

# ТЕОРІЯ ІНФОРМАЦІЇ ТА КОДУВАННЯ

## Модуль №1. Основи теорії інформації

### Теоретичні питання

1. Теорія інформації як наукова дисципліна. Предмет та зв'язок з іншими науками.
2. Поняття ентропії. Ентропія як міра невизначеності стану фізичної системи.
3. Властивості ентропії.
4. Ентропія бінарних систем.
5. Кількість інформації як міра знятої невизначеності. Формули Хартлі та Шеннона.
6. Об'єм даних. Співвідношення між об'ємом даних та кількістю інформації.
7. Ентропія складної системи. Випадок незалежних систем.
8. Часткова та повна умовна ентропія.
9. Властивості умовної ентропії.
10. Зв'язок ентропії об'єднання систем та умовної ентропії.
11. Канальні матриці. Інформаційна повнота характеристик каналу зв'язку.
12. Інформаційні втрати при передачі інформації та їх обчислення.
13. Обчислення кількості прийнятої інформації.
14. Модель джерела повідомлень. Надлишковість повідомлень.
15. Передача інформації каналами зв'язку. Пропускна здатність каналу. Теорема Шеннона.

### Типові задачі

1. Визначте ентропію системи, стан якої описується випадковою величиною  $X$  із розподілом  

$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
0.01	0.02	0.13	0.37	0.47
2. Ймовірності появи символів алфавіту дорівнюють відповідно 0,05; 0,45; 0,34; 0,16. Визначити кількість інформації, яка припадає на символ повідомлення, складеного за допомогою такого алфавіту.
3. Яка кількість інформації припадає на букву рівноймовірного алфавіту, що складається з 19, 26 або 34-х букв?
4. Визначити кількість інформації і об'єм даних за таких умов:
  - а) первинний алфавіт  $a_1, a_2, \dots, a_8$  рівноймовірний. Символи вторинного алфавіту ( $m_2 = 2$ ) комбінуються в рівномірні кодові комбінації;
  - б) первинний алфавіт містить вісім букв ( $m_1 = 8$ ). Букви алфавіту зустрічаються в повідомленні з ймовірностями:  $p_1 = 0.01$ ;  $p_2 = 0.09$ ;  $p_3 = p_4 = p_5 = p_6 = 0.15$ ;  $p_7 = 0.05$ ;  $p_8 = 0.25$ . Кодові комбінації у вторинному алфавіті ( $m_2 = 2$ ) рівномірні;
  - в) первинний алфавіт складається з семи букв  $a_1, a_2, \dots, a_7$  які зустрічаються з рівними ймовірностями в тексті, а вторинний алфавіт бінарний і повідомлення мають однакову довжину;
  - г) первинний алфавіт рівноймовірний  $m_1 = 8$ , а вторинні повідомлення побудовані з кодових комбінацій, що мають середню довжину шість двійкових символів.
5. Взаємодія двох систем описується матрицею  $P(X, Y)$ . Визначити безумовні ентропії систем, повні і часткові умовні ентропії та ентропію об'єднання систем.

$$P(X, Y) = \begin{pmatrix} 0.35 & 0.03 & 0 \\ 0.05 & 0.27 & 0.02 \\ 0 & 0 & 0.28 \end{pmatrix}$$

6. Канал зв'язку описано матрицею

$$P(X, Y) = \begin{pmatrix} 0.21 & 0.04 & 0 \\ 0 & 0.33 & 0.02 \\ 0 & 0.03 & 0.37 \end{pmatrix}$$

Визначити кількість інформації на 1 символ повідомлень та інформаційні втрати в каналі зв'язку. Знайти пропускну здатність каналу зв'язку, якщо джерело повідомлень працює зі швидкістю 100 символів за секунду

7. Канал зв'язку описано каналною матрицею  $P(Y/X)$ .

$$P(Y / X) = \begin{pmatrix} 0.67 & 0.31 & 0.02 \\ 0.07 & 0.77 & 0.16 \\ 0.15 & 0.03 & 0.82 \end{pmatrix}$$

Ймовірності появи символів передавача  $p(x_1)=0.35$ ,  $p(x_2)=0.25$ ,  $p(x_3)=0.4$ . Передано 150 символів. Знайти інформаційні втрати та кількість прийнятої інформації. Визначити швидкість передачі інформації, якщо час передачі одного символу  $\tau = 0,01$  мс. Порівняти результати із випадком рівноймовірних символів.

8. Ймовірності появи символів на виході передавача:  $p_1 = 0,25$ ;  $p_2 = 0,3$ ;  $p_3 = 0,35$ ;  $p_4 = 0,1$ . Тривалість передачі відповідних символів  $\tau_1 = 0,03$  мс;  $\tau_2 = 0,02$  мс;  $\tau_3 = 0,015$  мс,  $\tau_4 = 0,06$  мс. Визначити швидкість передачі повідомлень при відсутності завад.

9. Визначити пропускну здатність бінарного каналу зв'язку, здатного передавати 2500 символів 0 або 1 за одиницю часу, причому кожний із символів спотворюється з ймовірністю  $p = 0,002$ .

## РОЗРАХУНКОВА РОБОТА № 1

**Тема:** Розрахунок інформаційних характеристик каналу зв'язку

**Мета:** Набуття навичок розрахунку основних інформаційних характеристик каналу зв'язку.

### ЗАВДАННЯ ДЛЯ РОБОТИ:

**Дано:** Каналом зв'язку передаються повідомлення у вигляді послідовностей символів деякого алфавіту, імовірності яких задано:  $p(0)=\dots$ ;  $\dots$ ;  $p(\dots)=\dots$

Відома канална матриця  $P(y/x)$ , що визначає втрати інформації в каналі зв'язку.

#### Визначити:

1. Ентропію джерела інформації –  $H(X)$
2. Безумовну ентропію приймача інформації –  $H(Y)$ .
3. Повні умовні ентропії –  $H(Y/X)$ ,  $H(X/Y)$ .
4. Швидкість передачі інформації та пропускну спроможність каналу зв'язку, якщо час передачі одного символу первинного алфавіту дорівнює  $0,05V$  мс. ( $V$  – номер варіанта)
5. Втрати інформації в каналі зв'язку при передачі  $100V$  символів алфавіту ( $V$  – номер варіанта).
6. Середню кількість прийнятої інформації.

**Розрахунки та графічні побудови провести у середовищі Visual Studio, MS Office або MathCAD. Передбачити введення початкових даних, виведення проміжкових та кінцевих результатів, перевірку правильності розрахунків. У звіті навести розрахункові формули, лістинги, результати та зробити висновки.**



## Варіант 11

$p(0)=0,04; p(1)=0,6; p(2)=0,2; p(3)=0,01; p(4)=0,01;$   
 $p(5)=0,03; p(6)=0,02; p(7)=0,04; p(8)=0,03; p(9)=0,02.$

0.92	0.07	0.01	0	0	0	0	0	0	0
0.05	0.94	0.01	0	0	0	0	0	0	0
0	0.02	0.97	0.01	0	0	0	0	0	0
0	0.03	0.02	0.90	0.02	0.03	0	0	0	0
0	0	0	0.10	0.88	0.02	0	0	0	0
0	0	0	0.01	0.11	0.76	0.11	0.01	0	0
0	0	0	0	0.01	0.01	0.95	0.02	0.01	0
0	0	0	0	0	0	0.02	0.94	0.04	0
0	0	0	0	0	0	0	0.01	0.96	0.03
0	0	0	0	0	0	0	0.02	0.03	0.95

## Варіант 12

$p(0)=0,24; p(1)=0,51; p(2)=0,01; p(3)=0,03; p(4)=0,02;$   
 $p(5)=0,08; p(6)=0,01; p(7)=0,01; p(8)=0,04; p(9)=0,05$

0.90	0.03	0.07	0.01	0	0	0	0	0	0
0.01	0.98	0.01	0.01	0	0	0	0	0	0
0	0.02	0.95	0.03	0	0	0	0	0	0
0	0.01	0.02	0.95	0.01	0.01	0	0	0	0
0	0	0	0.01	0.94	0.05	0	0	0	0
0	0	0	0.01	0.01	0.94	0.03	0.01	0	0
0	0	0	0	0.01	0.02	0.90	0.06	0.01	0
0	0	0	0	0	0	0.02	0.95	0.03	0
0	0	0	0	0	0	0	0.07	0.91	0.02
0	0	0	0	0	0	0	0.01	0.01	0.98

## Варіант 13

$p(0)=0,05; p(1)=0,7; p(2)=0,01; p(3)=0,01; p(4)=0,01;$   
 $p(5)=0,1; p(6)=0,02; p(7)=0,03; p(8)=0,03; p(9)=0,04.$

0.95	0.01	0.04	0	0	0	0	0	0	0
0.02	0.95	0.03	0	0	0	0	0	0	0
0	0.01	0.95	0.04	0	0	0	0	0	0
0	0.01	0.02	0.95	0.01	0.01	0	0	0	0
0	0	0	0.01	0.90	0.09	0	0	0	0
0	0	0	0.01	0.01	0.94	0.03	0.01	0	0
0	0	0	0	0.01	0.01	0.90	0.07	0.01	0
0	0	0	0	0	0	0.08	0.90	0.02	0
0	0	0	0	0	0	0	0.02	0.95	0.03
0	0	0	0	0	0	0	0.02	0.01	0.97

## Варіант 14

$p(0)=0,14; p(1)=0,5; p(2)=0,2; p(3)=0,01; p(4)=0,01;$   
 $p(5)=0,03; p(6)=0,02; p(7)=0,05; p(8)=0,02; p(9)=0,02.$

0.90	0.00	0.10	0	0	0	0	0	0	0
0.00	0.89	0.11	0	0	0	0	0	0	0
0	0.02	0.95	0.03	0	0	0	0	0	0
0	0.01	0.01	0.96	0.01	0.01	0	0	0	0
0	0	0	0.01	0.98	0.01	0	0	0	0
0	0	0	0.01	0.10	0.86	0.02	0.01	0	0
0	0	0	0	0.01	0.00	0.98	0.01	0.00	0
0	0	0	0	0	0	0.02	0.98	0.02	0
0	0	0	0	0	0	0	0.01	0.99	0.00
0	0	0	0	0	0	0	0.01	0.00	0.99

## Варіант 15

$p(0)=0,08; p(1)=0,41; p(2)=0,01; p(3)=0,03; p(4)=0,05;$   
 $p(5)=0,24; p(6)=0,11; p(7)=0,01; p(8)=0,04; p(9)=0,02$

0.90	0.08	0.02	0	0	0	0	0	0	0
0.01	0.98	0.01	0	0	0	0	0	0	0
0	0.03	0.97	0	0	0	0	0	0	0
0	0.01	0.02	0.97	0	0	0	0	0	0
0	0	0	0.01	0.94	0.05	0	0	0	0
0	0	0	0.01	0.01	0.94	0.03	0.01	0	0
0	0	0	0	0.01	0.06	0.90	0.02	0.01	0
0	0	0	0	0	0	0.02	0.95	0.03	0
0	0	0	0	0	0	0	0.08	0.91	0.01
0	0	0	0	0	0	0	0.02	0.00	0.98

## Варіант 16

$p(0)=0,05; p(1)=0,7; p(2)=0,01; p(3)=0,01; p(4)=0,01;$   
 $p(5)=0,1; p(6)=0,02; p(7)=0,03; p(8)=0,03; p(9)=0,04.$

0.95	0.05	0.00	0	0	0	0	0	0	0
0.04	0.95	0.01	0	0	0	0	0	0	0
0	0.04	0.95	0.01	0	0	0	0	0	0
0	0.01	0.01	0.95	0.01	0.02	0	0	0	0
0	0	0	0.01	0.90	0.09	0	0	0	0
0	0	0	0.01	0.01	0.95	0.02	0.01	0	0
0	0	0	0	0.01	0.01	0.92	0.05	0.01	0
0	0	0	0	0	0	0.08	0.92	0.00	0
0	0	0	0	0	0	0	0.02	0.96	0.02
0	0	0	0	0	0	0	0.00	0.01	0.99

## Варіант 17

$p(0)=0,05; p(1)=0,12; p(2)=0,01; p(3)=0,01; p(4)=0,03;$   
 $p(5)=0,02; p(6)=0,05; p(7)=0,09; p(8)=0,08; p(9)=0,07;$   
 $p(A)=0,07; p(B)=0,3; p(C)=0,04; p(D)=0,03; p(E)=0,02;$   
 $p(F)=0,01.$

0.97	0.02	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.02	0.95	0.02	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0.01	0.98	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0.01	0.02	0.94	0.02	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.01	0.98	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.01	0.01	0.96	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0.01	0.02	0.94	0.02	0.01	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0.02	0.96	0.02	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.01	0.98	0.01	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.01	0.01	0.96	0.01	0.01	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.01	0.01	0.96	0.01	0.01	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0.01	0.01	0.95	0.02	0.01	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0.02	0.96	0.02	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0.01	0.02	0.94	0.02	0.01	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0.01	0.02	0.95	0.02	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0.01	0.03	0.96	0	0	0	0

## Варіант 18

$p(0)=0,3; p(1)=0,09; p(2)=0,08; p(3)=0,07; p(4)=0,02;$   
 $p(5)=0,05; p(6)=0,04; p(7)=0,12; p(8)=0,02; p(9)=0,01;$   
 $p(A)=0,08; p(B)=0,02; p(C)=0,05; p(D)=0,03; p(E)=0,01;$   
 $p(F)=0,01.$

0.98	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.01	0.97	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0.01	0.98	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0.01	0.02	0.94	0.02	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.01	0.98	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0.01	0.01	0.96	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0.01	0.02	0.94	0.02	0.01	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0.02	0.96	0.02	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.01	0.98	0.01	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0.01	0.01	0.96	0.01	0.01	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0.01	0.01	0.96	0.01	0.01	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0.01	0.01	0.95	0.02	0.01	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0.02	0.96	0.02	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0.01	0.02	0.94	0.02	0.01	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0.01	0.02	0.95	0.02	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0.01	0.01	0.98	0	0	0	0	0





## Приклад виконання розрахункової роботи № 1 з «Теорії інформації та кодування»

### Завдання

Через канал зв'язку передаються повідомлення, що представляють собою послідовність символів  $x_i, i = 1..n, n = 16$ . Ймовірності символів на виході джерела визначаються розподілом:

$$p(x_i) \left| \begin{array}{cccccccccccccccc} 0,22 & 0,13 & 0,05 & 0,06 & 0,12 & 0,05 & 0,01 & 0,11 & 0,02 & 0,04 & 0,08 & 0,02 & 0,02 & 0,01 & 0,05 & 0,01 \end{array} \right|$$

Канальна матриця, що визначає втрати інформації в каналі зв'язку, має вигляд:

$$p(Y/X) = \begin{array}{c} \left| \begin{array}{cccccccccccccccc} 0,91 & 0,07 & 0,01 & 0,01 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,02 & 0,98 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,02 & 0,97 & 0,01 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,02 & 0,02 & 0,94 & 0,02 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0,03 & 0,96 & 0,01 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,01 & 0,03 & 0,94 & 0,02 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0,04 & 0,96 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,01 & 0,98 & 0,01 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,01 & 0,01 & 0,96 & 0,01 & 0,01 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,01 & 0,96 & 0,03 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,01 & 0,01 & 0,95 & 0,02 & 0,01 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,02 & 0,96 & 0,02 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,02 & 0,96 & 0,02 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,95 & 0,05 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0,01 & 0,01 & 0,98 \end{array} \right| \end{array}$$

Знайти інформаційні характеристики каналу зв'язку. Час передачі одного символу первинного алфавіту дорівнює  $0,05^V$  мс. ( $V = 10$  – номер варіанта). Передано  $100^V$  символів алфавіту.

### Обчислення інформаційних характеристик:

– Для заданого розподілу ймовірностей символів на виході джерела:

$$\sum_i p(x_i) = 1$$

– Ентропія джерела інформації :

$$H(X) = -\sum_{i=1}^n p(x_i) \log_2 p(x_i) = 3.488 \text{ біт/символ}$$

– Ймовірності появи символів на вході приймача визначаються за допомогою умовних ймовірностей

$$p(y_j) = \sum_i p(x_i) p(y_j / x_i)$$

$$\begin{array}{cccccccccccccccc} p(y_1) & p(y_2) & p(y_3) & p(y_4) & p(y_5) & p(y_6) & p(y_7) & p(y_8) & p(y_9) & p(y_{10}) & p(y_{11}) & p(y_{12}) & p(y_{13}) & p(y_{14}) & p(y_{15}) & p(y_{16}) \\ 0,2028 & 0,145 & 0,0519 & 0,0627 & 0,1165 & 0,0515 & 0,0138 & 0,1064 & 0,02 & 0,0396 & 0,0774 & 0,0222 & 0,0198 & 0,0103 & 0,0478 & 0,0123 \end{array}$$

Прийняті повідомлення складають повну групу, тому їх сумарна ймовірність дорівнює:

$$\sum_j p(y_j) = 1$$

– Ентропія приймача інформації :

$$H(Y) = -\sum_{j=1}^n p(y_j) \log_2 p(y_j) = 3.527 \text{ біт/символ}$$

– Загальна умовна ентропія :

$$\begin{aligned} H(Y/X) &= \sum_{i=1}^n p(x_i) H(Y/x_i) = -\sum_{i=1}^n p(x_i) \sum_{j=1}^n p(y_j / x_i) \log_2 p(y_j / x_i) = \\ &= -\sum_{i=1}^n \sum_{j=1}^n p(x_i) p(y_j / x_i) \log_2 p(y_j / x_i) = 0.304 \end{aligned}$$

– Елементи матриці умовних ймовірностей  $P(X/Y)$  :

$$p(x_i / y_j) = \frac{p(x_i) p(y_j / x_i)}{p(y_j)}$$

