Vietnam National University Ho Chi Minh City, University of Science
Department of Information Technology

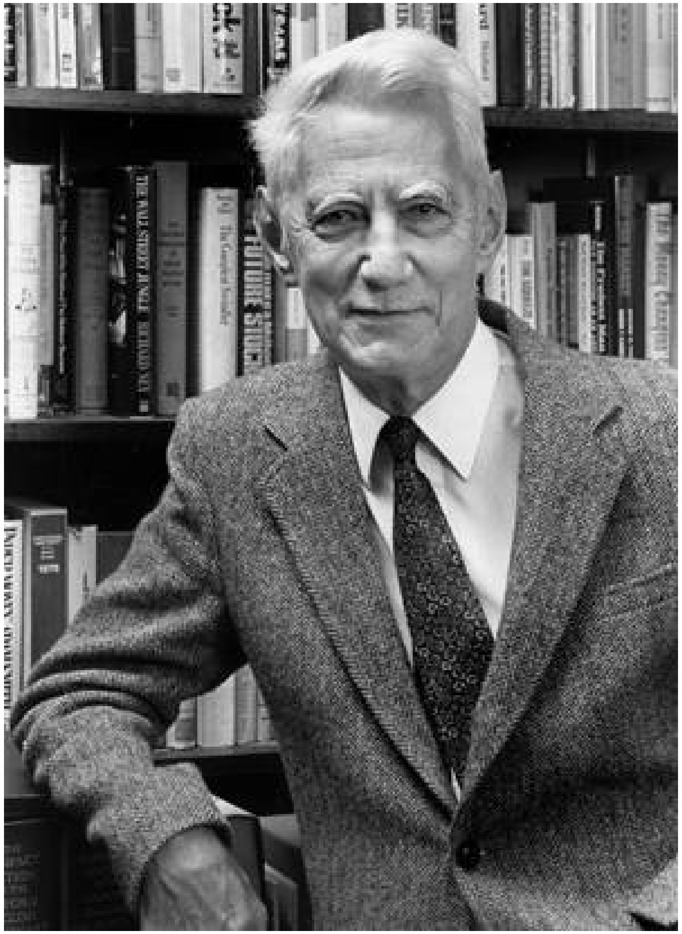
Topic 3:
Shannon theory

Assoc. Prof. Trần Minh Triết
PhD. Trương Toàn Thịnh

TRUONG DAI HOC KHOA HOC TU NHIEN
KHTN
TP. HO CHI MINH
cdio 4.0
fit@hcmus
KHOA CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN

Contents

- Introduction - Claude Shannon
- Perfect security
- Entropy
- Combination of crypto-systems



Claude E. Shannon (1916-2001)

http://www.wisdom.weizmann.ac.il/~naor/COURSE/GiladTsur_YossiOren_ShannonSecrecy.ppt

Probability

- Let X and Y be two random variables.
- Definition:
  - $p(x) = p(X = x)$ is a probability of X receiving value x
  - $p(y) = p(Y = y)$ is a probability of Y receiving value y
  - $p(x \mid y)$ is a probability of X receiving value x if Y receives value y (conditional probability)
- X and Y are independent random variables if only if $p(x, y) = p(x) \times p(y)$ for any value x of X and value y of Y

Probability

- Example: consider tossing 2 dices
  - We have result-space Ω = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),
    (2,1),(2,2),(2,3),(2,4),(2,5),(2,6),
    (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),
    (4,1),(4,2),(4,3),(4,4),(4,5),(4,6),
    (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),
    (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)}
  - |Ω| = 36 elements, for $w \in \Omega$ $w \mapsto X(w)$
  - Let X (based on Ω) be sum of two dices => $X(w) \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Probability

- Example: consider tossing 2 dices
  - We have result-space Ω = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),
    (2,1),(2,2),(2,3),(2,4),(2,5),(2,6),
    (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),
    (4,1),(4,2),(4,3),(4,4),(4,5),(4,6),
    (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),
    (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)}
  - $|\Omega| = 36$ elements, for $w \in \Omega$
  - Let X (based on Ω) be sum of 2 dices => $X(w) \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$
- Notation of mapping $X: \Omega \Rightarrow \mathbb{R}$

$$w \mapsto X(w)$$

- Consider $X(w) = 4 \Rightarrow$ Event of tossing 2 dices has 4 points
  - We have $\mathbf{Pr}[X(w) = 4] = 3/36$, due to {{1, 3}, {2, 2}, {3, 1}}, denote $\mathbf{Pr}[X = 4]$
  - So $\mathbf{Pr}[X = 2] + \mathbf{Pr}[X = 3] + \ldots + \mathbf{Pr}[X = 12] = 1 \Rightarrow \sum_x \mathbf{Pr}[X = x] = 1$

Probability

- Example: consider tossing 2 dices
  - We have result-space Ω = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),
    (2,1),(2,2),(2,3),(2,4),(2,5),(2,6),
    (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),
    (4,1),(4,2),(4,3),(4,4),(4,5),(4,6),
    (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),
    (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)}
  - |Ω| = 36 elements, for $w \in \Omega$
  - Let Y (based on Ω) be the result of tossing 2 dices with the same point => $Y(w) \in$ {"2 same points", "2 different points"}

$$w \mapsto Y(w)$$

  - Should change "2 same points" to 1, & "2 different points" to 0 => $Y(w)$

cdio 4.0
fit@hcmus

Probability

- Example: consider tossing 2 dices
  - We have result-space Ω = {(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),
    (2,1),(2,2),(2,3),(2,4),(2,5),(2,6),
    (3,1),(3,2),(3,3),(3,4),(3,5),(3,6),
    (4,1),(4,2),(4,3),(4,4),(4,5),(4,6),
    (5,1),(5,2),(5,3),(5,4),(5,5),(5,6),
    (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)}
  - $|\Omega|$ = 36 elements, for $w \in \Omega$
  - Let Y (based on Ω) be the result of tossing 2 dices with the same point => $Y(w) \in$ {“2 same points”, “2 different points”}
  - Should change “2 same points” to 1, & “2 different points” to 0 => $Y(w) \in \{0, 1\}$
- Notation of mapping $Y: \Omega \to \mathbb{R}$

$$w \mapsto Y(w)$$

- Let $Y(w) = 1$ => Event of tossing 2 dices with the same point
  - $\mathbf{Pr}[Y(w) = 1] = 6/36$, due to 6 cases {{1, 1}, …, {6, 6}}
  - So $\mathbf{Pr}[Y = 1] + \mathbf{Pr}[Y = 0] = 1$

Bayes theorem

- Let X and Y be two random variables

$$p(x, y) = p(x \mid y) \times p(y) = p(y \mid x) \times p(x)$$

- Bayes theorem

Apriori

$$\text{if } p(y) > 0 \Rightarrow p(x \mid y) = \frac{p(x) \times p(y \mid x)}{p(y)}$$

Aposteriori

- Corollary: X and Y are two independent ones $\Leftrightarrow p(x \mid y) = p(x)$, $\forall x, y$

Bayes theorem

- Reconsider: example of tossing 2 dices
  - We have result-space $\Omega = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),$
    …
    $(6,1),(6,2),(6,3),(6,4),(6,5),$
    $(6,6)\}$
  - $|\Omega| = 36$ elements, for $w \in \Omega$
  - X: Sum of points of 2 dices => $X(w) \in \{2, 3, 4, \ldots, 12\}$
  - Y: 2 dices with the same point => $Y(w) \in \{0, 1\}$
- Compute $\mathbf{Pr}[Y = 1|X = 4]$ (Probability of 4-point with 2 same faces)
  - $\mathbf{Pr}[X = 4] = 3/36$ => $\mathbf{Pr}[Y = 1|X = 4] = 1/3$ due to $\{(1, 3), (2, 2), (3, 1)\}$
- Compute $\mathbf{Pr}[X = 4|Y = 1]$ (Probability of 4-point with 2 same faces)
  - $\mathbf{Pr}[Y = 1] = 6/36$ => $\mathbf{Pr}[X = 4|Y = 1] = 1/6$ due to $\{(1, 1), (2, 2), \ldots, (6, 6)\}$
- So $\mathbf{Pr}[Y = 1|X = 4] \times \mathbf{Pr}[X = 4] = \mathbf{Pr}[X = 4|Y = 1] \times \mathbf{Pr}[Y = 1]$

Context of cryptography

- Some probabilistic notation for crypto-context
  - $p_P(x)$: Probability of appearing plaintext x
  - $p_K(k)$: Probability of choosing key k
  - $p_C(y)$: Probability of ciphertext receiving value y
- Note:
  - Notations p_P, p_K and p_C are the probabilities for each distinct set
  - It can be assumed that the key value k and the plaintext x are independent events
- From the probability distribution of plaintext and key on the set P and K, we can determine the conditional probability distribution of plaintext ???

Context of cryptography

- For each $k \in \mathcal{K}$, let $C(k) = \{e_k(x) \mid x \in \mathcal{P}\}$ be the set of cipher-text if encrypting plain-text $x \in \mathcal{P}$ with key $k \in \mathcal{K}$.
- So, we see that probability of cipher-text y is sum of probabilities of choosing k and $x = d_k(y)$.

$$p_C(y) = \sum_{k \in K | y \in C(k)} p_K(k) \times p_P(d_k(y))$$

- For each $y \in \mathcal{C}$ and $x \in \mathcal{P}$, $p_C(y \mid x)$ is probability of receiving cipher-text y when plain-text is x.

$$p_C(y \mid x) = \sum_{k \in K | x = d_k(y)} p_K(k)$$

Thực chất là xác suất chọn các khóa k

- Using Bayes theorem to compute $p_P(x \mid y)$

$$p_P(x \mid y) = \frac{p_P(x) \times p_C(y \mid x)}{p_C(y)} = \frac{p_P(x) \times \sum_{k \in K | x = d_k(y)} p_K(k)}{\sum_{k \in K | y \in C(k)} p_K(k) \times p_P(d_k(y))}$$

Example

- Let $P = \{a, b\}$ with $p_P(a) = 1/4$, $p_P(b) = 3/4$
- Let $K = \{k_1, k_2, k_3\}$ with $p_K(k_1) = 1/2$, $p_K(k_2) = p_K(k_3) = 1/4$
- Let $C = \{1, 2, 3, 4\}$
- Let E be a set of encryption rules
  - $e_{k1}(a) = 1$, $e_{k1}(b) = 2$
  - $e_{k2}(a) = 2$, $e_{k2}(b) = 3$
  - $e_{k3}(a) = 3$, $e_{k3}(b) = 4$

	a	b
k_1	1	2
k_2	2	3
k_3	3	4

- Let D be a set of decryption rules
  - $d_{k1}(1) = a$, $d_{k1}(2) = b$
  - $d_{k2}(2) = a$, $d_{k2}(3) = b$
  - $d_{k3}(3) = a$, $d_{k3}(4) = b$

	1	2	3	4
k_1	a	b		
k_2		a	b	
k_3			a	b

fit@hcmus

Example

- Compute $p_C(y)$
  - $p_C(y = 1) = ¼ \times ½ = 1/8$
  - $p_C(y = 2) = ½ \times ¾ + ¼ \times ¼ = 7/16$
  - $p_C(y = 3) = ¼ \times ¾ + ¼ \times ¼ = ¼$
  - $p_C(y = 4) = ¼ \times ¾ = 3/16$
- Condition probability of $p_P(x \mid y)$

1

- $p_P(x = a \mid y = 1) = \frac{p_P(x=a) \times p_C(y=1|x=a)}{p_C(y=1)} = \frac{\frac{1}{4}\times\frac{1}{2}}{\frac{1}{8}} = 1$
- $p_P(x = b \mid y = 1) = \frac{p_P(x=b) \times p_C(y=1|x=b)}{p_C(y=1)} = \frac{\frac{3}{4}\times 0}{\frac{1}{8}} = 0$

2

- $p_P(x = a \mid y = 2) = \frac{p_P(x=a) \times p_C(y=2|x=a)}{p_C(y=2)} = \frac{\frac{1}{4}\times\frac{1}{4}}{\frac{7}{16}} = \frac{1}{7}$
- $p_P(x = b \mid y = 2) = \frac{p_P(x=b) \times p_C(y=2|x=b)}{p_C(y=2)} = \frac{\frac{3}{4}\times\frac{1}{2}}{\frac{7}{16}} = \frac{6}{7}$

	a $p_P = ¼$	b $p_P = ¾$
k_1 $(p_K = ½)$	1	2
k_2 $(p_K = ¼)$	2	3
k_3 $(p_K = ¼)$	3	4

3

$p_P(x = a|y = 3) = \frac{p_P(x=a) \times p_C(y=3|x=a)}{p_C(y=3)} = \frac{\frac{1}{4}\times\frac{1}{4}}{\frac{1}{4}} = ¼$

$p_P(x = b|y = 3) = \frac{p_P(x=b) \times p_C(y=3|x=b)}{p_C(y=3)} = \frac{\frac{3}{4}\times\frac{1}{4}}{\frac{1}{4}} = ¾$

4

$p_P(x = a|y = 4) = \frac{p_P(x=a) \times p_C(y=4|x=a)}{p_C(y=4)} = \frac{\frac{1}{4}\times 0}{\frac{3}{16}} = 0$

$p_P(x = b|y = 4) = \frac{p_P(x=b) \times p_C(y=4|x=b)}{p_C(y=4)} = \frac{\frac{3}{4}\times\frac{1}{4}}{\frac{3}{16}} = 1$

Perfect security

- **Perfectly secure**?
- Significance: The attacker gets nothing from the ciphertext

$$\forall x \in \mathcal{P},\ \forall k \in \mathcal{K},\ p_P(x \mid c) = p_P(x),\ p_K(k \mid c) = p_K(k)$$

- Evaluate Shift-cipher
  - Assume 26 keys in Shift-cipher are randomly chosen with uniform probability (1/26)
  - With set of plaintext having any probability distribution, Shift-cipher achieve perfect security???
  - Let $P = C = K = Z_{26} = \{0, 1, 2, \ldots, 25\}$
  - $e_k(x) = (x + k)$ mod 26 and $d_k(y) = (y - k)$ mod 26

Perfect security (on Shift Cipher)

- Probability

$$p_C(y) = \sum_{k\in\mathbb{Z}_{26}} p_K(k)\times p_P(d_k(y)) = \sum_{k\in\mathbb{Z}_{26}} \frac{1}{26}\times p_P(y-k)$$
$$= \frac{1}{26}\times\sum_{k\in\mathbb{Z}_{26}} p_P(y-k)$$

- Given y, when changing k from 0 to 25, we receive all 26 values of $\mathbb{Z}_{26}$.

$$\sum_{k\in\mathbb{Z}_{26}} p_P(y-k) = \sum_{x\in\mathbb{Z}_{26}} p_P(x) = 1$$

- So, for all $y \in \mathbb{Z}_{26}$, we have $p_C(y) = 1/26$ (1)
- For (x, y), we have only one key $k \in \mathbb{Z}_{26}$ such that $y = x + k \bmod 26$. Hence, $p_C(y \mid x) = p_K(y - x \bmod 26) = 1/26$ (2)

Perfect security (on Shift Cipher)

- From (1) and (2), apply Bayes theorem we have:

$$p_{\mathrm{P}}(x \mid y) = \frac{p_{\mathrm{P}}(x) \times p_{\mathrm{C}}(y \mid x)}{p_{\mathrm{C}}(y)} = \frac{p_{\mathrm{P}}(x) \times \frac{1}{26}}{\frac{1}{26}} = p_{\mathrm{P}}(x) \text{ (satisfy standard)}$$

- Shift cipher is perfect security if randomly choosing a new k for each plain-text x.
- From Bayes theorem, we have $p_{\mathrm{P}}(x \mid y) = p_{\mathrm{P}}(x)$, $\forall x \in \mathcal{P}$, $\forall y \in \mathcal{C}$
  - This is similar to: $p_{\mathrm{C}}(y \mid x) = p_{\mathrm{C}}(y)$, $\forall x \in \mathcal{P}$, $\forall y \in \mathcal{C}$
  - Assume $p_{\mathrm{C}}(y) > 0$, $\forall y \in \mathcal{C}$ (All members of $\mathcal{C}$ are used)
  - Crypto-system is secure if $p_{\mathrm{C}}(y \mid x) > 0$, $\forall x \in \mathcal{P}$, $\forall y \in \mathcal{C} \Rightarrow |\mathcal{C}| \geq |\mathcal{P}|$
  - Due to $p_{\mathrm{C}}(y \mid x) > 0 \Rightarrow \exists k \in \mathcal{K}$: $e_k(x) = y \Rightarrow |\mathcal{K}| \geq |\mathcal{C}|$
- For the system to be perfect security, key-size used to encrypt must be at least equal to the size of the message to be encrypted: $|\mathcal{K}| \geq |\mathcal{P}|$

fit@hcmus

Vernam Cipher

- Is there a perfect secure crypto-system with $|\mathcal{K}| = |\mathcal{P}|$?
- Shannon theorem: Let $(\mathcal{P}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$ be a crypto-system with $|\mathcal{K}| = |\mathcal{P}| = |\mathcal{C}|$. So, it is perfect secure if and only if:
  - $\forall c \in \mathcal{C}, x \in \mathcal{P} \Rightarrow \exists k \in \mathcal{K}: e_k(x) = c$ (1)
  - $\forall k \in \mathcal{K}, p_K(k) = 1/|\mathcal{K}|$ (2)
- Proof: Let $(\mathcal{P}, \mathcal{K}, \mathcal{C}, \mathcal{E}, \mathcal{D})$ be a crypto-system with $|\mathcal{K}| = |\mathcal{P}| = |\mathcal{C}|$. Due to its perfect security, we have
  - $\forall x \in \mathcal{P}, p_P(x \mid c) = p_P(x)$ and Bayes theorem allows $p_C(c \mid x) = p_C(c)$
  - $\exists! k \in \mathcal{K}: e_k(x) = c$ for (x, c), due to $|\mathcal{K}| = |\mathcal{P}| = |\mathcal{C}|$
  - Fix c, for all x_i, let k_i be key such that $e_{k_i}(x_i) = c$
  - From Bayes theorem: $p_P(x_i \mid c) = \frac{p_C(c \mid x_i) \times p_P(x_i)}{p_C(c)} = \frac{p_K(k_i) \times p_P(x_i)}{p_C(c)}$
  - Due to $p_P(x_i \mid c) = p_P(x_i) \Rightarrow p_K(k_i) = p_C(c)$.

fit@hcmus

Vernam Cipher

- Gilbert Vernam (Bell Labs) proposed in 1919
  - A key is a "long enough" random sequence of values. So, $C = P \oplus K$
  - This method is proven to be perfect security
  - Limitation: the key is too long and cannot be reused
  - Advantage: simple
- Description:
  - Let integer $n \geq 1$, and $P = C = K = (\mathbb{Z}_2)^n$. For each $k \in (\mathbb{Z}_2)^n$, we let:
    - $e_k(x) = (x_1 + k_1, \ldots, x_n + k_n) \bmod 2$, where $x = (x_1, \ldots, x_n)$ and $k = (k_1, \ldots, k_n)$.
    - $d_k(y) = (y_1 + k_1, \ldots, y_n + k_n) \bmod 2$, where $y = (y_1, \ldots, y_n)$
- Note: operator (+ mod 2) is $\oplus$-bit

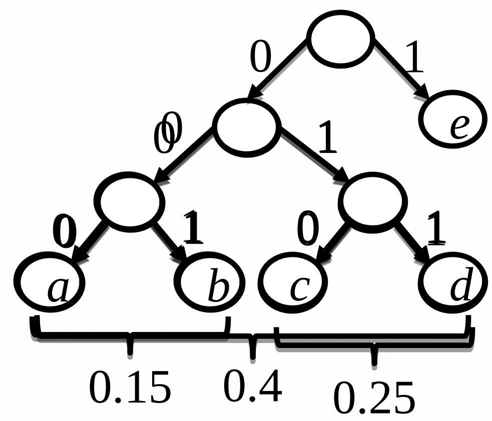
Informatic theory

- Some events are random but more common than others
- Some facts are more important than others
- Entropy is a measure of the uncertainty of a random variable, or the amount of information each event provides
- If X is a random variable receiving values in $\mathcal{X}$, so $H(X) = -\sum_{x\in\mathcal{X}} P[X=x]\log_2 P[X=x]$
- Note: $lim_{x\to 0}(x \times log_2 x) = 0$
  - $lim_{x\to 0}(x \times log_2 x) = \lim_{x\to 0}\frac{\log_2 x}{\frac{1}{x}}$
  - With L' Hopital, we have $\lim_{x\to 0}\frac{\frac{1}{x\ln 2}}{-\frac{1}{x^2}} = \lim_{x\to 0}\frac{-x^2}{x\ln 2} = \boxed{\lim_{x\to 0}\frac{x}{\ln 2} = 0}$

Entropy and Huffman encoding

- Recall the idea of Huffman encoding
- Example: we have $X = \{a, b, c, d, e\}$ with probabilities $p(a) = .05$, $p(b) = .10$, $p(c) = .12$, $p(d) = .13$ and $p(e) = .60$

a	*b*	*c*	*d*	*e*
.05	.1	.12	.13	.6
0	1			
.15		.12	.13	.6
		0	1	
.15		.25		.6
0		1		
.4				.6
0				1
1				



Huffman tree

x	***f*(*x*)**
a	000
b	001
c	010
d	011
e	1

Prefix-code

Entropy and Huffman encoding

- Average length to transmit information for an event

$$l(f) = \underbrace{0.05 \times 3}_{\text{'a'}} + \underbrace{0.1 \times 3}_{\text{'b'}} + \underbrace{0.12 \times 3}_{\text{'c'}} + \underbrace{0.13 \times 3}_{\text{'d'}} + \underbrace{0.6 \times 1}_{\text{'e'}} = 1.8$$

- Entropy:

$$H(X) = \overbrace{0.05 \times \log_2(0.05)}^{0.2161} + \overbrace{0.1 \times \log_2(0.1)}^{0.3322} + \overbrace{0.12 \times \log_2(0.12)}^{0.3671}$$
$$+ \underbrace{0.13 \times \log_2(0.13)}_{0.3842} + \underbrace{0.6 \times \log_2(0.6)}_{0.4422}$$
$$= 1.7402$$

- Result: $H(X) \le l(f) \le H(X) + 1$

Properties of Entropy

- Basic properties
  - $H(X) \geq 0$, equality occurs if and only if the variable X is constant
  - $H(X) \geq \log_2|X|$, equality occurs if and only if $p(X = x) = 1/|X|$
  - $H(X, Y) \geq H(X) + H(Y)$, '=' occurs $\Leftrightarrow$ X & Y are independent distribution
  - $H(X|Y) \geq H(X)$, equality occurs $\Leftrightarrow$ X & Y are independent distribution
  - **Chain Rule:** $H(X, Y) = H(X|Y) + H(Y)$
- Entropy of components of crypto-system
  - $H(C|K) = H(P)$
  - $H(C|P, K) = H(P|C, K) = 0$
  - $H(P, K) = H(P) + H(K)$
  - $H(C) \geq H(P)$
  - $H(C, P, K) = H(C, K) = H(P, K)$
  - $H(K|C) = H(K) + H(P) - H(C)$ và $H(K|C^n) = H(K) + H(P^n) - H(C^n)$

Comments

- There are 26! $\approx 10^{26}$ encryption rule (substitution) for English text (includes normal characters)
- Equivalent to 88-bit security $\Rightarrow$ why is it easy to be attacked in practice?
- Shannon: All approaches of mono-alphabetic cipher of English are easy to break if having 25 characters of cipher-text.

Language's unicity distance

- Spurious keys: if using shift-cipher, we have cipher-texts "WNAJW"
  - There may be 5 and 22 to decrypt "RIVER" and "ARENA"
  - One of them is wrong
- Introduction of random variables
  - Let $P \in \mathcal{P} = \{a, b, \ldots, z\}$ ($|\mathcal{P}| = 26$): set of characters
  - Let $P^2 \in \mathcal{P}^2 = \{aa, \ldots, zz\}$ ($|\mathcal{P}^2| = 26^2$): set of digraphs
  - …
  - Let $P^n \in \mathcal{P}^n = \{a\ldots a, \ldots, z\ldots z\}$ ($|\mathcal{P}^2| = 26^n$): set of *n*-graphs

Language’s unicity distance

- Some notations
  - $p(i)$ is probability of appearing of character “i”
  - $p_i(j)$ is probability of appearing of character ‘j’ when ‘i’ appears
  - $p(i, j)$ is probability of appearing of 2 characters ‘i’ and ‘j’
- Example: compute entropy of $P \in \mathcal{P} = \{a, b, \ldots, z\}$ ($|\mathcal{P}| = 26$)
  - $H(P) = -\sum_i p(i) \times \log_2 p(i) \approx 4.14$ bits/character (real data)
- Example: compute entropy of $P^2 \in \mathcal{P}^2 = \{aa, \ldots, zz\}$ ($|\mathcal{P}^2| = 26^2$)
  - $H(P^2) = -\sum_{i,j} p(i,j) \times \log_2 p(i,j) \approx 7.7$ bits
- Formular to compute entropy (for each character) of another language ‘L’: $H_L \equiv \lim_{n\to\infty} \frac{H(P^n)}{n}$
- Let $R_L = 1 - \frac{H_L}{\log_2 |\mathcal{P}|}$ (Rate of “spurious elements” of a language ‘L’)

Language's unicity distance

- Due to $H_L = \lim_{n\to\infty} \frac{H(P^n)}{n}$, $H(P^n) \approx n \times H_L = n \times (1 - R_L) \times \log_2|\mathcal{P}|$
- Due to $|\mathcal{P}| = |\mathcal{C}|$, $H(C^n) \le n \times H(C) \le n \times \log_2|\mathcal{C}| = n \times \log_2|\mathcal{P}|$
  - Where $\mathcal{P}$ and $\mathcal{C}$ are sets of plain-texts and cipher-texts
  - $\mathcal{K}$ is a set of keys
- We have $H(K|C^n) = H(K) + H(P^n) - H(C^n)$

$$\approx H(K) + n \times (1 - R_L) \times \log_2|\mathcal{P}| - n \times \log_2|\mathcal{P}|$$

$$= H(K) - n \times R_L \times \log_2|\mathcal{P}| \le \log_2|\mathcal{K}| - n \times R_L \times \log_2|\mathcal{P}|$$

- Crypto-system is broken when:

$$H(K|C^n) = 0 \Leftrightarrow \log_2|\mathcal{K}| - n \times R_L \times \log_2|\mathcal{P}| = 0 \Leftrightarrow n = \frac{\log_2|\mathcal{K}|}{R_L \times \log_2|\mathcal{P}|}$$

  - Means: entropy of random variable K when C^n is zero $\Rightarrow$ there is only one key to decrypt.

Language's unicity distance

- Unicity of crypto-system is n_0 such that a number of spurious-key are zero
- English case: ($|\mathcal{P}| = 26$, $R_L = 0.75$, $|\mathcal{K}| = 26!$ due to using substitution cipher)

$$n_0 = \frac{\log_2 |\mathcal{K}|}{R_L \times \log_2 |\mathcal{P}|} = \frac{\log_2 26!}{0.75 \times \log_2 26} \approx 25 \text{ (Language distance)}$$

- Mean that: need a cipher-text with at least length of 25 characters to ensure that there exists only one key

Comments

- **Data Compression**
  - Good compression – good encryption
  - Good encryption – bad compression
- **Combination of encryption approaches**
  - "Weighted sum" of crypto-systems
    - Create new crypto-systems from existing crypto-systems
    - Choose 2 crypto-systems with the same message space, use system A with probability p, use system B with probability $1 - p$.
  - Product cipher: sequentially apply successive encryption algorithms