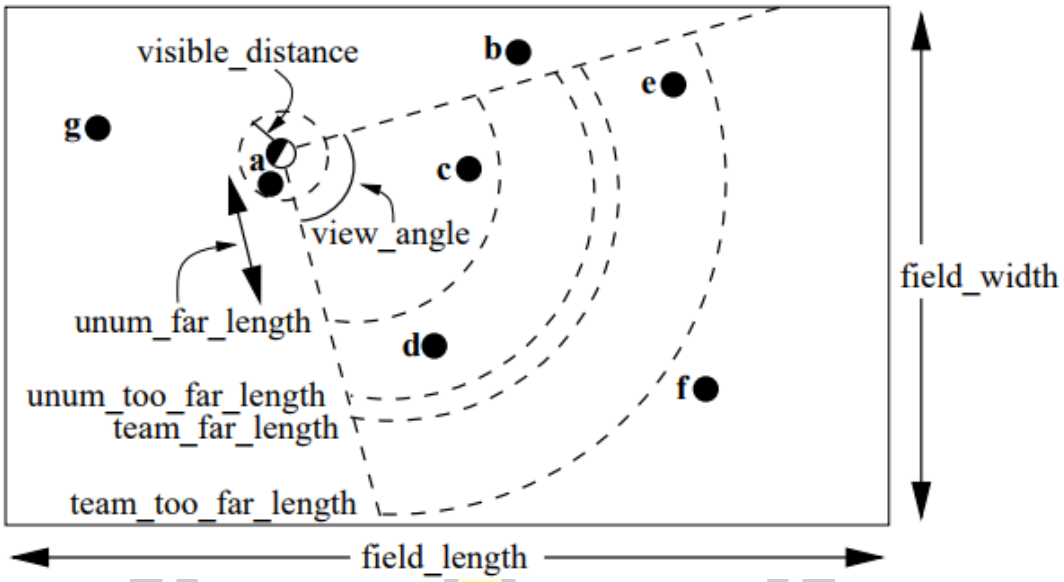


รูปที่ 2.1 โครงสร้างของ RoboCup soccer server (10)

2.2.3.1 ระยะการมองเห็นของเอเจนต์

ระยะการมองเห็นของเอเจนต์ เป็นค่าคุณสมบัติของการมองเห็นของเอเจนต์ซึ่งมีลักษณะที่เป็นกรวยดังรูปที่ 2.2

พหุน ปณ ฑิโต ชีเว



รูปที่ 2.2 ลักษณะการมองเห็นของเอเจนต์ (10)

ระยะของเอเจนต์จะมีประสิทธิภาพในการมองเห็น 5 ระยะ ที่แตกต่างกันดังนี้

1) ถ้าระยะทาง ≤ unum_far_length จะสามารถมองเห็นได้ทั้งหมายเลขและชื่อทีม (ช่วงระยะที่เอเจนต์ c อยู่)

2) ถ้า unum_far_length < ระยะทาง < unum_too_far_length จะสามารถมองเห็นชื่อทีม แต่การมองเห็นหมายเลขจะลดลง ตั้งแต่ 1 ไปจนถึง 0 (ช่วงระยะที่เอเจนต์ d อยู่)

3) ถ้าระยะทาง ≥ unum_too_far_length จะไม่สามารถมองเห็นหมายเลข (ช่วงระยะที่เอเจนต์ e และเอเจนต์ f อยู่)

4) ถ้าระยะทาง ≤ team_far_length จะสามารถมองเห็นชื่อทีม (ช่วงระยะที่เอเจนต์ c และเอเจนต์ d อยู่)

5) ถ้า team_far_length < ระยะทาง < team_too_far_length การมองเห็นชื่อทีมจะลดลง ตั้งแต่ 1 ไปจนถึง 0 (ช่วงระยะที่เอเจนต์ e อยู่)

2.2.3.2 การคำนวณระยะทางและทิศทางระหว่างเอเจนต์กับตำแหน่งเป้าหมาย

1) การคำนวณหาระยะทาง (10) จากสมการ 2.1 ถึง 2.3

$$p_{rx} = p_{xt} - p_{xo} \quad (2.1)$$

$$p_{ry} = p_{yt} - p_{yo} \quad (2.2)$$

$$Distance = \sqrt{p_{rx}^2 - p_{ry}^2} \quad (2.3)$$

โดยที่ (p_{xt}, p_{yt}) คือ ตำแหน่งของวัตถุเป้าหมาย

(p_{xo}, p_{yo}) คือ ตำแหน่งของเอเจนต์ที่ทำการตรวจจับ

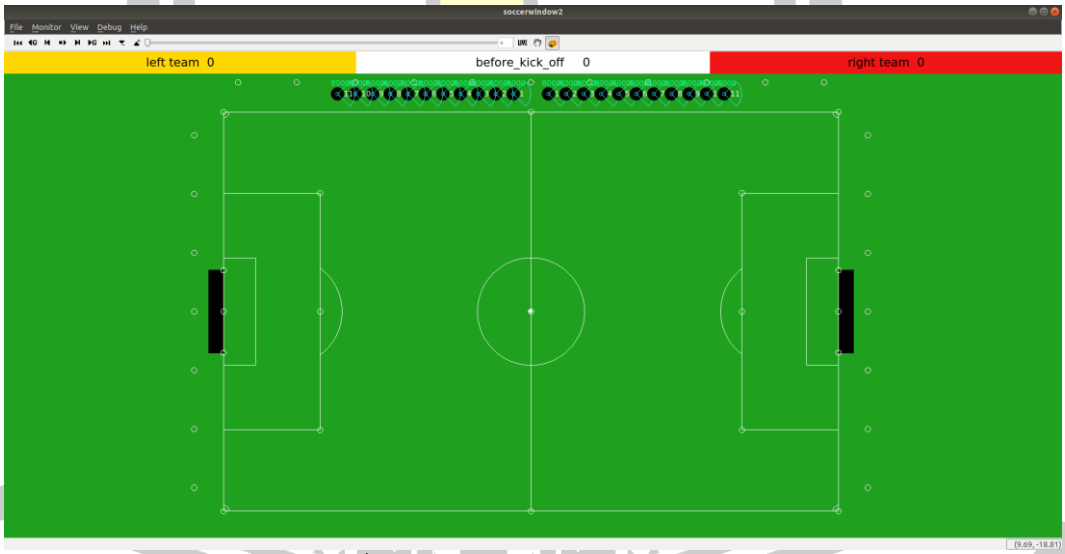
2) การคำนวณหาทิศทาง (10) จากสมการ 2.1 และ 2.2 และ 2.4

$$Direction = \arctan\left(\frac{p_{ry}}{p_{rx}}\right) - a_o \qquad (2.4)$$

โดยที่ a_o คือ ทิศทางของเอเจนต์ที่ทำการตรวจจับหันหน้า ซึ่งทิศทางการหันหน้าของเอเจนต์เป็นผลรวมของ BodyDir และ HeadDir ของเอเจนต์นั้นๆ การคำนวณการเคลื่อที่

2.2.4 soccerwindow2

soccerwindow2 เป็นซอฟต์แวร์ที่แสดงการทำงานต่างๆ ของระบบ rcssserver เพื่อให้มนุษย์สามารถมองเห็นและสั่งการได้ดัง ซึ่ง soccerwindow2 สามารถแสดงผลที่กำลังแข่งขันและบันทึกข้อมูลการแข่งขันเพื่อวิเคราะห์การแข่งขันที่ผ่านมาได้ดังรูปที่ 2.3



รูปที่ 2.3 ซอฟต์แวร์ soccerwindow2

2.3 บันทึกข้อมูลการแข่งขัน (RoboCup competition log file หรือ Log file)

บันทึกข้อมูลการแข่งขัน (RoboCup comprtition log file หรือ Log file) (1) เป็นไฟล์สำหรับเก็บข้อมูลสถานะต่างๆ ของระบบ โดย RoboCup 2D soccer simulation จะมีบันทึกข้อมูลการแข่งขัน 2 ชนิด ได้แก่ 1) rcl และ 2) rcg ซึ่งจะเก็บข้อมูลแตกต่างกัน โดย rcl เป็นบันทึกข้อมูลการแข่งขันที่เก็บสถานะจากเซิร์ฟเวอร์ เพื่อใช้ในการวิเคราะห์และปรับปรุงซอฟต์แวร์ และ rcg เป็น

บันทึกข้อมูลการแข่งขันที่เก็บตำแหน่งของลูกฟุตบอลและข้อมูลของต่างๆ เอเจนต์ ดังรูปที่ 2.4 เพื่อใช้ในการวิเคราะห์แผนการเล่น

```
(playmode 1 kick_off_l)
(team 1 tokA1 A_TSU_BI- 0 0)
(show 1 ((b) 0 0 0 0) ((l 1) 0 0x9 -49 0 0 0 89.978 0 (v h 180) (s
8000 1 1 130600) (c 0 0 396 0 1 397 1 0 0 0 0)) ((l 2) 7 0x1 -26 0 0
0 27.393 0 (v h 180) (s 8000 0.993412 1 130600) (c 0 0 386 0 1 387 1
0 0 0 0)) ((l 3) 12 0x1 -19 5 0 0 85.005 0 (v h 180) (s 8000 0.988083
1 130600) (c 0 0 385 0 1 386 1 0 0 0 0)) ((l 4) 4 0x1 -21.5 -13 0 0
-83.154 0 (v h 180) (s 8000 0.98783 1 130600) (c 0 0 384 0 1 385 1 0
0 0 0)) ((l 5) 6 0x1 -21.5 13 0 0 -30.065 0 (v h 180) (s 8000
0.956217 1 130600) (c 0 0 383 0 1 384 1 0 0 0 0)) ((l 6) 5 0x1 -19 -5
0 0 -60.22 0 (v h 180) (s 8000 0.980369 1 130600) (c 0 0 382 0 1 383
1 0 0 0 0)) ((l 7) 17 0x1 -8 -8.5 0 0 -113.086 0 (v h 180) (s 8000
0.960735 1 130600) (c 0 0 381 0 1 382 1 0 0 0 0)) ((l 8) 16 0x1 -8
8.5 0 0 -107.473 0 (v h 180) (s 8000 0.947507 1 130600) (c 0 0 380 0
```

รูปที่ 2.4 ตัวอย่างข้อมูลภายในบันทึกข้อมูลการแข่งขันนามสกุล rcg

คุณลักษณะของเอเจนต์แต่ละตัวมีดังรูปที่ 2.5 ซึ่งมีคุณลักษณะจำนวน 32 คุณลักษณะ โดยแบ่งเป็นคุณลักษณะของลูกฟุตบอล มีจำนวน 6 คุณลักษณะ สามารถอธิบายได้ดังตาราง 2.4 และคุณลักษณะของเอเจนต์ มีจำนวน 26 คุณลักษณะ สามารถอธิบายได้ดังตาราง 2.5

รูปที่ 2.5 คุณลักษณะของเอเจนต์แต่ละตัวในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ soccerwindow2

ตาราง 2.4 คำอธิบายของลูกฟุตบอลในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ soccerwindow2

ลำดับที่	คุณลักษณะ	ชนิดของข้อมูล	คำอธิบาย	ตัวอย่างข้อมูล
1	New position (x_{new})	float	ตำแหน่งของลูกฟุตบอล ณ ปัจจุบัน ในแนวแกน x	9.16
2	New position (y_{new})	float	ตำแหน่งของลูกฟุตบอล ณ ปัจจุบัน ในแนวแกน y	4.83
3	Old position (x_{old})	float	ตำแหน่งของลูกฟุตบอลก่อนหน้าในแนวแกน x	0.81
4	Old position (y_{old})	float	ตำแหน่งของลูกฟุตบอลก่อนหน้าในแนวแกน y	0.10
5	Velocity (x)	float	อัตราความเร็วในแนวแกน x	0.76
6	Velocity (y)	float	อัตราความเร็วในแนวแกน y	0.09

ตาราง 2.5 คำอธิบายของเอเจนต์แต่ละตัวในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ soccerwindow2

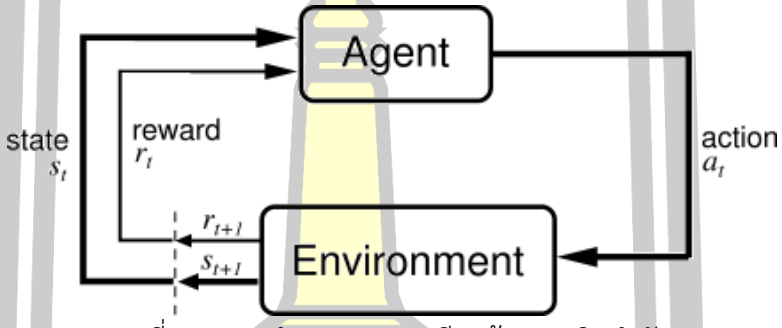
ลำดับที่	คุณลักษณะ	ชนิดของข้อมูล	คำอธิบาย	ตัวอย่างข้อมูล
1	ID	String	หลายเลขระบุตัวเอเจนต์	l 11
2	Type	int	ประเภทของเอเจนต์	17
3	State	String	บอกสถานะปัจจุบันของเอเจนต์	0x1
4	New position (x_{new})	float	ตำแหน่งของเอเจนต์ ณ ปัจจุบัน ในแนวแกน x	-0.385
5	New position (y_{new})	float	ตำแหน่งของเอเจนต์ ณ ปัจจุบัน ในแนวแกน y	0
6	Old position (x_{old})	float	ตำแหน่งของเอเจนต์ก่อนหน้าในแนวแกน x	-03723
7	Old position (y_{old})	float	ตำแหน่งของเอเจนต์ก่อนหน้าในแนวแกน y	-0.0981

ตาราง 2.5 คำอธิบายของเอเจนต์แต่ละตัวในแต่ละวินาที ที่ถูกนำมาแสดงผลในซอฟต์แวร์ soccerwindow2 (ต่อ)

ลำดับที่	คุณลักษณะ	ชนิดของข้อมูล	คำอธิบาย	ตัวอย่างข้อมูล
8	Body	float	มุมของตัวของเอเจนต์	171.877
9	Head	float	มุมของศีรษะของเอเจนต์	-90
10	View quality	Char	คุณภาพในการมองเห็น	h
11	View width	Int	ความกว้างของการมองเห็น	60
12	Stamina	Float	ค่าความแข็งแกร่งของเอเจนต์	8000
13	Effort	Float	ค่าความพยายาม	0.978243
14	Recovery	Float	อัตราการเรียกคืนค่าความแข็งแกร่งของเอเจนต์	1
15	Focus	String	เอเจนต์ที่กำลังครอบครองลูกฟุตบอล (มีเฉพาะเอเจนต์ที่ไม่ได้ครอบครองลูกฟุตบอล)	l 1
16	Kick	int	ค่าการเตะ	0
17	Dash	int	ค่าการชน	0
18	Turn	int	ค่าการหัน	151
19	Catch	int	ค่าการจับ (มีเฉพาะเอเจนต์ที่เป็นผู้รักษาประตู)	0
20	Move	int	ค่าการเคลื่อนที่	1
21	Turn neck	int	ค่าการหันของคอ	152
22	Change view	int	ค่าการเปลี่ยนมุมมอง	2
23	Say	int	ค่าการสื่อสาร	0
24	Tackle	int	ค่าการเข้าปะทะ	0
25	Point to	int	n/a	0
26	Attention to	int	n/a	1

2.4 การเรียนรู้แบบเสริมกำลัง (Reinforcement learning)

การเรียนรู้แบบเสริมกำลัง (Reinforcement learning) เป็นการให้เอเจนต์เรียนรู้จากสภาพแวดล้อมที่กำหนดให้ ซึ่งไม่ได้เป็นการเรียนรู้แบบมีผู้สอน ทำให้เอเจนต์ต้องเรียนรู้ด้วยตัวเองจากการได้รับรางวัลในการเลือกการกระทำในแต่ละครั้ง เพื่อสังเกตสภาพแวดล้อมปัจจุบันสำหรับหาการกระทำที่ดีที่สุดจากการกระทำที่มีโอกาสเป็นไปได้ทั้งหมด โดยพิจารณาจากการให้รางวัล และสังเกตในสถานะถัดไปดังรูปที่ 2.6 ซึ่งผลลัพธ์ที่เกิดขึ้นในสถานะปัจจุบันจะขึ้นอยู่กับผลลัพธ์ในสถานะก่อนหน้านี้ด้วย และการได้รับรางวัลอาจจะไม่เกิดขึ้นทันที แต่อาจจะเกิดขึ้นหลังจากการทำงานทั้งหมดเสร็จสิ้น



รูปที่ 2.6 การทำงานของการเรียนรู้แบบเสริมกำลัง

อัลกอริทึมของการเรียนรู้แบบเสริมกำลังที่ถูกใช้บ่อยครั้ง ได้แก่ Monte carlo Q-learning SARSA และ DQN เป็นต้น ซึ่งอัลกอริทึมที่ถูกใช้ในงานวิจัยนี้ คือ Q-learning เนื่องจากเป็นการสำรวจเพื่อค้นหาจุดหมายปลายทางที่ต้องการ เพื่อให้เอเจนต์ปรับตัวเข้ากับสภาพแวดล้อม เช่น การค้นหาว่าการกระทำแบบได้จะส่งผลให้เกิดการยิงประตู ถ้าการยิงประตูนั้นสามารถทำคะแนนได้ก็จะได้รางวัล แต่ถ้าการยิงประตูนั้นไม่สามารถทำคะแนนได้ก็จะถูกทำโทษ เป็นต้น จึงทำให้งานวิจัยเกี่ยวกับ RoboCup 2D soccer simulation เลือกใช้อัลกอริทึมนี้ (7)

2.4.1 Q-learning

Q-learning (7) เป็นการคำนวณหาผลประโยชน์ในแต่ละสถานะ ซึ่งจะพิจารณาจากสถานะและการกระทำ (ตารางเก็บสถานะและการกระทำ $Q(s,a)$) กล่าวคือถ้าเลือกการกระทำในสถานนั้นๆ สถานะถัดไปก็ให้แก้ไขค่าใน $Q(s,a)$ สามารถคำนวณดังสมการ 2.5 ใน Q-learning ทำงานโดยการสุ่มเลือกการกระทำ α จากสถานะ s จากนั้นก็ไปดูว่าสถานะถัดไปคืออะไร แล้วทำการแก้ไขค่าใน $Q(s,a)$ ทำต่อไปเรื่อยๆ จนบรรลุถึงเป้าหมาย

$$Q(s_t,a_t) \leftarrow Q(s_t,a_t) + \alpha\left[r_t + \gamma \max_a Q(s_{t+1},a) - Q(s_t,a_t)\right] \qquad (2.5)$$

โดยที่ s_t คือ สถานะปัจจุบัน

a_t คือ การกระทำของ s_t

r_t คือ รางวัลที่ได้รับจากการเลือกการกระทำ ในสถานะนั้นๆ

s_{t+1} คือ สถานะถัดไป

α คือ อัตราการเรียนรู้ $(0 \le \alpha < 1)$

γ คือ ปัจจัยส่วนลด $(0 \le \gamma < 1)$

2.5 การวัดประสิทธิภาพ (Performance Measurement)

การวัดประสิทธิภาพของการปรับปรุงแผนการเล่นของ RoboCup 2D soccer simulation เป็นการวัดประสิทธิภาพของแผนการเล่น (7) จะวัดประสิทธิภาพจากการตัดสินใจ ซึ่งจะการตัดสินใจกระทำสิ่งหนึ่งสิ่งใด เช่น การตัดสินใจส่งลูกฟุตบอล การตัดสินใจยิงประตู เป็นต้น โดยจะคำนวณได้จากสมการ 2.6

$$p = \frac{x}{n} \times 100 \qquad (2.6)$$

โดยที่ x คือ จำนวนเหตุการณ์ที่สนใจ เช่น การตัดสินใจยิงประตู จำนวนในการยิงประตู เป็นต้น

n คือ จำนวนเหตุการณ์ทั้งหมด

p คือ ร้อยละ

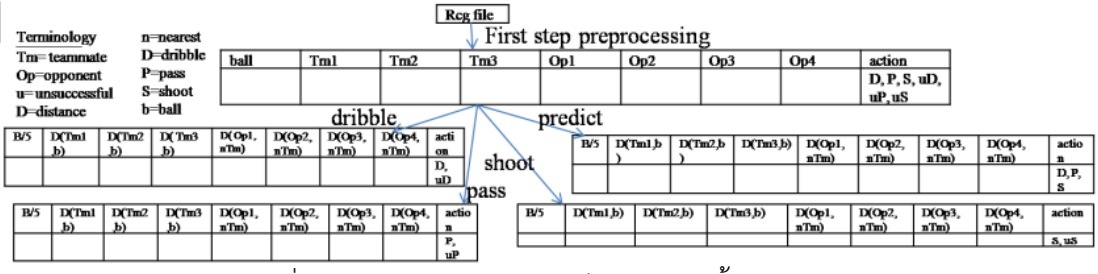
2.6 งานวิจัยที่เกี่ยวข้อง

Karimi และ Ahmazadeh (1) ทำการปรับปรุงแผนการเล่น และคาดการณ์พฤติกรรมของคู่แข่ง โดยใช้เทคนิคทางเหมืองข้อมูล ซึ่งอัลกอริทึมที่ใช้คือต้นไม้ตัดสินใจ C4.5 นำข้อมูลจากบันทึกข้อมูลการแข่งขัน 43 บันทึกข้อมูลการแข่งขัน ขั้นตอนการทำงานประกอบไปด้วย 1) เก็บรวบรวมข้อมูล 2) เตรียมข้อมูล 3) สร้างแบบจำลอง 4) วัดประสิทธิภาพ 5) สกัดกฎ 6) ทดสอบแผนการเล่น และ 7) เปรียบเทียบประสิทธิภาพของแผนการเล่น ก่อนจะนำข้อมูลไปสร้างแบบจำลองจะทำการเลือกข้อมูลของเอเจนต์มา 7 ตัว จากเอเจนต์ 22 ตัว ซึ่งจะเลือกจากการคำนวณระยะทางกับวิถีการผ่านของลูกฟุตบอล ลักษณะข้อมูลดังตาราง 2.6 แล้วนำมาสร้างต้นไม้ตัดสินใจ 4 ต้น ได้แก่ การเลี้ยงลูกฟุตบอลสำเร็จ การส่งลูกฟุตบอลสำเร็จ การยิงประตูสำเร็จ และ การคาดการพฤติกรรมของคู่แข่ง โดยมีลักษณะของข้อมูลดังรูปที่ 2.7 แล้วนำต้นไม้ตัดสินใจทั้ง 4 ต้น มาสกัดกฎ ซึ่งจะ

ทดสอบอีก 10 เกม ระหว่างทีม A และทีม B โดยที่ทีม A เป็นทีมที่ถูกเพิ่มกฎและทีม B เป็นทีมที่มีกฎตามเดิม พบว่าค่าเฉลี่ยในการทำประตูต่อเกมเพิ่มขึ้น 2.5 เท่า

ตาราง 2.6 คุณลักษณะข้อมูลที่ใช้ในการสร้างแบบจำลอง (1)

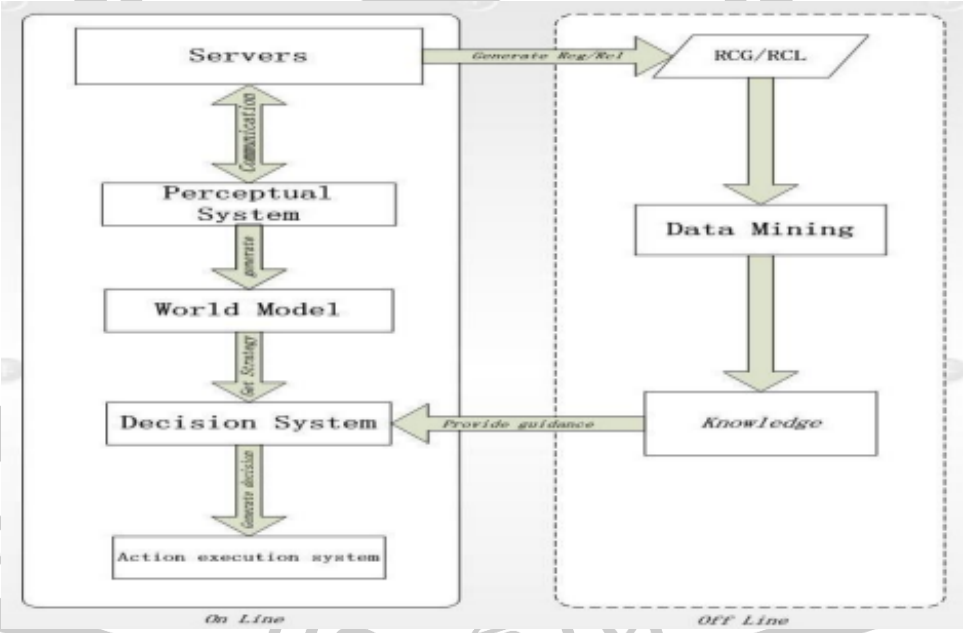
ลำดับที่	คุณลักษณะ	คำอธิบาย
1	ลูกฟุตบอล (x, y)	ตำแหน่งของลูกฟุตบอล
2	เอเจนต์ภายในทีม 1 (x, y)	ตำแหน่งของเอเจนต์ที่ครอบครองลูกฟุตบอล
3	เอเจนต์ภายในทีม 2 (x, y)	ตำแหน่งของเอเจนต์ที่ใกล้กับเอเจนต์ที่ครอบครองลูกฟุตบอลมากที่สุด
4	เอเจนต์ภายในทีม 3 (x, y)	ตำแหน่งของเอเจนต์ที่ใกล้กับวิถีการผ่านของลูกฟุตบอลครั้งสุดท้ายมากที่สุด
5	เอเจนต์ของทีมคู่แข่ง 1 (x, y)	ตำแหน่งของคู่แข่งที่ใกล้กับเอเจนต์ที่ครอบครองลูกฟุตบอลมากที่สุด
6	เอเจนต์ของทีมคู่แข่ง 2 (x, y)	ตำแหน่งของเอเจนต์คู่แข่งที่ใกล้กับวิถีการผ่านของลูกฟุตบอลครั้งสุดท้ายมากที่สุด
7	เอเจนต์ของทีมคู่แข่ง 3 (x, y)	ตำแหน่งของเอเจนต์คู่แข่งที่ใกล้จุดศูนย์กลางของวิถีการผ่านบอลมากที่สุด
8	เอเจนต์ของทีมคู่แข่ง 4 (x, y)	ตำแหน่งของเอเจนต์คู่แข่งที่ใกล้จุดศูนย์กลางของวิถีการผ่านบอลมากที่สุด
9	การกระทำ	การกระทำในสถานะการ ณ เวลานั้นๆ ได้แก่ การเลี้ยงลูกฟุตบอล การส่งลูกฟุตบอล การยิงประตู การเลี้ยงลูกฟุตบอลไม่สำเร็จ การส่งลูกฟุตบอลไม่สำเร็จ และการยิงประตูไม่สำเร็จ



รูปที่ 2.7 ลักษณะข้อมูลของต้นไม้ตัดสินใจทั้ง 4 ต้น (1)

Asali และคณะ (4) ได้ทำการวิเคราะห์รูปแบบการเล่นของทีมคู่แข่ง เพื่อโค้ชจะได้ปรับเปลี่ยนแผนการเล่น สำหรับรับมือกับทีมคู่แข่ง เนื่องจากระบบโค้ชของ RoboCup 2D soccer simulation มีการจำกัดการส่งข้อความไว้ที่ 128 ครั้ง แล้วแต่ละข้อมูลจะมีขนาดไม่เกิน 8 ไบต์ โดยใช้เทคนิคทางเหมืองข้อมูล ขั้นตอนการทำงานประกอบไปด้วย 1) เก็บรวบรวมข้อมูล 2) เตรียมข้อมูล 3) สร้างแบบจำลอง และ 4) วัดประสิทธิภาพ ข้อมูลที่นำมาใช้ ได้แก่ จุดตัดแกน x ของตำแหน่งคู่แข่ง 10 ตัว (ไม่รวมผู้รักษาประตู) และจุดตัดแกน x ของตำแหน่งลูกฟุตบอล ใช้อัลกอริทึม k-Nearest neighbors Multi support vector machine และ Neural network ในการสร้างแบบจำลอง โดยจะแบ่งเป็นข้อมูลชุดเรียนรู้ 70% ข้อมูลชุดตรวจสอบ 15% และข้อมูลชุดทดสอบ 15% ผลการทดลองพบว่า Neural network มีการวิเคราะห์รูปแบบการเล่นได้ดีที่สุด

Cheng และคณะ (11) ใช้เทคนิคทางเหมืองข้อมูล ในการปรับปรุงแผนการเล่นของทีม YuShan โดยใช้อัลกอริทึม Neural network ขั้นตอนการทำงานประกอบไปด้วย 1) เก็บรวบรวมข้อมูล 2) เตรียมข้อมูล 3) สร้างแบบจำลอง 4) สกัดกฎ และ 5) ทำระบบตัดสินใจดังรูปที่ 2.8 ซึ่งกฎที่ถูกสร้างจาก Neural network สามารถช่วยในการออกแบบแผนการเล่น โดยมีการทดสอบระหว่างแผนใหม่กับแผนเดิม พบว่าความแม่นยำในการส่งสูงขึ้นจาก 89.12% เป็น 91.20%



รูปที่ 2.8 ขั้นตอนการปรับปรุงแผนการเล่นของทีม YuShan (11)

Oliveira และคณะ (12) ได้ปรับปรุงระบบการตัดสินใจ เพื่อหาระยะเวลา ทิศทาง และระยะทางของการยิงประตูที่ดีที่สุด ใช้เทคนิคทางเหมืองข้อมูล โดยใช้อัลกอริทึม Neural network ประเภท Multi-layer perceptron เพื่อทำการเปรียบเทียบกับงานของ Kok และคณะ (13) ขั้นตอนการทำงานตามหลักการ CRISP-DM 1) เก็บรวบรวมข้อมูลและการพิจารณาข้อมูล 2) แปลงข้อมูลเพื่อ

นำไปวิเคราะห์ 3) สร้างแบบจำลอง 4) วัดประสิทธิภาพ ข้อมูลที่ถูกเลือกมาสร้างแบบจำลองมีทั้งหมด 10,691 เหตุการณ์ ซึ่งเป็นการแข่งขันระหว่างปี ค.ศ. 2005 ถึง ค.ศ. 2007 โดยจะแบ่งออกเป็นเหตุการณ์ที่ยิงประตูได้ 5,853 เหตุการณ์ และเหตุการณ์ที่ยิงประตูไม่ได้ 4,838 เหตุการณ์ (เกิดจากการป้องกันประตูของผู้รักษาประตูและการยิงออกนอกสนาม) มีลักษณะข้อมูลที่ใช้ในการสร้างแบบจำลองดังตาราง 2.7 เหตุการณ์ที่ใช้ในการสร้างแบบจำลองจะแบ่งเป็นชุดเรียนรู้ 50% ชุดตรวจสอบ 25% และชุดทดสอบ 25% การวัดปรสิทธิภาพของแบบจำลองจะใช้ Kolmogorov-smirnov similarity test และ ROC curve แล้วจะทำการแข่งขันอีก 100 แมตช์ ระหว่างทีมที่ใช้วิธีทางเหมืองข้อมูลของ Oliveira และคณะ กับทีมที่ใช้วิธีทางสถิติของ Kok และคณะ ซึ่งจะวัดประสิทธิภาพของแผนการเล่นด้านการยิงประตู ผลลัพธ์ที่ได้คือวิธีทางเหมืองข้อมูลของ Oliveira และคณะ มีการยิงประตูต่อเกมที่มากกว่าที่ 279 ลูก ต่อ 259 ลูก การยิงเข้าประตูมากกว่าที่ 96 ลูก ต่อ 54 ลูก และมีผลการแข่งขันที่ชนะมากกว่าที่ 50 แมตช์ ต่อ 24 แมตช์

ตาราง 2.7 คุณลักษณะข้อมูลที่ใช้ในการสร้างแบบจำลอง (12)

ลำดับที่	คุณลักษณะ	คำอธิบาย
1	มุมการมองของเอเจนต์ที่ยิงประตู	มุมมองของเอเจนต์ที่จะยิงประตูกับเส้นของประตูที่ซึ่งใช้ตัดสินใจในการยิงประตูว่าจะยิงไปทิศทางขวา ซ้าย หรือตรงกลางประตู
2	มุมมองที่ใช้ตัดสินใจในการยิงประตู	มุมของทิศทางที่ยิงประตู
3	ตำแหน่งของผู้รักษาประตูในแนวแกน x	ตำแหน่งของผู้รักษาประตูในแนวแกน x ณ ช่วงเวลานั้นๆ
4	ตำแหน่งของผู้รักษาประตูในแนวแกน y	ตำแหน่งของผู้รักษาประตูในแนวแกน y ณ ช่วงเวลานั้นๆ
5	ตำแหน่งของลูกฟุตบอลในแนวแกน x	ตำแหน่งของลูกฟุตบอลในแนวแกน x ณ ช่วงเวลานั้นๆ
6	ตำแหน่งของลูกฟุตบอลในแนวแกน y	ตำแหน่งของลูกฟุตบอลในแนวแกน y ณ ช่วงเวลานั้นๆ
7	มุมระหว่างลูกฟุตบอลกับผู้รักษาประตูและจุดหมายที่ต้องการยิงประตู	มุมที่ทำระหว่างลูกฟุตบอลกับผู้รักษาประตูและลูกฟุตบอลกับจุดหมายที่ต้องการยิงประตู

ตาราง 2.7 คุณลักษณะข้อมูลที่ใช้ในการสร้างแบบจำลอง (12) (ต่อ)

ลำดับที่	คุณลักษณะ	คำอธิบาย
8	ตำแหน่งของกองหลังของทีมคู่แข่งทั้งหมดในแนวแกน x	ตำแหน่งของกองหลังของทีมคู่แข่งทั้งหมดในแนวแกน x ณ ช่วงเวลานั้นๆ
9	ตำแหน่งของกองหลังของทีมคู่แข่งทั้งหมดในแนวแกน y	ตำแหน่งของกองหลังของทีมคู่แข่งทั้งหมดในแนวแกน y ณ ช่วงเวลานั้นๆ
10	ระยะเวลาในการยิงประตู	ช่วงระยะเวลาที่ยิงประตู
11	ผลลัพธ์	แบ่งเป็น 2 คลาส คือ เป็นประตูกับไม่เป็นประตู

Neri และคณะ (7) ได้ปรับปรุงพฤติกรรมการควบคุมลูกฟุตบอลของเอเจนต์ โดยใช้พื้นฐานของอัลกอริทึม Q-learning ซึ่งคาดหวังว่าเอเจนต์สามารถตัดสินใจได้ว่าจะเลือกเลี้ยงลูกฟุตบอล ส่งลูกฟุตบอล หรือยิงประตูได้ดีขึ้น ในงานนี้การยิงประตูจะสนใจในกรอบเขตโทษเพียงเท่านั้น ซึ่งในกรอบเขตโทษจะแบ่งเป็น 2 พื้นที่ดังรูปที่ 2.9 ซึ่งในพื้นที่ที่ 1 และพื้นที่ที่ 2 จะแยกการเรียนรู้แบบเสริมกำลังออกจากกัน เนื่องจากในพื้นที่ที่ 1 และพื้นที่ที่ 2 การตัดสินใจของเอเจนต์จะแตกต่างกัน โดยพื้นที่ที่ 1 คือ ระยะที่ใกล้ประตูมาก ซึ่งมีพื้นที่ระยะมากกว่า 5 เมตร จนถึงประตู และพื้นที่ที่ 2 คือ ระยะที่ใกล้ประตู ซึ่งมีพื้นที่ระยะไม่เกิน 5 เมตร คิดจากมุมของกรอบเขตโทษ ทดสอบประสิทธิภาพของแผนการเล่นจำนวน 10 เกม ผลการทดลองดังตาราง 2.8 ผลลัพธ์ที่ได้คือทีม GPR-2D ชนะ 4 เกม เสมอ 6 เกม ซึ่งเกมที่ชนะคิดเป็น 40% จากการแข่งขันทั้งหมด

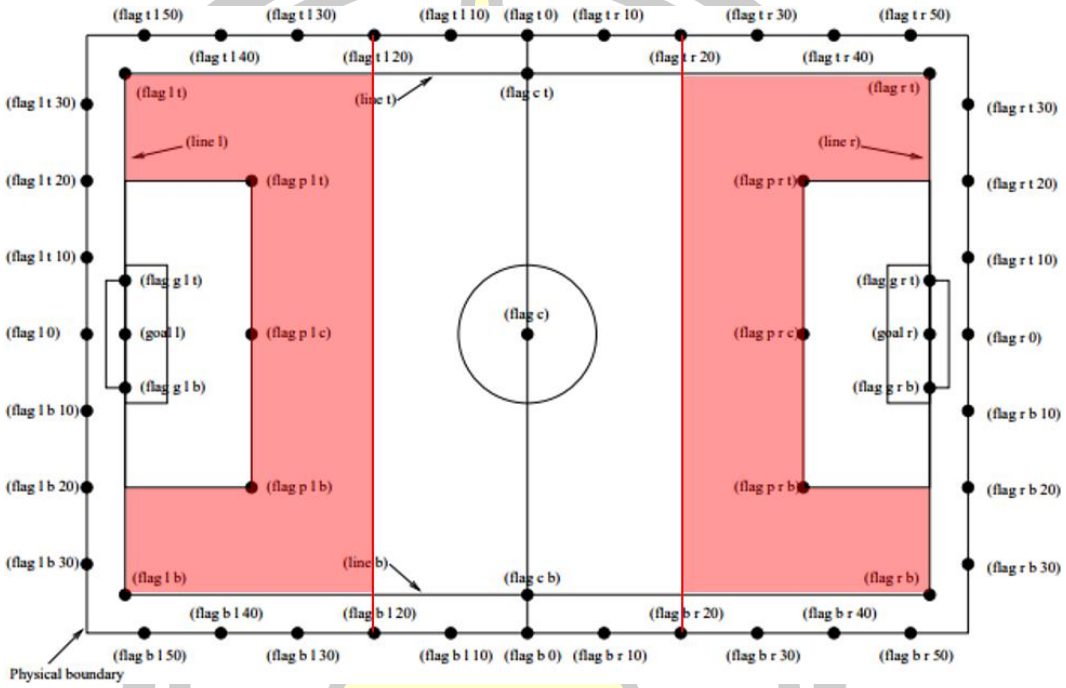
พหุน ปณ ฑิโต ชีเว

รูปที่ 2.9 พื้นที่ที่สนใจในการยิงประตู (พื้นที่หมายเลข 1 และ 2) (7)

ตาราง 2.8 ผลการทดลองระหว่างทีม Helios_Base กับทีม GPR-2D

การแข่งขันที่	**ทีม**	**ผลการแข่งขัน**	**ทีม**
1	Helios_Base	0 – 0	GPR-2D
2	Helios_Base	0 – 0	GPR-2D
3	Helios_Base	0 – 1	GPR-2D
4	Helios_Base	0 – 0	GPR-2D
5	Helios_Base	0 – 0	GPR-2D
6	Helios_Base	0 – 1	GPR-2D
7	Helios_Base	0 – 0	GPR-2D
8	Helios_Base	0 – 2	GPR-2D
9	Helios_Base	0 – 0	GPR-2D
10	Helios_Base	0 – 2	GPR-2D
Total	Helios_Base	0 – 6	GPR-2D

จากการทดลองของผู้วิจัยเอเจนต์สามารถยิงประตูจากนอกกรอบเขตในระยะทางที่ไกลที่สุด คือ ระยะ f [t|b] [l|r] [20] (ธงบอกตำแหน่ง) ดังรูปที่ 2.10 จึงกำหนดการยิงประตูจากระยะไกล (Long shot goal) ระยะจากนอกกรอบเขตโทษทั้งหมดจนถึงระยะ f [t|b] [l|r] [20] (ธงบอกตำแหน่ง) และการยิงประตูจากระยะใกล้ (Short shot goal) ระยะในกรอบเขตโทษ



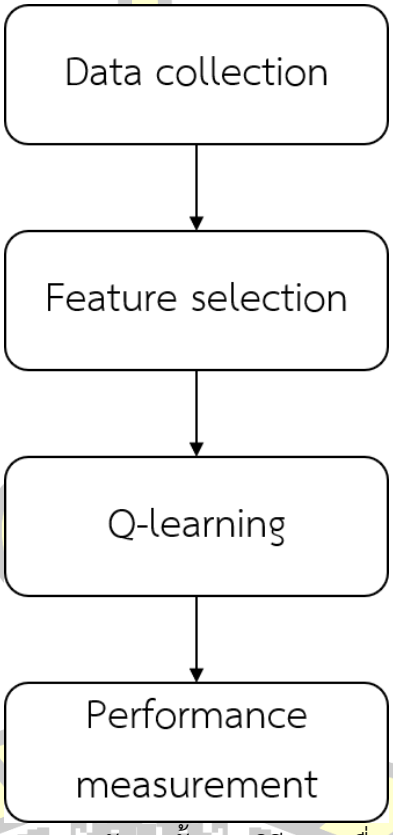
รูปที่ 2.10 ระยะนอกกรอบเขตโทษที่ไกลที่สุดที่สามารถยิงประตูได้ (พื้นที่สีทึบ)

พหุน ปณ ฑิโต ชีเว

บทที่ 3

วิธีดำเนินการวิจัย

ในบทนี้จะกล่าวถึงขั้นตอนต่างๆ ในการปรับปรุงขั้นตอนวิธีการเคลื่อนที่ของเกมรุก สำหรับการจำลองการแข่งขันหุ่นยนต์โรโบคัพ แบบ 2 มิติ โดยมีขั้นตอนต่างๆ ดังรูปที่ 3.1



รูปที่ 3.1 กระบวนการพัฒนาขั้นตอนวิธีการเคลื่อนที่ของเกมรุก

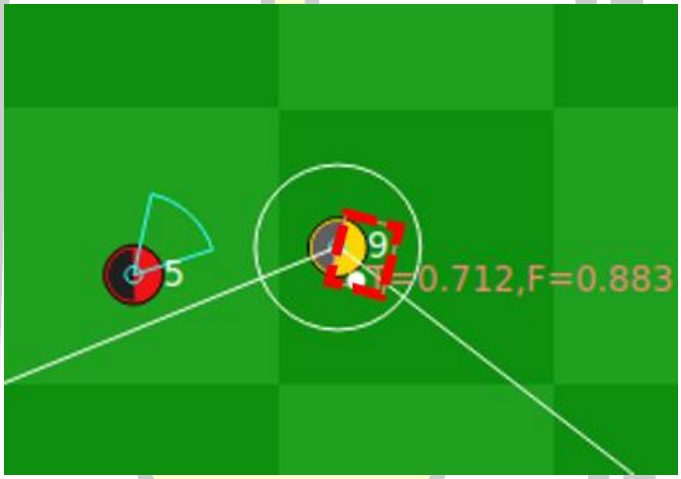
3.1 การเก็บรวบรวมข้อมูล (Data collection)

ลักษณะของข้อมูลที่นำมาใช้ในงานวิจัยจะเป็นข้อมูลการแข่งขันที่ถูกเซิร์ฟเวอร์สร้างขึ้น ซึ่งซอฟต์แวร์ที่ใช้คือ rcssserver ซึ่งเป็นซอฟต์แวร์มาตรฐานที่ใช้งานการแข่งขัน RoboCup 2D soccer simulation ลักษณะของข้อมูลจะตำแหน่งของลูกฟุตบอล ข้อมูลเอเจนต์ของทั้งสองทีม โดยข้อมูลดังกล่าวจะถูกเก็บลงในบันทึกข้อมูลการแข่งขันนามสกุล rcg ในงานวิจัยนี้มีเงื่อนไขในการเก็บข้อมูลดังนี้

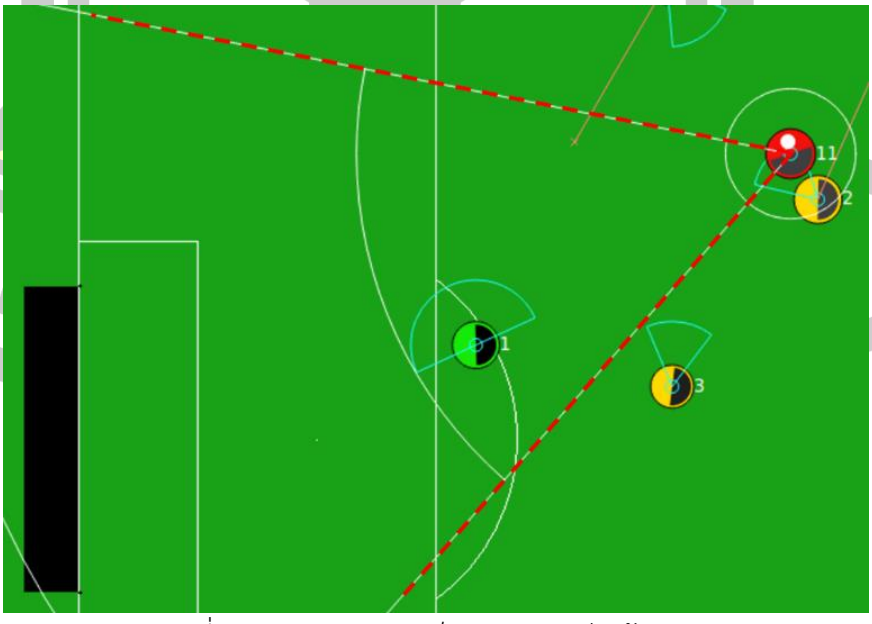
1) เป็นเหตุการณ์ที่มีโอกาสยิงประตูจากระยะไกล แต่ไม่เกิดการยิงประตู

2) เอเจนต์ที่ครอบครองลูกฟุตบอลนั้นจะต้องอยู่ในระยะ f [t|b] [l|r] [20] (ธงบอกตำแหน่ง) เป็นต้นไปดังรูปที่ 2.10

3) การเลี้ยงลูกฟุตบอลของเอเจนต์ที่ครอบครองลูกฟุตบอลจะต้องไม่อยู่ใกล้คู่แข่งมากเกินไปในระยะ ซึ่งถูกกำหนดจากค่าการเข้าประทะ (Tackle) ซึ่งแต่ละเอเจนต์จะมีค่าไม่เท่ากันดังรูปที่ 3.2 การส่งลูกฟุตบอลของเอเจนต์ที่ครอบครองลูกฟุตบอลจะต้องมองเห็นเอเจนต์ในทีมและเอเจนต์ในทีมจะต้องไม่อยู่ใกล้ระยะการเข้าประทะ (Tackle) ของคู่แข่ง ซึ่งแต่ละเอเจนต์จะมีค่าไม่เท่ากันดังรูปที่ 3.2 โดยระยะการประกบจากคู่แข่งนำมาจากบันทึกการแข่งขันและจะต้องมองเห็นเอเจนต์ภายในทีมดังรูปที่ 3.3



รูปที่ 3.2 ระยะการประกบตัวของเอเจนต์ (สี่เหลี่ยมประ)



รูปที่ 3.3 ระยะการมองเห็นของเอเจนต์ (เส้นประ)

3.2 การเลือกข้อมูล (Feature selection)

ลักษณะของข้อมูลที่ถูกเลือกมาใช้ในงานวิจัย ได้แก่ ตำแหน่งของเอเจนต์ มุมของตัวของเอเจนต์ มุมของศีรษะของเอเจนต์ และความกว้างของการมองเห็นของเอเจนต์ เนื่องจากงานวิจัยนี้เน้นไปที่การตัดสินใจยิงประตูจากระยะไกล ซึ่งเอเจนต์ที่จะสามารถยิงประตูได้จะต้องมีสถานะที่กำลังครอบครองลูกฟุตบอลอยู่ และต้องทราบตำแหน่งของเอเจนต์ที่อยู่ ณ ปัจจุบัน นอกจากนั้นทิศทางการเลี้ยงลูกฟุตบอล ส่งลูกฟุตบอล หรือยิงประตูของเอนต์นั้นจะมาจากมุมของตัว มุมมองการมองเห็นของเอเจนต์นั้นจะมาจากมุมของศีรษะและความกว้างของการมองเห็นของเอเจนต์ สามารถอธิบายได้ดังตาราง 3.1

ตาราง 3.1 คุณลักษณะที่ถูกนำมาใช้ในงานวิจัย

คุณลักษณะ	**คำอธิบาย**
New position (x_{new})	ตำแหน่งของเอเจนต์มีผลต่อการมองเห็นขอบเขตของธงบอกตำแหน่ง ในแนวแกน x
New position (y_{new})	ตำแหน่งของเอเจนต์มีผลต่อการมองเห็นขอบเขตของธงบอกตำแหน่ง ในแนวแกน y
Body	มุมของตัวทำให้ทราบว่าเอเจนต์กำลังจะเลี้ยงลูกฟุตบอล ส่งลูกฟุตบอล หรือยิงประตูไปในทิศทางใด
Head	มุมของศีรษะทำให้ทราบว่าเอเจนต์กำลังหันหน้าไปในทิศทางใด
View width	ความกว้างของการมองเห็นทำให้ทราบว่ามุมมองการมองเห็นของเอเจนต์สามารถมองเห็นอะไรได้บ้าง

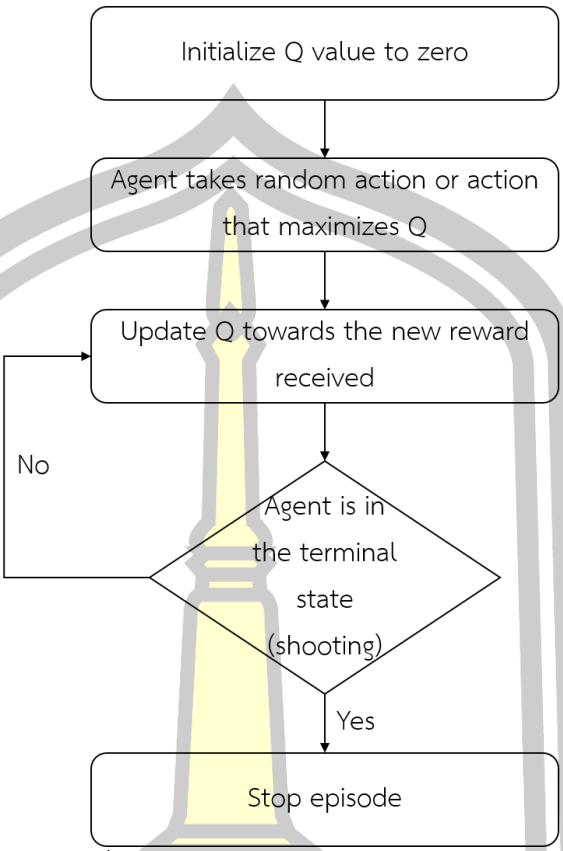
3.3 การเรียนรู้แบบเสริมกำลัง (Reinforcement learning)

การเรียนรู้แบบเสริมกำลัง โดยใช้การคำนวณแบบ Q-learning เพื่อหารางวัลในแต่ละสถานะ การทำงานของอัลกอริทึม Q-learning ดังตาราง 3.2 และรูปที่ 3.4

ตาราง 3.2 การทำงานของอัลกอริทึม Q-learning

ขั้นตอนที่ 1	สร้าง $Q(s,a)=0$ ตัวอย่างดังรูปที่ 3.5
ขั้นตอนที่ 2	กำหนดอัตราการเรียนรู้ $(0\le\alpha<1)$ และปัจจัยส่วนลด $(0\le\gamma<1)$ ดังตัวอย่างสมการ 3.1
ขั้นตอนที่ 3	สร้างรางวัลของสภาพแวดล้อมในเมทริกซ์ R ตัวอย่างดังรูปที่ 3.6
ขั้นตอนที่ 4	สุ่มเลือกการกระทำที่เป็นไปได้ในเมทริกซ์ R มา 1 การกระทำในสถานะนั้นๆ ที่มีโอกาสเป็นไปได้ ซึ่งในงานวิจัยนี้จะเลือกมาแต่ค่ารางวัลที่มีค่าเป็น 0 และ 100 เท่านั้น ตัวอย่างดังรูปที่ 3.7
ขั้นตอนที่ 5	เลือกค่ามากที่สุดจากเมทริกซ์ Q ของสถานะถัดไป ตัวอย่างดังรูปที่ 3.8
ขั้นตอนที่ 6	คำนวณจากสมการของ Q-learning ดังสมการ 3.2 สามารถแทนค่าได้ดังตัวอย่างสมการ 3.3 และ 3.4
ขั้นตอนที่ 7	นำเอาผลลัพธ์ที่คำนวณได้ดังตัวอย่างสมการ 3.5 ไปแก้ไขในเมทริกซ์ Q ตัวอย่างดังรูปที่ 3.9
ขั้นตอนที่ 8	ทำแบบต่อไปเรื่อยๆ จนได้ผลลัพธ์ คือ การยิงประตู ถ้าหากยิงประตูสำเร็จเอเจนต์ถึงจะได้รับค่ารางวัล

พหุน ปณฺฑิโต ชีเว



รูปที่ 3.4 ผังการทำงานของ Q-learning

3.3.1 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning

1) การกำหนดค่าตัวแปรของการทดลองที่ 1 ดังตาราง 3.3

ตาราง 3.3 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 1

ตัวแปร	ค่า	รายละเอียด
α	0.5	อัตราการเรียนรู้
γ	0.5	ปัจจัยส่วนลด
r	(-100, 0, 100)	รางวัลของการกระทำนั้นๆ - เป็น -100 ในกรณีที่ทำการกระทำนั้นๆ แล้วส่งผลเสีย ดังนี้ ■ เหตุการณ์ที่ไม่สามารถเลี้ยงลูกฟุตบอลได้ เนื่องจากถูกเอเจนต์คู่แข่งประกบอยู่ ■ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากเอเจนต์

ตาราง 3.3 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 1 (ต่อ)

ตัวแปร	ค่า	รายละเอียด
		ภายในทีมถูกเอเจนต์คู่แข่งประกบอยู่ ▪ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากไม่พบเอเจนต์ภายในทีมในทิศทางที่ศีรษะหันไป ▪ เหตุการณ์ที่ไม่สามารถยิงประตูได้ เนื่องจากทิศทางของศีรษะของเอเจนต์ที่ครอบครองลูกฟุตบอลไม่ได้หันไปหาประตู - เป็น 0 ในกรณีที่สามารถเลี้ยงและส่งลูกฟุตบอลได้ - เป็น 100 ในกรณีที่สามารถยิงประตูได้

2) การกำหนดค่าตัวแปรของการทดลองที่ 2 ดังตาราง 3.4

ตาราง 3.4 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 2

ตัวแปร	ค่า	รายละเอียด
α	0.5	อัตราการเรียนรู้
γ	0.5	ปัจจัยส่วนลด
r	(-100, 0, 100)	รางวัลของการกระทำนั้นๆ - เป็น -100 ในกรณีที่ทำการกระทำนั้นๆ แล้วส่งผลเสีย ดังนี้ ▪ เหตุการณ์ที่ไม่สามารถเลี้ยงลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ มีเอเจนต์คู่แข่งประกบอยู่ ▪ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ ไม่มีเอเจนต์ภายในทีมอยู่

ตาราง 3.4 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 2 (ต่อ)

ตัวแปร	ค่า	รายละเอียด
		▪ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ เอเจนต์ภายในทีมถูกเอเจนต์คู่แข่งประกบอยู่ ▪ เหตุการณ์ที่ไม่สามารถยิงประตูได้ เนื่องจากทิศทางของศีรษะของเอเจนต์ที่ครอบครองลูกฟุตบอลไม่ได้หันไปหาประตู - เป็น 0 ในกรณีที่สามารถเลี้ยงและส่งลูกฟุตบอลได้ - เป็น 100 ในกรณีที่สามารถยิงประตูได้

3.3.2 การออกแบบเมทริกซ์ Q และ R

1) การออกแบบเมทริกซ์ Q โดยกำหนดแถวเท่ากับจำนวนของสถานะ ซึ่งสถานะในการทดลองที่ 1 กำหนดจากวินาที โดย 1 สถานะต่อ 1 วินาที และการทดลองที่ 2 กำหนดสถานะจากการกระทำปัจจุบัน คือ การเลี้ยงลูกฟุตบอล นอกจากนั้นกำหนดหลักเท่ากับจำนวนของการกระทำทั้งหมด ซึ่งการทดลองที่ 1 กำหนดการกระทำดังนี้ การเลี้ยงลูกฟุตบอล การส่งลูกฟุตบอล และการยิงประตู และการทดลองที่ 2 กำหนดการกระทำดังนี้ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้าย การเลี้ยงลูกฟุตบอลไปทางข้างหน้า การเลี้ยงลูกฟุตบอลไปทางข้างหน้าขวา การส่งลูกฟุตบอลไปทางข้างหน้าซ้าย การส่งลูกฟุตบอลไปทางข้างหน้า การส่งลูกฟุตบอลไปทางข้างหน้าขวา การยิงประตูไปทางข้างหน้าซ้าย การยิงประตูไปทางข้างหน้า และการยิงประตูไปทางข้างหน้าขวา ซึ่งการกำหนดค่าเริ่มต้นในเมทริกซ์ Q จะต้องกำหนดเป็น 0 ทั้งหมด เนื่องจากเมทริกซ์ Q เป็นเมทริกซ์ที่มีไว้สำหรับดูพฤติกรรมของเอเจนต์ตั้งแต่เริ่มต้นจนบรรลุถึงเป้าหมายที่ต้องการ ดังตัวอย่างของการสร้างเมทริกซ์ Q ดังรูปที่ 3.5

2) การออกแบบเมทริกซ์ R โดยเป็นเมทริกซ์ที่เก็บรางวัลของการกระทำ โดยมีเงื่อนไขการกำหนดรางวัลของการทดลองที่ 1 ดังตาราง 3.5 และเงื่อนไขการกำหนดรางวัลของการทดลองที่ 2 ดังตาราง 3.6 ดังตัวอย่างของการสร้างเมทริกซ์ R ดังรูปที่ 3.6

ตาราง 3.5 เงื่อนไขการกำหนดรางวัลของการทดลองที่ 1

- รางวัลเป็น -100 ในกรณีที่ทำการกระทำนั้นๆ แล้วส่งผลเสีย ดังนี้
  - เหตุการณ์ที่ไม่สามารถเลี้ยงลูกฟุตบอลได้ เนื่องจากถูกเอเจนต์คู่แข่งประกบอยู่
  - เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากเอเจนต์ภายในทีมถูกเอเจนต์คู่แข่งประกบอยู่
  - เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากไม่พบเอเจนต์ภายในทีมในทิศทางที่ศีรษะหันไป
  - เหตุการณ์ที่ไม่สามารถยิงประตูได้ เนื่องจากทิศทางของศีรษะของเอเจนต์ที่ครอบครองลูกฟุตบอลไม่ได้หันไปหาประตู
- รางวัลเป็น 0 ในกรณีที่เหตุการณ์สามารถเลี้ยงและส่งลูกฟุตบอลได้
- รางวัลเป็น 100 ในกรณีที่เหตุการณ์สามารถยิงประตูได้

ตาราง 3.6 เงื่อนไขการกำหนดรางวัลของการทดลองที่ 2

- เป็น -100 ในกรณีที่ทำการกระทำนั้นๆ แล้วส่งผลเสีย ดังนี้
  - เหตุการณ์ที่ไม่สามารถเลี้ยงลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ มีเอเจนต์คู่แข่งประกบอยู่
  - เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ ไม่มีเอเจนต์ภายในทีมอยู่
  - เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ เอเจนต์ภายในทีมถูกเอเจนต์คู่แข่งประกบอยู่
  - เหตุการณ์ที่ไม่สามารถยิงประตูได้ เนื่องจากทิศทางของศีรษะของเอเจนต์ที่ครอบครองลูกฟุตบอลไม่ได้หันไปหาประตู
- เป็น 0 ในกรณีที่สามารถเลี้ยงและส่งลูกฟุตบอลได้
- เป็น 100 ในกรณีที่สามารถยิงประตูได้

ตัวอย่างการคำนวณแบบ Q-learning ดังรูปที่ 3.5 ถึงรูปที่ 3.9 และสมการ 3.1 ถึง 3.5 จากตาราง 3.2 ขั้นตอนที่ 1 จะทำการสร้างเมทริกซ์ Q ประกอบด้วย 6 สถานะ (State) และ 3 การกระทำ (Action) โดยให้มีค่าเริ่มต้นเท่ากับ 0 ดังรูปที่ 3.5

		Action		
		dribble	*pass*	*shoot*
State	1	0	0	0
	2	0	0	0
$Q =$	3	0	0	0
	4	0	0	0
	5	0	0	0
	6	0	0	0

รูปที่ 3.5 สร้างเมทริกซ์ Q

จากค่าบันทึกข้อมูลการแข่งขันที่มี 6 สถานะ (State) หรือ 6 วินาที จะมีการกำหนดค่ารางวัลของ 3 การกระทำ (Action) ได้ดังรูปที่ 3.6

		Action		
		dribble	*pass*	*shoot*
State	1	−100	0	100
	2	−100	0	−100
$R =$	3	0	−100	−100
	4	0	−100	100
	5	0	−100	−100
	6	0	0	−100

รูปที่ 3.6 สร้างเมทริกซ์ R

จากขั้นตอนที่ 4 ในตาราง 3.2 จะทำการสุ่มเลือกการกระทำที่มีโอกาสเป็นไปได้ทั้งหมด โดยสมมิตสุ่มได้การยิงประตู (Shooting) ซึ่งมีค่าเท่ากับ 100 ดังรูปที่ 3.7

		Action		
		dribble	*pass*	*shoot*
State	1	−100	0	100
	2	−100	0	−100
$R =$	3	0	−100	−100
	4	0	−100	100
	5	0	−100	−100
	6	0	0	−100

รูปที่ 3.7 การกระทำที่ถูกสุ่มเลือกที่มีโอกาสเป็นไปได้ (ในกรอบสี่เหลี่ยม)

จากนั้นจะทำการเลือกค่าที่มากที่สุดของสถานะถัดไปในเมทริกซ์ Q จากรูปที่ 3.8 จะพบว่ามีค่าเท่ากัน คือ 0

$Q =$

State	Action: *dribble*	*pass*	*shoot*
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0

รูปที่ 3.8 เลือกค่ามากที่สุดของสถานะถัดไปจากเมทริกซ์ Q

ขั้นตอนถัดไปจะดำเนินการคำนวณค่าในเมทริกซ์ Q ใหม่ โดยสามารถคำนวณจากสมการ 3.1 ถึง 3.5 แล้วนำผลลัพธ์ที่คำนวณได้มาแก้ไขในเมทริกซ์ Q ดังรูปที่ 3.9

$$\alpha = 0.5, \gamma = 0.5 \tag{3.1}$$

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha\left[r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\right] \tag{3.2}$$

$$Q(1, shoot) = Q(1, shoot) + 0.5 \cdot \left[R(1, shoot) + 0.5 \cdot Max\left[Q(2, dribble), Q(2, pass), Q(2, shoot)\right] - Q(1, shoot)\right] \tag{3.3}$$

$$Q(1, shoot) = 0 + 0.5 \cdot \left[100 + (0.5 \cdot 0) - 0\right] \tag{3.4}$$

$$Q(1, shoot) = 50 \tag{3.5}$$

$Q =$

State	Action: *dribble*	*pass*	*shoot*
1	0	0	50
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0

รูปที่ 3.9 แก้ไขเมทริกซ์ Q จากผลลัพธ์ที่คำนวณได้ (ในกรอบสี่เหลี่ยม)

พหุน ปณ ฑิโต ชีว

3.4 การวัดประสิทธิภาพ (Performance measurement)

งานวิจัยนี้จะวัดประสิทธิภาพของแผนการเล่นด้านปริมาณการตัดสินใจยิงประตูจากระยะไกล ซึ่งวัดประสิทธิภาพได้จากร้อยละของการตัดสินใจยิงประตูจากระยะไกล ซึ่งจะคำนวณได้จากสมการ 3.6

$$p = \frac{x}{n} \times 100 \qquad (3.6)$$

โดยที่ x คือ จำนวนเหตุการณ์ที่สนใจ เช่น การตัดสินใจยิงประตู จำนวนในการยิงประตู เป็นต้น

n คือ จำนวนเหตุการณ์ทั้งหมด

p คือ ร้อยละ

พหูน ปณ ฑิโต ชีเว

บทที่ 4

ผลการวิจัยและการอภิปราย

การทดลองงานวิจัยเรื่อง การปรับปรุงขั้นตอนวิธีการเคลื่อนที่ของเกมรุก สำหรับการจำลองการแข่งขันหุ่นยนต์โรโบคัพ แบบ 2 มิติ จะแบ่งออกเป็น 2 การทดลอง ได้แก่ 1) การออกแบบการกระทำ (เฉพาะการกระทำ) และ 2) การออกแบบการกระทำและทิศทาง (การกระทำและทิศทาง) ซึ่งในบทนี้จะแบ่งออกเป็น 3 ส่วน ได้แก่ ส่วนที่ 1 เครื่องมือและข้อมูลที่ใช้ในการทดลอง ส่วนที่ 2 วิธีการทดลอง ประสิทธิภาพของการทดลอง และการวิจารณ์ผลการทดลองที่ 1 และส่วนที่ 3 วิธีการทดลอง ประสิทธิภาพของการทดลอง และการวิจารณ์ผลการทดลองที่ 2

4.1 เครื่องมือและข้อมูลที่ใช้ในการทดลอง

4.1.1 เครื่องมือที่ใช้ในการทดลอง

เครื่องมือที่ใช้ในการทดลอง ได้แก่ 1) ด้านฮาร์ดแวร์ ประกอบด้วย คอมพิวเตอร์ตั้งโต๊ะ โดยมีคุณลักษณะดังนี้ ซีพียู Intel(R) Core(TM) i7-8700K แรม 16GB และฮาร์ดดิสก์ ประเภท M.2 ขนาด 256GB 2) ด้านระบบปฏิบัติการ ประกอบด้วย ระบบปฏิบัติการ Ubuntu desktop 18.04 LTS ด้านซอฟต์แวร์ ประกอบด้วย MATLAB R2017a และ Clion นอกจากนั้น soccerwindow2 ถูกใช้เป็นซอฟต์แวร์สำหรับแสดงผลของบันทึกข้อมูลการแข่งขัน

4.1.2 ข้อมูลที่ใช้ในการทดลอง

ข้อมูลที่ใช้ในการทดลองเป็นบันทึกข้อมูลการแข่งขันของการแข่งขันรายการ RoboCup ของปี ค.ศ. 2016 จำนวน 20 Log file มาจาก (14) โดยข้อมูลที่ถูกใช้มีกฎดังนี้

1) ข้อมูลของเอเจนต์ที่กำลังครอบครองลูกฟุตบอล ทั้งการกระทำและทิศทาง

2) ระยะต้องอยู่ระหว่างธง f [t|b] [l|r] [20] เป็นต้นไปดังรูปที่ 2.10 เนื่องจากภายในระยะนี้เอเจนต์ที่ครอบครองลูกฟุตบอลจะสามารถมองเห็นประตูและจำแนกทีมได้ดังรูปที่ 2.2

3) เป็นเหตุการณ์ที่มีโอกาสยิงประตูจากระยะไกล แต่ไม่เกิดการยิงประตู

1) <u>การเก็บรวบรวมข้อมูลของการทดลองที่ 1</u> เก็บมาจากบันทึกข้อมูลการแข่งขัน จำนวน 5 Log file ดังตาราง 4.1 โดยจะแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดสำหรับการเรียนรู้แบบเสริม

กำลังจำนวน 50 เหตุการณ์ และชุดสำหรับวัดประสิทธิภาพจำนวน 10 เหตุการณ์ เป็นเหตุการณ์ที่ไม่ซ้ำกับชุดสำหรับการเรียนรู้แบบเสริมกำลัง

ตาราง 4.1 เหตุการณ์ของการแข่งขันของการทดลองที่ 1

Log file ที่	**การแข่งขัน**	**จำนวนเหตุการณ์**
1	HELIOS2017 vs ITAndroids	16
2	HELIOS2017 vs MT2017	6
3	ITAndroids vs MT2017	8
4	MT2017 vs Nexus2D	19
5	Nexus2D vs HELIOS2017	11

คุณลักษณะที่ถูกนำมาใช้ในการทดลองนี้ ดังตาราง 4.2

ตาราง 4.2 คุณลักษณะที่ถูกนำมาใช้ในการทดลองที่ 1

คุณลักษณะ	**คำอธิบาย**
ตำแหน่งของเอเจนต์ที่ครอบครองลูกฟุตบอล	เอเจนต์จะต้องอยู่ในระยะดังรูปที่ 2.10
ทิศทางของศีรษะ	ทิศทางของศีรษะจะระบุว่าเอเจนต์หันหน้าไปทางทิศทางใด
มุมมองการมองเห็นของเอเจนต์	การกระทำแต่ละการกระทำของเอเจนต์จะต้องตัดสินใจจากมุมมองการมองเห็นของเอเจนต์
การกระทำ	การกระทำ ได้แก่ การเลี้ยงลูกฟุตบอล การส่งลูกฟุตบอล และการยิงประตู

2) การเก็บรวบรวมข้อมูลของการทดลองที่ 2 เก็บมาจากบันทึกข้อมูลการแข่งขันจำนวน 20 Log file ดังตาราง 4.3 โดยจะแบ่งข้อมูลออกเป็น 2 ชุด คือ ชุดสำหรับการเรียนรู้แบบเสริมกำลังจำนวน 196 เหตุการณ์ และชุดสำหรับวัดประสิทธิภาพจำนวน 50 เหตุการณ์ เป็นเหตุการณ์ที่แตกต่างจากเหตุการณ์ของการแข่งขัน สำหรับการเรียนรู้แบบเสริมกำลัง

ตาราง 4.3 เหตุการณ์ของการแข่งขันของการทดลองที่ 2

Log file ที่	**การแข่งขัน**	**จำนวนเหตุการณ์**
1	Alice vs Fifty-Storms	6
2	Alice vs HfutEngine2017	12

ตาราง 4.3 เหตุการณ์ของการแข่งขันของการทดลองที่ 2 (ต่อ)

Log file ที่	การแข่งขัน	จำนวนเหตุการณ์
3	CSU_Yunlu vs Alice	13
4	CSU_Yunlu vs Fifty-Storms	12
5	CYRUS vs WIT	10
6	Fifty-Storms vs HfutEngine2017	8
7	FRA-UNIted vs HillStone	10
8	FRA-UNIted vs Ziziphus	13
9	HELIOS2017 vs ITAndroids	16
10	HELIOS2017 vs MT2017	6
11	HfutEngine2017 vs CSU_Yunlu	19
12	HillStone vs Ziziphus	7
13	ITAndroids vs MT2017	4
14	ITAndroids vs Nexus2D	8
15	MT2017 vs Nexus2D	19
16	Nexus2D vs HELIOS2017	11
17	Oxsy vs FRA-UNIted	34
18	Oxsy vs HillStone	14
19	Ri-one vs CYRUS	4
20	Ziziphus vs Oxsy	19

คุณลักษณะที่ถูกนำมาใช้ในการทดลองนี้ ดังตาราง 4.4

ตาราง 4.4 คุณลักษณะที่ถูกนำมาใช้ในการทดลองที่ 2

คุณลักษณะ	คำอธิบาย
ตำแหน่งของเอเจนต์ที่ครอบครองลูกฟุตบอล	เอเจนต์จะต้องอยู่ในระยะดังรูปที่ 2.10
ทิศทางของตัว	ทิศทางของตัวจะระบุว่าเอเจนต์จะเลี้ยงลูกฟุตบอล ส่งลูกฟุตบอล หรือยิงประตูไปในทิศทางใด
ทิศทางของศีรษะ	ทิศทางของศีรษะจะระบุว่าเอเจนต์หันหน้าไปทาง

ตาราง 4.4 คุณลักษณะที่ถูกนำมาใช้ในการทดลองที่ 2 (ต่อ)

คุณลักษณะ	**คำอธิบาย**
	ทิศทางใด
มุมมองการมองเห็นของเอเจนต์	การกระทำแต่ละการกระทำของเอเจนต์จะต้องตัดสินใจจากมุมมองการมองเห็นของเอเจนต์
ทิศทางของการกระทำ	ทิศทางของการกระทำ ได้แก่ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้าย การเลี้ยงลูกฟุตบอลไปทางข้างหน้า การเลี้ยงลูกฟุตบอลไปทางข้างหน้าขวา การส่งลูกฟุตบอลไปทางข้างหน้าซ้าย การส่งลูกฟุตบอลไปทางข้างหน้า การส่งลูกฟุตบอลไปทางข้างหน้าขวา การยิงประตูไปทางข้างหน้าซ้าย การยิงประตูไปทางข้างหน้า และการยิงประตูไปทางข้างหน้าขวา
การกระทำ	การกระทำ ได้แก่ การเลี้ยงลูกฟุตบอล การส่งลูกฟุตบอล และการยิงประตู

4.2 การออกแบบ Q-learning สำหรับการทดลองที่ 1 (เฉพาะการกระทำ)

การกำหนดการกระทำ เป็นการกำหนดการกระทำของสถานะแต่ละสถานะ ซึ่งจะกำหนดการกระทำ 3 การกระทำ ได้แก่ การเลี้ยงลูกฟุตบอล การส่งลูกฟุตบอล และการยิงประตู ดังนั้นเมทริกซ์ Q จะเป็นดังรูปที่ 4.1 โดยการกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ดังตาราง 4.5 และการคำนวณรางวัล (Reward) ในแต่ละสถานะ (State) ได้ผลลัพธ์ดังรูปที่ 4.2

$$Q = \begin{array}{c} \\ \text{State} \end{array} \begin{array}{c} \text{Action} \\ \begin{array}{ccc} \textit{dribble} & \textit{pass} & \textit{shoot} \end{array} \end{array} \begin{array}{c} \\ \begin{array}{l} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \end{array} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

รูปที่ 4.1 เมทริกซ์ Q ของการทดลองที่มีการกำหนดการกระทำ (Action)

ตาราง 4.5 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 1

ตัวแปร	ค่า	รายละเอียด
γ	0.5	ปัจจัยส่วนลด
α	0.5	อัตราการเรียนรู้
r	(-100, 0, 100)	รางวัลของการกระทำนั้นๆ - เป็น -100 ในกรณีที่ทำการกระทำนั้นๆ แล้วส่งผลเสีย ดังนี้ ■ เหตุการณ์ที่ไม่สามารถเลี้ยงลูกฟุตบอลได้ เนื่องจากถูกเอเจนต์คู่แข่งประกบอยู่ ■ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากเอเจนต์ภายในทีมถูกเอเจนต์คู่แข่งประกบอยู่ ■ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากไม่พบเอเจนต์ภายในทีมในทิศทางที่ศีรษะหันไป ■ เหตุการณ์ที่ไม่สามารถยิงประตูได้ เนื่องจากทิศทางของศีรษะของเอเจนต์ที่ครอบครองลูกฟุตบอลไม่ได้หันไปหาประตู - เป็น 0 ในกรณีที่สามารถเลี้ยงและส่งลูกฟุตบอลได้ - เป็น 100 ในกรณีที่สามารถยิงประตูได้

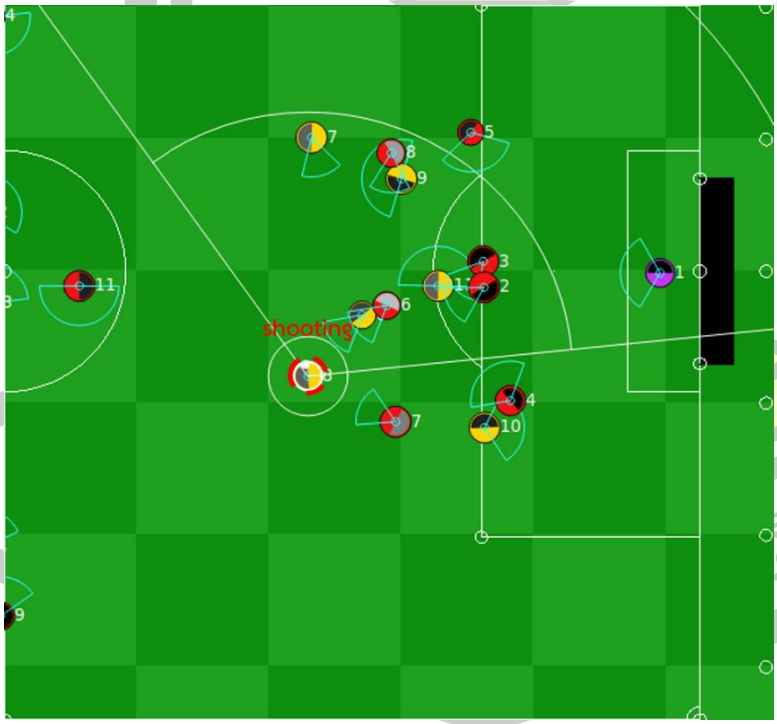
$Q =$

State	Action: *dribble*	*pass*	*shoot*
1	0	0	0
2	–0.5	0	2746508.31
3	–0.5	0	8400.21
4	0	–0.5	375.31
5	0	–0.5	50
6	0	0	1311.90

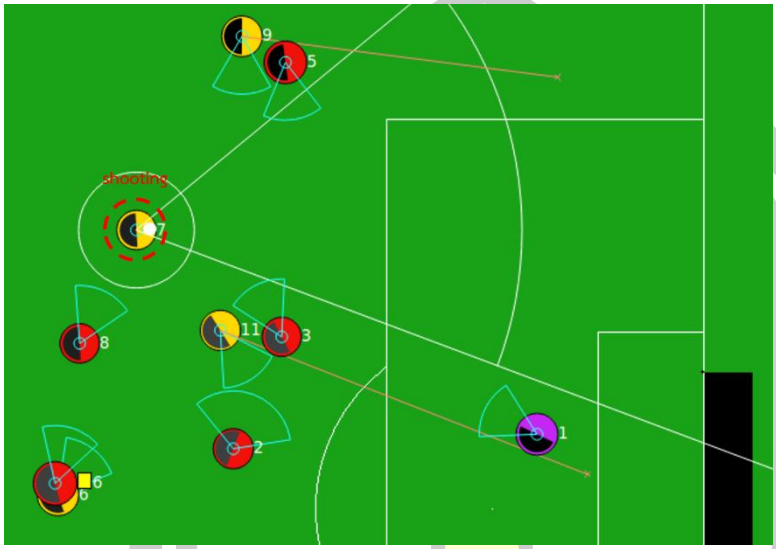
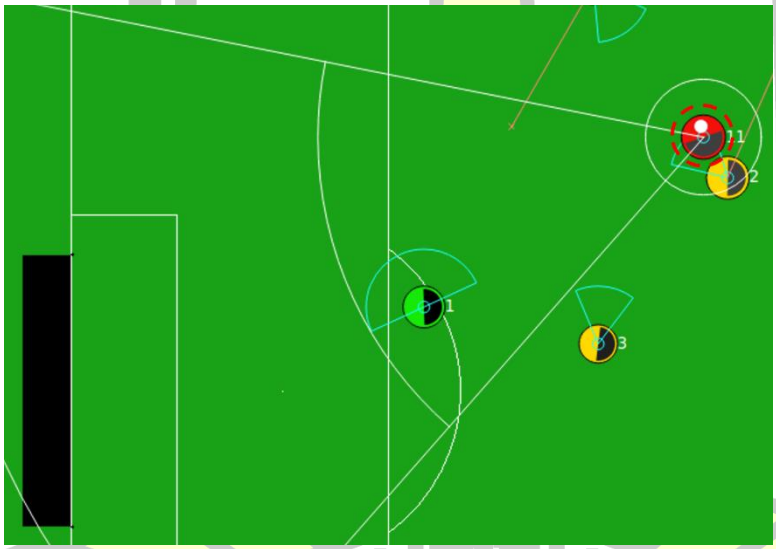
รูปที่ 4.2 เมทริกซ์ Q ที่ถูกปรับรางวัลของการทดลองที่ 1

ประสิทธิภาพของการทดลอง และการวิจารณ์ผลการทดลองที่ 1 จากการทดลองพบว่ามีการเพิ่มขึ้นของการตัดสินใจยิงประตู ซึ่งคิดเป็น 80% ตัวอย่างของผลการทดลองดังตาราง 4.6 แต่การตัดสินใจยิงประตูทั้ง 80% นั้น เป็นการตัดสินใจยิงประตูที่เข้ากรอบประตูเพียง 20% เนื่องจากเป็นการกระทำที่เอเจนต์ตัดสินใจทำโดยไม่รู้ทิศทาง ซึ่งเป็นการใช้โอกาสที่สิ้นเปลือง เพราะฉะนั้นจึงเป็นการติดสินใจยิงประตูที่ไม่แน่นอน

ตาราง 4.6 ตัวอย่างของผลการทดลองที่ 1

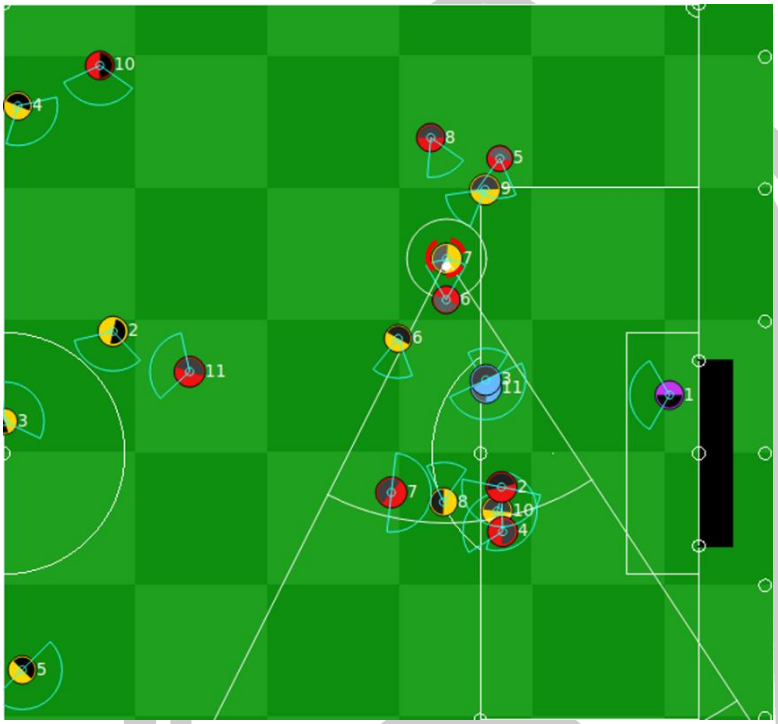
เหตุการณ์	**ผลลัพธ์**
	ตัดสินใจยิงประตู

ตาราง 4.6 ตัวอย่างของผลการทดลองที่ 1 (ต่อ)

เหตุการณ์	ผลลัพธ์
	ตัดสินใจยิงประตู
	ตัดสินใจไม่ยิงประตู

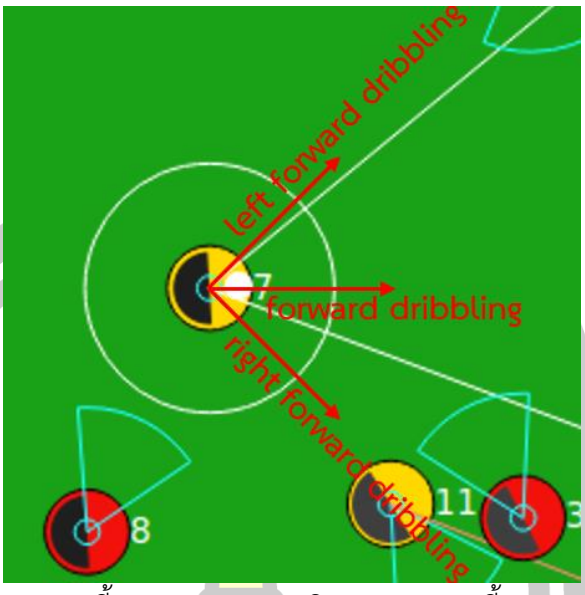
พหุน ปณ ฑิโต ชีเว

ตาราง 4.6 ตัวอย่างของผลการทดลองที่ 1 (ต่อ)

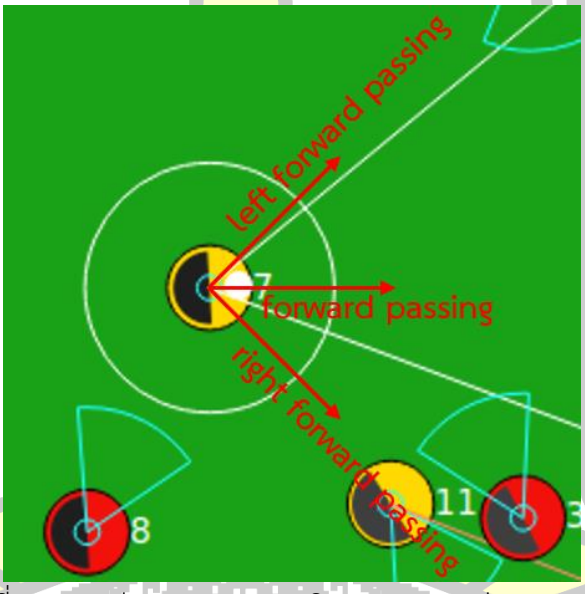
เหตุการณ์	ผลลัพธ์
	ตัดสินใจไม่ยิงประตู

4.3 การออกแบบ Q-learning สำหรับการทดลองที่ 2 (การกระทำและทิศทาง)

การกำหนดการกระทำและทิศทาง เป็นการกำหนดการกระทำของสถานะแต่ละสถานะ ซึ่งจะกำหนดการกระทำ 11 การกระทำ ได้แก่ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้าย การเลี้ยงลูกฟุตบอลไปทางข้างหน้า การเลี้ยงลูกฟุตบอลไปทางข้างหน้าขวา การส่งลูกฟุตบอลไปทางข้างหน้าซ้าย การส่งลูกฟุตบอลไปทางข้างหน้า การส่งลูกฟุตบอลไปทางข้างหน้าขวา การยิงประตูไปทางข้างหน้าซ้าย การยิงประตูไปทางข้างหน้า และการยิงประตูไปทางข้างหน้าขวา ดังรูปที่ 4.3 ถึงรูปที่ 4.5 เนื่องจากการทดลองที่กำหนดการกระทำเพียงอย่างเดียวทำให้เอเจนต์ตัดสินใจได้ว่าจะกระทำสิ่งใด แต่ไม่สามารถตัดสินใจได้ว่าจะกระทำไปในทิศทางใด เช่น ตัดสินใจที่จะยิงประตู แต่ไม่ทราบว่าจะยิงประตูไปในทิศทางใด เป็นต้น โดยปกติ RoboCup 2D soccer simulation ไม่ได้มีเพียงการกระทำเท่านั้น แต่มีทิศทางของการกระทำด้วย จึงเป็นเหตุผลที่นำทิศทางเข้ามาเป็นตัวตัดสินใจ ดังนั้นเมทริกซ์ Q จะเป็นดังรูปที่ 4.6 โดยการกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ดังตาราง 4.7 และการคำนวณรางวัล (Reward) ในแต่ละสถานะ (State) ได้ผลลัพธ์ดังรูปที่ 4.7

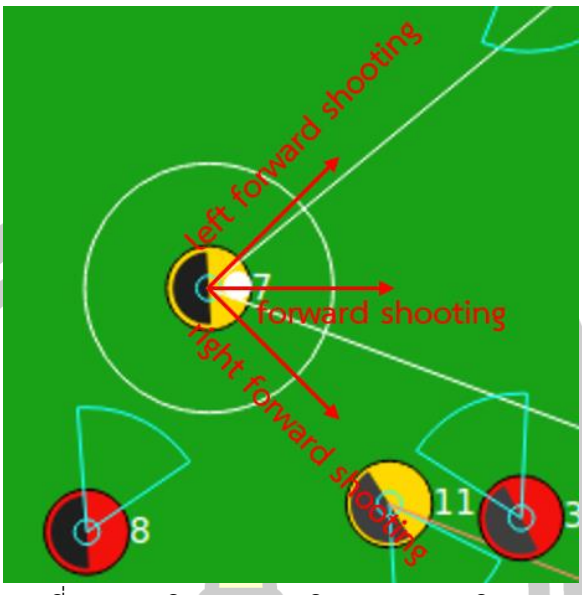


รูปที่ 4.3 การเลี้ยงลูกฟุตบอลและทิศทางของการเลี้ยงลูกฟุตบอล



รูปที่ 4.4 การส่งลูกฟุตบอลและทิศทางของการส่งลูกฟุตบอล

พหูน ปณ ฑิโต ชีเว



รูปที่ 4.5 การยิงประตูและทิศทางของการยิงประตู

$$Q = \begin{array}{c|ccccccccc} & & & & \text{Action} & & & & & \\ \text{State} & lfd & fd & rfd & lfp & fp & rfp & lfs & fs & rfs \\ \hline d & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

รูปที่ 4.6 เมทริกซ์ Q ของการทดลองที่มีการกำหนดการกระทำและทิศทาง (Action และ direction)

โดยที่ d คือ การเลี้ยงลูกฟุตบอล

lfd คือ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้าย

fd คือ การเลี้ยงลูกฟุตบอลไปทางข้างหน้า

rfd คือ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าขวา

lfp คือ การส่งลูกฟุตบอลไปทางข้างหน้าซ้าย

fp คือ การส่งลูกฟุตบอลไปทางข้างหน้า

rfp คือ การส่งลูกฟุตบอลไปทางข้างหน้าขวา

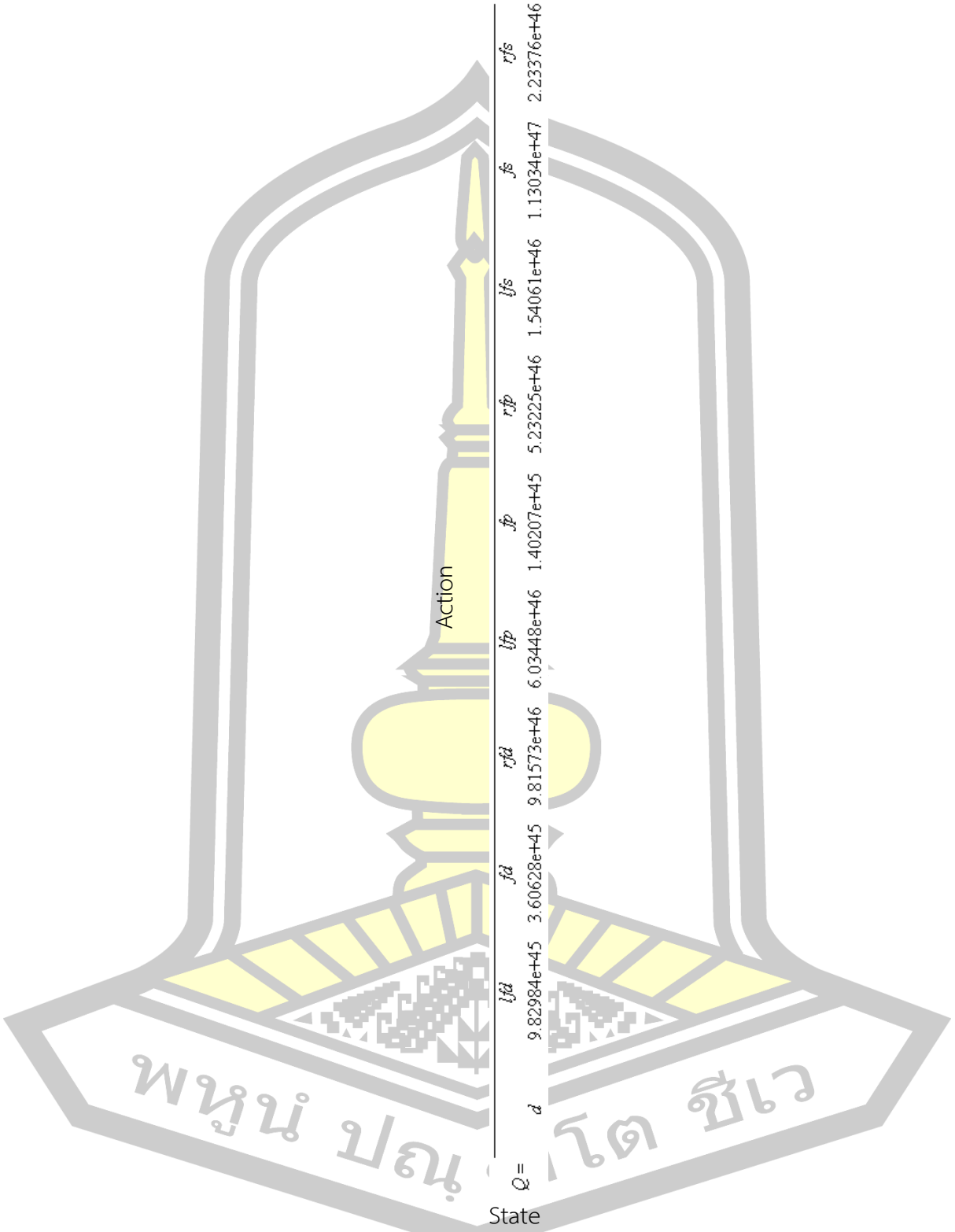
lfs คือ การยิงประตูไปทางข้างหน้าซ้าย

fs คือ การยิงประตูไปทางข้างหน้า

rfs คือ การยิงประตูไปทางข้างหน้าขวา

ตาราง 4.7 การกำหนดค่าตัวแปร สำหรับการคำนวณแบบ Q-learning ของการทดลองที่ 2

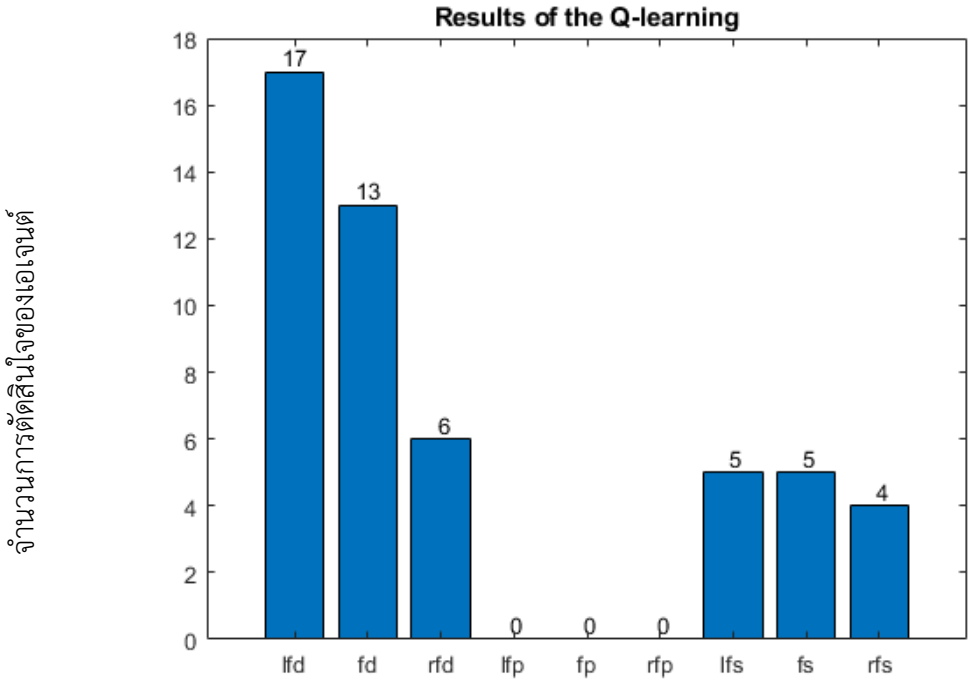
ตัวแปร	ค่า	รายละเอียด
γ	0.5	ปัจจัยส่วนลด
α	0.5	อัตราการเรียนรู้
r	(-100, 0, 100)	รางวัลของการกระทำนั้นๆ - เป็น -100 ในกรณีที่ทำการกระทำนั้นๆ แล้วส่งผลเสีย ดังนี้ ■ เหตุการณ์ที่ไม่สามารถเลี้ยงลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ มีเอเจนต์คู่แข่งประกบอยู่ ■ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ ไม่มีเอเจนต์ภายในทีมอยู่ ■ เหตุการณ์ที่ไม่สามารถส่งลูกฟุตบอลได้ เนื่องจากทิศทางนั้นๆ เอเจนต์ภายในทีมถูกเอเจนต์คู่แข่งประกบอยู่ ■ เหตุการณ์ที่ไม่สามารถยิงประตูได้ เนื่องจากทิศทางของศีรษะของเอเจนต์ที่ครอบครองลูกฟุตบอลไม่ได้หันไปหาประตู - เป็น 0 ในกรณีที่สามารถเลี้ยงและส่งลูกฟุตบอลได้ - เป็น 100 ในกรณีที่สามารถยิงประตูได้



รูปที่ 4.7 เมทริกซ์ Q ที่ถูกปรับรางวัลของการทดลองที่ 2

ประสิทธิภาพของการทดลอง และการวิจารณ์ผลการทดลองที่ 2 จากการทดลองพบว่ามีการเพิ่มขึ้นของการตัดสินใจยิงประตู ซึ่งคิดเป็น 28% นอกจากนั้นการตัดสินใจยิงประตูทั้ง 28% ล้วน

เป็นการตัดสินใจยิงประตูที่เข้ากรอบประตูทั้งหมด ดังตาราง 4.8 และรูปที่ 4.8 โดยพบว่าเอเจนต์ตัดสินใจเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้ายมากที่สุด มีจำนวน 17 เหตุการณ์ ซึ่งคิดเป็น 34% ส่วนการตัดสินใจยิงประตูของเอเจนต์จะตัดสินใจยิงประตูไปทางข้างหน้าซ้ายและข้างหน้ามากที่สุด มีจำนวน 5 เหตุการณ์ ซึ่งคิดเป็น 10% และจะตัดสินใจยิงประตูไปทางข้างหน้าขวาน้อยที่สุด มีจำนวน 4 เหตุการณ์ ซึ่งคิดเป็น 8% แต่เอเจนต์ไม่ตัดสินใจส่งลูกฟุตบอลเลย นอกจากการกระทำและทิศทางแล้วการยิงประตูต้องนำค่าการเตะของเอเจนต์และค่าความเร็วของลูกฟุตบอลมาคำนวณด้วย



การกระทำและทิศทางที่เอเจนต์ตัดสินใจ

รูปที่ 4.8 ผลลัพธ์การทดสอบของการทดลองที่ 2

จากรูปที่ 4.8 โดยที่ *lfd* คือ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้าย

fd คือ การเลี้ยงลูกฟุตบอลไปทางข้างหน้า

rfd คือ การเลี้ยงลูกฟุตบอลไปทางข้างหน้าขวา

lfp คือ การส่งลูกฟุตบอลไปทางข้างหน้าซ้าย

fp คือ การส่งลูกฟุตบอลไปทางข้างหน้า

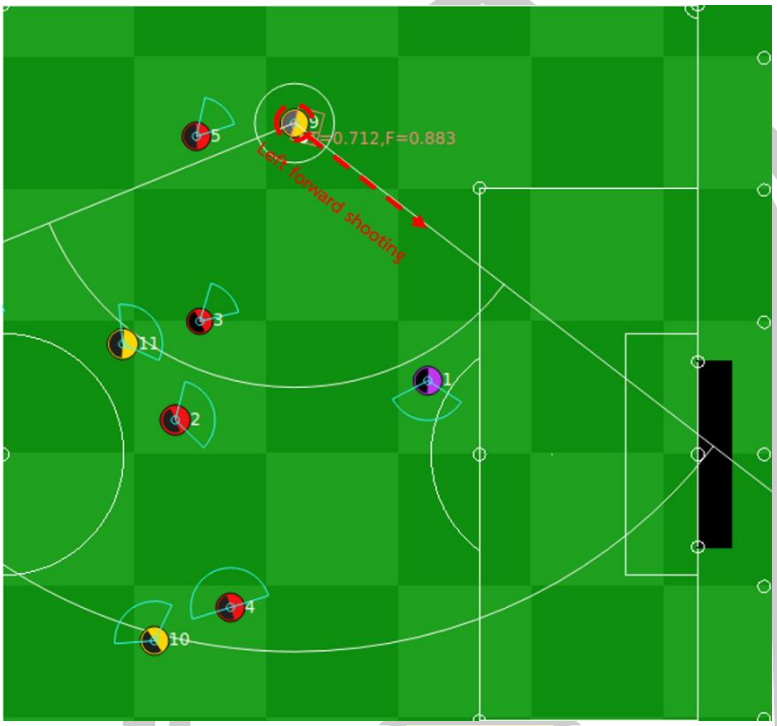
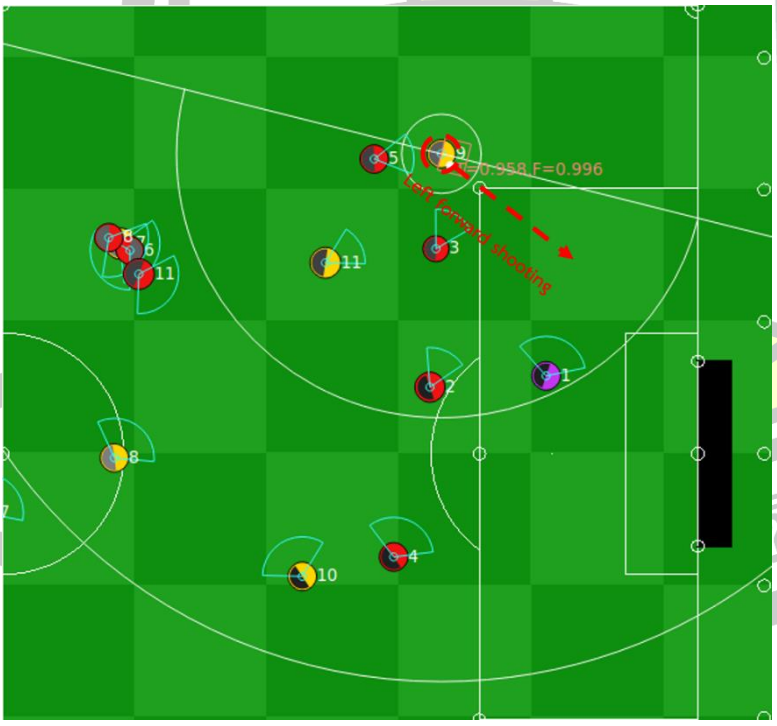
rfp คือ การส่งลูกฟุตบอลไปทางข้างหน้าขวา

lfs คือ การยิงประตูไปทางข้างหน้าซ้าย

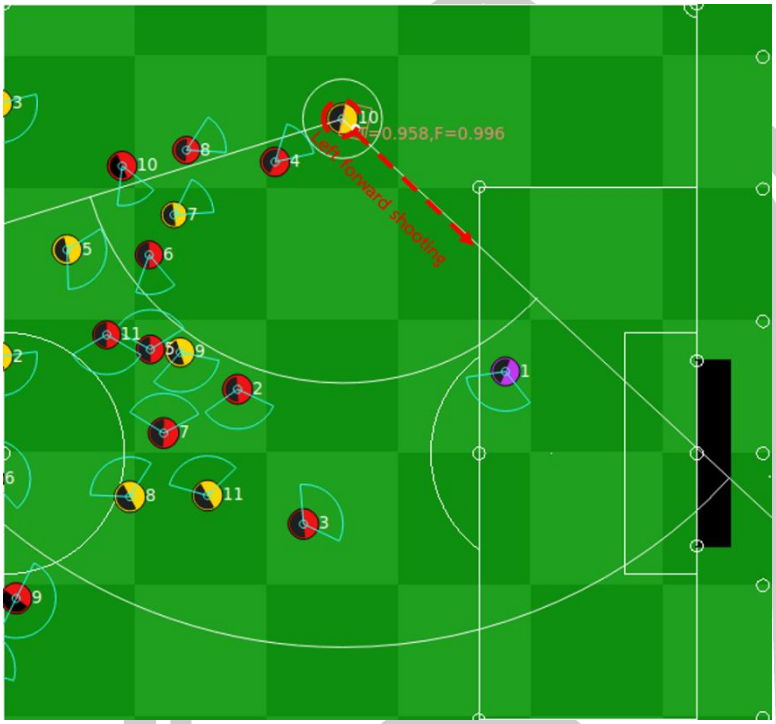
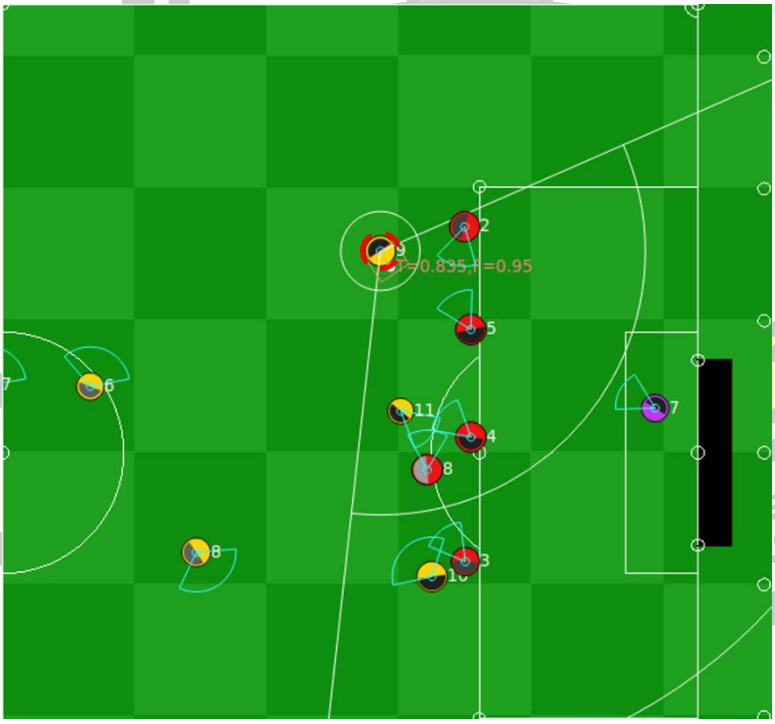
fs คือ การยิงประตูไปทางข้างหน้า

rfs คือ การยิงประตูไปทางข้างหน้าขวา

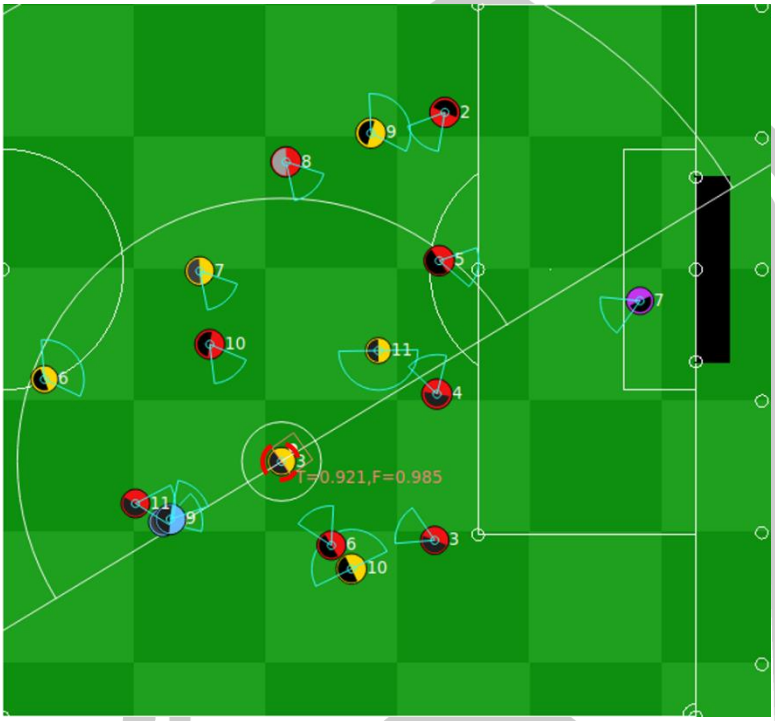
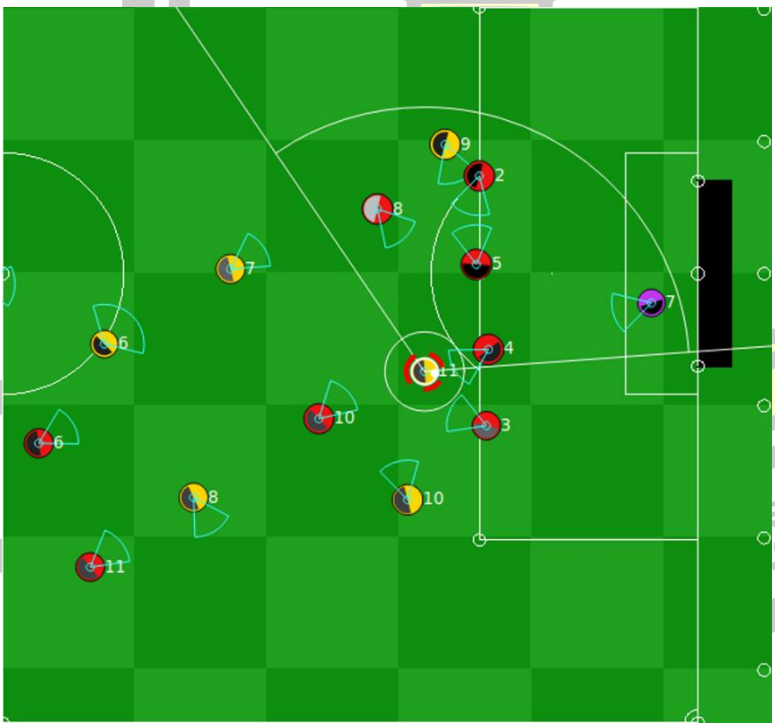
ตาราง 4.8 ตัวอย่างของผลการทดลองที่ 2

เหตุการณ์	ผลลัพธ์
	ตัดสินใจยิงประตู
	ตัดสินใจยิงประตู

ตาราง 4.8 ตัวอย่างของผลการทดลองที่ 2 (ต่อ)

เหตุการณ์	ผลลัพธ์
	ตัดสินใจยิงประตู
	ตัดสินใจไม่ยิงประตู

ตาราง 4.8 ตัวอย่างของผลการทดลองที่ 2 (ต่อ)

เหตุการณ์	ผลลัพธ์
	ตัดสินใจไม่ยิงประตู
	ตัดสินใจไม่ยิงประตู

ผลลัพธ์จากตาราง 4.8 เหตุการณ์ที่เจนต์ไม่ตัดสินใจยิงประตูเนื่องจากเอเจนต์พบการกระทำอื่นๆ ที่ส่งผลดีกว่าการยิงประตู เช่น การเลี้ยงลูกฟุตบอลหรือการส่งลูกฟุตบอล

ตาราง 4.9 เปรียบเทียบผลการทดลองที่ 1 และ 2

	การทดลองที่ 1	**การทดลองที่ 2**
การตัดสินใจยิงประตูจากระยะไกล	80%	20%
การเข้ากรอบประตู	20%	28%

จากตาราง 4.9 การทดลองที่ 1 พบว่าเอเจนต์ตัดสินใจยิงประตูจากระยะไกลเพิ่มขึ้น 80% แต่เข้ากรอบประตูเพียง 20% เนื่องจากเอเจนต์ไม่ทราบว่าควรจะยิงประตูไปในทิศทางใด การทดลองที่ 2 พบว่าเอเจนต์ตัดสินใจยิงประตูจากระยะไกลเพื่อขึ้น 28% และเข้ากรอบประตูทั้งหมด

พหูน ปณ ฑิโต ชีเว

บทที่ 5

บทสรุป

งานวิจัยนี้เป็นงานวิจัยเพื่อพัฒนาอัลกอริทึมสำหรับการเคลื่อนที่ของเกมรุก โดยการเพิ่มโอกาสการยิงประตูจากระยะไกล สำหรับ RoboCup 2D soccer simulation แบบอัตโนมัติ สามารถสรุปผล อภิปรายผล ปัญญาที่พบและแนวทางการแก้ไขปัญหา และแนวทางการพัฒนาดังต่อไปนี้

5.1 สรุปผล

งานวิจัยนี้เก็บรวบรวมข้อมูลการแข่งขันมาจาก (14) เป็นบันทึกข้อมูลการแข่งขันของการแข่งขันรายการ RoboCup ของปี ค.ศ. 2016 จำนวน 20 Log file โดยใช้การเรียนรู้แบบเสริมกำลังสำหรับคาดการณ์การยิงประตูของเอเจนต์ ซึ่งแบ่งการทดลองออกเป็น 2 การทดลอง ได้แก่ 1) การทดลองโดยการกำหนดการกระทำ และ 2) การทดลองการกำหนดการกระทำและทิศทาง และการวัดประสิทธิภาพของแผนการเล่นด้านปริมาณการตัดสินใจยิงประตูจากระยะไกล ซึ่งวัดประสิทธิภาพได้จากร้อยละของการตัดสินใจยิงประตูจากระยะไกล โดยผลของการทดลองที่ 1 (เฉพาะการกระทำ) ได้ผลลัพธ์เป็นการตัดสินใจยิงประตูจากระยะไกลเพิ่มขึ้น 80% แต่การตัดสินใจยิงประตูทั้ง 80% นั้นเป็นการตัดสินใจยิงประตูที่เข้ากรอบประตูเพียง 20% จึงเป็นการตัดสินใจยิงประตูที่ไม่แน่นอน เนื่องจากเอเจนต์ไม่ทราบว่าควรจะยิงประตูไปในทิศทาง และผลของการทดลองที่ 2 (การกระทำและทิศทาง) ได้ผลลัพธ์เป็นการตัดสินใจยิงประตูจากระยะไกลเพิ่มขึ้น 28% ซึ่งล้วนเป็นการตัดสินใจยิงประตูที่เข้ากรอบประตูทั้งหมด โดยเป็นการตัดสินใจยิงประตูที่นำเอาทั้งทิศทางและการกระทำมาคำนวณหาความเป็นไปได้ที่จะตัดสินใจยิงประตูจากระยะไกล เพื่อเพิ่มประสิทธิภาพของการยิงประตู นอกจากนั้นยังพบว่าเอเจนต์ตัดสินใจเลี้ยงลูกฟุตบอลไปทางข้างหน้าซ้ายมากที่สุด มีจำนวน 17 เหตุการณ์ ซึ่งคิดเป็น 34% ส่วนการตัดสินใจยิงประตูของเอเจนต์จะตัดสินใจยิงประตูไปทางข้างหน้าซ้ายและข้างหน้ามากที่สุด มีจำนวน 5 เหตุการณ์ ซึ่งคิดเป็น 10% และจะตัดสินใจยิงประตูไปทางข้างหน้าขวาน้อยที่สุด มีจำนวน 4 เหตุการณ์ ซึ่งคิดเป็น 8% แต่เอเจนต์ไม่ตัดสินใจส่งลูกฟุตบอลเลย

5.2 แนวทางการพัฒนา

ในงานวิจัยนี้พิจารณาจากมุมของตัวของเอเจนต์ มุมการมองเห็นของเอเจนต์ และระยะห่างระหว่างเอเจนต์กับประตู ซึ่งเป็นเพียงการคาดการณ์โอกาสการยิงประตูจากระยะไกลเท่านั้น ซึ่งในการยิงประตูจะต้องพิจารณาจากหลายๆ ส่วนดังนี้

1) ค่าความเร็วของลูกฟุตบอล
2) ค่าการเตะของเอเจนต์ที่ครอบครองลูกฟุตบอล
3) ค่าการเคลื่อนที่ของที่เอเจนต์ที่เป็นผู้รักษาประตู
4) ค่าการจับของที่เอเจนต์ที่เป็นผู้รักษาประตู

ค่าที่กล่าวในข้างต้นนี้มีผลต่อการยิงประตูของเอเจนต์ เนื่องจากค่าความเร็วของลูกฟุตบอลมีผลต่อความเร็วของลูกฟุตบอลที่ถูกยิงออกไป ค่าการเตะมีผลต่ออัตราการเคลื่อนที่ของลูกฟุตบอลที่ถูกยิงออกไป และค่าการเคลื่อที่และค่าการจับของผู้รักษาประตูมีผลต่อการป้องกันประตู นอกจากนั้นการสื่อสารระหว่างโค้ชและเอเจนต์ที่ครอบครองลูกฟุตบอลก็มีผลต่อการตัดสินใจยิงประตูของเอเจนต์ด้วย ซึ่งพิจารณาเป็นปัจจัยเสริมสำหรับการยิงประตูที่แม่นยำมากขึ้น

พหุน ปณ ฑิโต ชีเว

บรรณานุกรม

1. Karimi M, Ahmadzadeh M. Mining RoboCup Log Files to Predict Own and Opponent Action2014.

2. Lin Z, Li W, Li M, editors. Modeling decision and cooperation of multi-agent system using Petri net. 2009 4th International Conference on Computer Science & Education; 2009 25-28 July 2009.

3. Mota L, Reis L. Setplays: Achieving Coordination by the appropriate Use of arbitrary Pre-defined Flexibl Plans and Inter-Robot Communication2007. 13 p.

4. Asali E, Valipour M, Zare N, Afshar A, Katebzadeh M, Dastghaibyfard G, editors. Using Machine Learning approaches to detect opponent formation. 2016 Artificial Intelligence and Robotics (IRANOPEN); 2016 9-9 April 2016.

5. Objective [cited 2016 Oct 5]. Available from: http://www.robocup.org/about-robocup/objective/.

6. Mota L, Lau N, Reis LP, editors. Co-ordination in RoboCup's 2 D simulation league: Setplays as flexible, multi-robot plans. 2010 IEEE Conference on Robotics, Automation and Mechatronics; 2010 28-30 June 2010.

7. Neri JRF, Zatelli MR, Santos CHFd, Fabro JA, editors. A Proposal of QLearning to Control the Attack of a 2D Robot Soccer Simulation Team. 2012 Brazilian Robotics Symposium and Latin American Robotics Symposium; 2012 16-19 Oct. 2012.

8. ฟุตบอล [updated 2016 Mar 3; cited 2016 Apr 25]. Available from: https://th.wikipedia.org/wiki/ฟุตบอล.

9. Robertson J. คู่มือลูกหนัง: เจาะลึกแผนฟุตบอลสมัยใหม่ที่กุนซือระดับเทพนิยม [updated 2015; cited 2016 Apr 25]. Available from: http://www.fourfourtwo.com/th/features/khuumuueluukhnang-ecchaaaluekaephnfutblsmayaihmthiikunchuueradabethphniym#:KB0hdo0FIYzD2A.

10. Chen M, Dorer K, Foroughi E, Heintz F, Huang Z, Kapetanakis S, et al. Users Manual RoboCup Soccer Server for Soccer Server Version 7.07 and later 2003 Feb 11.

11. Cheng Z, Liu Q, Guo F, Tong F. YuShan Soccer 2 D Simulation Team Description Paper for RoboCup 2 0 1 2 . The 1 6 th annual RoboCup International Symposium; 2012 Jun 18-24; Mexico2012.

12. Oliveira R, Adeodato P, Carvalho A, Viegas I, Diego C, Ing-Ren T, editors. A data mining approach to solve the goal scoring problem. 2 0 0 9 International Joint Conference on Neural Networks; 2009 14-19 June 2009.

13. Kok J, de Boer R, Vlassis N, Groen FCA, editors. Towards an Optimal Scoring Policy for Simulated Soccer Agents. RoboCup 2002: Robot Soccer World Cup VI; 2003 2003//; Berlin, Heidelberg: Springer Berlin Heidelberg.

14. RoboCup Soccer - 2D Simulation League [cited 2016 Apr 25]. Available from: https://archive.robocup.info/Soccer/Simulation/2D/.

ประวัติผู้เขียน

ชื่อ	นายวรัญญู เกียรติจินดารัตน์
วันเกิด	วันที่ 7 มกราคม พ.ศ. 2536
สถานที่เกิด	อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม
สถานที่อยู่ปัจจุบัน	บ้านเลขที่ 163 ถนนเฉลิมพระเกียรติ ร.9 ตำบลตลาด อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม รหัสไปรษณีย์ 44000
ตำแหน่งหน้าที่การงาน	กรรมการ
สถานที่ทำงานปัจจุบัน	บริษัท เทอร์มินอล เอธ จำกัด บ้านเลขที่ 160-161 ถนนเฉลิมพระเกียรติ ร.9 ตำบลตลาด อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม รหัสไปรษณีย์ 44000
ประวัติการศึกษา	พ.ศ. 2551 มัธยมศึกษาตอนต้น โรงเรียนสารคามพิทยาคม อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม พ.ศ. 2554 มัธยมศึกษาตอนปลาย โรงเรียนสารคามพิทยาคม อำเภอเมืองมหาสารคาม จังหวัดมหาสารคาม พ.ศ. 2558 ปริญญาวิทยาศาสตรบัณฑิต (วท.บ.) สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม พ.ศ. 2562 ปริญญาวิทยาศาสตรมหาบัณฑิต (วท.ม.) สาขาวิชาวิทยาการคอมพิวเตอร์ คณะวิทยาการสารสนเทศ มหาวิทยาลัยมหาสารคาม
ทุนวิจัย	ทุนอุดหนุนการวิจัยจากงบประมาณเงินรายได้ ประจำปี 2562 มหาวิทยาลัยมหาสารคาม
ผลงานวิจัย	Kiatjindarat, W., Jareanpon, C. (2019). An Action and Direction Design of Q-learning for RoboCup 2D Soccer Simulation. In B-Con PLAZA, 24th International Symposium on Artificial Life and Robotics. (pp. 14-18).

พหุน ปณ ฑิโต ชีเว