

**สำนักหอสมุดกลาง พระจอมเกล้าลาดกระบัง**

**โปรแกรมซื้อขายค่าเงินอัตโนมัติ**

**FOREX ROBOT TRADING**



T119273



เลขหมู่.....  
เลขทะเบียน **119273**  
วัน,เดือน,ปี. **- 6 S.ค. 2554**

b. **12.24.05**  
i.....

**ปริญญาบัตรนี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต  
สาขาวิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์  
สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง**

ปีการศึกษา 2553  
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้ในเพียงห้องที่ออกให้เท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ปริญญาโทบริหารศึกษาศาสตร์ 2553

สาขาวิชาวิศวกรรมคอมพิวเตอร์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง โปรแกรมซื้อขายค่าเงินอัตโนมัติ

FOREX ROBOT TRADING

ผู้จัดทำ

1. นายกรวิชญ์ พัววรานุเคราะห์ รหัสนักศึกษา 50010031
2. นายพงศภัค ถิ่นสันติสุข รหัสนักศึกษา 50011009
3. นายวุฒิพงษ์ ฉันทร์มิตรกุล รหัสนักศึกษา 50011516



*Yungh* *Keer*  
(รศ.ดร.บุญธีร์ เครือตราฐ)

อาจารย์ที่ปรึกษา

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# โปรแกรมซื้อขายค่าเงินอัตโนมัติ

นายกรวิชัย	พัวรานุเคราะห์	50010031
นายพงศภัค	ถีนสันติสุข	50011009
นายวุฒิพงษ์	ฉันทร์มิตรกุล	50011516
รศ. ดร. บุญธีร์	เครือตราชู	อาจารย์ที่ปรึกษา
ปีการศึกษา 2553		

## บทคัดย่อ

การเรียนรู้แบบเสริมกำลังเป็นเทคนิคการเรียนรู้ของคอมพิวเตอร์แบบหนึ่ง ซึ่งเป็นการเรียนรู้จากประสบการณ์ในอดีต ใช้ประสบการณ์ในอดีตเป็นตัวตัดสินใจการกระทำในปัจจุบัน ได้ทดลองนำการเรียนรู้แบบเสริมกำลังนี้ มาประยุกต์ใช้กับการซื้อขายอัตราแลกเปลี่ยนค่าเงิน ซึ่งเป็นเรื่องที่น่าสนใจและท้าทายความสามารถ ถ้าหากเราสามารถคิดอัลกอริทึมที่ทำให้ได้กำไรอย่างสม่ำเสมอทุกช่วงเวลาแล้ว มันจะมีประโยชน์ต่อผู้ที่สนใจเป็นอย่างมาก



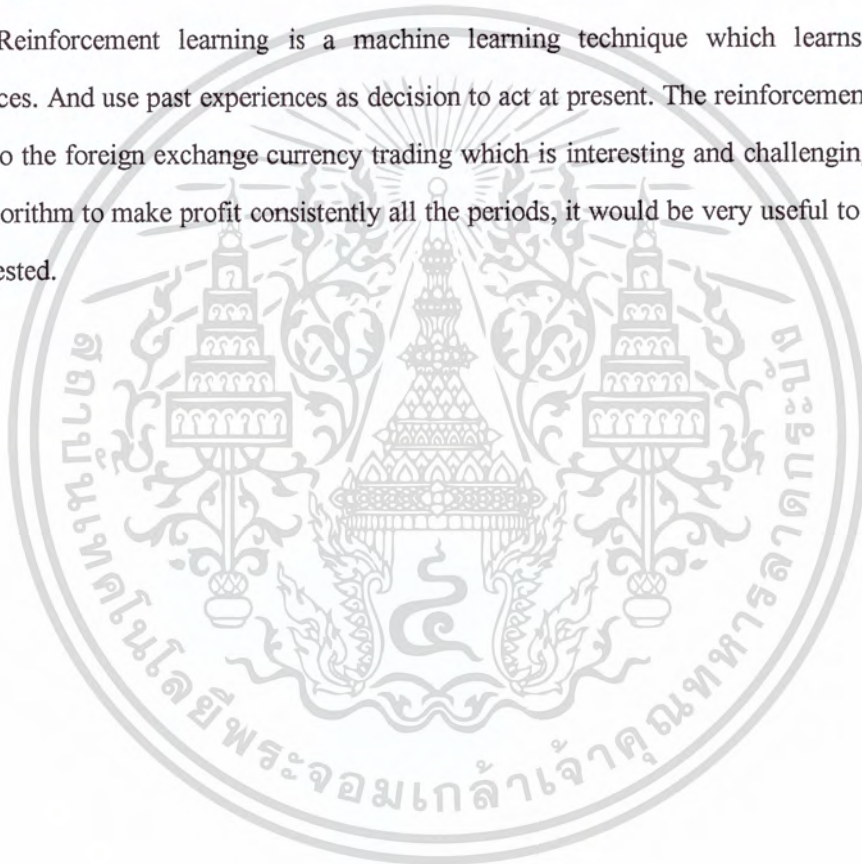
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# FOREX ROBOT TRADING

Mr. Kornravit	Puavaranukhor	50011031
Mr. Pongsapak	Tinsuntesuk	50011009
Mr. Vutipong	Shanmitgul	50011516
Assoc.Prof. Boontee	Kruatrachue	Advisor
Academic Year 2010		

## ABSTRACT

Reinforcement learning is a machine learning technique which learns from past experiences. And use past experiences as decision to act at present. The reinforcement learning is applied to the foreign exchange currency trading which is interesting and challenging. If we can think algorithm to make profit consistently all the periods, it would be very useful to people who are interested.



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## กิตติกรรมประกาศ

ขอขอบพระคุณอาจารย์ที่ปรึกษา รศ.ดร.บุญธิร์ เครือตราชู ที่กรุณาสละเวลาให้คำแนะนำ และคำปรึกษาโครงการนี้มาโดยตลอด และขอขอบพระคุณ คุณพ่อและคุณแม่ ผู้เป็นทั้งที่รักและให้กำลังใจให้โอกาสการศึกษา

กรวิชญ์ พัวรานูเคราะห์  
พงศภัค ถิ่นสันติสุข  
วุฒิพงษ์ ฉันทร์มิตรกุล



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญ

หน้า

บทคัดย่อภาษาไทย .....	I
บทคัดย่อภาษาอังกฤษ .....	II
กิตติกรรมประกาศ.....	III
สารบัญ.....	IV
สารบัญตาราง.....	VI
สารบัญรูป.....	VII

บทที่ 1 บทนำ.....	1
1.1 ความเป็นมาของปัญหา .....	1
1.2 วัตถุประสงค์ของรายงาน .....	1
1.3 ขอบเขตของโครงการ .....	1
1.4 ประโยชน์ที่คาดว่าจะได้รับ .....	1
1.5 ส่วนประกอบของรายงาน .....	1

บทที่ 2 การเรียนรู้แบบเสริมกำลัง.....	3
2.1 แนวคิดของการเรียนรู้แบบเสริมกำลัง .....	3
2.2 การโต้ตอบระหว่างเอเจนต์กับสภาพแวดล้อม .....	3
2.3 กระบวนการตัดสินใจแบบมาร์คอฟ .....	5
2.4 ฟังก์ชันมูลค่า.....	6
2.5 การประเมินค่าตอบแทน .....	7
2.6 อัลกอริทึม Monte Carlo.....	8

บทที่ 3 ตัวอย่างอัลกอริทึม Monte Carlo .....	10
3.1 อธิบายตัวอย่าง .....	10
3.2 สรุป .....	19

บทที่ 4 เครื่องมือสำหรับการวิเคราะห์และชี้วัดทางเทคนิค (Technical Indicator) .....	20
4.1 RSI (Relative Strength Index - เครื่องมือดัชนีกำลังสัมพัทธ์) .....	20

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญ (ต่อ)

	หน้า
4.2 MACD (MOVING AVERAGES CONVERGENCE/ DIVERGENCE) .....	22
4.3 OsMA (Moving Average of Oscillator – ค่าเฉลี่ยเคลื่อนที่ของกรอบการแกว่งตัว) ...	26
4.4 Alligator .....	28
4.5 Moving Average (เส้นค่าเฉลี่ย) .....	29
<b>บทที่ 5 โปรแกรม Meta Trader Alpari UK4.....</b>	<b>31</b>
5.1 การสมัครสมาชิก.....	31
5.2 การทำสัญญาซื้อขาย.....	31
5.3 การปิดสัญญาซื้อขาย.....	32
<b>บทที่ 6 อัลกอริทึม.....</b>	<b>33</b>
6.1 อัลกอริทึม.....	33
6.2 อัลกอริทึม เพิ่มเติม.....	38
6.3 อัลกอริทึม การ ดับเบิ้ลล็อต .....	40
<b>บทที่ 7 Meta Quotes Language 5.....</b>	<b>41</b>
7.1 ส่วนประกอบพื้นฐานของ Meta Trader Qoutes Language 5 .....	41
7.2 ส่วนการทำงานหลักของโปรแกรมที่สร้างจาก MQL 5 .....	47
<b>บทที่ 8 การทดลองและผลการทดลอง .....</b>	<b>49</b>
8.1 การทดลอง .....	49
<b>บทที่ 9 สรุปผลการทดลอง .....</b>	<b>52</b>
9.1 สรุปผลทดลอง .....	52
9.2 แนวทางการพัฒนาต่อ .....	52
<b>บรรณานุกรม .....</b>	<b>53</b>

# สารบัญตาราง

ตาราง

หน้า

7.1 เปรียบเทียบความแตกต่างกันของแต่ละภาษา..... 42



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# สารบัญรูป

รูป	หน้า
2.1 การโต้ตอบระหว่างเอเจนต์และสภาพแวดล้อมในการเรียนรู้แบบเสริมกำลัง .....	4
2.2 แบบคัพค่าผลตอบแทนของ (a) $V_{\pi}$ และ (b) $Q_{\pi}$ .....	7
2.3 ชูโดโค้ดของอัลกอริทึม Monte Carlo ด้วยวิธีการเรียนรู้แบบอนโพลีซี.....	9
3.1 สัญลักษณ์ต่างๆที่ควรทราบในตัวอย่างอัลกอริทึม Monte Carlo .....	10
3.2 อธิบายโครงสร้างของ state ต่างๆ.....	11
3.3 อธิบาย episode ที่ 1 .....	12
3.4 อธิบาย episode ที่ 1 หลัง update .....	13
3.5 อธิบาย episode ที่ 2 .....	14
3.6 อธิบาย episode ที่ 2 (ต่อ).....	15
3.7 อธิบาย episode ที่ 1,2 หลัง update แล้ว .....	16
3.8 อธิบาย episode ที่ 3 .....	17
3.9 อธิบาย episode ที่ 1,2,3 หลัง update แล้ว .....	19
4.1 การปรับตัวของค่า RSI รวมถึงขอบเขต oversold และ overbought.....	22
4.2 RSI จากโปรแกรม Terminal.....	22
4.3 MACD.....	23
4.4 Negative Divergence.....	25
4.5 Positive Divergence .....	25
4.6 MACD จาก โปรแกรม Terminal .....	26
4.7 การเปลี่ยนแปลงของตลาด และ OsMA .....	27
4.8 OsMA จาก โปรแกรม Terminal .....	27
4.9 Alligator .....	28
4.10 Aligator จากโปรแกรม Terminal.....	29
4.11 MA .....	29
5.1 การสมัครสมาชิก.....	31
5.2 การซื้อขาย .....	31
5.3 การปิดสัญญา.....	32
6.1 state ที่กำหนดโดยใช้ indicator .....	33
6.2 จำนวนค่าที่เป็นไปได้ทั้งหมดของ Trend .....	34

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## สารบัญรูป (ต่อ)

รูป	หน้า
6.3 การกำหนด episode .....	35
6.4 ข้อมูลที่เก็บ 1 วัน .....	36
6.5 ข้อมูลที่เก็บ 31 วัน .....	36
6.6 จุดเริ่มต้นที่จะนำมาคำนวณ .....	37
6.7 ตัวอย่างการคำนวณ โดยให้ความสำคัญตามระยะห่างของวัน .....	37
6.8 ซูโดโค้ด .....	38
6.9 จำลองการซื้อขายในการเทรดเพื่อหาตัวที่ดีที่สุด .....	39
6.10 ความแตกต่างด้านระยะเวลาซื้อขายของตัว 60%,70% และ 80% .....	39
7.1 MetaTraderTerminal .....	41
7.2 รายละเอียดของชนิดข้อมูลพื้นฐาน .....	43
7.3 การเก็บข้อมูลของ string.....	43
7.4 การแปลงชนิดข้อมูล เส้นทึบแสดงการ ไม่สูญเสียข้อมูล เส้นประแสดงการสูญเสียข้อมูล .....	44
7.5 ลำดับการแปลงชนิดข้อมูล .....	44
7.6 ตัวอย่าง event handling functions.....	47
8.1 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 0.01 ตั้งแต่วันที่ 1-1-2000 ถึง 26-8-2010.....	49
8.2 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 2-3-2000 ถึง 26-8-2010.....	49
8.3 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 7-1-2000 ถึง 26-8-2010.....	49
8.4 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 1 ตั้งแต่วันที่ 1-2-2000 ถึง 26-8-2010.....	50
8.5 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 4-1-2000 ถึง 26-8-2010.....	50
8.6 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 1 ตั้งแต่วันที่ 1-1-2000 ถึง 26-8-2010.....	50
8.7 ผลการทดลองโดยใช้อัลกอริทึมเพิ่มเติมและดับเบิ้ลล็อต .....	51

# บทที่ 1

## บทนำ

### 1.1 ความเป็นมาของปัญหา

การทำให้คอมพิวเตอร์มีความฉลาด เป็นสิ่งที่มนุษย์คิดอยากจะทำมาเป็นเวลานาน ตั้งแต่สมัยที่มนุษย์เริ่มมีคอมพิวเตอร์ในยุคแรกๆ ซึ่งเป็นต้นกำเนิดของปัญญาประดิษฐ์ หรือ AI (Artificial Intelligence) เป็นความฉลาดเทียมที่สร้างขึ้นให้กับสิ่งที่ไม่มีชีวิต การเรียนรู้ของเครื่อง (Machine learning) เป็นสาขาหนึ่งของปัญญาประดิษฐ์ ซึ่งพัฒนาเทคนิควิธี เพื่อให้คอมพิวเตอร์สามารถเรียนรู้โดยเน้นวิธีการสร้างโปรแกรมคอมพิวเตอร์จากการวิเคราะห์ชุดข้อมูล การเรียนรู้ของเครื่องจึงเกี่ยวข้องกับอย่างมากกับสถิติศาสตร์ ในปัจจุบันมีการศึกษาค้นคว้าการเรียนรู้ของเครื่องเพื่อนำมาใช้ให้เกิดประโยชน์อย่างต่อเนื่อง

การทำให้เครื่องคอมพิวเตอร์สามารถเรียนรู้ได้จึงเป็นสิ่งที่น่าสนใจในปัจจุบัน ซึ่งการเรียนรู้ของเครื่องสามารถแบ่งออกได้เป็นหลายประเภท หนึ่งในนั้นคือการเรียนรู้แบบเสริมกำลัง (Reinforcement learning) ซึ่งเป็นวิธีการเรียนรู้แบบใหม่ที่มีความน่าสนใจเนื่องจาก เป็นการเรียนรู้ที่มีการโต้ตอบระหว่างผู้เรียนกับสภาพแวดล้อม ผ่านการลองผิดลองถูก และเรียนรู้การกระทำที่ให้ผลระยะยาว นับว่าเป็นการเรียนรู้ที่เหมือนกับธรรมชาติจึงนำทฤษฎีการเรียนรู้แบบเสริมกำลังนี้มาประยุกต์ใช้กับอัตราซื้อขายค่าเงิน

### 1.2 วัตถุประสงค์ของรายงาน

เพื่อนำการเรียนรู้แบบเสริมกำลังมาประยุกต์ใช้กับอัตราซื้อขายค่าเงิน

### 1.3 ขอบเขตของโครงการ

นำการเรียนรู้แบบเสริมกำลังมาใช้กับอัตราซื้อขายค่าเงินแล้วได้กำไรอย่างต่อเนื่อง

### 1.4 ประโยชน์ที่คาดว่าจะได้รับ

- 1) สามารถนำการเรียนรู้แบบเสริมกำลังมาประยุกต์ใช้กับปัญหาต่างๆ ได้อย่างเหมาะสม
- 2) สามารถคิดวิเคราะห์และแก้ปัญหาต่างๆอย่างมีประสิทธิภาพ
- 3) สามารถทำงานเป็นทีมร่วมกับผู้อื่น

### 1.5 ส่วนประกอบของรายงาน

เอกสารนี้เป็นรายงานเล่มนี้ประกอบด้วยเนื้อหา 8 บท ดังนี้ ศึกษานี้ศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

บทที่ 1 กล่าวถึงความเป็นมาของปัญหาวัตถุประสงค์ของ โครงการขอบเขตของโครงการ  
ประโยชน์ที่คาดว่าจะได้รับและส่วนประกอบของรายงานฉบับนี้

บทที่ 2 กล่าวถึงทฤษฎีพื้นฐานและอัลกอริทึม Monte Carloของการเรียนรู้แบบเสริมกำลัง

บทที่ 3 กล่าวถึงตัวอย่างการเรียนรู้แบบเสริมกำลัง โดยใช้อัลกอริทึม Monte Carlo

บทที่ 4 กล่าวถึงเครื่องมือสำหรับวิเคราะห์และชี้วัดทางเทคนิค (Technical Indicator)

บทที่ 5 กล่าวถึงโปรแกรม MetaTrader 4 Alpari UK

บทที่ 6 อธิบาย อัลกอริทึม

บทที่ 7 กล่าวถึงโปรแกรม MetaTrader 5 Alpari UK

บทที่ 8 การทดลองและผลการทดลอง

บทที่ 9 สรุปผลการทดลอง



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 2

# การเรียนรู้แบบเสริมกำลัง

### 2.1 แนวคิดของการเรียนรู้แบบเสริมกำลัง

การเรียนรู้ผ่านการโต้ตอบกับสภาพแวดล้อมรอบตัวเรา คือสิ่งแรกปรากฏกับเรา เมื่อเราเรียนรู้ธรรมชาติเหมือนเด็กที่ออกไปวิ่งเล่นนอกบ้าน พวกเขาไม่มีครูคอยแนะนำ แต่พวกเขามีการโต้ตอบโดยตรงกับธรรมชาติ แบบฝึกหัดนี้สร้างข้อมูลที่มีค่าเกี่ยวกับเหตุและผลกระทบบนพื้นฐานของผลที่ตามมาของการกระทำ และเกี่ยวกับการทำอย่างไรจึงจะบรรลุเป้าหมาย ตลอดชีวิตของเรา การโต้ตอบกับสิ่งต่างๆ เป็นแหล่งสำคัญของความรู้เกี่ยวกับสภาพแวดล้อมรอบตัวเราและตัวเรา ไม่ว่าเราจะเรียนรู้การขับรถหรือทำอะไรก็ตาม สภาพแวดล้อมของเราตอบสนองต่อสิ่งที่เราทำและเราต้องการหาผลของสิ่งที่เกิดขึ้นจากการกระทำของเรา การเรียนรู้จากการติดต่อกับปฏิสัมพันธ์คือแนวคิดพื้นฐานของทฤษฎีเกือบจะทั้งหมดของการเรียนรู้และปัญญา

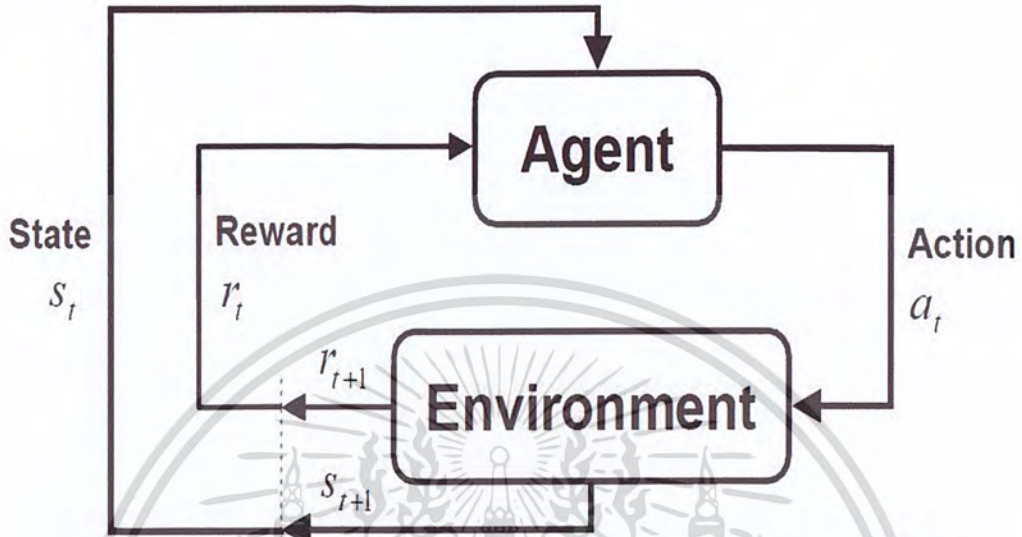
การเรียนรู้แบบเสริมกำลังคือการเรียนรู้สิ่งที่ทำ สถานการณ์นี้ควรทำอะไร ดังนั้นจึงใช้ตัวเลขรางวัลที่ได้รับเป็นสัญญาณบอกการกระทำ ผู้เรียนรู้ไม่ได้เป็นคนบอกการกระทำที่จะทำโดยตรง ดังเช่นในรูปแบบส่วนใหญ่ของการเรียนรู้ของเครื่อง (machine learning) แต่แทนที่ด้วยการต้องดูว่าการกระทำไหน ได้ผลรางวัลมากที่สุดจากความพยายามของพวกเขา ในส่วนใหญ่ของกรณีที่น่าสนใจและท้าทาย การกระทำอาจจะไม่มีผลต่อรางวัลที่ได้รับในทันที แต่มีผลกับสถานการณ์ถัดไปและรางวัลที่จะตามมาในภายหลังทั้งหมด การทดลองค้นหาข้อผิดพลาดและให้รางวัลอย่างซ้ำๆ เป็นสองคุณสมบัติเด่นที่สุดของการเรียนรู้แบบเสริมกำลัง

### 2.2 การโต้ตอบระหว่างเอเจนต์กับสภาพแวดล้อม

โจทย์ปัญหาเกี่ยวกับการเรียนรู้แบบเสริมกำลังเปรียบเสมือนกับ โครงสร้างพื้นฐานของปัญหาเกี่ยวกับการเรียนรู้ผ่านการโต้ตอบเพื่อให้บรรลุเป้าหมายที่ต้องการระบบที่ทำการตัดสินใจและเรียนรู้เรียกว่าเอเจนต์ (agent) ซึ่งเป็นระบบที่โต้ตอบกับสภาพแวดล้อม (environment) หรือระบบที่อยู่ภายนอกเอเจนต์ การโต้ตอบระหว่างกันของเอเจนต์กับสภาพแวดล้อมจะเป็นไปอย่างต่อเนื่องจากการเลือกการกระทำของเอเจนต์และการตอบสนองต่อการกระทำเหล่านั้นของสภาพแวดล้อมด้วยการให้ข้อมูลเกี่ยวกับสถานการณ์ปัจจุบัน

นอกจากนี้สภาพแวดล้อมยังเป็นสิ่งที่กำหนดรางวัลที่เอเจนต์จะได้รับซึ่งเป็นผลของการตัดสินใจของเอเจนต์ตลอดอายุการเรียนรู้ของเอเจนต์รางวัลที่มีค่าสูงที่สุดคือสิ่งที่เอเจนต์พยายาม

เรียนรู้และปรับตัวเพื่อให้ได้มารายละเอียดทั้งหมดของสภาพแวดล้อมเป็นสิ่งที่กำหนดทาสก์ (task) ซึ่งเป็นหนึ่งในองค์ประกอบของโจทย์ปัญหาเกี่ยวกับการเรียนรู้แบบเสริมกำลัง



รูป 2.1 การโต้ตอบระหว่างเอเจนต์และสภาพแวดล้อมในการเรียนรู้แบบเสริมกำลัง

รูป 2.1 แสดงแผนภาพการโต้ตอบระหว่างเอเจนต์กับสภาพแวดล้อมตามลำดับขั้นเวลา  $t = 0, 1, 2, 3, \dots$  ที่แต่ละลำดับขั้นเวลา  $t$  เอเจนต์จะได้รับข้อมูลเกี่ยวกับสถานะ (state) ของสภาพแวดล้อมซึ่งแทนด้วยสัญลักษณ์  $s_t \in S$  เมื่อ  $S$  คือเซตของสถานะทั้งหมดที่เป็นไปได้และเลือกการกระทำ (action) ตามสถานะนั้นซึ่งแทนด้วยสัญลักษณ์  $a_t \in A(s_t)$  เมื่อ  $A(s_t)$  คือเซตของการกระทำที่สามารถทำได้ในสถานะ  $s_t$  ในลำดับขั้นเวลาขั้นถัดไปซึ่งเป็นผลจากการกระทำของเอเจนต์เอเจนต์จะได้รับรางวัล (reward) ในรูปของค่าตัวเลข  $r_{t+1} \in \mathbb{R}$  และสถานะของสภาพแวดล้อมจะเปลี่ยนไปเป็น  $s_{t+1}$  การตัดสินใจเลือกการกระทำของเอเจนต์ในแต่ละลำดับขั้นเวลาจะเป็นไปตามฟังก์ชันซึ่งกำหนดค่าความเป็นไปได้ในการเลือกการกระทำในสถานะต่างๆของสภาพแวดล้อมฟังก์ชันนี้เรียกว่า โพลีซี (policy) ซึ่งแทนด้วยสัญลักษณ์  $\pi$  โดย  $\pi_{(s,a)}$  หมายถึงค่าความน่าจะเป็นที่  $a_t = a$  เมื่อ  $s_t = s$  วิธีการต่างๆในการเรียนรู้แบบเสริมกำลังจะกำหนดการปรับเปลี่ยนโพลีซีของเอเจนต์ผ่านประสบการณ์ของเอเจนต์เองเป้าหมายของเอเจนต์คือการทำให้ได้มาซึ่งค่าคาดหวังของผลตอบแทน (expected return) ที่สูงที่สุดในที่นี้ผลตอบแทน  $R_t$  หมายถึงฟังก์ชันของลำดับของรางวัลที่ได้รับหลังจากลำดับขั้นเวลา  $t$  ซึ่งสามารถเขียนได้ในรูปแบบของลำดับเลขคณิต  $r_{t+1}, r_{t+2}, r_{t+3}, \dots$  ฟังก์ชัน  $R_t$  สามารถเขียนให้อยู่ในรูปทั่วไปได้ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

$$R_t = \sum_{k=0}^T \gamma^k r_{t+k+1} \quad (2.1)$$

เมื่อ  $\gamma$  เป็นค่าอัตราลดค่า (discount rate) และ  $T$  เป็นค่าของลำดับขั้นเวลาขั้นสุดท้ายในกรณีที่ทำสำเนาของเอเจนต์สามารถแบ่งเป็นเอพิโซด (episode) ได้  $\gamma$  จะมีค่าเป็น 1 และ  $T$  จะต้องมียกเว้นค่าจำกัดซึ่งไม่เท่ากับ  $\infty$  ในกรณีที่สำเนาของเอเจนต์เป็นแบบต่อเนื่อง  $\gamma$  จะมีค่าอยู่ในช่วง  $(0, 1)$  และ  $T$  จะมีค่าเข้าใกล้  $\infty$  เนื่องจากเป็นทาสก์ที่ไม่มีสถานะจบและไม่มีการเริ่มต้นเอพิโซดใหม่

ตัวอย่างที่เป็นแบบเอพิโซด เช่น การแข่งขันเทนนิส อาจจะทำให้แต้มหนึ่งของการเล่นเป็นหนึ่งเอพิโซด ภายในเอพิโซดก็อาจจะมีสถานะว่าตอนนี้อยู่ที่จุดไหนของคอร์ด และการกระทำก็คือการตีลูกไปลงจุดไหน เมื่อมีผลแพ้ชนะในแต้มนั้นก็จบเอพิโซดแล้วก็นำค่ารีเทิร์นนี้ไปคูณแต่ละสถานะแล้วก็เริ่มเอพิโซดใหม่

ตัวอย่างที่เป็นแบบต่อเนื่อง เช่น หุ่นยนต์ที่คอยประคองไม้ที่ตั้งอยู่บนตัวมันไม่ให้ล้ม สถานะก็คือไม้เอียงไปทางไหนมากเท่าไร ส่วนการกระทำก็คือต้องเลื่อนตัวหุ่นยนต์ไปทางไหนระยะทางเท่าไรจึงจะทำให้ไม้ที่ตั้งอยู่ไม่ล้ม กรณีแบบนี้จะไม่มีวันจบไม่มีเอพิโซดดังนั้นค่าก็ต้องปรับไปเรื่อยๆ โดยค่า  $\gamma$  จะลดลงไปตามระยะทาง ซึ่งก็คือส่วนที่กำลังทำอยู่จะได้รับรีเทิร์นไปมากกว่าส่วนที่อยู่หลังๆ ที่ผ่านมานานแล้วพอไกลมากก็จะได้รีเทิร์นที่เข้าใกล้ศูนย์หรือเรียกได้ว่าไม่ได้รับรีเทิร์นเลย

### 2.3 กระบวนการตัดสินใจแบบมาร์คอฟ

โดยทั่วไปทาสก์ในการเรียนรู้แบบเสริมกำลังส่วนใหญ่สามารถอธิบายได้ในรูปแบบกระบวนการตัดสินใจแบบมาร์คอฟ (Markov decision process: MDP) ซึ่งประกอบด้วยเซตของสถานะที่เป็นไปได้เซตของการกระทำทั้งหมดที่เป็นไปได้ความน่าจะเป็นของการเปลี่ยนสถานะ (transition probabilities) และค่าคาดหวังของรางวัลที่จะได้รับโดยทันที (expected immediate reward) ความน่าจะเป็นของการเปลี่ยนสถานะจากสถานะ  $s$  ไปสถานะ  $s'$  เมื่อเลือกการกระทำ  $a$  สามารถเขียนในรูปทั่วไปได้ดังนี้

$$P_a(s, s') = Pr(s_{t+1} = s' | s_t = s, a_t = a) \quad (2.2)$$

และค่าคาดหวังของรางวัลที่จะได้รับโดยทันทีหลังจากเปลี่ยนสถานะจากสถานะ  $s$  ไปสถานะ  $s'$  จากการเลือกการกระทำ  $a$  สามารถเขียนในรูปทั่วไปได้ดังนี้

$$R_a(s, s') = E(r_{t+1} | s_t = s, a_t = a, s_{t+1} = s') \quad (2.3)$$

MDP มีประโยชน์ในการกำหนดรายละเอียดทั้งหมดเกี่ยวกับสภาพแวดล้อมในการเรียนรู้แบบเสริมกำลังและมีส่วนสำคัญในการทำความเข้าใจอัลกอริทึมที่เกี่ยวกับการเรียนรู้แบบเสริมกำลังเนื่องจาก MDP สามารถอธิบายการเปลี่ยนแปลงต่างๆที่เกิดขึ้นในการเรียนรู้แบบเสริมกำลังได้เป็นอย่างดี

## 2.4 ฟังก์ชันมูลค่า

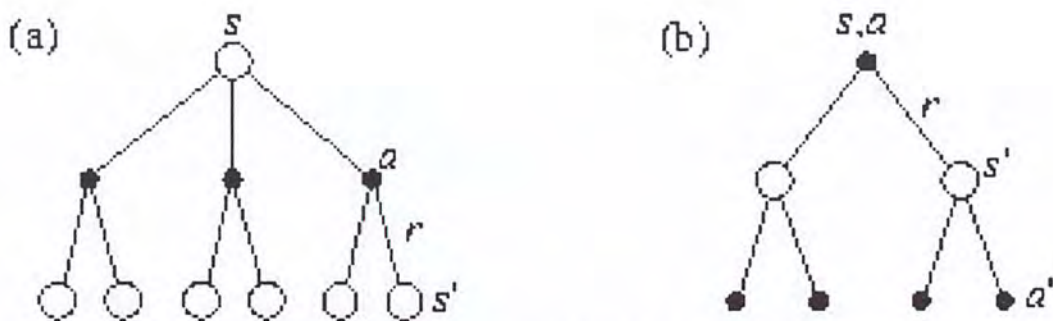
ฟังก์ชันมูลค่า (value function) คือฟังก์ชันที่ใช้ในอัลกอริทึมที่เกี่ยวกับการเรียนรู้แบบเสริมกำลังในการประเมินมูลค่าหรือประโยชน์ของการอยู่ในสถานะที่กำหนดหรือประโยชน์ของการเลือกการกระทำใดๆในสถานะที่กำหนดซึ่งเป็นฟังก์ชันของสถานะหรือฟังก์ชันของกลุ่มของสถานะและการกระทำซึ่งมูลค่าในที่นี้กำหนดโดยค่าคาดหวังของผลตอบแทนและแปรเปลี่ยนไปตามโพลีซีที่ใช้มูลค่าของสถานะ  $s \in S$  ภายใต้โพลีซี  $\pi$  ซึ่งแทนด้วยสัญลักษณ์  $V_{\pi(s)}$  หมายถึงค่าคาดหวังของผลตอบแทนซึ่งนับเริ่มต้นจากสถานะ  $s$  และดำเนินไปตามโพลีซี  $\pi$  ซึ่งเป็นตัวกำหนดความน่าจะเป็น  $\pi_{(s,a)}$  ในการเลือกการกระทำ  $a \in A(s)$  เมื่ออยู่ในสถานะ  $s$  การคำนวณค่า  $V_{\pi(s)}$  สามารถเขียนในรูปของสมการสำหรับ MDP ได้ดังนี้

$$V^\pi(s) = E_\pi(\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s) \quad (2.4)$$

เมื่อ  $E_\pi$  หมายถึงค่าคาดหวังของผลตอบแทนเมื่อกำหนดให้เอเจนต์ดำเนินตามโพลีซี  $\pi$  และ  $t$  หมายถึงลำดับขั้นเวลาใดๆในกรณีที่หาสภาวะสิ้นสุดมูลค่าของสถานะสิ้นสุดจะมีค่าเป็นศูนย์เสมอฟังก์ชัน  $V_\pi$  นี้เรียกว่าฟังก์ชันมูลค่าของสถานะตามโพลีซี  $\pi$  (state-value function for policy  $\pi$ ) ในทำนองเดียวกันมูลค่าของการเลือกการกระทำ  $a$  ในสถานะ  $s$  ภายใต้โพลีซี  $\pi$  ซึ่งแทนด้วยสัญลักษณ์  $Q_{\pi(s,a)}$  หมายถึงค่าคาดหวังของผลตอบแทนซึ่งนับเริ่มต้นจากสถานะ  $s$  ผ่านตัดสินใจเลือกการกระทำ  $a$  และดำเนินไปตามโพลีซี  $\pi$  การคำนวณค่า  $Q_{\pi(s,a)}$  สามารถเขียนในรูปของสมการสำหรับ MDP ได้ดังนี้

$$Q^\pi(s, a) = E_\pi(R_t | s_t = s, a_t = a) = E_\pi(\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a) \quad (2.5)$$

ฟังก์ชัน  $Q_\pi$  นี้เรียกว่าฟังก์ชันมูลค่าของการกระทำตามโพลีซี  $\pi$  (action-value function for policy  $\pi$ )



รูป 2.2 แบบจำลองค่าผลตอบแทนของ (a)  $V_\pi$  และ (b)  $Q_\pi$

รูป 2.2 แสดงค่าผลตอบแทนสำหรับ (a)  $V_\pi$  และ (b)  $Q_\pi$  โหนดบนสุดคือ โหนดรูท ด้านล่างคือความเป็นไปได้ของการกระทำและสถานะทั้งหมดที่ปรากฏ จะเห็นว่าค่าของ  $V_\pi$  จะเป็นตัวบอกว่าคุณค่าที่คาดหวังของสถานะนั้นดีหรือเปล่า ส่วนค่า  $Q_\pi$  จะเป็นตัวบอกว่าคุณค่าที่สถานะนี้ทำการกระทำนี้ดีหรือเปล่า

## 2.5 การประเมินค่าตอบแทน

คุณสมบัติที่สำคัญที่สุดของการเรียนรู้แบบเสริมกำลัง แตกต่างจากการเรียนรู้ประเภทอื่นๆ คือใช้ข้อมูลที่ได้การฝึกอบรมมาประเมินการกระทำมากกว่าการตั้งให้กระทำในสิ่งที่คุณนี่คือสิ่งที่สร้างความจำเป็นในการสำรวจสำหรับการทดลองและค้นหาข้อผิดพลาดเพื่อการกระทำที่ดีโดยทั่วไป วิธีการตามแนวทางการเรียนรู้แบบเสริมกำลังจะมุ่งเน้นไปที่การแก้ปัญหาสองเรื่องที่มีความสัมพันธ์กันคือปัญหาการทำนายค่า (prediction problem) ซึ่งเป็นการเรียนรู้ฟังก์ชันมูลค่าเมื่อดำเนินไปตามโพลีซีใดๆ และปัญหาการควบคุม (control problem) ซึ่งเป็นการปรับเปลี่ยนหรือค้นหาโพลีซีเพื่อให้ได้มาซึ่งโพลีซีที่เหมาะสมที่สุดซึ่งสำคัญสำหรับการได้มาซึ่งผลตอบแทนที่สูงที่สุดในการเรียนรู้แบบเสริมกำลังคือการปรับสมดุลระหว่างการสำรวจและการใช้ประโยชน์จากความรู้อยู่เดิมซึ่งเป็นผลจากการตัดสินใจเลือกการกระทำของเอเจนต์วิธีการเลือกการกระทำแบบเอพซิลอน-กริดิ (ε-greedy) เป็นวิธีหนึ่งที่เป็นที่นิยมที่สุดในวิธีการตัดสินใจเลือกการกระทำในการเรียนรู้แบบเสริมกำลังซึ่งใช้การเลือกการกระทำที่มีมูลค่าสูงที่สุดเป็นส่วนใหญ่เพื่อใช้ประโยชน์จากความรู้อยู่เดิมแต่ในบางครั้งจะใช้การสุ่มเลือกการกระทำเพื่อสำรวจการกระทำอื่นๆ ไปด้วยโดยกำหนดให้ความน่าจะเป็นในการสุ่มเลือกการกระทำที่มีมูลค่าสูงที่สุดมีค่าเป็น  $1 - \epsilon + \epsilon / A(s)$  และ ความน่าจะเป็นในการสุ่มเลือกการกระทำอื่นๆ มีค่าเป็น  $\epsilon / A(s)$  เมื่อ A หมายถึงจำนวนการกระทำที่เป็นไปได้และ  $\epsilon$  เป็นอัตราการสำรวจซึ่งมักกำหนดด้วยค่าน้อยๆ แต่ไม่เท่ากับศูนย์

นอกจากวิธีการเลือกการกระทำแบบเอพซิลอน-กริดิแล้ววิธีการเลือกการกระทำที่สำคัญอีกวิธีหนึ่งคือการเลือกการกระทำแบบซอฟท์แมกซ์ (softmax) ข้อดีของวิธีซอฟท์แมกซ์คือความน่าจะเป็นในการสุ่มเลือกการกระทำที่มีค่าสูงจะมีค่ามากถึงแม้ว่าการกระทำนั้นจะไม่ใช่อการกระทำที่มีมูลค่าสูงที่สุดทำให้ความน่าจะเป็นในการเลือกการกระทำที่มีค่าน้อยมีน้อยกว่าความน่าจะเป็น

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ในการเลือกการกระทำที่มีค่าใกล้เคียงกับการกระทำที่มีค่าสูงที่สุดซึ่งต่างจากวิธีการเลือกการกระทำแบบเอพซิลอน-กริดที่ใช้การสุ่มเลือกการกระทำอื่นๆที่ไม่ใช่การกระทำที่มีค่าสูงที่สุดด้วยความน่าจะเป็นที่เท่าๆกัน โดยทั่วไปวิธีการแบบซอฟต์แวร์แมกซ์ที่นิยมที่สุดจะการใช้การแจกแจงแบบกิบส์ (Gibbs distribution) หรือที่เรียกโดยทั่วไปว่าการแจกแจงโบลทซ์มันน์ (Boltzmann distribution) ซึ่งกำหนดความน่าจะเป็นในการเลือกการกระทำ  $a$  ณ เวลา  $t$  ดังสมการ

$$Pr(a_t = a) = \frac{e^{Q_t(a)/\tau}}{\sum_{b=1}^n e^{Q_t(b)/\tau}} \quad (2.6)$$

เมื่อ  $Q(a)$  เป็นฟังก์ชันมูลค่าของการกระทำและ  $\tau$  หมายถึงพารามิเตอร์ที่เรียกว่าอุณหภูมิ (temperature) ซึ่งมีค่าเป็นบวกค่าอุณหภูมิที่สูงจะส่งผลให้ความน่าจะเป็นในการเลือกการกระทำต่างๆมีค่าเท่าๆกันอุณหภูมิที่ต่ำจะส่งผลให้เกิดความแตกต่างของความน่าจะเป็นในการเลือกการกระทำที่มีมูลค่าแตกต่างกันเมื่อค่าพารามิเตอร์  $\tau$  มีค่าเข้าใกล้ศูนย์การเลือกการกระทำจะให้ผลเหมือนกับการเลือกการกระทำแบบกริด

## 2.6 อัลกอริทึม Monte Carlo

อัลกอริทึม Monte Carlo เป็นหนึ่งในอัลกอริทึมสำคัญที่ใช้ในการเรียนรู้แบบเสริมกำลัง โดยอัลกอริทึมนี้ค่าฟังก์ชันมูลค่า ( $V_\pi$  หรือ  $Q_\pi$ ) ของแต่ละสถานะในเอพิโสดจะถูกคำนวณใหม่เมื่อจบเอพิโสด โดยเกิดจากการเฉลี่ยค่าฟังก์ชันมูลค่าของสถานะนั้น (ซึ่งก็คือค่าผลตอบแทน) กับทุกเอพิโสดที่เคยเกิดสถานะนั้น โดยประสบการณ์จะได้รับก็ต่อเมื่อจบเอพิโสดนั้นแล้วเท่านั้น โดยวิธีการคำนวณค่าผลตอบแทนสามารถแบ่งออกได้เป็นสองประเภทคือ first visit และ every visit วิธีการของ first visit ก็คือภายในเอพิโสดเดียวกันถ้าเกิดมีสถานะเดียวกันทำซ้ำกันกี่ครั้งก็ตามจะนำสถานะนั้นมาคำนวณเหมือนเกิดขึ้นครั้งเดียว ส่วนวิธีการ every visit จะนำมาคำนวณตามจำนวนสถานะที่เกิดขึ้นจริงเลย

อัลกอริทึม Monte Carlo ยังแบ่งวิธีการเรียนรู้ได้อีกสองชนิดคือ วิธีการเรียนรู้แบบออน โพลีซี (on-policy learning) ซึ่งใช้การเรียนรู้ฟังก์ชันมูลค่าของ โพลีซีที่เลือกใช้ในการตัดสินใจและวิธีการเรียนรู้แบบออฟ โพลีซี (off-policy learning) ซึ่งใช้การเรียนรู้ฟังก์ชันมูลค่าสำหรับ โพลีซีแบบหนึ่งขณะที่ตัดสินใจด้วยโพลีซีอีกแบบหนึ่งที่แตกต่างกัน รูป 2.3 แสดงชุด โค้ดของอัลกอริทึม Monte Carlo ด้วยวิธีการเรียนรู้แบบออน โพลีซี ที่ใช้โพลีซีชนิดเอพซิลอน-กริด

เริ่มต้นสำหรับทุก  $s \in S, a \in A(s)$  :

กำหนดค่า  $Q(s, a)$  ตามต้องการ

กำหนดค่า  $\text{Return}(s, a)$  เป็นลิสต์ที่ว่างเปล่า

กำหนด  $\pi$  เป็นชนิดเอพิลอน-กริดิ

ทำซ้ำตลอดไป:

สร้างเอพิโซดมาหนึ่งเอพิโซด โดยใช้  $\pi$

สำหรับทุกคู่  $(s, a)$  ที่เกิดขึ้น ในเอพิโซด:

$R \leftarrow$  ค่าผลตอบแทน (เอาเฉพาะที่เกิดขึ้นครั้งแรกของ  $(s, a)$ )

นำ  $R$  ใส่ต่อท้ายเข้าไปใน  $\text{Return}(s, a)$

$Q(s, a) \leftarrow \text{average}(\text{Return}(s, a))$

สำหรับแต่ละ  $s$  ในเอพิโซด:

$a^* \leftarrow \text{argmax} Q(s, a)$

สำหรับทุก  $a \in A(s)$ :

$$\pi(s, a) \leftarrow \begin{cases} 1 - \epsilon + \epsilon/|A(s)| & \text{if } a = a^* \\ \epsilon/|A(s)| & \text{if } a \neq a^* \end{cases}$$

รูป 2.3 ชูโดไลด์ของอัลกอริทึม Monte Carlo ด้วยวิธีการเรียนรู้แบบออนโพลีซี

จากรูปจะเห็นว่าเมื่อมีการทำเอพิโซดใหม่ก็จะมีการปรับค่า  $Q(s, a)$  และเมื่อ  $Q(s, a)$  เปลี่ยนก็มีการปรับค่าโพลีซีใหม่ และค่าโพลีซีนีก็จะนำไปใช้เลือก  $Q(s, a)$  ใหม่ในเอพิโซดถัดไป ทำซ้ำแบบนี้ไปเรื่อยๆ จนสุดท้ายค่าที่ได้จะลู่เข้าสู่จุดสมดุล

## บทที่ 3

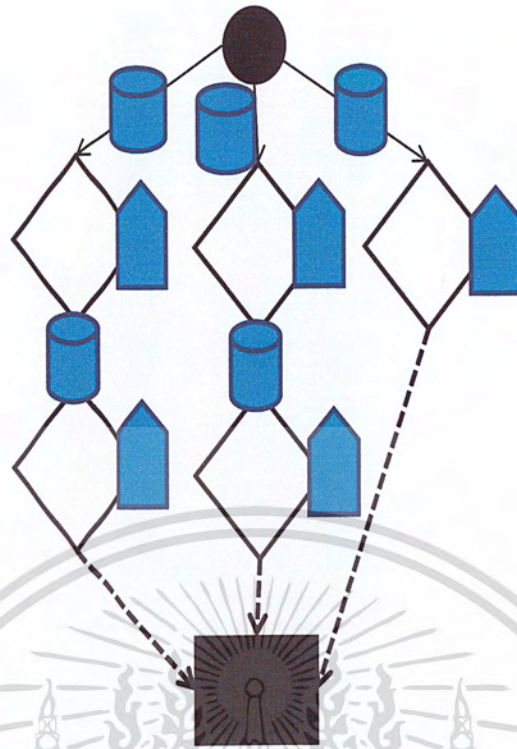
### ตัวอย่างอัลกอริทึม Monte Carlo

#### 3.1 อธิบายตัวอย่าง

เราจะใช้วิธีการเรียนรู้แบบ Reinforcement Learning มาหาว่าการตัดสินใจในเส้นทางต่างของชีวิตแบบไหนดีที่สุดโดยวัดจากเงินที่ได้ โดยเราจะมีตัวอย่างให้ 3 ตัวอย่าง episode ที่ 1 แสดงการเรียนรู้แล้วตายก่อนได้ทำงาน episode ที่ 2 แสดงการเรียนรู้แล้วได้ทำงาน episode ที่ 3 แสดงการเรียนรู้ในที่มีคนตายแต่ได้ทำงาน โดยในตัวอย่างจะใช้สัญลักษณ์ ในรูปดังนี้



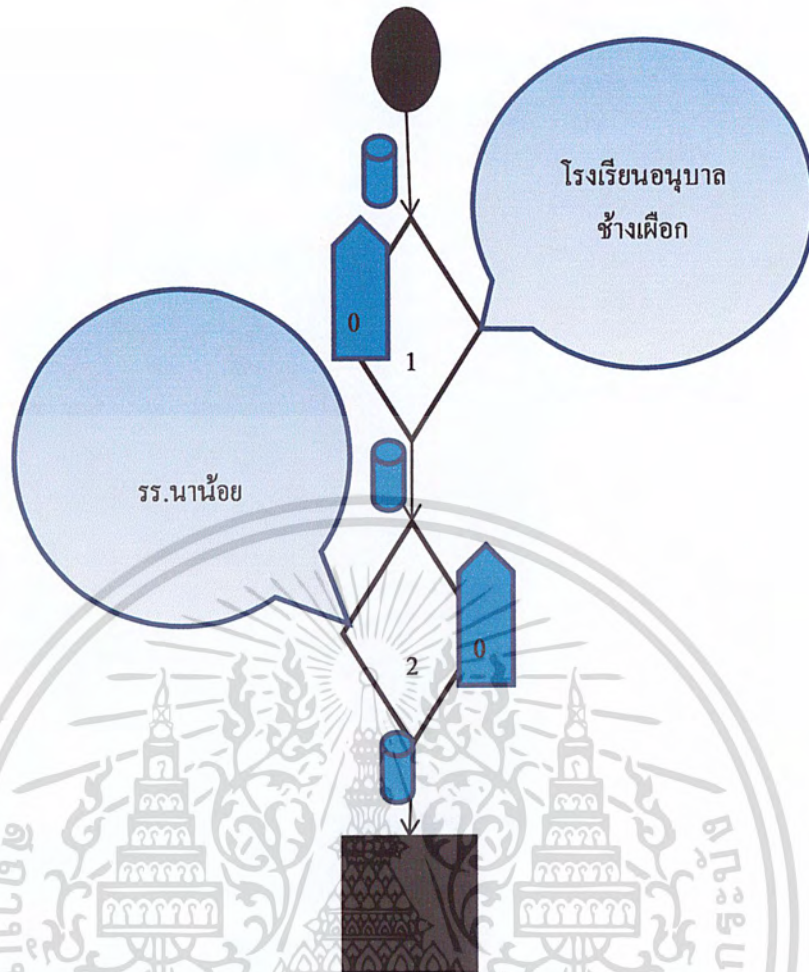
เราสรุปได้ว่าทุก episode เริ่มต้นด้วยการเกิดเหมือนกันและจบที่การตายเหมือนกันจะเป็นไปตามรูป 3.2



รูป 3.2 อธิบายโครงสร้างของ state ต่างๆ



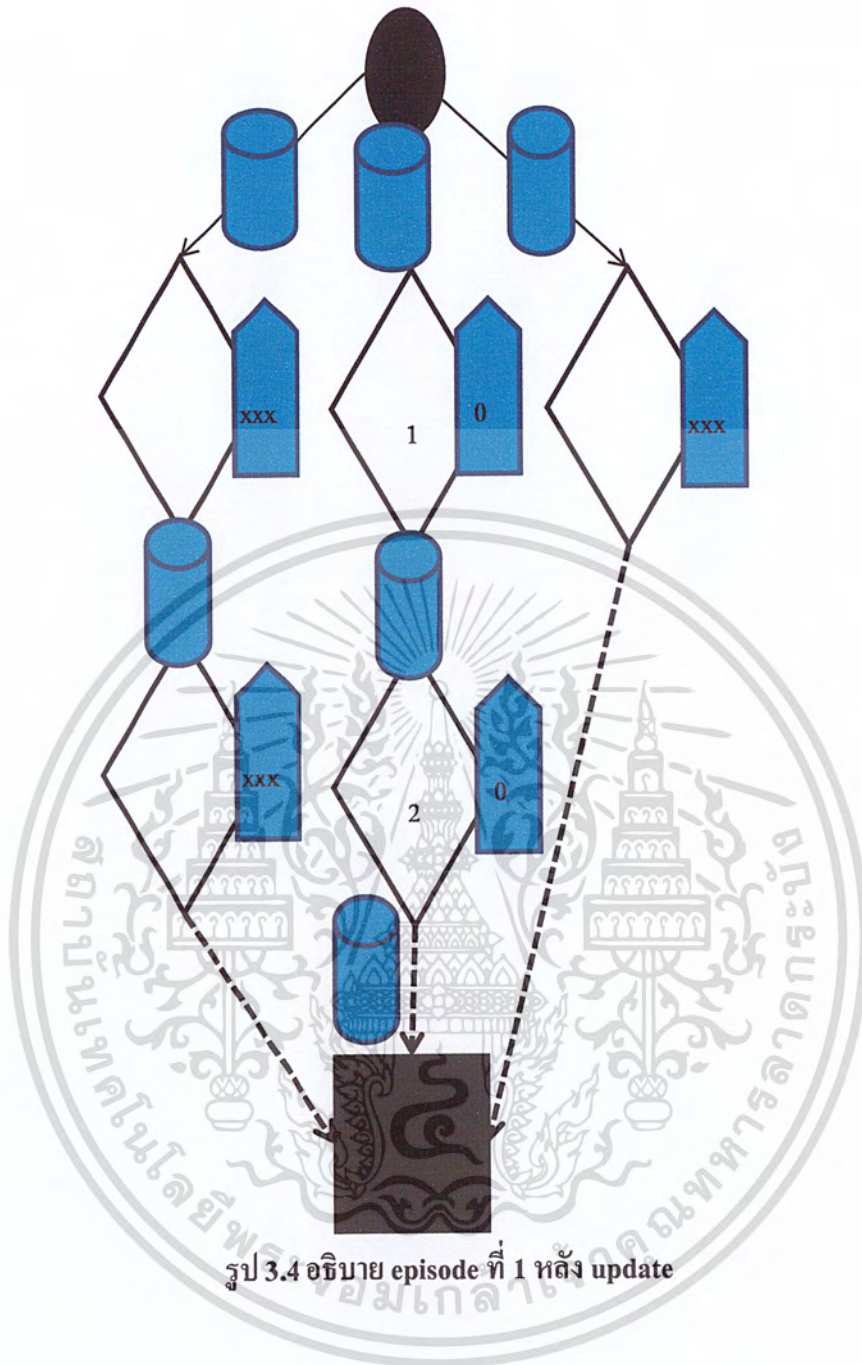
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.3 อธิบาย episode ที่ 1

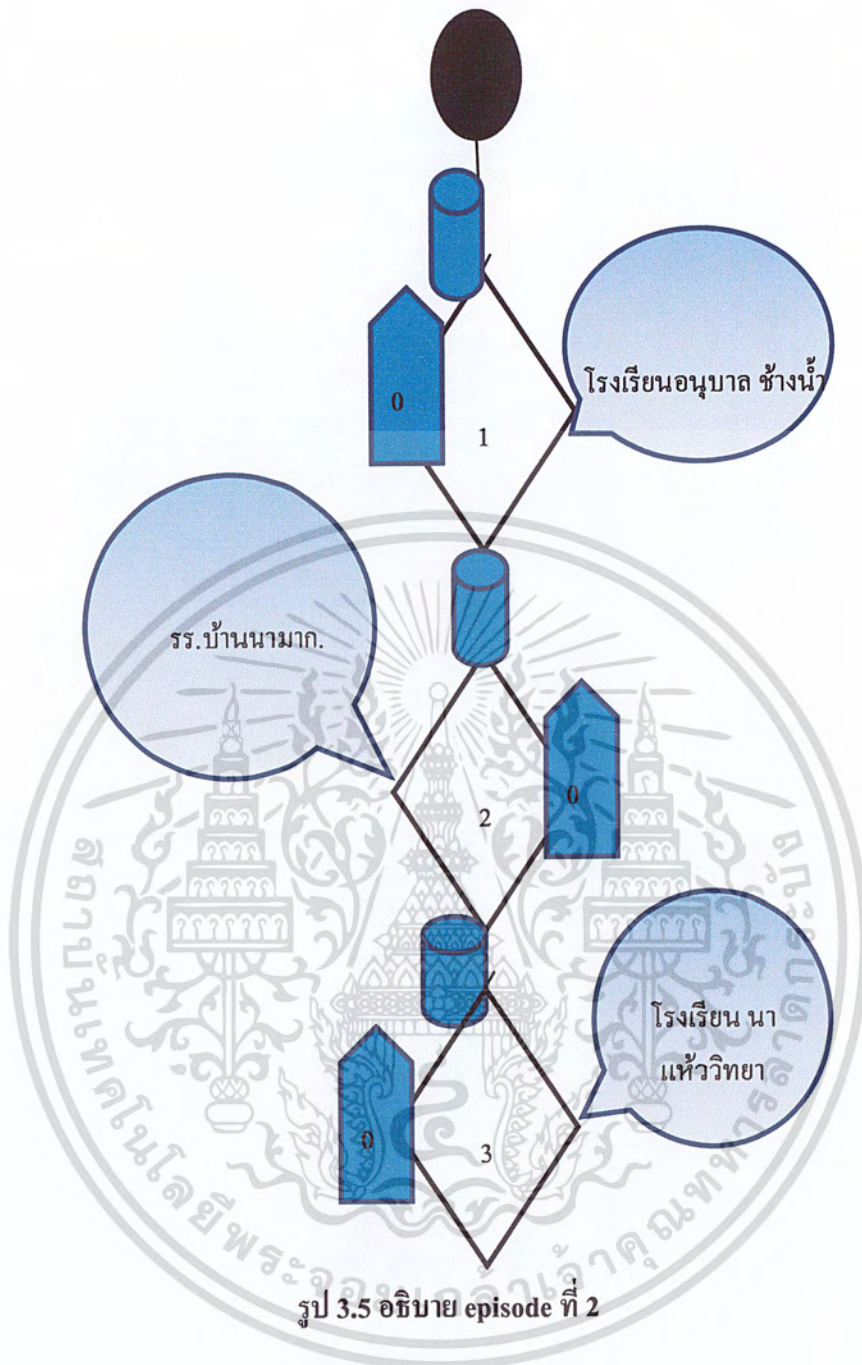
จากรูป 3.3 จะเห็นได้ว่าค่า  $Q$  ของรูปจะมีการเปลี่ยนจากค่าเริ่มต้นมาเป็นค่า 0 โคนมีวิธีคำนวณดังนี้ โดยทุกตัวต้อง + ด้วย 0 จากการตาย  $Q$  (เกิดมา, โรงเรียนอนุบาล ช้างเผือก) จาก state ที่ 0  $\text{Return} = 0$  (จากการเรียนอนุบาล) + 0 (จากการเรียนประถม),  $Q$  (โรงเรียนอนุบาล ช้างเผือก, ไปเรียนที่รร.บ้านน้อย)  $\text{Return} = 0$  (จากการเรียนประถม),  $Q$  (ไปเรียนที่รร.บ้านน้อย, ตาย)  $\text{Return} = 0$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

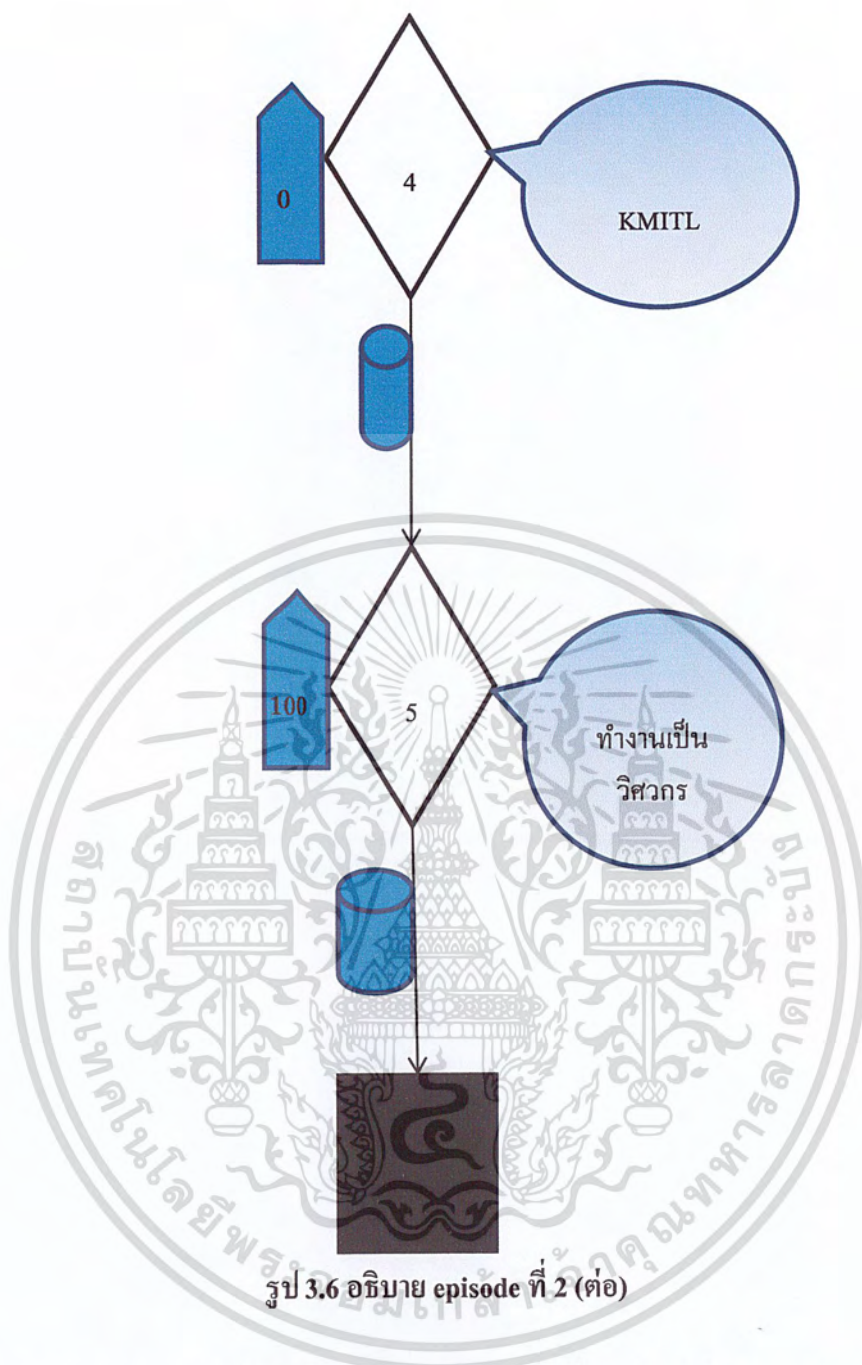


รูป 3.4 อธิบาย episode ที่ 1 หลัง update

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



จากรูป 3.5 และ 3.6 จะเห็นได้ว่าค่า Q ของรูปจะมีการเปลี่ยนจากค่าเริ่มต้นมาเป็นค่า 0 โคนมีวิธีคำนวณดังนี้ โดยทุกตัวต้อง + ด้วย 0 จากการตาย

Q จาก state ที่ 0  $\text{return} = 0$  (จากการเรียนอนุบาล) + 0 (จากการเรียนประถม) + 0 (จากการเรียนมัธยม) + 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100

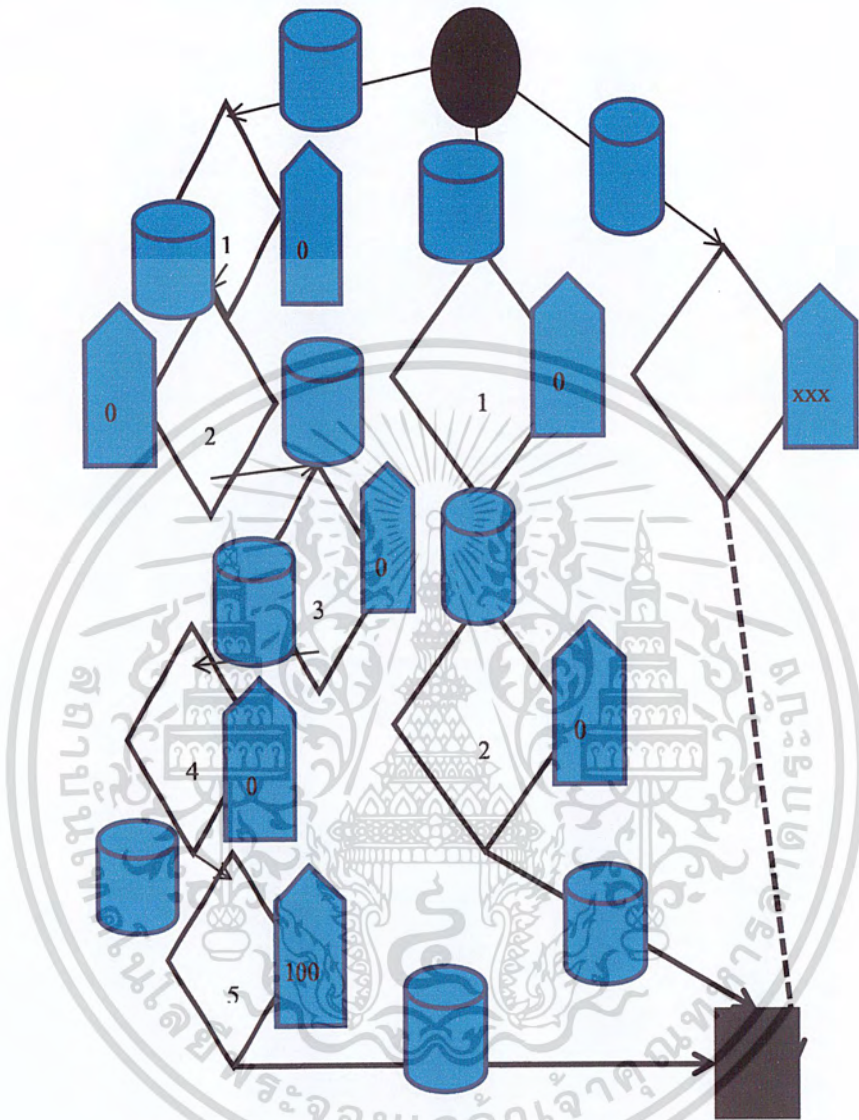
Q จาก state ที่ 1  $\text{return} = 0$  (จากการเรียนประถม) + 0 (จากการเรียนมัธยม) + 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100

Q จาก state ที่ 2  $\text{return} = 0$  (จากการเรียนมัธยม) + 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

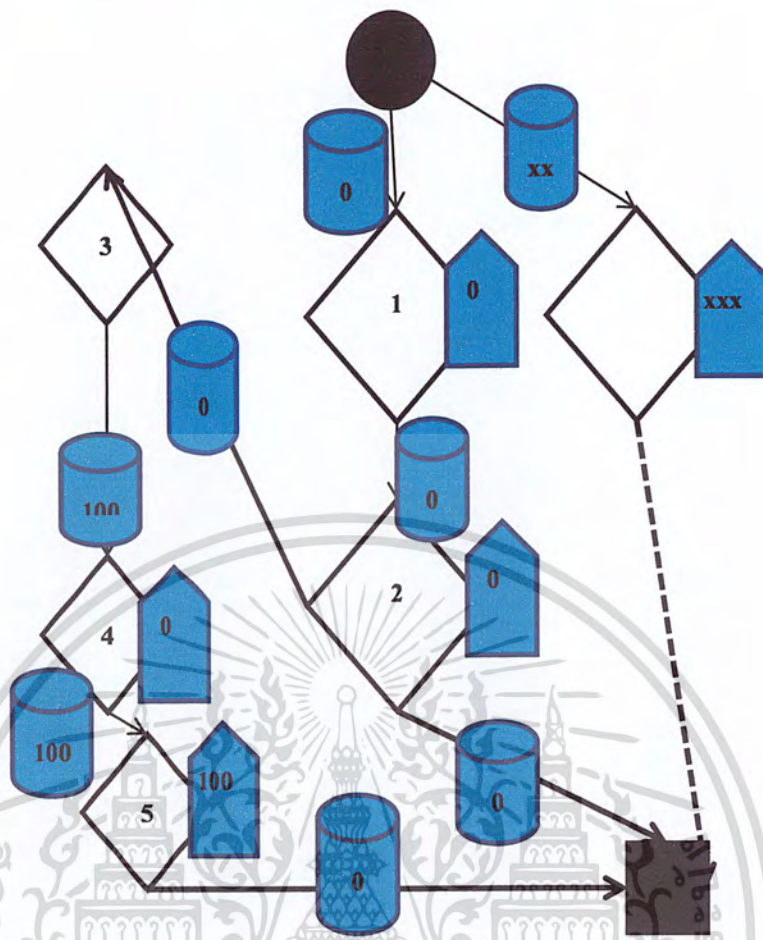
Q จาก state ที่ 3 return = 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100

Q จาก state ที่ 5 return 0



รูป 3.7 อธิบาย episode ที่ 1,2 หลัง update แล้ว

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.8 อธิบาย episode ที่ 3

จากรูป 3.8 จะเห็นได้ว่าค่า Q ของรูปจะมีการเปลี่ยนจากค่าเริ่มต้นมาเป็นค่า 0 โดยมีวิธีคำนวณ ดังนี้ โดยทุกตัวต้อง + ด้วย 0 จากการตาย

Q จาก state ที่ 0 return = 0 (จากการเรียนอนุบาล) + 0 (จากการเรียนประถม) + 0 (จากการเรียนมัธยม) + 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100 ซึ่งแต่เดิมค่าของมันเป็น 0 อยู่และมีการ update มาแล้ว 1 ครั้ง ในการ update ครั้งที่ 2 นี้เราจะนำครั้งแรกกับครั้งที่ 2 มาเฉลี่ยกัน  $0+100/2=50$

Q จาก state ที่ 1 return = 0 (จากการเรียนประถม) + 0 (จากการเรียนมัธยม) + 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100 ซึ่งแต่เดิมค่าของมันเป็น 0 อยู่และมีการ update มาแล้ว 1 ครั้ง ในการ update ครั้งที่ 2 นี้เราจะนำครั้งแรกกับครั้งที่ 2 มาเฉลี่ยกัน  $0+100/2=50$

Q จาก state ที่ 2 return = 0 (จากการเรียนมัธยม) + 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100

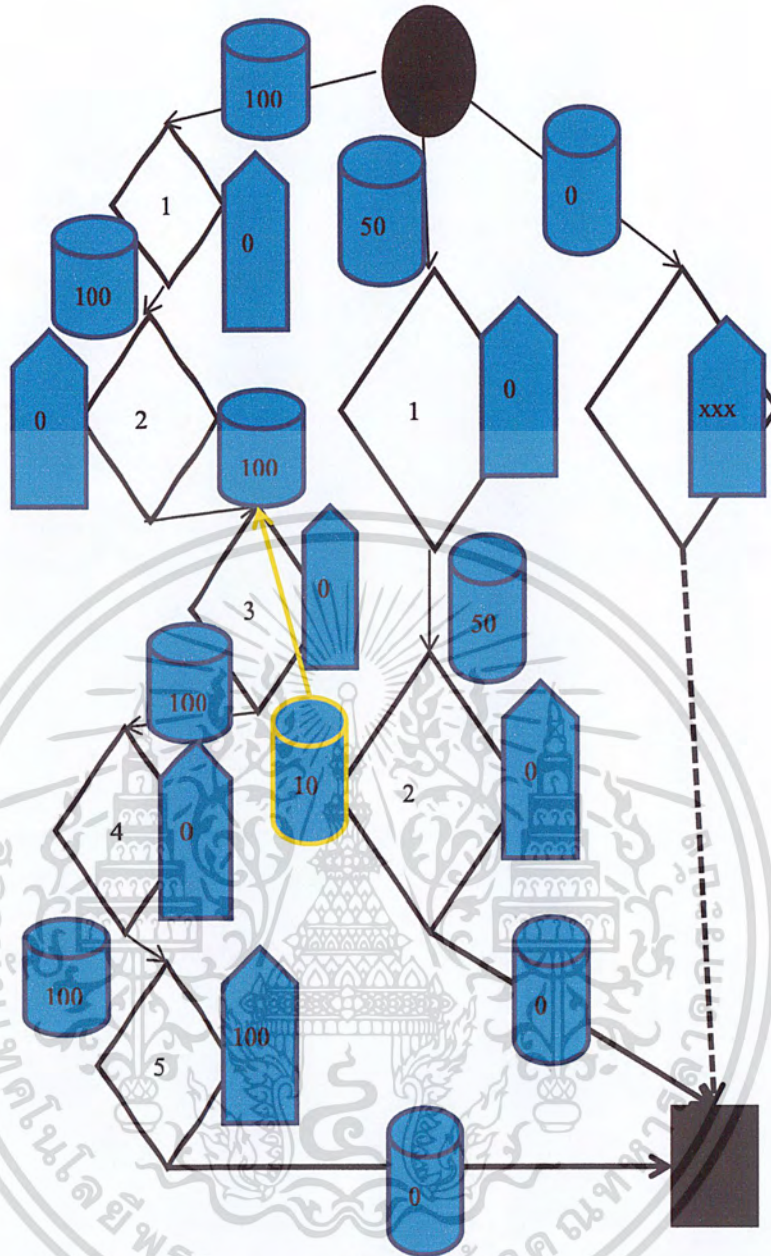
Q จาก state ที่ 3 return = 0 (จากการเรียนมหาลัย) + 100 (จากการทำงาน) = 100 ซึ่งแต่เดิมค่าของมันเป็น 0 อยู่และมีการ update มาแล้ว 1 ครั้ง ในการ update ครั้งที่ 2 นี้เราจะนำครั้งแรกกับครั้งที่ 2 มาเฉลี่ยกัน  $100+100/2=100$

Q จาก state ที่ 4 return = 100 (จากการทำงาน) = 100 ซึ่งแต่เดิมค่าของมันเป็น 0 อยู่และมีการ update มาแล้ว 1 ครั้ง ในการ update ครั้งที่ 2 นี้เราจะนำครั้งแรกกับครั้งที่ 2 มาเฉลี่ยกัน  
 $100+100/2=100$

Q จาก state ที่ 5 return = 0 ซึ่งแต่เดิมค่าของมันเป็น 0 อยู่และมีการ update มาแล้ว 1 ครั้ง ในการ update ครั้งที่ 2 นี้เราจะนำครั้งแรกกับครั้งที่ 2 มาเฉลี่ยกัน  $0+0/2=0$



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
 ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 3.9 อธิบาย episode ที่ 1,2,3 หลัง update แล้ว

### 3.2 สรุป

จากตัวอย่างจะเห็นได้ว่าค่า  $Q$  นั้นที่สูงสุดอยู่ที่ 100 ซึ่งมีความหมายว่าการเลือกการกระทำนั้น นำเราไปสู่ผลรางวัลเฉลี่ย 100 โดย ค่า  $Q$  50 หมายความว่ามันจะนำไปสู่ผลรางวัลเฉลี่ย 50 เพราะ คน มา 21 ทางนี้เคย ตายไปก่อน ได้รางวัล 1 คน โดยการเลือกโดยทั่วไปแล้วจะเลือกการกระทำ 100 แต่ มีโอกาสเลือกการกระทำอื่นได้เช่นกันขึ้นกับค่า  $\epsilon$  ที่กำหนดไว้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 4

# เครื่องมือสำหรับการวิเคราะห์และชี้วัดทางเทคนิค (Technical Indicator)

### 4.1 RSI (Relative Strength Index - เครื่องมือดัชนีกำลังสัมพัทธ์)

เป็น Indicator ที่เกิดจากการให้ความสำคัญกับค่าข้อมูลย้อนหลังล่าสุดในขอบเขตช่วงเวลาหนึ่งของตลาด ซึ่งอาจจะมีทั้งทิศทาง การเปลี่ยนแปลงที่เป็นบวกและลบรวมกัน หรืออย่างใดอย่างหนึ่ง โดย RSI จะทำการแยกคำนวณค่าที่เป็นบวกและลบออกจากกันตามแต่ละช่วงเวลา ซึ่งค่าของราคาของอัตราแลกเปลี่ยนที่นำมาคำนวณนั้นจะเป็นราคาปิดของการซื้อขายของช่วงเวลานั้น และ RSI ยังสามารถวัดปริมาณการผลักดันและกระตุ้นการเคลื่อนที่ของราคาในลักษณะไม่เรียบ (Momentum Oscillator) ซึ่งจะแสดงออกมาในรูปของความเร็วในการเปลี่ยนแปลงของราคาควบคู่ไปกับขนาดการเปลี่ยนแปลงของทิศทางของการเคลื่อนที่ในอนาคต ตามรูปแบบของ RSI นั้น ค่าของ RSI นั้นอาจจะอนุมานได้ว่าแปรผันตามทิศทางรวมของการเปลี่ยนแปลงระหว่างช่วงเวลาต่างๆ ในขอบเขตที่กำหนด ซึ่งโดยปกติทั่วไปขอบเขตช่วงเวลาที่ใช้นั้นคือ 14 วัน และค่าที่คำนวณได้จะทำการถูกทำการเปลี่ยนแปลงให้อยู่ในช่วงขอบเขตของค่า 0-100 และมีการกำหนดเพิ่มเติมให้ค่าที่ตำแหน่ง 70 ขึ้น ไปเป็นขอบเขตบน ค่าที่ตำแหน่ง 30 ลงมาเป็นขอบเขตล่าง

#### 4.1.1 วิธีการคำนวณ

$$RS = U/D \quad (4.1)$$

U = ค่ารวมของช่วงเวลาที่ขึ้นบวก

D = ค่ารวมของช่วงเวลาที่ลงลบ

$$RSI = 100 - \frac{100}{1 + RS} \quad (4.2)$$

หรือ

$$RSI = 100 \times \frac{U}{U + D} \quad (4.3)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เมื่อกำหนดให้

$$U = \sum_{i=1}^n u_i \quad (4.4)$$

$$D = \sum_{i=1}^n d_i \quad (4.5)$$

โดยการหาค่า  $u_i$  และ  $d_i$  นั้นมีช่วงเวลาอยู่ 3 ช่วงเวลา

1) ช่วงเวลาที่ปัจจุบันมากกว่าอดีต

$$u_i = \text{ราคาปิดช่วงเวลาปัจจุบัน} - \text{ราคาปิดช่วงเวลาก่อนหน้า} \quad (4.6)$$

$$d_i = 0 \quad (4.7)$$

2) ช่วงเวลาที่ปัจจุบันน้อยกว่าอดีต

$$d_i = \text{ราคาปิดช่วงเวลาก่อนหน้า} - \text{ราคาปิดช่วงเวลาปัจจุบัน} \quad (4.8)$$

$$u_i = 0 \quad (4.9)$$

3) ช่วงเวลาที่ปัจจุบันเท่ากับอดีต

$$d_i = 0 \quad (4.10)$$

$$u_i = 0 \quad (4.11)$$

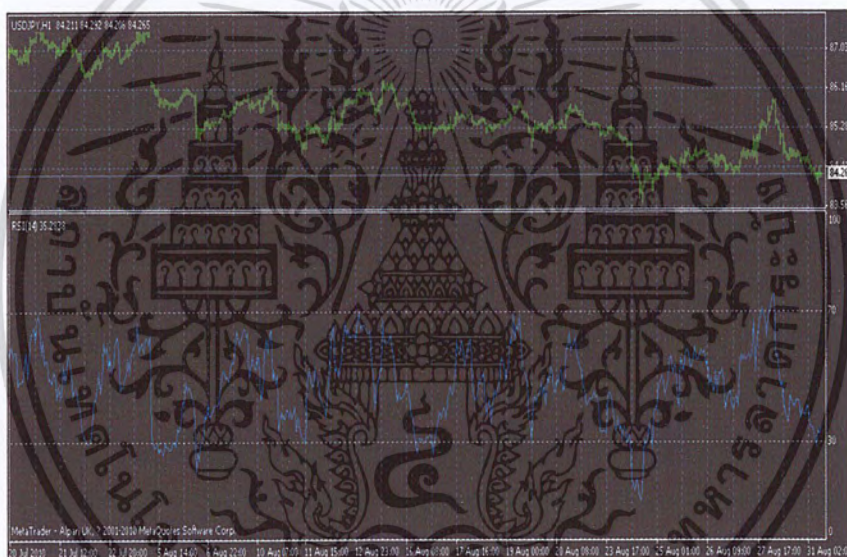
#### 4.1.2 การวิเคราะห์

โดยปกติค่าของการเปลี่ยนแปลงของกราฟแสดงค่า RSI ที่เวลาต่างๆ นั้นจะเคลื่อนไหวอยู่ในช่วงขอบเขตที่ค่า RSI อยู่ระหว่าง 30 – 70 ถ้าหากค่า RSI มีค่ามากกว่า 70 จะเกิดจากการที่ตลาดมีการซื้อมากเกินไป (OVERBOUGHT) ซึ่งทำให้ช่วงเวลาต่อมามีโอกาสที่ราคารวมไปถึงค่า RSI จะอยู่ที่จุดสูงสุด ณ ช่วงเวลานั้น และมีโอกาสเปลี่ยนแปลงปรับตัวลดลงสูง ส่วนถ้าค่าต่ำกว่า 30 ก็กลับกัน เกิดจากการขายที่มากเกินไป (OVERSOLD) ทำให้ตลาดมีโอกาปรับตัวขึ้นสูง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 4.1 การปรับตัวของค่า RSI รวมถึงขอบเขต oversold และ overbought



รูป 4.2 RSI จากโปรแกรม Terminal

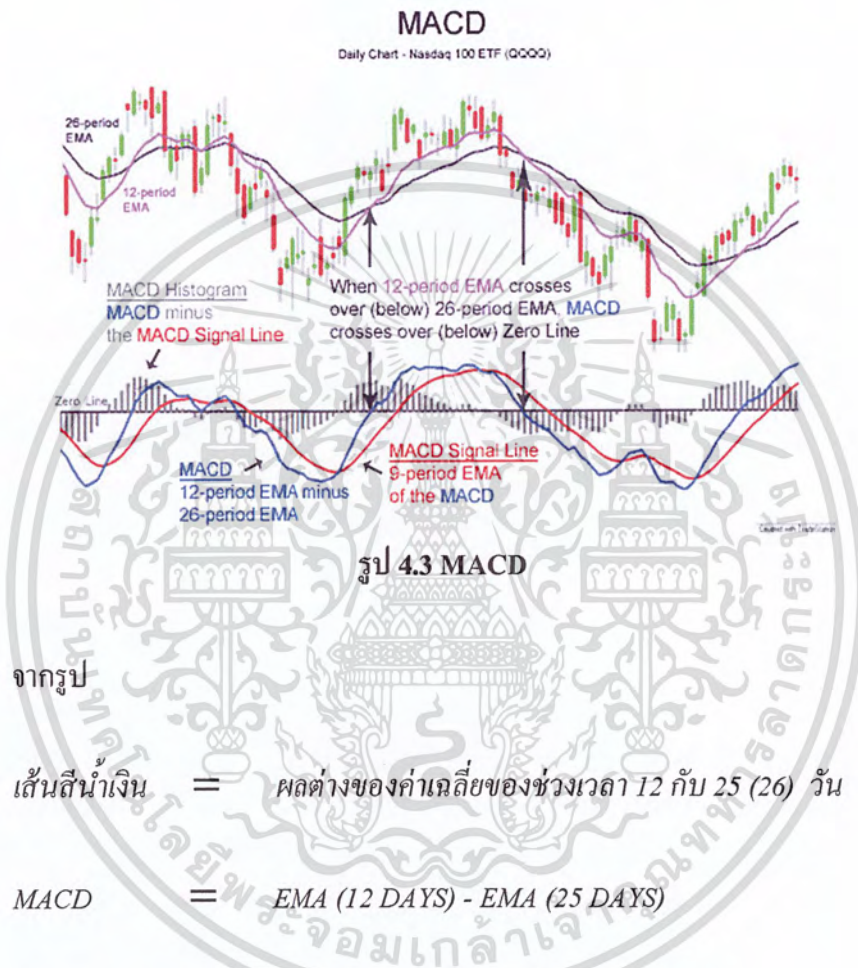
## 4.2 MACD (MOVING AVERAGES CONVERGENCE/ DIVERGENCE)

เป็นเครื่องมือวิเคราะห์ทางเทคนิคที่สร้างขึ้น และพัฒนาโดย GERALD APPEL ในปี ค.ศ.1979 ซึ่งเป็นเครื่องมือที่เคลื่อนที่ไปในทิศทางเดียวกับราคา (TREND FOLLOWING) สามารถใช้วัดระดับ (DEGREE) ตลาดว่าเป็นตลาด BULL หรือตลาด BEAR รวมถึงช่วงเวลาการคงอยู่ของทิศทาง การเคลื่อนที่นั้น และท้ายที่สุดยังแสดงถึงจุดเปลี่ยนแปลงของทั้งทิศทางและ momentum ที่เกิดขึ้นด้วย

### 4.2.1 วิธีการคำนวณ

เส้น MACD สร้างขึ้นโดยใช้ความต่างระหว่างเส้นค่าเฉลี่ยเคลื่อนที่ 2 เส้น โดยที่เส้นค่าเฉลี่ยเคลื่อนที่เส้นหนึ่ง ใช้ระยะเวลาในการคำนวณยาวกว่าเส้นค่าเฉลี่ยอีกเส้นหนึ่ง และเส้นเอกสารเป็นเอกสารที่ส่งแนวโน้มสำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านราคาไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ค่าเฉลี่ยเคลื่อนที่ 2 เส้นนี้ นิยมใช้เส้นค่าเฉลี่ยเคลื่อนที่แบบ EXPONENTIAL ส่วนจำนวนวันที่นำมาหาค่าเฉลี่ย ก็อาจเปลี่ยนแปลงได้ แต่ที่นิยมใช้กันทั่วไปคือ 12 วัน และ 25 (หรือ 26 วัน) มีข้อสังเกตว่า เส้นค่าเฉลี่ยระยะยาวนี้ จะมีระยะเวลายาวนาน กว่าเส้นค่าเฉลี่ยระยะสั้นประมาณ 1 เท่า การให้สัญญาณซื้อขายที่นิยมวิธีหนึ่งของ MACD คือ การใช้สัญญาณ (SIGNAL LINE) ตัดกับเส้น MACD



จากรูป

$$\text{เส้นสีน้ำเงิน} = \text{ผลต่างของค่าเฉลี่ยของช่วงเวลา 12 กับ 25 (26) วัน} \quad (4.12)$$

$$\text{MACD} = \text{EMA (12 DAYS) - EMA (25 DAYS)} \quad (4.13)$$

$$\text{เส้นสีแดง} = \text{ค่าเฉลี่ยของเส้น MACD line ใน 9 วัน} \quad (4.14)$$

โดยที่

$$\text{SIGNAL LINE} = \text{EMA 9 DAYS OF MACD} \quad (4.15)$$

$$\text{EMA} = \text{EXPONENTIAL MOVING AVERAGE} \quad (4.16)$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

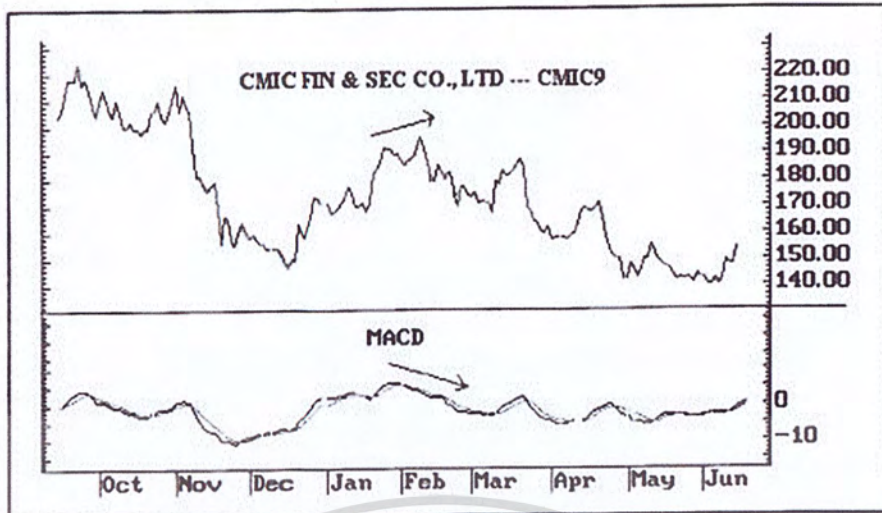
#### 4.2.2 หลักการวิเคราะห์

- 1) ถ้า MACD มีค่าเป็นบวก แสดงว่าราคาอยู่ในแนวโน้มขึ้นระยะกลาง
- 2) ถ้า MACD มีค่าเป็นลบ แสดงว่าราคาอยู่ในแนวโน้มลงระยะกลาง
- 3) ถ้า MACD มีค่าเป็นบวก และตัดเส้นสัญญาณ (SIGNAL LINE) ขึ้นไป แสดงว่าราคามีแนวโน้มสูงขึ้น เป็นสัญญาณซื้อ (BUY SIGNAL)
- 4) ถ้า MACD มีค่าเป็นลบ และตัดเส้นสัญญาณ (SIGNAL LINE) ลงมา แสดงว่าราคามีแนวโน้มลดลง เป็นสัญญาณขาย (SELL SIGNAL)
- 5) ถ้า MACD มีค่าเป็นบวก แต่ตัดเส้นสัญญาณ (SIGNAL LINE) ลงมา แสดงว่าราคากำลังมีแนวโน้มชะลอการลงหรือปรับตัวขึ้นช่วงสั้น
- 6) ถ้า MACD มีค่าเป็นลบ แต่ตัดเส้นสัญญาณ (SIGNAL LINE) ขึ้นไป แสดงว่าราคากำลังมีแนวโน้มชะลอการลงหรือปรับตัวขึ้นช่วงสั้น
- 7) ถ้า MACD มีค่าเป็นบวก และอยู่ในระดับสูงใกล้เคียงกับยอดเก่า แสดงว่าราคามีโอกาสที่จะทรงตัวหรือปรับตัวลดลง
- 8) ถ้า MACD มีค่าเป็นลบ และอยู่ในระดับต่ำใกล้เคียงกับฐานเก่า แสดงว่าราคามีโอกาสที่จะทรงตัวหรือปรับตัวสูงขึ้น
- 9) ถ้า MACD และเส้นสัญญาณ (SIGNAL LINE) มีค่าเป็นบวก แสดงว่าตลาดเป็นตลาด BULL
- 10) ถ้า MACD และเส้นสัญญาณ (SIGNAL LINE) มีค่าเป็นลบ แสดงว่าตลาดเป็นตลาด BEAR

#### 4.2.3 การแยกตัวออกจากกันของ MACD กับดัชนีราคา (DIVERGENCE)

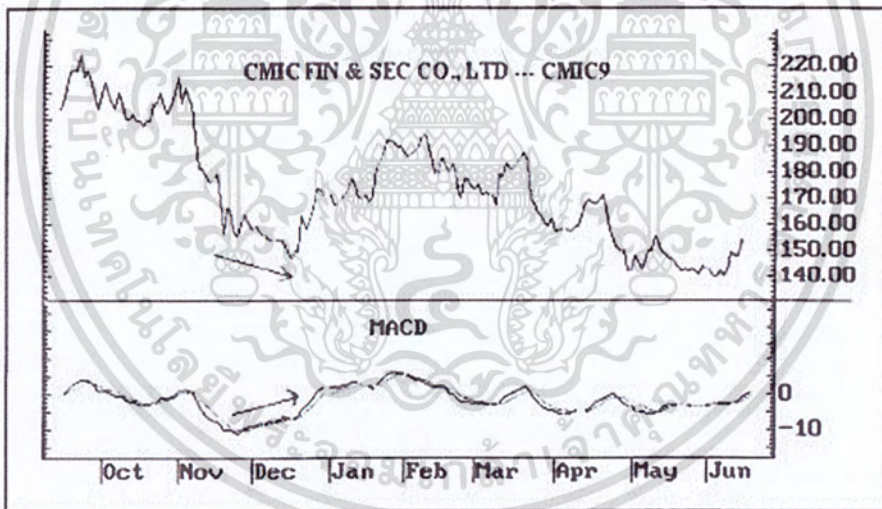
DIVERGENCE คือ การแยกตัวออกจากกันของ MACD กับราคามี 2 ลักษณะคือ

- 1) NEGATIVE DIVERGENCE จะเกิดขึ้นเมื่อ MACD มีการปรับตัวลงสวนทางกับการสูงขึ้นของดัชนีราคา เป็นการเตือนว่าราคาอาจมีการปรับตัวลง



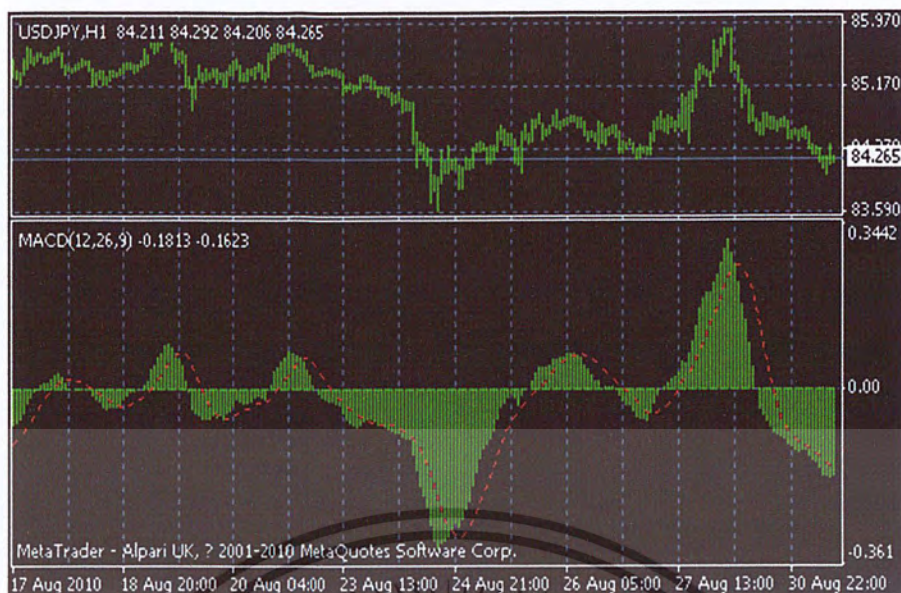
รูป 4.4 Negative Divergence

- 2) POSITIVE DIVERGENCE จะเกิดขึ้นเมื่อ MACD มีการปรับตัวสูงขึ้นสวนทางกับการลดลงของดัชนีราคา เป็นการบอกว่าการลดลงของราคาใกล้สิ้นสุด



รูป 4.5 Positive Divergence

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 4.6 MACD จาก โปรแกรม Terminal

#### 4.3 OsMA (Moving Average of Oscillator – ค่าเฉลี่ยเคลื่อนที่ของกรอบการแกว่งตัว)

สืบเนื่องจากการวิเคราะห์ที่ใช้งาน MACD นั้นยังให้ประสิทธิภาพได้จำกัดอยู่ โดยค่าที่ได้นั้น ยังมีผลกระทบที่เกิดจากการรบกวนที่ไม่อยู่ในรูปแบบโดยรวม เพื่อให้ผลลัพธ์สุดท้ายมีรูปแบบสอดคล้องกับการเคลื่อนที่โดยรวมของการเปลี่ยนแปลง ซึ่ง OsMA นั้นเกิดจากผลต่างของ MACD และ Signal โดย Signal นั้นจะเป็นตัวคั่นกรองอีกชั้นหนึ่งของ MACD

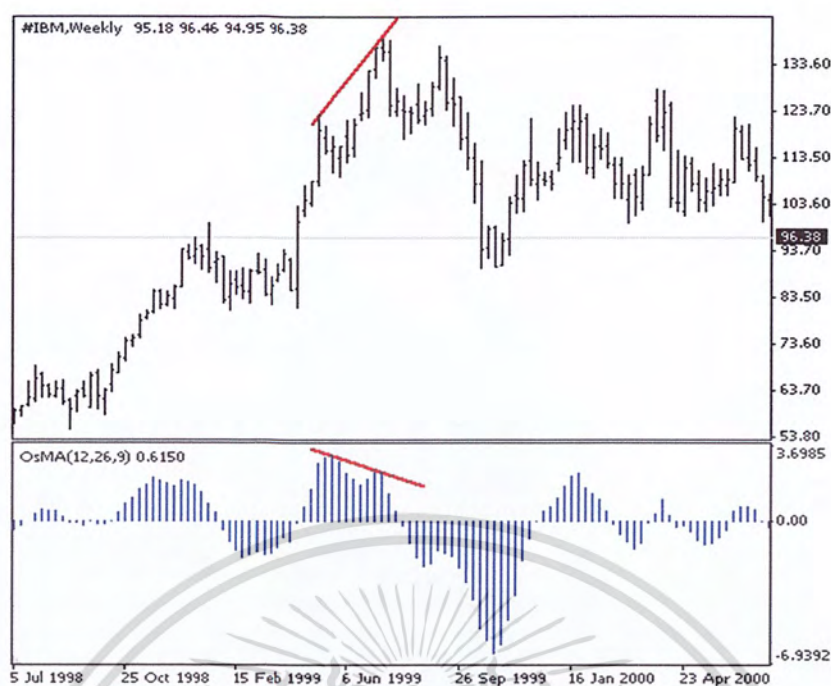
$$OsMA = MACD - Signal \quad (4.17)$$

##### 4.2.1 วิธีการวิเคราะห์

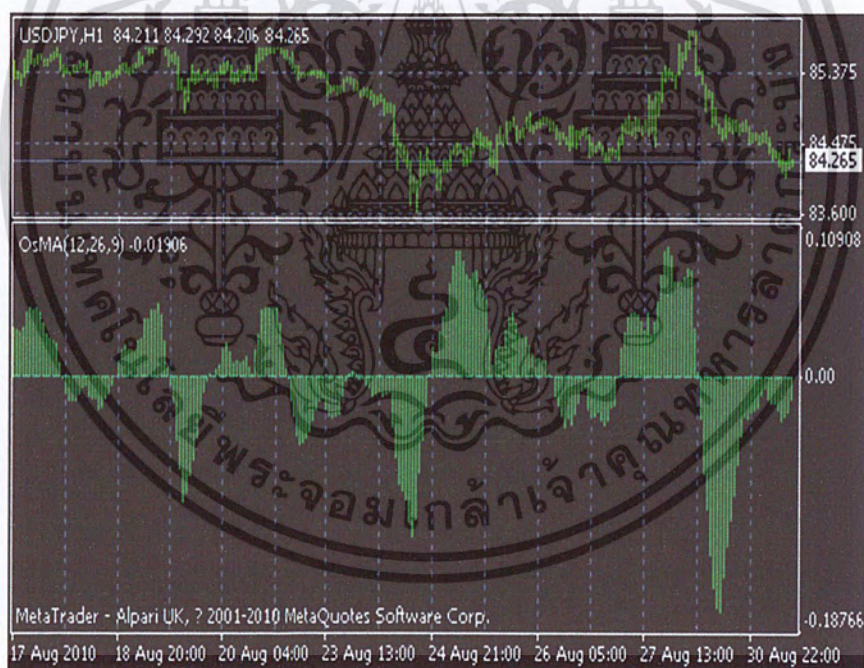
- 1) เมื่อค่า OsMA หยุคปรับตัวลดลงเปลี่ยนเป็นเริ่มปรับตัวเพิ่มขึ้นนั้น เป็นสัญญาณให้ทำการซื้อ
- 2) เมื่อค่า OsMA หยุคปรับตัวเพิ่มขึ้นเปลี่ยนเป็นเริ่มปรับตัวลดลงนั้น เป็นสัญญาณให้ทำการขาย

โดยความแตกต่างที่เกิดขึ้นกับการเปลี่ยนแปลงของ OsMA นั้น ให้สัญญาณการเปลี่ยนแปลงที่อาจจะเกิดขึ้นได้นำเชื่อถือในระดับหนึ่ง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 4.7 การเปลี่ยนแปลงของตลาด และ OsMA



รูป 4.8 OsMA จาก โปรแกรม Terminal

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

#### 4.4 Alligator

สร้างขึ้นบนพื้นฐานของ moving average โดยรวม moving average ที่มีเสถียรภาพ 3 แบบเข้าด้วยกัน ซึ่งจะช่วยให้ประเมินค่าการเปลี่ยนแปลงของราคา และแสดงความเชื่อมโยงกันของราคา ณ ตำแหน่งต่างๆ ในแต่ละช่วงเวลา



รูป 4.9 Alligator

จากรูป 4.9

The alligator lips (เส้นเขียว) คือ วัดค่าขนาด 5 ช่วงเวลา (bar) ของ moving average และอยู่ห่างกัน 3 bars

The alligator teeth (เส้นแดง) คือ วัดค่าขนาด 8 ช่วงเวลา (bar) ของ moving average และอยู่ห่างกัน 5 bars

The alligator jaws (เส้นน้ำเงิน) คือ วัดค่าขนาด 13 ช่วงเวลา (bar) ของ moving average และอยู่ห่างกัน 8 bars

##### 4.4.1 วิธีการวิเคราะห์

Alligator ช่วยในการตัดสินใจในแนวโน้มทิศทางการเคลื่อนที่ในปัจจุบันและอนาคต

1) ทิศทางการเปลี่ยนแปลงเมื่อราคาเพิ่มขึ้น (uptrend) เส้นทั้ง 3 เส้นต้องจัดเรียงตัวกันอย่างเหมาะสมโดยที่ เส้นน้ำเงิน (Jaws) , แดง (Teeth) , เขียว (Lips) เรียงลำดับจากล่างไป บนตามลำดับ

2) ทิศทางการเปลี่ยนแปลงเมื่อราคาลดลง (downtrend) ตรงกันข้ามกับ uptrend เส้นน้ำเงิน (Jaws) , แดง (Teeth) , เขียว (Lips) เรียงลำดับจากบนไปล่างตามลำดับ

3) ทิศทางการเปลี่ยนแปลงคงที่การเคลื่อนที่ของเส้นทั้ง 3 เคลื่อนตัวโดยสานกัน ไปหรือเคลื่อนที่ในลำดับการเรียงตัวรูปแบบอื่นๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษานั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 4.10 Aligator จากโปรแกรม Terminal

#### 4.5 Moving Average (เส้นค่าเฉลี่ย)

นับเป็นเครื่องมือสำคัญ เครื่องมือหนึ่งในการวิเคราะห์ทางเทคนิคโดยมีส่วนช่วยในการมองเห็นถึงแนวโน้มการ เคลื่อนที่ของราคาหุ้นรวมถึงจุดที่เปลี่ยนแนวโน้ม เพื่อเป็นสัญญาณซื้อขาย รวมถึงแนวรับแนวต้านของราคาหุ้น ในช่วงเวลาต่างๆซึ่งกลยุทธ์การลงทุนนั้น จะกำหนดเส้นค่าเฉลี่ยในจำนวนวันที่แตกต่างกัน ขึ้นอยู่กับ ระยะเวลาการลงทุนของแต่ละบุคคลหรือรอบการเคลื่อนที่ของหุ้นตัวนั้น ว่าการกำหนดด้วยเส้นค่าเฉลี่ยเท่าใดที่น่าจะได้ผลตอบแทนสูงที่สุด



รูป 4.11 MA

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สัญญาณซื้อ (ลูกศรสีเขียว) คือเมื่อราคาเคลื่อนขึ้นและทะลุผ่านเส้นค่าเฉลี่ยตามช่วงระยะเวลาต่างๆ เช่น 5 วัน , 10 วันเส้นค่าเฉลี่ยระยะสั้นตัดขึ้นเหนือเส้นค่าเฉลี่ยระยะยาว เรียกว่า Golden Cross

สัญญาณขายคือเมื่อราคาเคลื่อนลงและทะลุผ่านเส้นค่าเฉลี่ยตามช่วงระยะเวลาต่างๆเช่น 5 วัน 10 วันเส้นค่าเฉลี่ยระยะสั้นตัดลงต่ำกว่าเส้นค่าเฉลี่ยระยะยาวเรียกว่า Dead cross



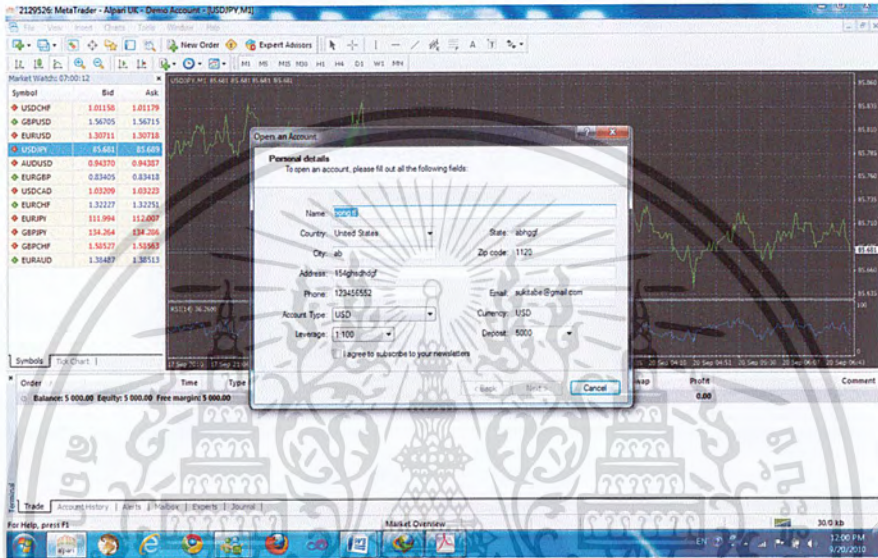
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 5

# โปรแกรม Meta Trader Alpari UK4

### 5.1 การสมัครสมาชิก

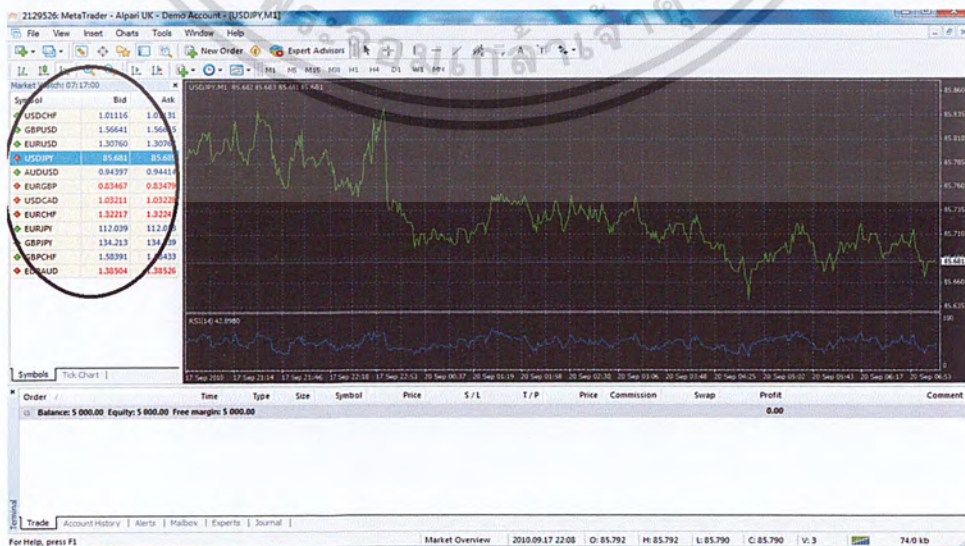
การสมัครให้เข้าไปที่ File -> Open an Account -> next -> next ตามรูป 5.1



รูป 5.1 การสมัครสมาชิก

### 5.2 การทำสัญญาซื้อขาย

ทำตามรูป 5.2 คือเลือกเงินที่จะซื้อหรือขาย

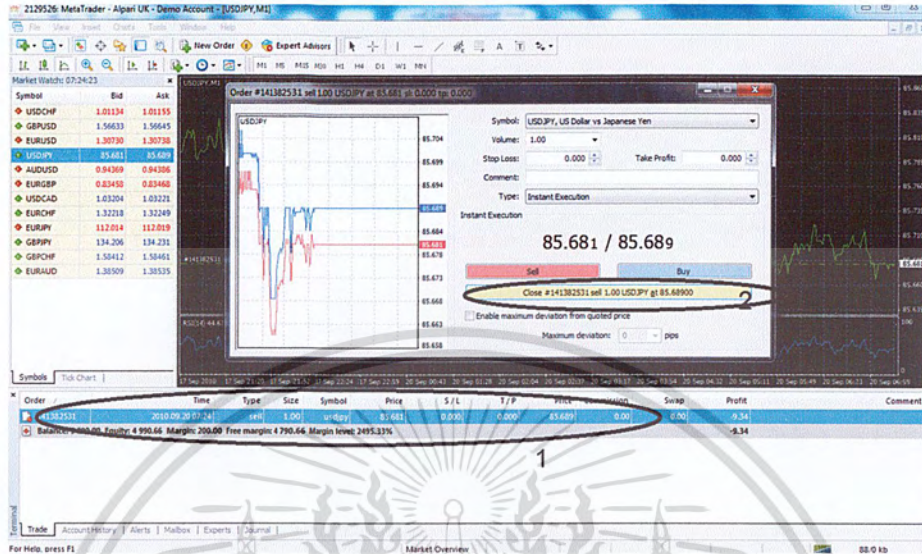


รูป 5.2 การซื้อขาย

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 5.3 การปิดสัญญาซื้อขาย

ทำตามรูป เลือกสัญญาที่ต้องการ เลือก close



รูป 5.3 การปิดสัญญา

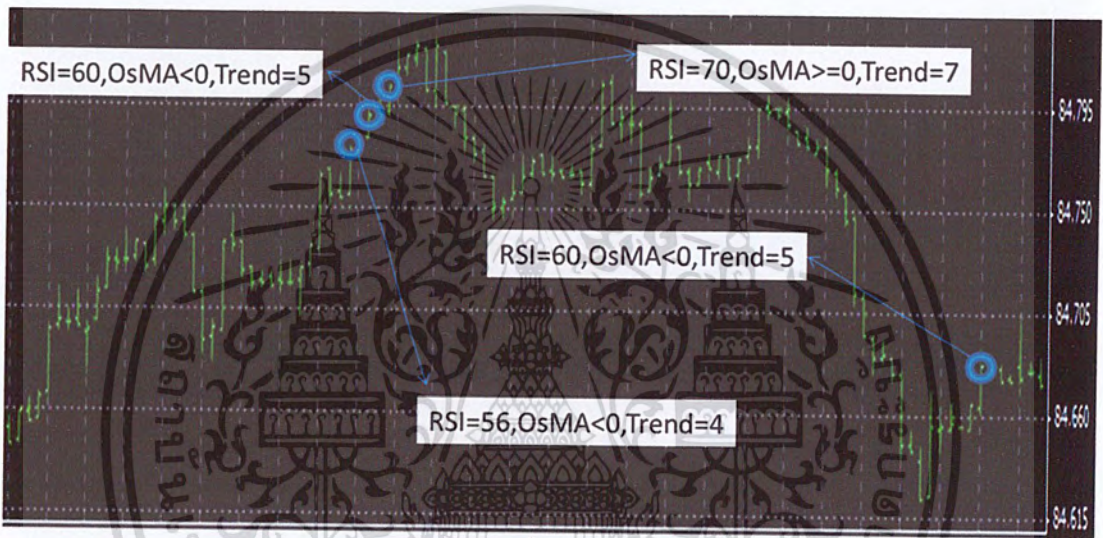
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

# บทที่ 6

## อัลกอริทึม

### 6.1 อัลกอริทึม

อัลกอริทึมของโปรแกรมนี้ใช้ทฤษฎี Reinforcement learning มาประยุกต์ใช้ โดยใช้การเก็บสถิติมาช่วยในการตัดสินใจเลือกว่าจะทำการซื้อการขายหรือจะปล่อยให้เฉยๆ โดยไม่ทำอะไร การเก็บสถิติจะเก็บ โดยใช้อินดิเคเตอร์เป็นตัวบอก state ของจุดต่างๆที่เกิดขึ้นบนกราฟ ดังรูป 6.1

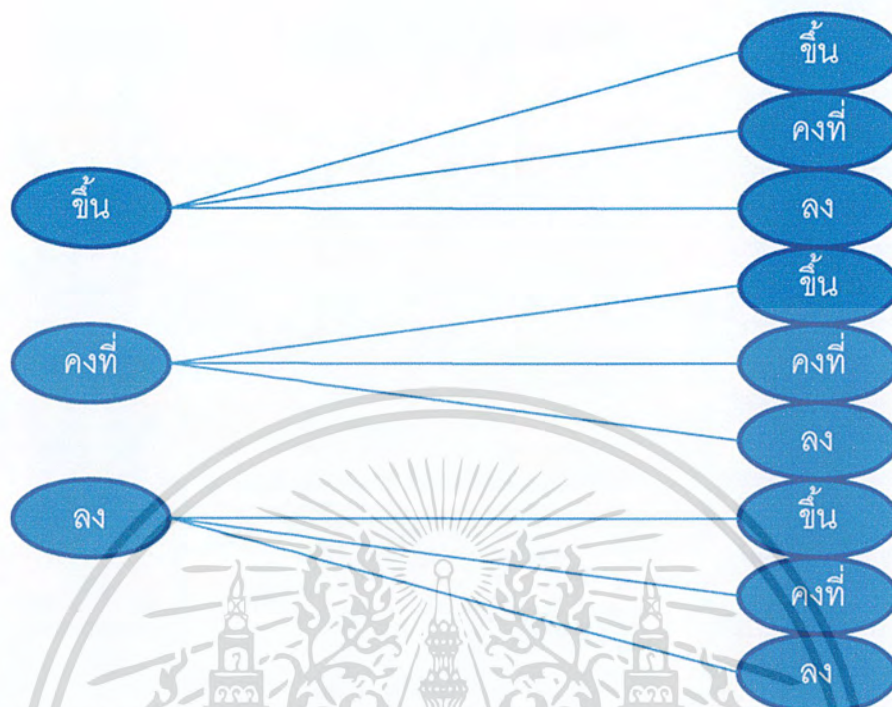


รูป 6.1 state ที่กำหนดโดยใช้ indicator

จากภาพแกนนอนเป็นแกนเวลา แกนตั้งเป็นแกนที่บอกอัตราแลกเปลี่ยนค่าเงินอินดิเคเตอร์จะ  
ใช้ดังนี้

- 1) RSI แบ่งเป็น 101 ค่าคือ 0 ถึง 100
- 2) OsMA แบ่งเป็น 2 ค่าคือ  $\geq 0$  และ  $< 0$
- 3) Trend เป็น indicator ที่เขียนขึ้นมาเอง โดยใช้ Moving Average มาคิดในระยะสั้นและยาว โดยในระยะสั้นใช้ค่าเฉลี่ยในช่วง 5 นาที ส่วนระยะยาวใช้ค่าเฉลี่ยในช่วง 30 นาทีวิธีคำนวณก็คือถ้าค่าปัจจุบันมากกว่าค่าเฉลี่ยก็จะนับว่าขึ้น น้อยกว่าก็จะนับว่าลง แต่ถ้ามีค่าประมาณค่าเฉลี่ยจะนับว่าคงตัว (โดยไม่มากหรือน้อยกว่าค่าคงที่ที่กำหนด) คิดของทั้งระยะสั้นและยาวทั้งหมด จะมี 9 ค่า ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

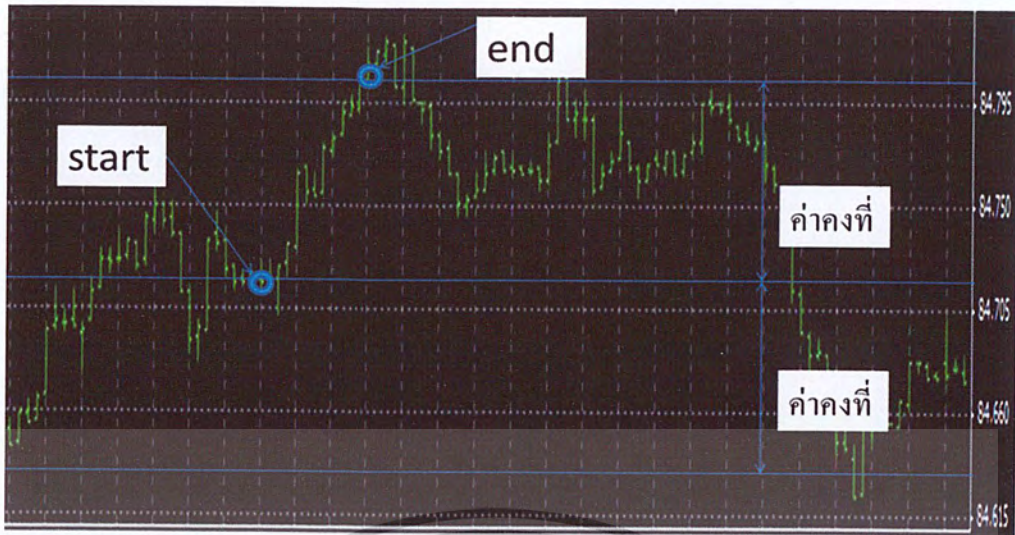


รูป 6.2 จำนวนค่าที่เป็นไปได้ทั้งหมดของ Trend

state ที่มีทั้งหมดก็จะเท่ากับ  $101 * 2 * 9 = 1818$  state

ค่าที่เก็บในแต่ละ state นั้นคือค่าที่บอกว่า ที่ state นั้นต่อไปจะขึ้นหรือลง ซึ่งตัดสินใจจากค่าคงที่ค่าหนึ่ง โดยทำการวัดว่าถ้าค่าที่ state นั้นมากกว่าหรือน้อยกว่าค่าคงที่นั้นก็ให้ตัดจบ episode นั้นและทำการเก็บข้อมูล ดังภาพ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 6.3 การกำหนด episode

การเก็บข้อมูลจะเก็บเป็นจำนวนครั้งที่เกิดขึ้นแบ่งเป็นจำนวนครั้งที่ขึ้นกับจำนวนครั้งที่ลง ทำการเก็บข้อมูลแบบนี้กับทุกๆจุดที่เกิดขึ้นบนกราฟ ส่วนการนำสถิติที่เก็บไปใช้คำนวณในการซื้อขาย จะทำเมื่อมีข้อมูลใหม่มา จากนั้นนำข้อมูลในอดีตซึ่งเก็บเป็นสถิติที่เก็บตรงกับจุดที่เกิดขึ้นมาใช้ในการเลือกว่าจะซื้อขายหรือปล่อยผ่านไปเฉยๆ ซึ่งทำการคำนวณโดยสมการ

$$\text{Percent } n = \frac{n}{(p+n)} * 100 \quad (6.1)$$

$$\text{Percent } p = \frac{p}{(p+n)} * 100 \quad (6.2)$$

p คือ จำนวนครั้งของข้อมูลที่จบ episode แล้วราคาเพิ่มขึ้น

n คือ จำนวนครั้งของข้อมูลที่จบ episode แล้วราคาลดลง

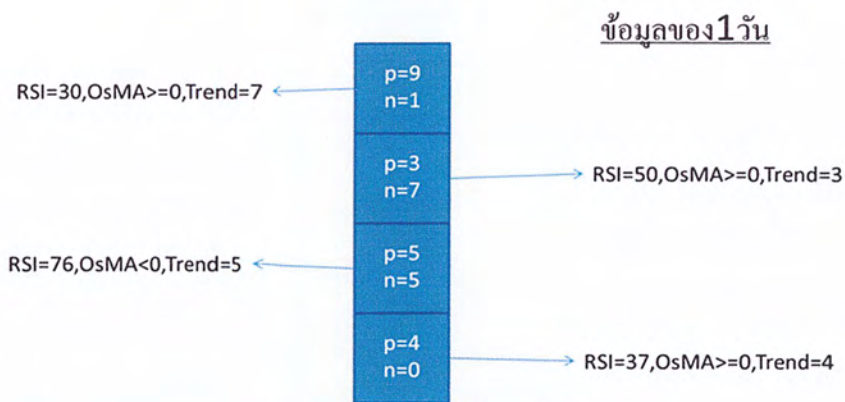
Percent p คือ เปอร์เซนต์ของข้อมูล p

Percent n คือ เปอร์เซนต์ของข้อมูล n

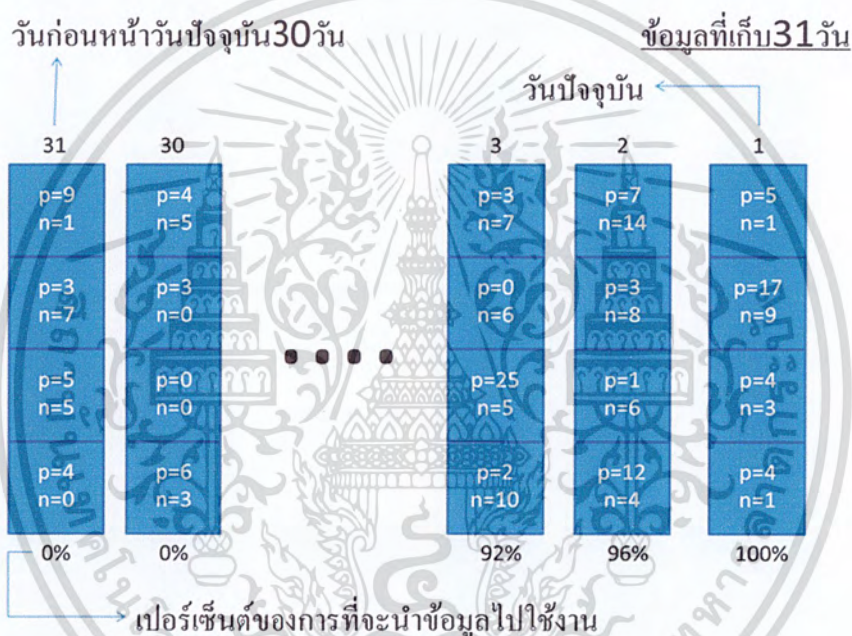
จะเลือกซื้อก็ต่อเมื่อ Percent p มีค่ามากกว่า 80 และจะเลือกขายก็ต่อเมื่อ Percent n มีค่ามากกว่า 80 ส่วนการปิดสัญญาซื้อขายนั้นจะตั้งไว้ที่ค่าคงที่ค่าหนึ่ง

ในการเลือกเก็บข้อมูลนั้นจะเก็บข้อมูลของแต่ละวันแยกกันและเก็บเฉพาะข้อมูลของ 31 วันล่าสุดนับจากวันปัจจุบัน โดยข้อมูลของแต่ละวันจะเก็บจากการจบ episode ของแต่ละ state ที่มีการจบ episode ในวันนั้นและทำการให้ความสำคัญของข้อมูล โดยแบ่งการนำข้อมูลมาใช้คำนวณตั้งแต่วันปัจจุบันใช้ 100% วันก่อนหน้าหนึ่งวันใช้ 96% วันก่อนหน้าสองวันใช้ 92% ลดไปที่ละ 4% ไปเรื่อยๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 6.4 ข้อมูลที่เก็บ 1 วัน



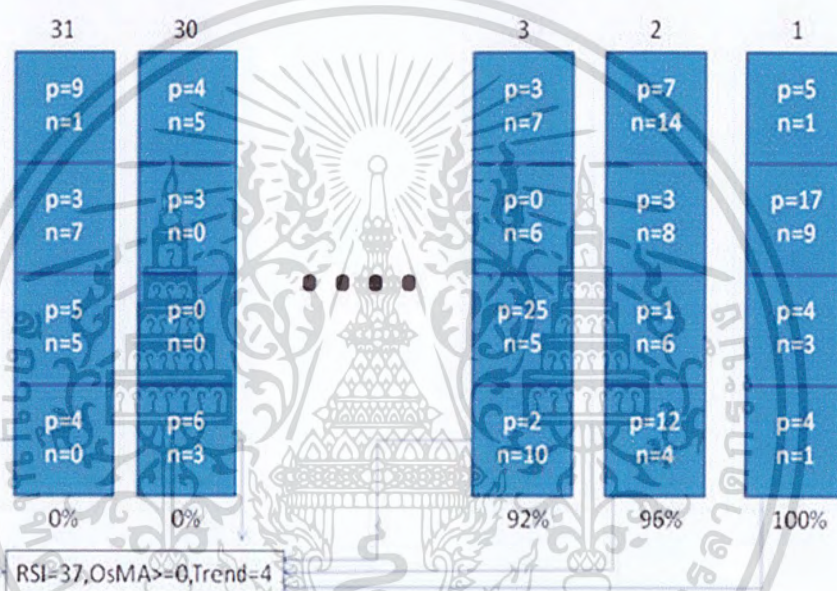
รูป 6.5 ข้อมูลที่เก็บ 31 วัน

ตัวอย่างการคำนวณการซื้อขาย สมมุติเราอยู่ที่จุดขณะนี้ RSI = 37 , OsMA >= 0 , Trend = 4

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 6.6 จุดเริ่มต้นที่จะนำมาคำนวณ



รูป 6.7 ตัวอย่างการคำนวณโดยให้ความสำคัญตามระยะห่างของวัน

#### ตัวอย่าง 6.1 การคิดเปอร์เซ็นต์

$$p = \left( \frac{4}{4+1} \right) * 100/100 + \left( \frac{12}{12+4} \right) * 96/100 + \left( \frac{2}{2+10} \right) * 92/100$$

$$+ \left( \frac{6}{6+3} \right) * 0/100 + \left( \frac{4}{4+0} \right) * 0/100$$

$$n = \left( \frac{1}{4+1} \right) * 100/100 + \left( \frac{4}{12+4} \right) * 96/100 + \left( \frac{10}{2+10} \right) * 92/100$$

$$+ \left( \frac{3}{6+3} \right) * 0/100 + \left( \frac{0}{4+0} \right) * 0/100$$

$$p = 0.8 + 0.72 + 0.1533 + 0 + 0 = 1.6733$$

$$n = 0.2 + 0.24 + 0.7667 + 0 + 0 = 1.2067$$

$$\text{Percent } p = \frac{1.6733}{1.6733 + 1.2067} * 100 = 58.1\%$$

$$\text{Percent } n = \frac{1.2067}{1.6733 + 1.2067} * 100 = 41.9\%$$

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

สมมุติว่าเราเก็บข้อมูลของทั้ง 31 วันแล้วมีข้อมูลค้างภาพ โดยถือว่าวันที่ไม่มีในภาพนั้นมีข้อมูล เป็นศูนย์ วิธีการคำนวณก็คือ จะคำนวณ โดยให้ความสำคัญของข้อมูลตามความห่างของวันค้างภาพ จากนั้นคำนวณเป็นเปอร์เซ็นต์ จากผลที่ออกมาได้ 58.1% กับ 41.9% ก็ขึ้นอยู่กับว่าอยากจะเอาข้อมูล ที่เปอร์เซ็นต์ไปใช้เท่าไร ในการทดลองจะใช้เปอร์เซ็นต์ที่มากกว่าหรือเท่ากับ 80%

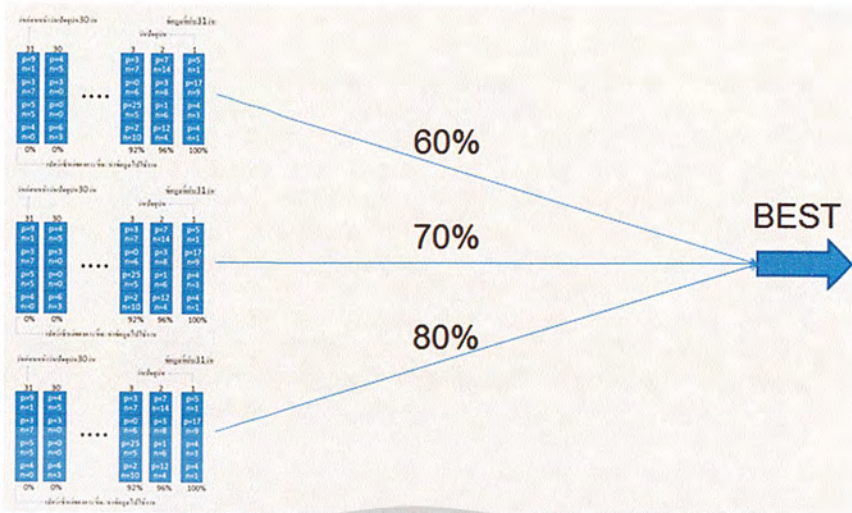
**เข้าสู่ฟังก์ชัน start**  
 ถ้าเริ่มเปลี่ยนวันจากวันปัจจุบันเป็นวันใหม่  
 วันปัจจุบัน<-วันใหม่  
 ลบค่าที่มีทั้งหมดของวันใหม่ออก  
 อาเรย์ค่าเงิน<-ค่าเงินปัจจุบันใส่ต่อท้าย  
 อาเรย์RSI<-RSIปัจจุบันใส่ต่อท้าย  
 อาเรย์OSMA<-OSMAปัจจุบันใส่ต่อท้าย  
 อาเรย์Trend<-Trendปัจจุบันใส่ต่อท้าย  
 สำหรับทุกค่าที่อยู่ในอาเรย์ค่าเงิน  
 ถ้าค่าเงินปัจจุบัน-ค่าที่อยู่ในอาเรย์ค่าเงิน>=2  
 อาเรย์ข้อมูลขึ้น<-เพิ่มขึ้น1 ที่ตำแหน่งที่มีค่าRSI, OSMA, Trend ตรงกับตำแหน่งอาเรย์ค่าเงินนี้ที่วันปัจจุบัน  
 ลบค่าที่อยู่ในอาเรย์ค่าเงินออก  
 ถ้าค่าที่อยู่ในอาเรย์ค่าเงิน-ค่าเงินปัจจุบัน>=2  
 อาเรย์ข้อมูลลง<-เพิ่มขึ้น1 ที่ตำแหน่งที่มีค่าRSI, OSMA, Trend ตรงกับตำแหน่งอาเรย์ค่าเงินนี้ที่วันปัจจุบัน  
 ลบค่าที่อยู่ในอาเรย์ค่าเงินออก  
 เปอร์เซนต์<-100  
 วันที่เอามาคำนวณ<-วันปัจจุบัน  
 ทำซ้ำตลอดไป  
 จำนวนค่าที่ขึ้น<-อาเรย์ข้อมูลขึ้นที่ตำแหน่งที่มีค่าRSI, OSMA, Trend ของวันที่เอามาคำนวณ\*เปอร์เซนต์/100  
 จำนวนค่าที่ลง<-อาเรย์ข้อมูลลงที่ตำแหน่งที่มีค่าRSI, OSMA, Trend ของวันที่เอามาคำนวณ\*เปอร์เซนต์/100  
 เปอร์เซนต์<-ลดลง4  
 วันที่เอามาคำนวณ<-ลดลง1  
 ถ้าเปอร์เซนต์<=0  
 หยุดการทำซ้ำ  
 ถ้า(จำนวนค่าที่ขึ้น/(จำนวนค่าที่ขึ้น+จำนวนค่าที่ลง))\*100>=80  
 ทำการซื้อสัญญา  
 ถ้า(จำนวนค่าที่ลง/(จำนวนค่าที่ขึ้น+จำนวนค่าที่ลง))\*100>=80  
 ทำการขายสัญญา

รูป 6.8 ชูโคโค้ด

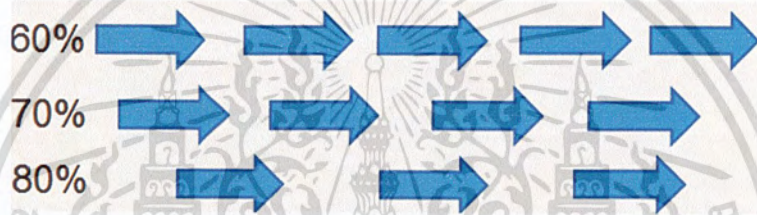
## 6.2 อัลกอริทึม เพิ่มเติม

จาก อัลกอริทึม เดิม นั้น ได้มีการพัฒนาต่อเพื่อให้มีความสามารถในการทำกำไร ได้ดียิ่งขึ้น โดยการทำให้มีหลายๆ อัลกอริทึม นั้น เทรน ไปพร้อมๆ กัน แล้วเลือกตัวที่ดีที่สุด ในขณะนั้น มาใช้ซื้อขายจริง ในการทดลองจะใช้สามตัวคือตัวที่เลือกการซื้อขายเมื่อมี % การ ได้กำไร จากข้อมูลในอดีตเป็น 60% , 70% และ 80% โดยความแตกต่างของแต่ละตัวคือ 60% จะเข้าทำการซื้อขายเร็วกว่า 70% และ 80% ในการเทรนนั้นจะเป็นแบบจำลองการซื้อขายเสมือนกับซื้อขายจริงๆ และเลือกตัวที่ดีที่สุด ในขณะนั้นมาทำการซื้อขายจริงๆ

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 6.9 จำลองการซื้อขายในการเทรดเพื่อหาตัวที่ดีที่สุด



รูป 6.10 ความแตกต่างด้านระยะเวลาซื้อขายของตัว 60%,70% และ 80%

ตัวอย่าง 6.2 วิธีการคำนวณหาตัวที่ดีที่สุด

ตัว60% -> เสีย เสีย เสีย ได้ เสีย เสีย เสีย ได้ เสีย เสีย ได้ เสีย ได้ เสีย ได้  
 ผลลัพธ์ =  $(3+3+2+1+1)/5 = 2$   
 ตัว70% -> เสีย เสีย ได้ เสีย เสีย เสีย เสีย ได้ เสีย ได้ ได้  
 ผลลัพธ์ =  $(2+4+1+0)/4 = 1.75$   
 ตัว80% -> เสีย ได้ เสีย เสีย ได้  
 ผลลัพธ์ =  $(1+2)/2 = 1.5$

การคำนวณนี้เป็นการคำนวณหาว่าเสียไปทั้งหมดกี่ครั้งต่อครั้งที่ได้มาทั้งหมด คำนวณตัวไหนที่ผลลัพธ์ออกมาน้อยสุดจะเป็นตัวที่ดีที่สุด จากตัวอย่างจะเห็นว่าที่ตัว 80% ได้ค่าน้อยที่สุดในตัวอย่างนี้จึงเลือกตัว 80% ไปใช้ในการซื้อขายจริง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 6.3 อัลกอริทึม การ ดับเบิ้ลลีสต์

หลังจากได้ศึกษาการเขียน ea มาในระยะหนึ่งแล้วก็ได้มีคนในเว็บที่ได้ไปรับเขียนมีการพูดถึงเรื่อง ดับเบิ้ลลีสต์ จึงได้นำมาลองเขียนดูหลักการก็ง่ายๆคือ ถ้าซื้อหรือขายแล้วขาดทุนจะทำการเริ่มสัญญาใหม่ โคนที่มีขนาดเป็น 2 เท่าของสัญญาเดิมเช่น 1 2 4 8 16 ... แต่ถ้าถูกแล้วก็จะไปเริ่มที่ 1 ใหม่อีกครั้งหนึ่ง

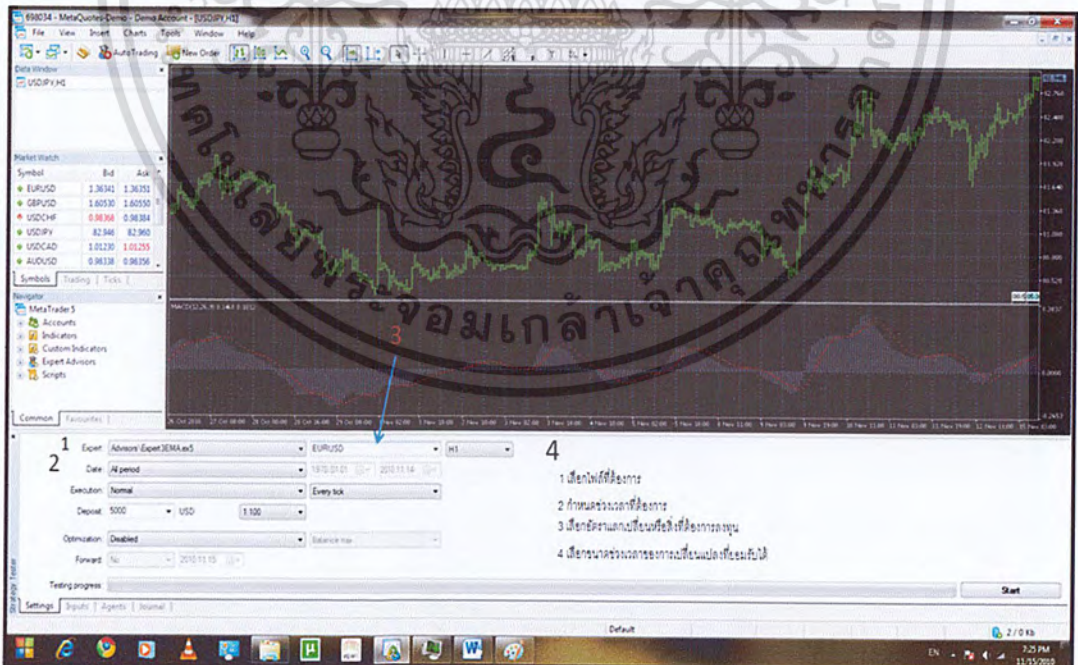


เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ดัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## Meta Quotes Language 5

### 7.1 ส่วนประกอบพื้นฐานของ Meta Trader Qoutes Language 5

เป็นเครื่องมือที่ใช้ในสร้างเครื่องมือในการกำหนดกลยุทธ์ วางแผน สำหรับการซื้อขาย รวมไปถึงการสร้างเครื่องมือชี้วัดสถานการณ์ของข้อมูลสถิติที่ผ่านมา เพื่อกำหนดทิศทางและตัวช่วยในการตัดสินใจในการลงทุน ทำให้สามารถประเมินพอร์ตการลงทุนในการลงทุนและยังสามารถใช้ร่วมกันเพื่อประสิทธิภาพในการปรับเปลี่ยนการลงทุนตามสถานการณ์ต่าง ซึ่งแตกต่างจากการไม่มีเครื่องมือเหล่านี้ช่วย และท้ายที่สุดยังสามารถสร้างเครื่องมือที่เป็นรูปธรรม(วัตถุ)ในการวัดค่าต่างๆภายในกราฟการเปลี่ยนแปลง เพื่อวิเคราะห์และหาค่าต่างๆเพื่อเปรียบเทียบกับค่าที่เหมาะสม ส่วนในด้านการใช้งานนั้นตัว โปรแกรมที่ถูกสร้างขึ้นมาจะต้องใช้ควบคู่กับ โปรแกรม MetaTrader Terminal 5 โดยในส่วนนี้ในกรณีที่มีตัว Expert Advisor หรือ Custom Indicators เวอร์ชัน 4 นั้นไม่สามารถนำมาใช้งานร่วมกับ Meta Trader Terminal 5 ได้โดยตรง ต้องทำการแปลงในส่วนของตัวโปรแกรมให้เป็นเวอร์ชัน 5 ก่อนจึงจะสามารถใช้งานร่วมกันได้ ซึ่งก็มีเครื่องมือที่ช่วยในการแปลงเวอร์ชัน



รูป 7.1 MetaTraderTerminal

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

เป็นภาษาที่เป็นตัวช่วยสำหรับการสร้างเครื่องมือในการลงทุน โดยแนวคิดพื้นฐานของตัวภาษานั้นได้ใช้หลักการของ Object Oriented programming และได้ยังได้ใช้รูปแบบบางส่วนที่คล้ายคลึงกับภาษา C++ โดยสิ่งที่แตกต่างจาก MQL 4 ซึ่งสร้างจากพื้นฐานของภาษา C ที่เห็นได้ชัดเจน คือ Enumerations Structures Classes Event handling

### 7.1.1 โครงสร้างภาษาของ MQL5 (Syntax)

มีรายละเอียดของความสามารถในการใช้งานคล้ายกันกับภาษา C++

ตาราง 7.1 เปรียบเทียบความแตกต่างกันของแต่ละภาษา

	MQL 5	C++	C
การอ้างอิงตำแหน่งในหน่วยความจำ	ไม่มี	มี	มี
Anonymous Enumeration	ไม่มี	มี	มี
Parameters ของ Constructor	Constructor ต้องไม่มี Parameters	มีหรือไม่มีก็ได้	ไม่มี Constructor
Multiple Inheritance	ไม่มี	มี	ไม่มี Inheritance

### 7.1.2 ชนิดของข้อมูล (Data Type)

เนื่องจากข้อมูลทุกชนิดนั้นจำเป็นต้องถูกกำหนดชนิดของข้อมูลนั้น โดยมีความแตกต่างกันตามวัตถุประสงค์และชนิดของข้อมูลที่แตกต่างกันนั้นมีผลต่อความเร็วในการประมวลผลคำสั่งต่างๆ เช่น Integer นั้นจะถูกประมวลผลได้เร็วที่สุด double นั้นอาจจะต้องใช้หลายหน่วยประมวลผลในการทำงาน เพราะความซับซ้อนของตัวข้อมูล

#### 7.1.2.1 Built-in Datatype

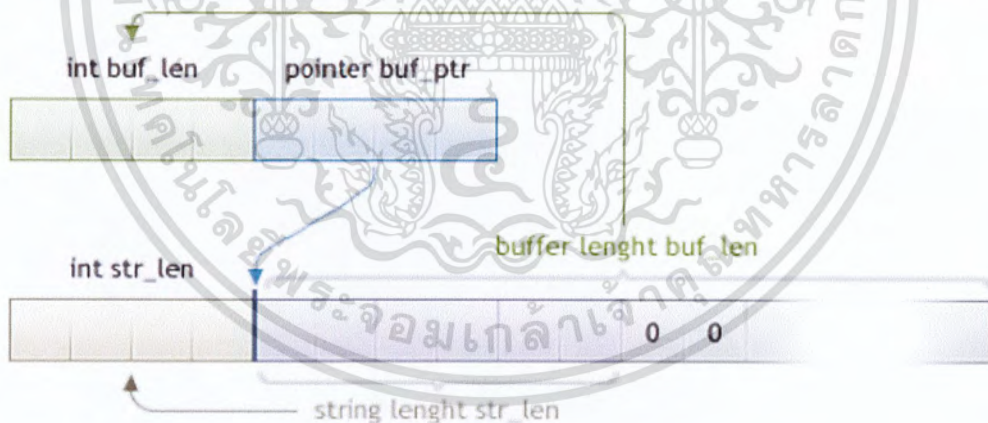
เป็นชนิดของข้อมูลพื้นฐานที่ถูกกำหนดมาแล้วภายในตัวภาษา MQL 5

Type	Size in Bytes	Minimum Value	Maximum Value	C++ Analog
<u>char</u>	1	-128	127	char
<u>uchar</u>	1	0	255	unsigned char, BYTE
<u>bool</u>	1	0(false)	1(true)	bool
<u>short</u>	2	-32 768	32 767	short, wchar_t
<u>ushort</u>	2	0	65 535	unsigned short, WORD
<u>int</u>	4	-2 147 483 648	2 147 483 647	int
<u>uint</u>	4	0	4 294 967 295	unsigned int, DWORD
<u>color</u>	4	-1	16 777 215	int, COLORREF
<u>long</u>	8	-9 223 372 036 854 775 808	9 223 372 036 854 775 807	__int64
<u>ulong</u>	8	0	18 446 744 073 709 551 615	unsigned __int64
<u>datetime</u>	8	0 (1970.01.01 0:00:00)	32 535 244 799 (3000.12.31 23:59:59)	__time64_t

รูป 7.2 รายละเอียดของชนิดข้อมูลพื้นฐาน

### 7.1.2.2 String Type

ชนิดของข้อมูลที่ไว้สำหรับเก็บตัวอักษรที่เรียงต่อกัน ซึ่งค่าสุดท้ายนั้นเท่ากับศูนย์ บ่งบอกถึงจุดสิ้นสุดของข้อมูล โดยตัวแปร string นั้นเก็บข้อมูลในรูปของ structure แบ่งเป็น 3 ส่วน คือเก็บขนาด buffer ตำแหน่งเริ่มต้นของ buffer ที่เก็บ string และขนาดจริงของชุดข้อมูล



รูป 7.3 การเก็บข้อมูลของ string

### 7.1.2.3 Structure

เป็นชนิดของข้อมูลที่เกิดจากการประกอบกันของชนิดของข้อมูลต่างๆ ยกเว้น void type ขึ้นเป็นชนิดของข้อมูลชนิดใหม่

### 7.1.2.4 Class

1) Key word ที่ต้องใช้คำว่า class

2) คำกำหนดสำหรับการเข้าถึงข้อมูลเริ่มต้น สำหรับ class เป็น private แต่

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการวิจัยเท่านั้น ไม่อนุญาตให้เผยแพร่หรือใช้ประโยชน์ด้านการค้า ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

structure เป็น public

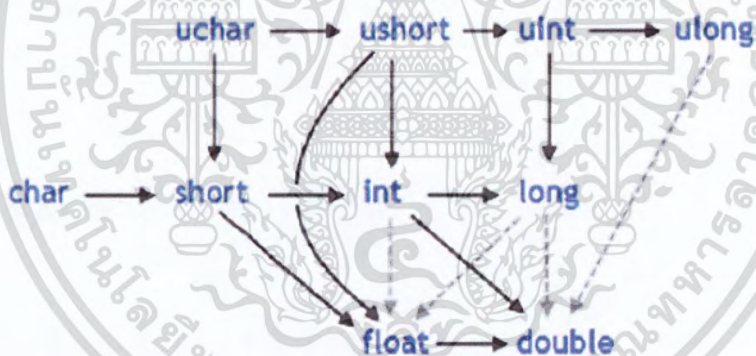
- 3) class จะมีตาราง virtual functions แต่ structure จะไม่มีในส่วนนี้
- 4) class object สามารถใช้ new operator ได้ แต่ structure object ไม่ได้
- 5) class สามารถสืบทอด inherited จาก class ได้เท่านั้น และ structure ก็สามารถสืบทอดได้จาก structure เท่านั้น

### 7.1.1.3 Dynamic Array Object

Array ที่มีขนาดเปลี่ยนแปลงตามการใช้งาน ที่สามารถประกาศการใช้งานได้ 4 มิติ เมื่อทำการประกาศการใช้งาน dynamic array ในส่วนของ compiler นั้นจะสร้าง สำหรับเก็บรายละเอียดที่ถูกต้องของ dynamic array ที่สร้างขึ้นมา และ dynamic array จะถูกคืนค่าหน่วยความจำ เมื่อการทำงานออกนอกขอบเขตของส่วนที่ประกาศ

### 7.1.1.4 TypeCasting

บ่อยครั้งในการ ใช้ข้อมูลที่มีชนิดของข้อมูลที่แตกต่างกันนั้น สามารถที่จะแปลงรูปแบบระหว่างชนิดข้อมูลได้ แต่ไม่ทั้งหมด โดยการแปลงชนิดข้อมูลทั้งอาจทำให้เกิดการสูญเสียของข้อมูลบางส่วน หรืออาจจะไม่ก็ได้



รูป 7.4 การแปลงชนิดข้อมูล เส้นทึบแสดงการไม่สูญเสียข้อมูล เส้นประแสดงการสูญเสียข้อมูล



รูป 7.5 ลำดับการแปลงชนิดข้อมูล

สำหรับการแปลงข้อมูลนั้นมีอยู่ 2 แบบ คือ การแปลงโดยอัตโนมัติ (Explicit) และการแปลงที่กำหนดเอง (Implicit)

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 7.1.1.5 Object Pointers

ในการสร้าง object จาก class หรือ structure นั้น จะใช้ new operator คล้ายกับ C++ แต่ตัวแปรที่ได้จาก object นั้นไม่ได้อ้างอิงไปที่ตำแหน่งในหน่วยความจำ โดยค่าของตัวแปรนั้นจะเป็นค่าที่กำหนดอยู่ใน object

### 7.1.1.6 การส่งผ่านค่า และ this

การส่งผ่านค่าข้อมูลมีอยู่ 2 แบบ คือ pass by value และ pass by reference โดยถ้าหากเป็น structure และ class รวมไปถึง array นั้นจะส่งผ่านค่าด้วยวิธี pass by reference ซึ่งการส่งค่าแบบ pass by reference จะต้องใช้สัญลักษณ์ &(ampersand) หน้าชื่อตัวแปร สำหรับ pass by reference นั้น ค่าที่ส่งไปให้ตัวแปรในฟังก์ชันนั้นจะเป็นตำแหน่งของข้อมูล ซึ่งเป็นค่าคงที่ในส่วนของ keyword this นั้นเป็นตัวแปรที่เก็บค่าตำแหน่งของ object ที่สามารถอ้างอิงถึง object นั้นๆ โดย MQL 5 นั้นจะ ไม่มีการคืนค่าเป็น object แต่จะคืนค่าเป็นตำแหน่ง หรือ reference ด้วยการใส่คำสั่ง GetPointer(this)

### 7.1.1.7 การส่งค่าไปยังตัวแปรในฟังก์ชัน

ในการรับค่าของฟังก์ชันนั้น จะทำการรับค่าในลำดับที่สลับกับลำดับที่ประกาศไว้ในฟังก์ชัน

### 7.1.1.8 Function Overloading

บ่อยครั้งที่มีการตั้งชื่อฟังก์ชันที่แสดงถึงการทำงานที่เป็นไปในรูปแบบเดียวกัน แต่ใช้งานแตกต่างกันในเชิงของรายละเอียดวัตถุประสงค์การใช้งานของแต่ละฟังก์ชัน แต่โดยรวมนั้นเหมือนกัน สามารถที่จะใช้การ overloading มาช่วย เพื่อทำให้สร้างฟังก์ชันโดยการตั้งชื่อที่เหมือนกัน แต่มีรายละเอียดในส่วนของ parameters ที่แตกต่างกัน โดยการ overloading ฟังก์ชันนั้นทำได้โดยการกำหนดให้จำนวนของ parameters มีจำนวนที่แตกต่างกัน หรือ มีชนิดข้อมูลของ parameter ที่แตกต่างกันสำหรับ อัลกอริทึม ในการเลือกใช้ฟังก์ชันที่ทำการ overload นั้น compiler จะทำการตรวจสอบตามลำดับดังนี้

- 1) ฟังก์ชันที่ใช้งานนั้นมีรูปแบบตรงกันหมดตามฟังก์ชันที่ประกาศไว้หรือไม่
- 2) ถ้าไม่ตรงตามข้อ 1) จะทำการแปลง parameters ไว้แล้วตรวจสอบว่าตรงกับฟังก์ชันไหน เช่น float เป็น double และ bool , char , short , enum เป็น int
- 3) ถ้ายังไม่ตรงตามข้อ 2) จะทำการแปลงชนิดของ parameter ตามขนาดของชนิดข้อมูล เช่น bool , char , uchar มีขนาด 1 byte

แต่แน่นอนการทำงานแล้วให้ผลลัพธ์ตามข้อ 1) นั้นจะให้ฟังก์ชันที่ตรงตามความต้องการมากที่สุด และ compiler ยังไม่สามารถเลือกฟังก์ชันที่ถูกแปลงมาแล้ว มีลักษณะที่กำกวมคือ เป็นไปได้หลายรูปแบบ ดังนั้นไม่ควรที่จะเลือกใช้งานฟังก์ชันโดยวิธีการที่แปลงชนิดของข้อมูลเพื่อเลือกใช้ฟังก์ชัน

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

### 7.1.1.9 การเรียกใช้งานฟังก์ชันจากภายนอก

ในกรณีที่ฟังก์ชันที่ต้องการใช้งานนั้นถูกประกาศอยู่ในไฟล์อื่น นอกเหนือจากไฟล์ที่กำลังใช้งานอยู่ ดังนั้นถ้าหากเรียกใช้งานฟังก์ชันนั้นเลยจะเกิด error เพราะฉะนั้นจึงต้องใช้ keyword `#import` อ้างอิงไปยังไฟล์ต้นทาง

ในส่วนของการเรียกใช้งานจากไฟล์ ex 5 นั้นตัวฟังก์ชันภายในไฟล์นั้นจะต้องมี keyword `export` กำหนดอยู่ที่ฟังก์ชัน จึงจะสามารถใช้งานได้

### 7.1.1.10 Event handling Functions

ใน MQL 5 นั้นมีฟังก์ชันที่ถูกกำหนดการใช้งานตาม event ที่เกิดขึ้น โดยที่ event นั้นเป็นส่วนที่ถูกกำหนดมาอยู่แล้วใน MQL 5 ว่าเมื่อเกิดเหตุการณ์อย่างนี้จะเกิด event อย่างนี้แล้วจะส่ง event ไปเก็บไว้ใน message queue รอการใช้งานฟังก์ชันที่มีในโปรแกรมตามลำดับ ถ้าหาก message queue เต็ม event ที่ถูกส่งมาใหม่จะเข้าทำการแทนที่ event ที่อยู่ใน queue นานที่สุดสำหรับการใช้งานฟังก์ชันที่ใช้จัดการ event นั้นจะต้องถูกกำหนดไว้ในโปรแกรมให้มีรูปแบบถูกต้องตามที่กำหนดมาใน MQL5 มิฉะนั้น event บางตัวที่เรียกใช้งานฟังก์ชันที่ไม่ถูกต้องอาจจะไม่มีฟังก์ชันใดๆถูกเรียกขึ้นมาใช้งาน

Type	Function name	Parameters	Application	Comment
int	<a href="#">OnInit</a>	none	Expert Advisors and indicators	<a href="#">Init</a> event handler. It allows to use the void return type.
void	<a href="#">OnDeinit</a>	const int reason	Expert Advisors and indicators	<a href="#">Deinit</a> event handler.
void	<a href="#">OnStart</a>	none	scripts	<a href="#">Start</a> event handler.
int	<a href="#">OnCalculate</a>	const int rates_total, const int prev_calculated, const datetime &Time[], const double &Open[], const double &High[], const double &Low[], const double &Close[], const long &TickVolume[], const long &Volume[], const int &Spread[]	indicators	<a href="#">Calculate</a> event handler for all prices.
int	<a href="#">OnCalculate</a>	const int rates_total, const int prev_calculated, const int begin, const double &price[]	indicators	<a href="#">Calculate</a> event handler on the single data array. Indicator cannot have two event handlers simultaneously.
				In this case the only one event handler will work on the data array.
void	<a href="#">OnTick</a>	none	Expert Advisors	<a href="#">NewTick</a> event handler. While the event of a new tick receipt is being processed, no other events of this type are received.
void	<a href="#">OnTimer</a>	none	Expert Advisors and indicators	<a href="#">Timer</a> event handler.
void	<a href="#">OnTrade</a>	none	Expert Advisors	<a href="#">Trade</a> event handler.
double	<a href="#">OnTester</a>	none	Expert Advisors	<a href="#">Tester</a> event handler.
void	<a href="#">OnChartEvent</a>	const int id, const long &lparam, const double &dparam, const string &sparam	Expert Advisors and indicators	<a href="#">ChartEvent</a> event handler.
void	<a href="#">OnBookEvent</a>	const string &symbol_name	Expert Advisors	<a href="#">BookEvent</a> event handler.

รูป 7.6 ตัวอย่าง event handling functions

## 7.2 ส่วนการทำงานของโปรแกรมที่สร้างจาก MQL 5

หลังจากโปรแกรมที่สร้างขึ้นถูกนำไปใช้ควบคู่กับกราฟแสดงการเปลี่ยนแปลงของราคาค่าเงินแล้ว นั่นก็คือโปรแกรมนั้นๆ ก็จะทำงานอยู่ภายในส่วนของ client terminal สำหรับการซื้อขายแล้ว ถัดจากนั้นจะทำการกำหนดค่าเริ่มต้นให้ตัวแปรที่เป็นแบบ global ที่ถูกสร้างขึ้นมาจาก class ที่สร้างเอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น เมื่อนุญาตให้เข้าไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

ขึ้นเพื่อกำหนดข้อมูลและการทำงานต่างๆที่ต้องการ โดยจะกำหนดค่าเริ่มต้นผ่านสิ่งที่เรียกว่า constructor หลังจากนั้นตัวโปรแกรมจะทำการรอรับ event ที่เกิดขึ้นและจะถูกเรียกใช้งานตามฟังก์ชันที่เรียกใช้ภายในโปรแกรม โดยประเภทของโปรแกรมนั้นสามารถแบ่งตามจุดประสงค์และรูปแบบของการใช้งานได้ 3 ประเภท

- 1) Expert Advisor คือ กลไกในการซื้อขาย ที่สร้างขึ้นให้สามารถอ้างอิงกับการเปลี่ยนแปลงของอัตราแลกเปลี่ยนที่เปลี่ยนแปลงอยู่เสมอ และจะทำงานเมื่อได้รับ event ที่เกิดขึ้นในระบบ สำหรับ Expert Advisor นั้น โดยปกติจะมีฟังก์ชันสำหรับรับ event อยู่ 3 ส่วนของกระบวนการ คือ ส่วนการเริ่มต้นการทำงาน ส่วน event ที่เกิดจากการเปลี่ยนแปลงของอัตราค่าเงิน และส่วนของการสิ้นสุดการทำงาน
- 2) Custom Indicator คือ Indicator ที่สร้างเพิ่มขึ้นมาจาก Indicator ที่มีให้อยู่แล้วในระบบ ไว้ใช้สำหรับการวิเคราะห์บางอย่างในสถานการณ์เฉพาะช่วงเวลาหนึ่ง
- 3) Script คือ โปรแกรมที่ถูกสร้างขึ้นเพื่อการทำงานในขอบเขตงานที่จำกัดหนึ่งๆ โดยไม่สนใจประมวลผลข้อมูลที่ได้รับจาก event ภายนอก ยกเว้น OnStart()



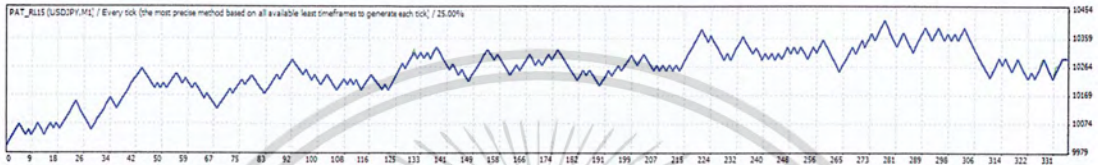
เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## บทที่ 8

### การทดลองและผลการทดลอง

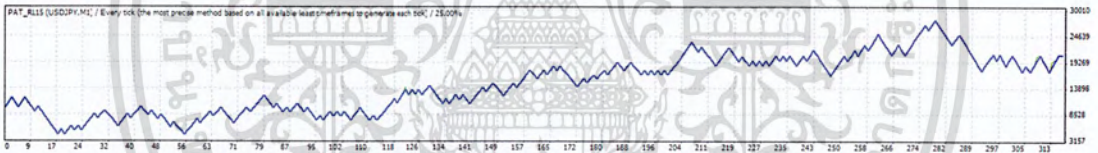
#### 8.1 การทดลอง

ทำการทดลองตามอัลกอริทึม และใช้ค่าคงที่ job episode เท่ากับ 2 ทำการซื้อขายที่ละสัญญา โดยใช้จำนวน lots เท่ากับ 0.01 ตั้งแต่วันที่ 1-1-2000 ถึง 26-8-2010 ได้ผลลัพธ์ดังนี้



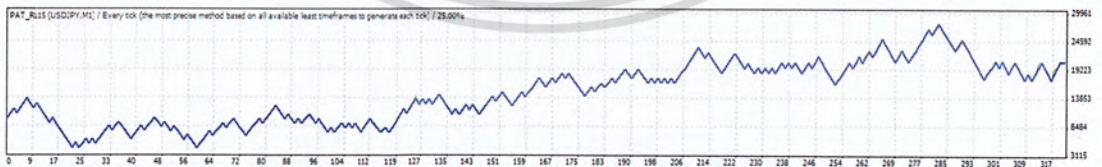
รูป 8.1 จำนวนเงินของเรทที่มี โดยใช้จำนวน lots เท่ากับ 0.01 ตั้งแต่วันที่ 1-1-2000 ถึง 26-8-2010

ทำการซื้อขายที่ละสัญญา โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 2-3-2000 ถึง 26-8-2010 ได้ผลลัพธ์ดังนี้



รูป 8.2 จำนวนเงินของเรทที่มี โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 2-3-2000 ถึง 26-8-2010

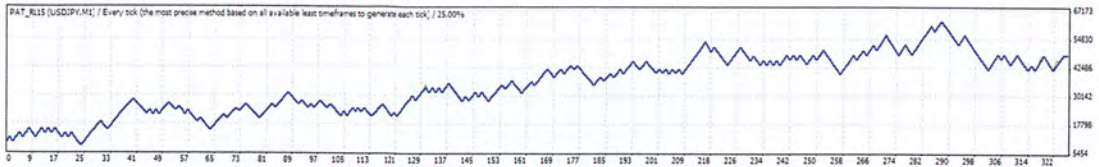
ทำการซื้อขายที่ละสัญญา โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 7-1-2000 ถึง 26-8-2010 ได้ผลลัพธ์ดังนี้



รูป 8.3 จำนวนเงินของเรทที่มีโดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 7-1-2000 ถึง 26-8-2010

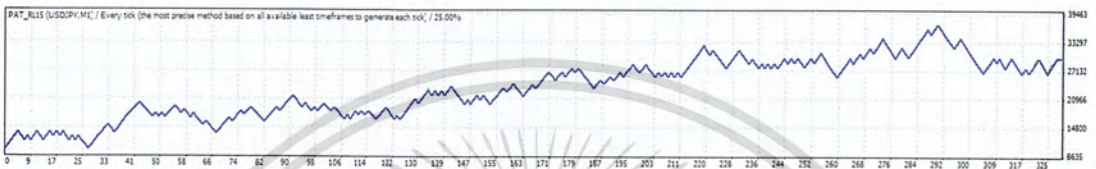
ทำการซื้อขายที่ละสัญญา โดยใช้จำนวน lots เท่ากับ 1 ตั้งแต่วันที่ 1-2-2000 ถึง 26-8-2010 ได้ผลลัพธ์ดังนี้

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้



รูป 8.4 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 1 ตั้งแต่วันที่ 1-2-2000 ถึง 26-8-2010

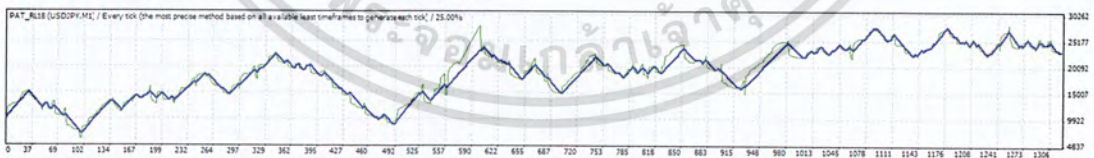
ทำการซื้อขายทีละสัญญา โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 13-1-2000 ถึง 26-8-2010 ได้ผลลัพธ์ดังนี้



รูป 8.5 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 0.5 ตั้งแต่วันที่ 4-1-2000 ถึง 26-8-2010

จากการทดลองจะเห็นได้ว่าผลการทดลองในแต่ละครั้งลักษณะของกราฟจะออกมาคล้ายๆกัน ไม่ต่างกันมาก ความชันของกราฟค่อนข้างดีมีแนวโน้มที่จะเพิ่มขึ้นเรื่อยๆแต่จากการวิเคราะห์ดูแล้ว การซื้อขายทีละสัญญาอาจจะมีปัญหาการเลื่อนของช่วงเวลาทำให้วันที่เริ่มเล่นต่างกันผลลัพธ์ก็อาจจะต่างกันได้ ซึ่งถ้าการซื้อขายนั้นเลื่อนไปโดนช่วงต่อการซื้อขายระหว่างที่จบสัญญาแล้วซื้อสัญญาใหม่ไม่ถูกจุด อาจจะทำให้กราฟออกมาขาดทุนก็เป็นได้ จึงทำการทดลองซื้อขายทีละหลายๆสัญญา

ทำการซื้อขายวันละสัญญา โดยใช้จำนวน lots เท่ากับ 0.1 ตั้งแต่วันที่ 1-1-2000 ถึง 26-8-2010 ได้ผลลัพธ์ดังนี้



รูป 8.6 จำนวนเงินของเราที่มี โดยใช้จำนวน lots เท่ากับ 1 ตั้งแต่วันที่ 1-1-2000 ถึง 26-8-2010

จากกราฟจะเห็นได้ว่าแนวโน้มส่วนใหญ่จะเพิ่มขึ้นคล้ายๆเล่นทีละสัญญา แต่มีช่วงที่ลงมากอยู่บางช่วง ซึ่งดูจากช่วงที่ลงหนักๆแล้วจะเห็นว่าช่วงที่ลงนั้นหักล้างกำไรที่ได้มาหมดไปเลย และถ้าเกิดไปเริ่มตรงจุดที่อยู่บนสุดก่อนจะถึงช่วงที่ลงหนักจะทำให้เงินที่มีอยู่เสียไปจนหมดก่อนที่จะกลับมาได้กำไรในช่วงหลัง

เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้าไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้

## Closed Transactions:

Ticket	Open Time	Type	Size	Item	Price	S / L	T / P	Close Time	Price	Commission	Taxes	Swap	Profit
175650793	2011.02.01 06:48	balance		Deposit									10 000.00
175813736	2011.02.01 16:09	buy	0.10	usdjpy	81.878	81.678	82.078	2011.02.01 16:28	81.678	0.00	0.00	0.00	-24.49
175826976	2011.02.01 16:43	buy	0.20	usdjpy	81.648	81.448	81.848	2011.02.01 17:05	81.448	0.00	0.00	0.00	-49.11
175838870	2011.02.01 17:08	buy	0.40	usdjpy	81.465	81.265	81.665	2011.02.02 17:18	81.665	0.00	0.00	0.10	97.96
176112907	2011.02.02 17:19	sell	0.10	usdjpy	81.665	81.865	81.465	2011.02.03 15:45	81.865	0.00	0.00	-0.26	-24.43
176394614	2011.02.03 15:46	buy	0.20	usdjpy	81.862	81.662	82.062	2011.02.03 17:07	81.662	0.00	0.00	0.00	-48.98
176431402	2011.02.03 17:08	buy	0.40	usdjpy	81.630	81.430	81.830	2011.02.03 18:11	81.430	0.00	0.00	0.00	-98.24
176453138	2011.02.03 18:12	buy	0.80	usdjpy	81.441	81.241	81.641	2011.02.03 20:03	81.641	0.00	0.00	0.00	195.98
176477778	2011.02.03 20:08	buy	0.10	usdjpy	81.646	81.446	81.846	2011.02.04 14:30	81.446	0.00	0.00	0.00	-24.56
176626862	2011.02.04 14:33	buy	0.20	usdjpy	81.545	81.345	81.745	2011.02.04 14:37	81.745	0.00	0.00	0.00	48.93
176635174	2011.02.04 14:41	buy	0.10	usdjpy	81.859	81.659	82.059	2011.02.04 14:45	81.659	0.00	0.00	0.00	-24.49
176641037	2011.02.04 14:51	buy	0.20	usdjpy	81.821	81.621	82.021	2011.02.04 17:17	82.021	0.00	0.00	0.00	48.77
176691666	2011.02.04 17:18	buy	0.10	usdjpy	82.024	81.824	82.224	2011.02.04 18:15	82.224	0.00	0.00	0.00	24.32
176708778	2011.02.04 18:18	buy	0.10	usdjpy	82.198	81.998	82.398	2011.02.04 18:50	82.398	0.00	0.00	0.00	24.27
176719439	2011.02.04 19:05	buy	0.10	usdjpy	82.409	82.209	82.609	2011.02.04 20:58	82.209	0.00	0.00	0.00	-24.33
176731215	2011.02.04 20:59	buy	0.20	usdjpy	82.234	82.034	82.434	2011.02.07 08:54	82.434	0.00	0.00	0.05	48.52
176789542	2011.02.07 09:28	buy	0.10	usdjpy	82.363	82.163	82.563	2011.02.08 08:52	82.163	0.00	0.00	0.02	-24.34
177068005	2011.02.08 08:55	buy	0.20	usdjpy	82.164	81.964	82.364	2011.02.08 16:42	81.964	0.00	0.00	0.00	-48.80
177231378	2011.02.08 16:44	buy	0.40	usdjpy	81.929	81.729	82.129	2011.02.08 19:24	82.129	0.00	0.00	0.00	97.41
177276692	2011.02.08 19:30	sell	0.10	usdjpy	82.166	82.366	81.966	2011.02.08 20:51	82.366	0.00	0.00	0.00	-24.28
177297849	2011.02.08 20:55	buy	0.20	usdjpy	82.378	82.178	82.578	2011.02.09 09:55	82.578	0.00	0.00	0.00	48.44
										0.00	0.00	-0.09	218.55
													<b>Closed P/L: 218.46</b>

รูป 8.7 ผลการทดลองโดยใช้อัลกอริทึมเพิ่มเติมและดับเบิ้ลล็อต

จากผลการทดลองโดยใช้อัลกอริทึมเพิ่มเติมและดับเบิ้ลล็อตจะเห็นได้ว่าได้กำไรดีมาก โดยผลการทดลองนี้ทดลอง 8 วันได้กำไรมา 218 จาก 10000 ที่จริงเริ่มจาก 5000 ก็พอ ถือว่าได้กำไรค่อนข้างดีมากแต่อาจจะมีปัญหาคือการทดลองนี้มีจำนวนวันน้อยจึงอาจจะมีโอกาสขาดทุนได้ ต้องดูต่อไปในระยะยาว

## บทที่ 9

# สรุปผลการทดลอง

### 9.1 สรุปผลทดลอง

เนื่องจากพบปัญหาของตัวโปรแกรมที่ทำการจำลองการเล่นนั้นมีข้อมูลไม่ต่อเนื่องทำให้ไม่สามารถวัดอะไรได้มาก จึงทำการทดลองเล่นจริงตามเวลาปัจจุบันทำให้ผลการทดลองที่มีอยู่มีไม่มากนัก การทดลองที่ตามอัลกอริทึมเพิ่มเติมรวมกับอัลกอริทึมในการดับเบิลคือพบว่าผลการทดลองออกมาใช้ได้มีส่วนใหญ่ แต่การทดลองมีจำนวนวันไม่มากทำให้ยังไม่สามารถรับประกันได้ว่าจะได้กำไรแน่นอนตลอดไป

จากการทดลองเราพบว่าโปรแกรมที่เราเขียนขึ้นนั้นมีทั้งได้กำไรและขาดทุนโดย agent แต่ละตัวนั้นจะได้กำไรหรือขาดทุนนั้นหากวิเคราะห์ดูจะพบว่าการ ดับเบิลคือ จะทำให้มีโอกาสได้กำไรมากขึ้นแต่ถ้ามองให้ลึกแล้วก็มาพร้อมด้วยโอกาสเสียที่มากขึ้นด้วย โดยการทดลองนี้นั้นชี้ให้เห็นความเป็นจริงที่ว่าค่าเงินในอดีตนั้นไม่ได้ส่งผลโดยตรงกับราคาในอนาคตมากนัก จากการที่ได้เข้าไปหาข้อมูลถึงเรื่องนี้เราจึงขออธิบายว่า agent ที่เราสร้างนั้นเราคาดหวังว่าจะได้กำไรได้ถึง 60% ก็นับว่าดีแล้ว แต่หากท่านมีข่าวสารที่เชื่อถือได้ท่านสามารถนำข่าวสารนั้นมาเลือกประกอบการณ์ให้ agent ได้อย่างเหมาะสม โอกาสได้กำไรของท่านก็จะมีความมากขึ้น โดยที่ไม่ต้องเหนื่อยเล่นเองเราคาดหวังว่าผู้ที่ต้องการศึกษาในเรื่องนี้หรือคิดที่จะเล่นได้ใช้ข้อมูลที่เราศึกษามาให้เป็นประโยชน์ไม่มากนัก

### 9.2 แนวทางการพัฒนาต่อ

แนวทางการปรับปรุงพัฒนาเพิ่มเติมมีดังนี้ สามารถทำการทดลองเปลี่ยนจำนวน state ให้เหมาะสมกับอัลกอริทึมมากยิ่งขึ้น และทดลองเปลี่ยนอินดิเคเตอร์ให้เข้ากับอัลกอริทึมมากยิ่งขึ้น และในส่วนของอัลกอริทึมเพิ่มเติมที่ใช้กับอัลกอริทึมดับเบิลคือสามารถทำการเปลี่ยนโดยเลือกอัลกอริทึมที่คิดว่ามีความสามารถในการได้กำไรสูงมาทดลองใช้แทนและปรับค่าพารามิเตอร์ต่างๆ ให้เหมาะสม

## บรรณานุกรม

Richard, S. and Andrew, G. 1998. **Reinforcement Learning: An Introduction**. Cambridge : MIT Press.

เอกพล อนันตพรกิจ. 2552. “การเลือกเซตย่อยสอดคล้องขนาดเล็กที่สุดโดยใช้การเรียนรู้แบบเสริมกำลัง.” วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรม คอมพิวเตอร์ บัณฑิตวิทยาลัย, สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง.

Wikipedia. 2010. **Reinforcement learning**. [Online].

Available : [http://en.wikipedia.org/wiki/Reinforcement\\_learning](http://en.wikipedia.org/wiki/Reinforcement_learning).



เอกสารนี้เป็นเอกสารที่สงวนไว้สำหรับการใช้งานเพื่อการศึกษาเท่านั้น ไม่อนุญาตให้นำไปใช้ประโยชน์ด้านการค้า  
ไม่ว่ากรณีใดๆ ทั้งสิ้น อีกทั้งห้ามมิให้ตัดแปลงเนื้อหา และต้องอ้างอิงถึงเจ้าของเอกสารทุกครั้งที่มีการนำไปใช้