บทที่ 8

ทฤษฎีเข้ารหัส (Coding Theory)

Coding theory is dealing with the error-prone process of transmitting data across noisy channels, so that errors that occur can be corrected.

การส่งคลื่นวิทยุผ่านอวกาศมายังโลก เป็นระยะทางหลายล้านกิโลเมตร ต้องผ่านสิ่งแวดล้อมที่ไม่พึงประสงค์หลายประการ เช่น สัญญาณรบกวน การลดทอน ความแปรปรวน ฯลฯ แต่คลื่นวิทยุต้นทางมีกำลังส่งเพียงแค่ไม่กี่วัตต์เท่านั้น **ทฤษฎีการเข้ารหัส** (Coding Theory) ว่าด้วยขั้นตอนวิธีทางคณิตศาสตร์และวิทยาการคอมพิวเตอร์ เพื่อการส่งข้อมูลผ่านช่องสัญญาณ ที่มีสัญญาณรบกวน และการกู้คืนข้อมูลข่าวสารที่ด้านรับ รวมถึงกระบวนการที่ทำให้ข่าวสารสามารถอ่านได้ง่าย

การเข้ารหัส

เพื่อความสอดคล้องของเนื้อหาสำหรับการศึกษาทฤษฎีการเข้ารหัส (Coding) ในที่นี้ กำหนดให้

- ข้อมูลอยู่ในรูปของเลขฐานสอง (Binary Digits หรือ Bits)
- ข้อมูลส่งไปตามสายสัญญาณ ที่มีสัญญาณรบกวนแบบสุ่ม (Random Noise)
- ผู้ส่งไม่สามารถ ทำนายค่าของสัญญาณรบกวน ณ เวลาใดๆ ได้ แต่รู้อัตราของสัญญาณรบกวน (เช่น เป็น dB เทียบกับสัญญาณ)
- การเข้ารหัสที่มี ภูมิต้านทาน ต่อสัญญาณรบกวน รหัสต้องมีความยาว (จำนวน Bits) มากกว่าข้อมูลต้นฉบับ

กรอบการทำงานของระบบเข้ารหัสข้อมูลสามารถแสดงได้ดังแผนผังต่อไปนี้

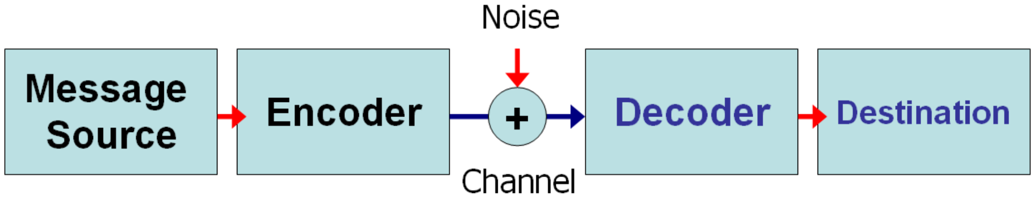

FIGURE 8.1 แผนผังของระบบเข้ารหัสประกอบด้วย แหล่งกำเนิดข่าวสาร ส่วนเข้ารหัส ช่องสัญญาณ (พิจารณาผลของสัญญาณรบกวน) ส่วนถอดรหัส และปลายทางรับข้อมูล

แนวคิดของข่าวสาร
ข่าวสาร (Information) หมายถึง **ความรู้ใหม่** ซึ่งอาจนำเสนอแนวคิดได้จากเหตุการณ์ดังต่อไปนี้

พิจารณากล่องบรรจุลูกบอล 3 กล่อง ดังต่อไปนี้

ลูกบอลสีดำ 1 ลูก และ ลูกบอลสีขาว 1 ลูก

ลูกบอลสีดำ 9 ลูก และ ลูกบอลสีขาว 1 ลูก

ลูกบอลสีดำ 999 000 ลูก และ ลูกบอลสีขาว 1 000 ลูก

จากทฤษฎีของความน่าจะเป็น พบว่า โอกาสที่จะ **ทำนาย** (*a priori*) ว่า *ลูกบอลที่จะหยิบออกมาเป็นสีใด* ในกรณีที่ 1 (ความน่าจะเป็นพอๆ กัน) <u>ยากกว่า</u> กรณีที่ 2 และ 3 (บอลน่าจะเป็นสีดำมากกว่า)

เมื่อเราหยิบลูกบอลออกจากกล่องแล้ว (*a posteriori*) สำหรับกรณี 1 ไม่ว่าหยิบได้สีขาวหรือดำ **เราจะได้ความรู้ใหม่เสมอ** (คาดเดาไม่ได้) ทว่าสำหรับกรณีที่ 2 ถ้าหยิบได้สีดำ **เราไม่ได้ความรู้ใหม่** (คาดว่าน่าจะเป็นสีดำอยู่แล้ว) แต่ถ้าหยิบได้สีขาว เราได้ความรู้ใหม่มากมาย (คาดไม่ถึงมาก่อน) ดังนั้นอาจกล่าวได้ว่าข่าวสาร หรือความรู้ใหม่ มีความสัมพันธ์กับความไม่แน่นอน (Uncertainty)

ความแปรปรวน
ความแปรปรวน (Entropy) สามารถอธิบายได้โดย กำหนดการทดลอง **E** ที่มี ผลลัพธ์ m กรณี ได้แก่ X = $\{x_1, x_2, \ldots, x_m\}$

ให้ผลลัพธ์แต่ละตัว x_i มีค่าความน่าจะเป็นที่จะเกิดขึ้น p_i โดยที่ ผลรวมของ p_i สำหรับทุกๆ กรณี ($\Sigma_i p_i$) มีค่าเท่ากับ **1** แล้ว ความไม่แน่นอนของการทดลอง **E** สามารถวัดได้ด้วย Entropy ซึ่งนิยามดังต่อไปนี้

$$H_b(X) = -\sum_{i=1}^{m} p_i \log_b p_i$$

เมื่อ $b = 2$ หน่วยของ Entropy จะมีค่าเป็น *bit*

$b = \mathrm{e}$ หน่วยของ Entropy จะมีค่าเป็น *nat*

$b = 10$ หน่วยของ Entropy จะมีค่าเป็น *decit*

ความแปรปรวน และข่าวสาร
ปริมาณของข่าวสาร (Information) ของการทดลอง **E** เมื่อกำหนด $X = \{ x_i \}$ นิยามโดยสมการ

$$I_b(x_i) = -\log_b p_i$$

จากนิยามข้างต้นอาจสรุปได้ว่า Entropy คือ Expected Value หรือ ค่าเฉลี่ยถ่วงน้ำหนัก (mean) ของ Information ดังนี้

$$H_b(X) = E[I(x_i)]$$
$$= -\sum_{i=1}^{m} p_i \log_b p_i$$

ข่าวสารหลายตัวแปร
สมมติตัวแปรสุ่ม X = {$x_1, x_2, \ldots, x_m$ }ให้ผลลัพธ์แต่ละตัว x_iมีค่าความน่าจะเป็นที่จะเกิดขึ้น p_i โดยที่ ผลรวมของ p_i สำหรับทุกๆ กรณี ($\Sigma_i p_i$) มีค่าเท่ากับ **1** ในทำนองเดียวกัน และสมมติตัวแปรสุ่ม Y = {y_1, $y_2, \ldots, y_n$ }ให้ผลลัพธ์แต่ละตัว y_k มีค่าความน่าจะเป็นที่จะเกิดขึ้น q_k โดยที่ ผลรวมของ q_k สำหรับทุกๆ กรณี ($\Sigma_k q_k$) มีค่าเท่ากับ **1**

ดังนั้นสำหรับตัวแปรสุ่ม (X, Y) = { (x_i, y_k) }มีความน่าจะเป็น p_{ik} = P { X=x_i, Y=y_k } Entropy ของตัวแปรสุ่ม (X, Y) นิยามด้วยสมการ

$$H(X,Y) = -\sum_{i=1}^{m}\sum_{k=1}^{n} p_{ik}\log p_{ik}$$

สำหรับระบบที่ประกอบด้วยตัวแปรสุ่ม 2 ตัว (X, Y) Entropy ของตัวแปรสุ่ม Y เมื่อกำหนดตัวแปรสุ่ม X (Conditional Entropy) นิยามได้ดังนี้

$$H(Y|X = x_i) = -\sum_{k=1}^{n} p_{k|i}\log p_{k|i}$$
$$H(Y|X) = \sum_{i=1}^{m} p_i H(Y|X = x_i)$$

ความสัมพันธ์ของค่า Entropy ในรูปแบบต่างๆ จึงเขียนได้ดังนี้ (ความสัมพันธ์บรรทัดสุดท้าย เป็นสมการก็ต่อเมื่อ X และ Y เป็นอิสระต่อกัน)

$$H(X,Y) = H(X) + H(Y|X)$$
$$H(X,Y) \le H(X) + H(Y)$$
$$H(X,Y) = H(X) + H(Y)$$

ช่องทางการส่งผ่านข่าวสาร

การส่งผ่านข่าวสารในการสื่อสาร สามารถแสดงในรูปแบบจำลองทางคณิตศาสตร์ของช่องสัญญาณ (Transmission Channel) ได้ดังนี้

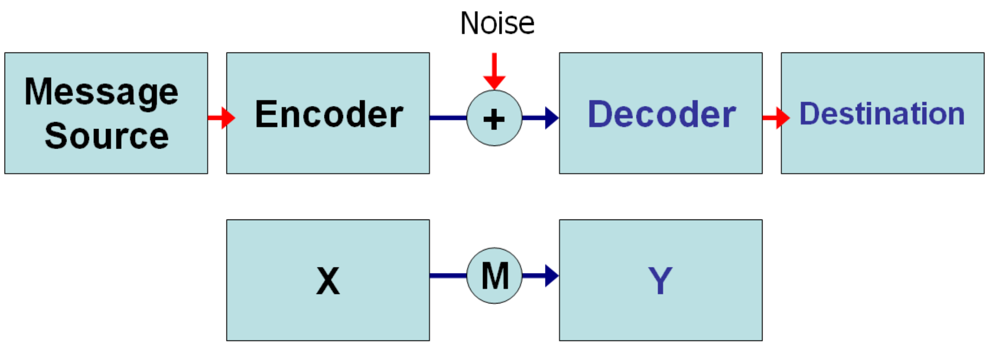

FIGURE 8.2 แบบจำลองการส่งข้อมูลจาก X ไปยัง Y ผ่านช่องสัญญาณ M ซึ่งแสดงในรูปแบบจำลองของการเข้ารหัสข้อมูล

จากแผนผังดังรูป Input ได้แก่ X ให้กำเนิดสัญลักษณ์จากภาษา A = $\{x_1, x_2, \ldots, x_m\}$

Output ได้แก่ Y ให้กำเนิดสัญลักษณ์จากภาษา B = $\{y_1, y_2, \ldots, y_n\}$

ช่องสัญญาณ Transmission Matrix [**T**] นิยามด้วย $t_{ik} = p_{k|i}$

สามารถเขียนความสัมพันธ์เชิงเส้นของ Input และ Output (**Y** = **TX**) ได้ดังนี้

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} p_{1|1} & p_{1|2} & \cdots & p_{1|m} \\ p_{2|1} & p_{2|2} & \cdots & p_{2|m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n|1} & p_{n|2} & \cdots & p_{n|m} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

ระบบสมการข้างต้นอธิบายถึงความน่าจะเป็นของ Output เมื่อทราบความน่าจะเป็นของ Input และสมบัติของช่องสัญญาณ อย่างเช่น ความน่าจะเป็นของการเกิด y_1 มีค่าเท่ากับผลรวมของ ความน่าจะเป็นของการเกิด y_1 ถ้าทราบว่าเกิด x_i (สมบัติของช่องสัญญาณ) คูณกับความน่าจะเป็นของการเกิด x_i

ช่องสัญญาณเลขฐานสอง

ช่องสัญญาณเลขฐานสอง (Binary Channel: BC) เป็น Transmission Channel รูปแบบหนึ่ง ซึ่งสามารถนิยามโดยอ้างอิงกับแบบจำลองข้างต้นได้ดังนี้

Input ได้แก่ X ให้กำเนิดสัญลักษณ์จากภาษา A = $\{0, 1\}$

Output ได้แก่ Y ให้กำเนิดสัญลักษณ์จากภาษา B = $\{0, 1\}$

สามารถแสดงความสัมพันธ์ระหว่าง X และ Y ด้วยระบบสมการของ Transmission Channel ดังนี้

$$\begin{aligned} \begin{bmatrix} p(Y=0) \\ p(Y=1) \end{bmatrix} &= \begin{bmatrix} p_{y=0|x=0} & p_{y=0|x=1} \\ p_{y=1|x=0} & p_{y=1|x=1} \end{bmatrix} \begin{bmatrix} p(X=0) \\ p(X=1) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} p(X=0) \\ p(X=1) \end{bmatrix} \Rightarrow Y = X \end{aligned}$$

ถ้า Transmission Matrix เป็น Identity Matrix จะได้ว่าช่องสัญญาณนั้นไม่ถูกรบกวนด้วยสัญญาณรบกวน (Noise) ทำให้สัญญาณปลายทางเหมือนกันทุกประการกับสัญญาณต้นทาง

ถ้าช่องสัญญาณถูกรบกวน ซึ่งมีความน่าจะเป็นในการเกิดความผิดพลาด = q จะได้ว่า

$$\begin{bmatrix} p(Y=0) \\ p(Y=1) \end{bmatrix} = \begin{bmatrix} 1-q & q \\ q & 1-q \end{bmatrix} \begin{bmatrix} p(X=0) \\ p(X=1) \end{bmatrix}$$

เรียกช่องสัญญาณที่มีแบบจำลองข้างต้นว่า Binary Symmetric Channel (BSC) เนื่องจากความผิดพลาดมีค่าความน่าจะเป็นเท่ากัน (เท่ากับ q) ทั้งในกรณีที่ Input มีค่าเป็น 0 และ 1 (สมมาตรกัน)

ข่าวสารร่วม

ข่าวสารร่วม (Mutual Information: MI) **ของการส่งข้อมูลจากแหล่งกำเนิด** X **ผ่าน** M **ไปยังด้านรับ** Y **นิยามโดย**

$$I(X;Y) = H(X) - H(X|Y)$$
$$= H(Y) - H(Y|X)$$

ความหมายของสมการ (MI)

Mutual Information: "ช่องสัญญาณทำให้เราแปลกใจได้เพียงใด"

ปริมาณข่าวสาร (ความรู้ใหม่ I) หาได้จากความแตกต่างระหว่าง ความไม่แน่นอนของสัญญลักษณ์ที่ Output เมื่อเราไม่ทราบอะไรเลย กับ ความไม่แน่นอนเมื่อเราทราบข้อมูลเกี่ยวกับสัญลักษณ์ **ที่ส่งมา**

ปริมาณข่าวสาร (ความรู้ใหม่ I) หาได้จากความแตกต่างระหว่าง ความไม่แน่นอนของสัญญลักษณ์ที่ Input เมื่อเราไม่ทราบอะไรเลย กับ ความไม่แน่นอนเมื่อเราทราบข้อมูลเกี่ยวกับสัญลักษณ์ **ที่รับได้**

ความจุของช่องสัญญาณ

ความจุของช่องสัญญาณเมื่อพิจารณาสัญญาณรบกวน (Capacity of a Noisy Channel) **สามารถคำนวณได้จาก** Mutual Information **ที่มากที่สุดที่เป็นไปได้**

$$C = \max_{p(X)} I(X;Y)$$

ความหมายของสมการ (C)

Channel Capacity: "ช่องสัญญาณต้องรองรับข้อมูลได้มากเพียงใด"

เมื่อกำหนดสิ่งต่อไปนี้ ความน่าจะเป็นของสัญลักษณ์แต่ละตัวที่ input *p (X)*

คุณสมบัติเชิงสัญญาณรบกวนของช่องสัญญาณ (Transmission Matrix) *T*

สามารถคำนวณค่าต่อไปนี้ได้ H (X), H (Y), H (X, Y), H (Y | X) และ I (X, Y)

C คือค่า I (X; Y) ที่มากที่สุด สำหรับค่า *p (X)* ที่เป็นไปได้ค่าหนึ่ง

แบบฝึกหัด

1. อธิบายหลักการของทฤษฎีการเข้ารหัสมาพอสังเขป

2. กำหนดแหล่งกำเนิด $X = \{ A, B, C, D \}$ ซึ่งสัญลักษณ์แต่ละตัวมีความน่าจะเป็นดังนี้: $P(A) = 0.5$ $P(B) = 0.25$ $P(C) = 0.125$ และ $P(D) = 0.125$ คำนวณหาค่า Entropy $H(X)$

3. จากข้อ 2 หากสัญลักษณ์แต่ละตัวมีความน่าจะเป็นเท่ากัน คำนวณหาค่า $H(X)$ พร้อมทั้งอภิปราย เปรียบเทียบผลลัพธ์ของทั้งสองกรณีในแง่ของ Information

4. กำหนดแหล่งกำเนิด $Y = \{ 0, 1 \}$ ซึ่งสัญลักษณ์แต่ละตัวมีความน่าจะเป็นเท่ากัน และถ้าตัวแปรสุ่ม Y เป็นอิสระจาก X ในข้อ 2 คำนวณหา Joint Entropy $H(X, Y)$

5. จากข้อ 2 และ 4 ในกรณีที่ X และ Y ไม่เป็นอิสระจากกัน กำหนดให้ $P(y_k|A) = P(y_k|B) = 0.40$ และ $P(y_k|C) = P(y_k|D) = 0.10$ คำนวณหา Conditional Entropy $H(Y|X)$

กำหนดแหล่งกำเนิด $X = \{ 00, 01, 10, 11 \}$ ซึ่งสัญลักษณ์แต่ละตัวมีความน่าจะเป็นเท่ากัน ส่งผ่านช่องทางการสื่อสาร (Binary Channel) M ไปยังปลายทาง Y โดยที่ M ประกอบด้วยสัญญาณรบกวนมีผลให้ ความน่าจะเป็นที่ ส่งผ่านข้อมูลถูกต้อง ข้อมูลผิดพลาด 1 บิต และข้อมูลผิดพลาด 2 บิต เป็น 0.8 0.2 และ 0.0 ตามลำดับ ตอบคำถามข้อ 6 – 9

6. เขียนแผนผังแสดงระบบการสื่อสารผ่าน Binary Channel พร้อมทั้ง Transmission Matrix (M)

7. คำนวณหาความน่าจะเป็นของสัญลักษณ์แต่ละตัวที่รับได้ ณ ปลายทาง Y

8. คำนวณหา Entropy ของแหล่งกำเนิด $H(X)$ ปลายทาง $H(Y)$ และ Conditional Entropy $H(X|Y)$

9. คำนวณหา Mutual Information ของการส่งผ่านข้อมูล $I(X; Y)$ และ Channel Capacity (C)

10. อภิปรายความหมายของสมการของ Mutual Information มาพอสังเขป