**การฝึกพฤติกรรมของตัวละครที่ผู้เล่นไม่ได้ควบคุมในเกมโดยวิธีการเรียนแบบเสริมกำลัง**

**NON-PLAYER CHARACTER BEHAVIOR TRAINING IN GAME USING REINFORCEMENT LEARNING**

**โดย**

**อาจารย์ที่ปรึกษา**

**ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรวิทยาศาสตรบัณฑิต**

**สาขาวิชาเทคโนโลยีสารสนเทศ**

**คณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**ภาคเรียนที่ 1 ปีการศึกษา 2562**

**NON-PLAYER CHARACTER BEHAVIOR TRAINING IN GAME USING REINFORCEMENT LEARNING**

**SURACHET YAITAMMASAN**

**AKARAPON AKARASURI**

**A PROJECT SUBMITTED IN PARTIAL FULFILLMENT**

**OF THE REQUIREMENT FOR THE DEGREE OF**

**BACHELOR OF SCIENCE PROGRAM IN INFORMATION TECHNOLOGY**

**FACULTY OF INFORMATION TECNOLOGY**

**KING MONGKUT’S INSTITUTE OF TECHNOLOGY LADKRABANG**

**1/2019**

**COPYRIGHT 2019**

**FACULTY OF INFORMATION TECHNOLOGY**

**KING MONGKUT’S INSTITUTE OF TECHNOLOGY LADKRABANG**

**ใบรับรองปริญญานิพนธ์ ประจำปีการศึกษา 2562**

**คณะเทคโนโลยีสารสนเทศ**

**สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

**เรื่อง การฝึกพฤติกรรมของตัวละครที่ผู้เล่นไม่ได้ควบคุมในเกมโดยวิธีการ เรียนแบบเสริมกำลัง**

**NON-PLAYER CHARACTER BEHAVIOR TRAINING IN GAME USING REINFORCEMENT LEARNING**

**ผู้จัดทำ**

1. **นาย สุรเชษฐ์ ใหญ่ธรรมสาร รหัสนักศึกษา 59070180**
2. **นาย อัครพล อัครสูริย์ รหัสนักศึกษา 59070189**

**.........................................อาจารย์ที่ปรึกษา**

(.......................................)

**.........................................อาจารย์ที่ปรึกษาร่วม**

(.......................................)

**ใบรับรองโครงงาน (PROJECT)**

**เรื่อง**

**การฝึกพฤติกรรมของตัวละครที่ผู้เล่นไม่ได้ควบคุมในเกมโดยวิธีการเรียนแบบเสริมกำลัง**

**NON-PLAYER CHARACTER BEHAVIOR TRAINING IN GAME USING REINFORCEMENT LEARNING**

**นาย สุรเชษฐ์ ใหญ่ธรรมสาร รหัสนักศึกษา 59070180**

**นาย อัครพล อัครสูริย์ รหัสนักศึกษา 59070189**

ขอรับรองว่ารายงานฉบับนี้ ข้าพเจ้าไม่ได้คัดลอกมาจากที่ใด

รายงานฉบับนี้ได้รับการตรวจสอบและอนุมัติให้เป็นส่วนหนึ่งของ

การศึกษาวิชาโครงงาน หลักสูตรวิทยาศาสตรบัณฑิต (เทคโนโลยีสารสนเทศ)

ภาคเรียนที่ 1 ปีการศึกษา 2562

……………………………….

(นาย สุรเชษฐ์ ใหญ่ธรรมสาร)

……………………………….

(นาย อัครพล อัครสูริย์)

**หัวข้อวิทยานิพนธ์** การฝึกพฤติกรรมของตัวละครที่ผู้เล่นไม่ได้ควบคุมในเกมโดยวิธีการ เรียนรู้แบบเสริมกำลัง

**นักศึกษา** นาย สุรเชษฐ์ ใหญ่ธรรมสาร **รหัสนักศึกษา** 59070180

นาย อัครพล อัครสูริย์ **รหัสนักศึกษา** 59070189

**ปริญญา** วิทยาศาสตร์บัณฑิต

**สาขาวิชา** เทคโนโลยีสารสนเทศ

**พ.ศ. 2562**

**อาจารย์ที่ปรึกษา**

**อาจารย์ที่ปรึกษาร่วม**

**บทคัดย่อ**

**Project Title** NON-PLAYER CHARACTER BEHAVIOR TRAINING IN GAME USING REINFORCEMENT LEARNING

**Student** Surachet Yaitammasan **Student ID** 59070180

Akarapon Akarasuri **Student ID** 59070189

**Degree** Bachelor of Science

**Program** Information Technology

**Academic Year** 2019

**Project Advisor**

**Project Advisor (Co)**

**ABSTRACT**

**กิตติกรรมประกาศ**

**.......................................................................................................................................................................................................................................................................................................................................................................................................................................................................................**

ชื่อ...........นามสกุล..........

ชื่อ...........นามสกุล..........

[สารบัญรูป 5](#_Toc21179241)

[บทที่ 1 6](#_Toc21179242)

[บทนำ 6](#_Toc21179243)

[1.1 ที่มาและความสำคัญ 6](#_Toc21179244)

[1.2 วัตถุประสงค์ 6](#_Toc21179245)

[1.3 ขอบเขตของโครงงาน 6](#_Toc21179246)

[1.4 ประโยชน์ที่คาดว่าจะได้รับ 6](#_Toc21179247)

[บทที่ 2 7](#_Toc21179248)

[ทฤษฏีและงานวิจัยที่เกี่ยวข้อง 7](#_Toc21179249)

[2.1 ทฤษฏี 7](#_Toc21179250)

[2.1.1 การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning) 7](#_Toc21179251)

[2.1.2 โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network) 8](#_Toc21179252)

[2.1 งานวิจัยที่เกี่ยวข้อง 9](#_Toc21179253)

[บทที่ 3 10](#_Toc21179254)

[วิธีการดำเนินงาน 10](#_Toc21179255)

[3.1. โปรแกรมหรือซอฟต์แวร์ที่ใช้ในการพัฒนา 10](#_Toc21179256)

[3.2. ขั้นตอนการดำเนินงาน 10](#_Toc21179257)

# **สารบัญรูป**

หน้า

รูปที่

2.1 โครงสร้างการทำงานของการเรียนรู้แบบเสริมกำลัง………………………………...16

2.2 โครงสร้างของโครงข่ายประสาทแบบคอนโวลูชัน………………………………….16

# **บทที่ 1**

# **บทนำ**

บทนี้จะกล่าวถึงที่มาและความสำคัญ รวมไปถึงวัตถุประสงค์และประโยชน์ที่คาดว่าจะได้รับจากวิจัยครั้งนี้เพื่อทราบถึงจุดมุ่งหมายที่แท้จริงของผู้วิจัยโดยที่มีรายละเอียดของการดำเนินงานและขอบเขตรวมไปถึงอุปกรณ์ที่ใช้เพื่อเป็นประโยชน์แก่ผู้สนใจในการศึกษางานวิจัยนี้โดยมี รายละเอียดดัง ต่อไปนี้

## **ที่มาและความสำคัญ**

Reinforcement Learning เป็นหนึ่งในแขนงของ Machine Learning ที่ถูกนำมาใช้กับอุตสาหกรรมวิดีโอเกมมากขึ้น เช่น AlphaGo เป็นต้น ผู้จัดทำมีความประสงค์ในการพัฒนาองค์ความรู้ที่เกี่ยวข้อง เพื่อพัฒนา Algorithm ที่สามารถฝึกตัวละครในเกมที่กำหนดเพื่อเพิ่มขีดความสามารถในการเล่นเกมให้เทียบเคียงกับมนุษย์ ซึ่งองค์ความรู้ที่กำหนดได้สามารถนำไปประยุกต์ใช้ได้ในหลากหลายสาขาในอนาคต เช่น Robot contol เป็นต้น

## **วัตถุประสงค์**

### 1. เพื่อพัฒนาระบบปัญญาประดิษฐ์กับวีดิโอเกมด้วยวิธีการ Reinforcement Learning

### 2. เพื่อพัฒนาองค์ความรู้ด้านการ Reinforcement Learning

## **ขอบเขตของโครงงาน**

### 1. ออกแบบวิธีการด้วยวิธีการ Reinforcement Learning สำหรับพัฒนาพฤติกรรมของ NPC

### 2. เปรียบเทียบผลของ NPC ที่ใช้Algorithm ต่างกันของ Reinforcement Learning

### 3. วิเคราะห์ผลของเกมที่นำ NPC ที่ผ่านการพัฒนาโดย Reinforcement Learning และ NPC เดิม

## **ประโยชน์ที่คาดว่าจะได้รับ**

อุตสาหกรรมเกมในปัจจุบัน ความสามารถของเกมและ AI ยังคงมาจากการเขียนสคริปต์จากผู้พัฒนาเกม ผู้จัดทำโครงการจึงต้องการสร้างพื้นฐานของ AI ที่ทำการเรียนรู้โดยการลองผิดลองถูก ซึ่งเป็นรากฐานของ Reinforcement Learning เพื่อเป็นรากฐานในการนำ Reinforcement Learning ไปใช้เพื่อเพิ่มความท้าทายของตัวเกมต่อไป

# **บทที่ 2**

# **ทฤษฏีและงานวิจัยที่เกี่ยวข้อง**

# **ทฤษฏี**

## **การเรียนรู้แบบเสริมกำลัง (Reinforcement Learning)**

การเรียนรู้แบบเสริมกำลังเป็นส่วนหนึ่งของการเรียนรู้ด้วยเครื่อง (Machine Learning) ที่ทำการโดยนำปัญญาประดิษฐ์นำมาอยู่ภายใต้สภาพแวดล้อมหนึ่งที่ปัญญาประดิษฐ์ทำการตัดสินใจในการกระทำหนึ่ง ที่มาจากการสุ่มหรือ เลือกการกระทำจากข้อมูลที่อยู่ภายใต้สิ่งแวดล้อมที่กำหนด ซึ่งปัญญาประดิษฐ์ของการเรียนรู้แบบเสริมกำลังมีเป้าหมายคือ เลือกการกระทำที่ทำให้รับรางวัลที่ดีที่สุดในการแก้ปัญหาหนึ่ง ผ่านการลองผิดลองถูกของตัวปัญญาประดิษฐ์

การเรียนรู้แบบเสริมกำลังส่วนใหญ่จะถูกใช้ในอุตสาหกรรมหุ่นยนต์ และอุตสาหกรรมวิดีโอเกม เช่น ใช้การเรียนรู้แบบเสริมกำลังสร้างปัญญาประดิษฐ์ในการเล่นเกมคอมพิวเตอร์ Starcraft 2 หรือ สร้างปัญญาประดิษฐ์ควบคุมการทำงานของแขนกล

ซึ่งองค์ประกอบของการเรียนรู้แบบเสริมกำลังมีทั้งหมด 5 ส่วน

### **เอเจนต์ (Agent)**

ปัญญาประดิษฐ์ที่อยู่ภายใต้สภาพแวดล้อมและการทำงานของการเรียนรู้แบบเสริม กำลัง ซึ่งภายในเอเจนต์หนึ่งตัวจะมีส่วนประกอบภายใน 1 ประเภทหรือมากกว่า ซึ่งมีทั้งหมด 3 องค์ประกอบดังนี้

#### **Policy**

เป็นวิธีการที่จะให้ไปถึงเป้าหมายที่ต้องการเมื่ออยู่ในสถานะที่ต่างกันออกไป

#### **Value Function**

เป็นค่าที่ใช้วัดผลจากการกระทำในสถานะต่าง ๆ เพื่อวัดผลว่าถ้าทำการกระทำหนึ่ง ณ สถานะปัจจุบันให้ผลดีต่อรางวัลในอนาคตอย่างไร

#### **Model**

โมเดลเป็นการทำนายว่าในสภาพแวดล้อมจะเกิดอะไรขึ้นต่อไปทั้งสถานะและรางวัลที่จะได้จากการกระทำ

### **สภาพแวดล้อม (Environment)**

เป็นพื้นที่ทำการนำปัญญาประดิษฐ์ทำภารกิจที่กำหนดตามสภาพแวดล้อมที่กำหนด

### **สถานะ (State)**

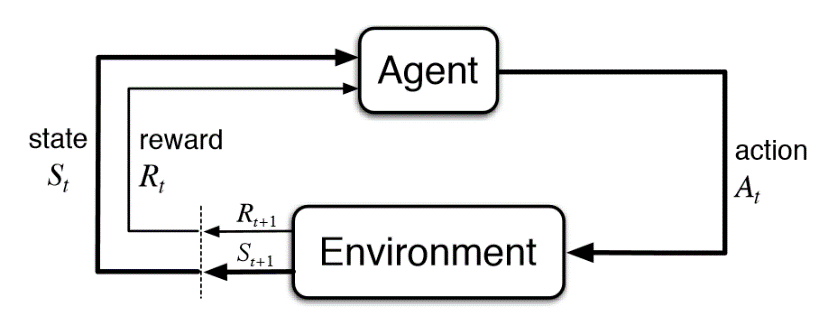
เป็นสถานะของสภาพแวดล้อมในช่วงเวลาต่างๆที่ปัญญาประดิษฐ์สามารถรับรู้เพื่อตัดสินใจเลือกการกระทำในแต่ละช่วงเวลา

### **การกระทำ (Action)**

เอเจนต์จะเลือกการกระทำที่ทำการตัดสินใจจากสถานะก่อนหน้าเข้าไปยังสภาพแวดล้อมเพื่อแสดงสถานะถัดไปและรางวัลที่เลือกการกระทำ

### **รางวัล (Reward)**

เป็นรางวัลที่ได้จากการกระทำเพื่อแสดงว่าการกระทำที่เลือกไปดีต่อกับสภาพแวดล้อมและเป้าหมายของสภาพแวดล้อมได้ดีเพียงใด



รูป 2.1 โครงสร้างการทำงานของการเรียนรู้แบบเสริมกำลัง

## **โครงข่ายประสาทแบบคอนโวลูชัน (Convolutional Neural Network)**

โครงข่ายประสาทแบบคอนโวลูชัน เป็นประเทหนึ่งของ Deep Learning ที่ทำการจำลองการมองเห็นของมนุษย์โดยการแบ่งเป็นส่วนย่อย และนำมารวมเป็นภาพรวมเพื่อแยกประเภทหรือหมวดหมู่ มักจะใช้ในการประมวลผลภาพสำหรับการฝึกปัญญาประดิษฐ์จำแนกประเภทด้วยภาพ ทำการแบ่งภาพเป็นว่านย่อย ๆ ในการจดจำรูปแบบในแต่ละกลุ่มของรูปหนึ่งรูป เพื่อจำแนกคุณลักษณะ เพื่อการจำแนกค่ารับเข้าในรูปแบบของรูปภาพได้

รูป 2.2 โครงสร้างของโครงข่ายประสาทแบบคอนโวลูชัน

องค์ประกอบของโครงข่ายคอนโวลูชันมีดังนี้

* + - 1. **Convolutional Layer**

เป็นชั้นที่ทำการสแกนค่ารับเข้าซึ่งเป็นรูปภาพ เพื่อแยกองค์ประกอบของรูป เช่น สี รูปทรง ขอบของภาพ

* + - 1. **Pooling Layer**

เป็นการลดขนาดของข้อมูลมีขนาดเล็กลงโดยที่รายละเอียดยังคงเดิม

ซึ่งทั้ง Convolutional Layer และ Pooling Layer จะทำงานคู่กันซึ่งจะทำงานหลายๆครั้งเพื่อจำแนกได้ครบทุกรูปแบบ

* + - 1. **Fully-Connected Layer**

เป็นชั้นที่มีค่านำเข้าเป็นข้อมูลจากการกระทำของข้อมูลจากชั้นก่อนหน้าสำหรับนำมาคำนวณเพื่อจำแนกประเภทจากข้อมูลที่ได้มา

## **งานวิจัยที่เกี่ยวข้อง**

กลุ่ม Deepmind (2556)ได้ทำการสร้างโมเดลสำหรับการเรียนรู้ โดยใช้โครงข่ายคอนโวลูชัน และ ใช้รูปแบบการเรียนรู้ Q-Learning ของการเรียนรู้แบบเสริมกำลัง ซึ่งเรียกว่า Deep Q Networkโดยใช้รูปภาพสำหรับค่านำเข้า และค่าส่งออกเป็นกราฟแสดงผลของรางวัลที่ได้จากการกระทำ และนำมาใช้โดยการนำเกมจากเครื่อง อาตาริ 2600 (Atari 2600) จำนวน 6 เกม และได้นำโมเดลมาใช้กับการเล่นเกม ซึ่งผลที่ได้คือมีทั้งหมด 3 เกมที่มีคะแนนที่มากกว่ามนุษย์

Yitao Liang, Marlos C. Machado, Erik Tavitie และ Michael Bowling (2559)ได้นำความสำเร็จของกลุ่ม Deepmind ในการทำโมเดล Deep Q Network ที่เป็นจุดเริ่มต้นของการเรียนรู้แบบเสริมกำลัง จึงสร้างโมเดลมาเปรียบเทียบกับโมเดลที่ Deepmind ได้ทำไว้โดยที่มีขนาดของโมเดลที่ขนาดที่เล็กลง เพื่อนำมาหาข้อดีข้อเสียและวัดผลกับโมเดล Deep Q Network ที่ทาง Deepmind ได้ทำไว้ โดยที่ผลสรุปที่ได้คือ Deep Q Network มีการประมวลผลที่เสถียรกว่า ผลการการเล่นเกมมีผลที่ดีกว่า แต่โมเดลที่สร้างมีบางเกมที่เอาชนะโมเดลจาก Deepmind ได้ดีกว่าในเกมประเภทที่เสียพลังชีวิตได้ง่าย เพราะ Deep Q Network จะทำการหยุดเรียนรู้ชั่วคราวเมื่อเอเจนต์เสียพลังชีวิตภายในเกม โมเดลที่สร้างจะเหมาะกับเกมที่เน้นการแก้ปัญหา หรือเกมที่เป้าหมายของเกมที่ไม่ใช่เล่นเพื่อให้ได้คะแนนสูงสุด

# **บทที่ 3**

# **วิธีการดำเนินงาน**

## **โปรแกรมหรือซอฟต์แวร์ที่ใช้ในการพัฒนา**

### ภาษาไพทอน สำหรับการเขียนโครงสร้างของงานวิจัย ซึ่งประกอบไปด้วย library ดังนี้

#### Gym

#### Gym-retro

#### Numpy

#### Mapplotlib

#### Tensorflow

### ไฟล์เกม Kaboom สำหรับการสร้าง สภาพแวดล้อม

## **ขั้นตอนการดำเนินงาน**

### คิดหัวข้อโครงงานเพื่อนำเสนออาจารย์ที่ปรึกษาโครงงาน

### ศึกษาและค้นคว้าการทำงานของไลบรารี่ต่างๆที่นำมาใช้ในการทำโครงงาน

### ศึกษาและค้นคว้าข้อมูลที่งานวิจัยที่เกี่ยวข้องกับ การเรียนรู้แบบเสริมกำลัง เพื่อนำมาประยุกต์ใช้งานกับโครงงานของผู้จัดทำ

### จัดทำการสร้างสภาพแวดล้อมของตัวเกมที่จะใช้ในการนำเสนองานวิจัย และใช้สภาพแวดล้อมที่สร้างขึ้นเพื่อทำการสอนให้คอมพิวเตอร์เรียนรู้และนำมาวัดผลการทดลอง

### จัดทำการสอนให้คอมพิวเตอร์เรียนรู้และตัดสินใจด้วยตัวเอง โดยใช้อัลกอริทึมที่เกี่ยวกับการเรียนรู้แบบเสริมกำลัง โดยทำการเรียนรู้ภายใต้สภาพแวดล้อมที่กำหนด

### จัดทำรูปแบบการนำเสนอของผลที่ได้จากการเรียนรู้ของคอมพิวเตอร์ ในรูปแบบกราฟแบบแผนภูมิเส้นและทำรายงานสรุปผลการทดลอง

### **บรรณานุกรม**