

ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Факультет прикладної математики та інформатики

Теорія інформації та кодування

ПРАКТИЧНЕ ЗАВДАННЯ №1

Тема: «Дискретні джерела повідомлень»

Виконав:

Ст. Лук'янчук Денис

Група ПМІ-23

2025

Тема: «Дискретні джерела повідомлень»

Мета роботи: Дослідити та обчислити ключові характеристики ентропії, надлишковості та взаємної інформації дискретних джерел інформації, таких як X і Y. Завдання спрямоване на оцінку ступеня невизначеності джерел, їхньої взаємозалежності та ефективності передачі даних, використовуючи математичні моделі теорії інформації. Це дозволяє зрозуміти, як оптимізувати процеси кодування та передачі інформації в інформаційних системах.

Хід роботи

Завдання 1.1. Обчислити ентропію та надлишковість дискретного джерела інформації з алфавітом $X = \{x_1, x_2, x_3, x_4, x_5\}$ (об'єм алфавіту $k = 5$). Значення ймовірностей $p(x_i)$, $i = 1, 2, 3, 4, 5$ появі символів для різних варіантів наведені у табл. 1.1.

Таблиця 1.1

Варіант	$P(x_1)$	$P(x_2)$	$P(x_3)$	$P(x_4)$	$P(x_5)$
14	0.02	0.01	0.1	0.83	0.04

1. Формула: $H(x) = - \sum_{i=1}^k p(x_i) \log_2 p(x_i)$, $k = 5$

$$H(x) = -(0,02 \cdot \log_2 0,02 + 0,01 \log_2 0,01 + 0,1 \log_2 0,1 + 0,83 \log_2 0,83 + 0,04 \log_2 0,04) = 0,113 + 0,066 + 0,332 + 0,222 + 0,186 = 0,919.$$

$H(x) = 0,919$ Дим - Ентропія

2. Формула: $S(x) = 1 - \frac{H(x)}{\log_2 k}$, $k = 5$

$$S(x) = 1 - \frac{0,919}{\log_2 5} = 1 - \frac{0,919}{2,322} = 1 - 0,396 = 0,604$$

$S(x) = 0,604$ - Надлишковість

Завдання 1.2. Матрицю сумісних імовірностей $P(X, Y)$ появи двох символів на виходах джерел X та Y для різних варіантів наведено у табл. 1.2. Обчислити ентропію $H(X, Y)$ об'єднання двох джерел, загальні умовні ентропії $H(X|Y)$ і $H(Y|X)$, а також середню взаємну кількість інформації $I(X; Y)$. Визначити, яке з цих двох джерел має більшу надлишковість та чи є джерела статистично незалежними. Отримані результати подати у вигляді таблиці:

$H(X, Y)$	$H(X Y)$	$H(Y X)$	$I(X; Y)$	$H(X)$	$\rho(X)$	$H(Y)$	$\rho(Y)$
1,231	0,275	0,183	0,773	1,048	0,339	0,956	0,397

Таблиця 1.2

Варіант	$P(X, Y)$
14	0.28 0.01 0.0 0.01 0.67 0.0 0.01 0.01 0.01

$$1. \quad H(X, Y) = - \sum_i \sum_j p(x_i, y_j) \log_2 p(x_i, y_j) - \text{решина}$$

$$H(X, Y) = - (0,28 \log_2 0,28 + 0,01 \log_2 0,01 + 0,0 \log_2 0,0 + 0,01 \log_2 0,01 + 0,67 \log_2 0,67 + 0,0 \log_2 0,0 + 0,01 \log_2 0,01 + 0,01 \log_2 0,01) = 0,514 + 0,066 + 0,066 + 0,587 + 0,066 + 0,066 + 0,066 = 1,231 \text{ бтм}$$

$$\underline{H(X, Y) = 1,231 \text{ бтм}} - \text{спільна ентропія}$$

2. Формула: $H(x) = - \sum_i p(x_i) \log_2 p(x_i)$

$$p(x_1) = 0,28 + 0,01 + 0,0 = 0,29$$

$$p(x_2) = 0,01 + 0,67 + 0,0 = 0,68$$

$$p(x_3) = 0,01 + 0,01 + 0,01 = 0,03$$

$$H(x) = -(-0,518 - 0,378 - 0,152) = 0,518 + 0,378 + 0,152 = 1,048 \text{ бит}$$

$H(x) = 1,048 \text{ бит} - \text{энтропия } H(x)$

3. Формула: $H(y) = - \sum_j p(y_j) \log_2 p(y_j)$

$$p(y_1) = 0,28 + 0,01 + 0,01 = 0,3$$

$$p(y_2) = 0,01 + 0,67 + 0,01 = 0,69$$

$$p(y_3) = 0,0 + 0,0 + 0,01 = 0,01$$

$$H(y) = -(-0,521 - 0,369 - 0,066) = 0,521 + 0,369 + 0,066 = 0,956$$

$H(y) = 0,956 \text{ бит} - \text{энтропия } H(y)$

4. Формула: $I(x; y) = H(x) + H(y) - H(x, y)$

$$I(x; y) = 1,048 + 0,956 - 1,231 = 0,773 \text{ бит}$$

$I(x; y) = 0,773 \text{ бит} - \text{Взаимная информация}$

Дисперсия статистической величины, то

$$I(x; y) > 0$$

5. Формули : $H(x|y) = H(x,y) - H(y)$
 $H(y|x) = H(x,y) - H(x)$

$$H(x|y) = 1,231 - 0,956 = 0,275 \text{ біт}$$

$$H(y|x) = 1,231 - 1,048 = 0,183 \text{ біт}$$

$H(x|y) = 0,275 \text{ біт}$ — зменшення умовній ентропії
 $H(y|x) = 0,183 \text{ біт}$

6. Формула : $P(x) = 1 - \frac{H(x)}{\log_2 k}$, $k = 3$

$$P(y) = 1 - \frac{H(y)}{\log_2 k}$$

$$P(x) = 1 - \frac{1,048}{\log_2 3} = 1 - \frac{1,048}{1,585} = 0,339$$

$$P(y) = 1 - \frac{0,956}{\log_2 3} = 1 - \frac{0,956}{1,585} = 0,397$$

$P(x) = 0,339$ — надлишковість $P(x), P(y)$
 $P(y) = 0,397$

$P(y) > P(x) - P(y)$ має більшу надлишковість

Завдання 1.3. Для двох дискретних джерел X та Y числові значення безумовних $p(x_i)$ та умовних імовірностей $p(y_i | x_i)$ появі символів на виході джерела Y для різних варіантів наведено у табл. 1.3. Обчислити ентропію $H(X,Y)$ об'єднання двох джерел та середню взаємну кількість інформації $I(X;Y)$. Визначити, яке з цих двох джерел має більшу надлишковість. Результати обчислень помістити у таблицю:

$H(X,Y)$	$H(X Y)$	$H(Y X)$	$I(X;Y)$	$H(X)$	$\rho(X)$	$H(Y)$	$\rho(Y)$
1,823	0,475	1,297	0,051	0,526	0,668	1,348	0,850

Таблиця 1.3

Варіант	$p(x_1)$	$p(x_2)$	$p(x_3)$	$P(Y X)$
14	0.04	0.91	0.05	0.79 0.13 0.08 0.18 0.18 0.64 0.24 0.2 0.56

1. Формула: $P(x_i; y_j) = P(x_i) \cdot P(y_j | x_i)$

$$P(x_1, y_1) = 0,04 \cdot 0,49 = 0,0316$$

$$P(x_1, y_2) = 0,04 \cdot 0,13 = 0,0052$$

$$P(x_1, y_3) = 0,04 \cdot 0,08 = 0,0032$$

$$P(x_2, y_1) = 0,91 \cdot 0,18 = 0,1638$$

$$P(x_2, y_3) = 0,91 \cdot 0,08 = 0,0728$$

$$P(x_2, y_2) = 0,91 \cdot 0,18 = 0,1638$$

$$P(x_3, y_1) = 0,05 \cdot 0,24 = 0,0120$$

$$P(x_3, y_2) = 0,05 \cdot 0,2 = 0,01$$

$$P(x_3, y_3) = 0,05 \cdot 0,56 = 0,028$$

$$P(x_3, y_2) = 0,05 \cdot 0,56 = 0,028$$

Формула: $H(x, y) = - \sum_i \sum_j P(x_i, y_j) \log_2 P(x_i, y_j)$

$$H(x, y) = 0,158 + 0,039 + 0,024 + 0,428 + 0,428 + 0,455 + \\ + 0,074 + 0,066 + 0,145 = 1,823 \text{ бит}$$

$H(x, y) = 1,823 \text{ бит} - \text{столбец entropy}$

2. Формула: $H(x) = - \sum_i P(x_i) \log_2 P(x_i)$

$$P(x_1) = 0,0316 + 0,0052 + 0,0032 = 0,04$$

$$P(x_2) = 0,1638 + 0,5824 + 0,1638 = 0,91$$

$$P(x_3) = 0,012 + 0,01 + 0,028 = 0,05$$

$$H(x) = - (-0,186 - 0,124 - 0,216) = 0,186 + 0,124 + 0,216 = 0,526 \text{ бит}$$

$H(x) = 0,526 \text{ бит} - \text{столбец } H(x)$

3. Формула: $H(y) = - \sum_j p(x_j) \log_2 p(x_j)$

$$p(y_1) = 0,0316 + 0,1638 + 0,012 = 0,2074$$

$$p(y_2) = 0,0052 + 0,1638 + 0,01 = 0,179$$

$$p(y_3) = 0,0032 + 0,5824 + 0,028 = 0,6136$$

$$H(y) = -(-0,471 - 0,444 - 0,433) = 0,471 + 0,444 + 0,433 = 1,348 \text{ бит}$$

$H(y) = 1,348 \text{ бит} - \text{емкость } H(y)$

4. Формула: $I(x; y) = H(x) + H(y) - H(x, y)$

$$I(x; y) = 0,526 + 1,348 - 1,823 = 0,051 \text{ бит}$$

$I(x; y) = 0,051 \text{ бит} - \text{Взаимная информация}$

5. Формула: $H(x|y) = H(x, y) - H(y)$

$$H(y|x) = H(x, y) - H(x)$$

$$H(x|y) = 1,823 - 1,348 = 0,475 \text{ бит}$$

$$H(y|x) = 1,823 - 0,526 = 1,297 \text{ бит}$$

$H(x|y) = 0,475 \text{ бит} - \text{запасные условия емкости}$

$$H(y|x) = 1,297 \text{ бит}$$

$$6. \text{ Формули: } P(x) = 1 - \frac{H(x)}{\log_2 k}, \quad k = 3$$

$$P(y) = 1 - \frac{H(y)}{\log_2 k},$$

$$P(x) = 1 - \frac{0,526}{1,585} = 0,668$$

$$P(y) = 1 - \frac{1,348}{1,585} = 0,850$$

$P(x) = 0,668$ — надлишковість $P(x), P(y)$

$$\underline{P(y) = 0,850}$$

$P(y) = 0,850 > P(x) = 0,668$ — $P(y)$ має більшу надлишковість

Висновок: У результаті виконання лабораторної роботи було проведено аналіз і обчислення основних характеристик ентропії, надлишковості та взаємної інформації дискретних джерел інформації X і Y для різних варіантів. Визначено ступінь невизначеності джерел, їхню взаємозалежність та ефективність передачі даних за допомогою математичних моделей теорії інформації. Обчислення спільної ентропії $H(X, Y)$, окремих ентропій $H(X)$ і $H(Y)$, умовної ентропії $H(Y|X)$ та взаємної інформації $I(X; Y)$ дозволили оцінити, наскільки одне джерело зменшує невизначеність іншого та чи є вони статистично незалежними. Аналіз показав, що в розглянутих прикладах джерела X і Y зазвичай демонструють статистичну залежність, коли $I(X; Y) > 0$, що свідчить про наявність взаємозв'язку між ними. Отримані результати підтверджують практичне застосування теорії інформації для оптимізації процесів кодування та передачі даних, а також демонструють важливість оцінки ентропії для підвищення ефективності інформаційних систем.