МІНІСТЕРСТВО ОСВІТИ ТА НАУКИ УКРАЇНИ
ЛЬВІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені ІВАНА ФРАНКА

Кафедра програмування

Практичне завдання № 1

ДИСКРЕТНІ ДЖЕРЕЛА ПОВІДОМЛЕНЬ

з курсу "Теорія інформації"

Виконав:
студент групи ПМІ-41
Шипка Олена

Варіант 7

Оцінка

Прийняв:
доц. Рикалюк Р.Є.
ас. Жировецький В.В.

Львів 2020

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 появи символів наступні:

p(x₁)	***p(x₂)***	***p(x₃)***	***p(x₄)***	***p(x₅)***
0,35	0,15	0,05	0,25	0,20

Для обчислення ентропії джерела скористаємося формулою

$$H(X) = -\sum_{i=1}^{k} p_i log_2(p_i)$$

$H(x) = -\left(0{,}35\, log_2(0{,}35) + 0{,}15\, log_2(0{,}15) + 0{,}05\, log_2(0{,}05) + 0{,}25\, log_2(0{,}25) + 0{,}2\, log_2(0{,}2)\right) = 0{,}530100610490415 + 0{,}410544839124931 + 0{,}216096404744368 + 0{,}5 + 0{,}464385618977472 = 2{,}12112747333719$

Надлишковість дискретного джерела обчислюється за формулою $\rho_X = 1 - \frac{H(X)}{H_{max}(X)}$, де $H_{max}(X) = log_2 k$.

Обчислимо максимальну ентропію $H_{max}(X) = log_2 5 = 2{,}32192809488736$.

Підставимо тепер максимальну ентропію у формулу для знаходження надлишковості $\rho_X = 1 - \frac{2{,}12112747333719}{2{,}32192809488736} = 0{,}0864801205482275$.

Відповідь:

$\boldsymbol{H(X)}$	$\boldsymbol{\rho_X}$
2,12112747333719	0,0864801205482275

Завдання 1.2 Матриця сумісних ймовірностей P(X,Y) появи двох символів на виходах джерел X та Y: $P(X,Y) = \begin{pmatrix} 0.39 & 0.005 & 0.005 \\ 0.005 & 0.19 & 0.005 \\ 0.005 & 0.005 & 0.39 \end{pmatrix}$.

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.

Проведемо згортку по i для того, щоб знайти $p(y)$,

$$p(y_1) = 0{,}39 + 0{,}005 + 0{,}005 = 0{,}4$$

$$p(y_2) = 0{,}005 + 0{,}19 + 0{,}005 = 0{,}2$$

$$p(y_3) = 0{,}005 + 0{,}005 + 0{,}39 = 0{,}4$$

та по j, щоб знайти $p(x)$.

$$p(x_1) = 0{,}39 + 0{,}005 + 0{,}005 = 0{,}4$$

$$p(x_2) = 0{,}005 + 0{,}19 + 0{,}005 = 0{,}2$$

$$p(x_3) = 0{,}005 + 0{,}005 + 0{,}39 = 0{,}4$$

Використовуючи формули, описані в завданні 1.1, обчислимо ентропію джерел X та Y, а також надлишковість джерел X та Y.

$H(X) = -\big(0{,}4\, log_2(0{,}4) + 0{,}2\, log_2(0{,}2) + 0{,}4\, log_2(0{,}4)\big) =$
$(0{,}528771237954945 + 0{,}464385618977472 + 0{,}528771237954945\ ) =$
$1{,}52192809488736$

$$\rho_X = 1 - \frac{1{,}52192809488736}{1{,}58496250072116} = \ 0{,}0397702821392389$$

Очевидно, що $H(X) = H(Y)$, та $\rho_X = \rho_Y$.

Обчислимо ентропію об'єднання двох джерел, використовуючи формулу

$$H(X,Y) = -\sum_{i=1}^{k}\sum_{j=1}^{k} p(x_i, y_j) log_2 p(x_i, y_j))$$

$H(X,Y) = -(0{,}39\, log_2 0{,}39 + 0{,}005\, log_2 0{,}005 + 0{,}005\, log_2 0{,}005 +$
$0{,}005\, log_2 0{,}005 + 0{,}19\, log_2 0{,}19 + 0{,}005\, log_2 0{,}005 + 0{,}005\, log_2 0{,}005 +$
$0{,}005\, log_2 0{,}005 + 0{,}39\, log_2 0{,}39) =$

$0{,}529797048655866 + 0{,}0382192809488736 + 0{,}0382192809488736 +$
$0{,}0382192809488736 + 0{,}455226448502916 + 0{,}0382192809488736 +$
$0{,}0382192809488736 + 0{,}0382192809488736 + 0{,}529797048655866 =$
$1{,}74413623150789$

Скориставшись формулою $p(y|x) = \frac{p(x,y)}{p(x)}$ побудуємо матрицю умовних ймовірностей

$$P(Y|X) = \begin{pmatrix} \frac{0.39}{0.4} & \frac{0.005}{0.4} & \frac{0.005}{0.4} \\ \frac{0.005}{0.2} & \frac{0.19}{0.2} & \frac{0.005}{0.2} \\ \frac{0.005}{0.4} & \frac{0.005}{0.4} & \frac{0.39}{0.4} \end{pmatrix} = \begin{pmatrix} 0.975 & 0.0125 & 0.0125 \\ 0.025 & 0.95 & 0.025 \\ 0.0125 & 0.0125 & 0.975 \end{pmatrix}$$

Обчислимо умовну ентропію джерела Y стосовно джерела X за формулою

$$H(Y|X) = \sum_{i=1}^{k} p(x_i) H(Y|x_i)$$

Для цього спершу обчислимо часткові умовні ентропії

$$H(Y|x_i) = -\sum_{j=1}^{k} p\big(y_j|x_i\big) log_2(p\big(y_j|x_i\big))$$

$H(Y|x_1) = -(0{,}975\, log_2 0{,}975 \;+\; 0{,}0125\, log_2 0{,}0125 \;+\; 0{,}0125\, log_2 0{,}0125) = 0{,}0356127291244862 \;+\; 0{,}079024101186092 \;+\; 0{,}079024101186092 = 0{,}19366093149667$

$H(Y|x_2) = \;-(0{,}025\, log_2 0{,}025 \;+\; 0{,}95\, log_2 0{,}95 \;+\; 0{,}025\, log_2 0{,}025\,) = 0{,}133048202372184 \;+\; 0{,}0703005523715881 \;+\; 0{,}133048202372184 = 0{,}336396957115956$

$H(Y|x_3) = -(0{,}0125\, log_2 0{,}0125 \;+\; 0{,}0125\, log_2 0{,}0125 + \; 0{,}975\, log_2 0{,}975\,) = 0{,}079024101186092 \;+\; 0{,}079024101186092 + \; 0{,}0356127291244862 \;= 0{,}19366093149667$

Тепер підставимо отримані значення у формулу для знаходження загальної умовної ентропії

$H(Y|X) = 0{,}4 * 0{,}19366093149667 \;+\; 0{,}2 * 0{,}336396957115956 \;+\; 0{,}4 * 0{,}19366093149667 + \;=\; 0{,}0774643725986681 \;+\; 0{,}0672793914231912 \;+\; 0{,}0774643725986681 = \; 0{,}222208136620527$

Обчислимо середню взаємну кількість інформації за формулою

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$I(X;Y) = \; 1{,}52192809488736 \;+\; 1{,}52192809488736 \;-\; 1{,}74413623150789 \;= 1{,}29971995826683$

Оскільки $I(X;Y) \neq 0$, то джерела є статистично залежними.

Обчислимо умовну ентропію джерела X стосовно джерела Y за формулою

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

$$H(X|Y) = 1{,}52192809488736 - 1{,}29971995826683 = 0{,}22220813662053$$

Відповідь:

$\boldsymbol{H(X,Y)}$	$\boldsymbol{H(X|Y)}$	$\boldsymbol{H(Y|X)}$	$\boldsymbol{I(X;Y)}$
1,74413623150789	0,22220813662053	0,222208136620527	1,29971995826683

$\boldsymbol{H(X)}$	$\boldsymbol{H(Y)}$	$\boldsymbol{\rho_X}$	$\boldsymbol{\rho_Y}$
1,52192809488736	1,52192809488736	0,0397702821392389	0,0397702821392389

Завдання 1.3 Для двох дискретних джерел X та Y чисельні значення безумовних $p(x_i)$ та умовних ймовірностей $p(y_j|x_i)$ появи символів на виході джерела Y для різних варіантів наступні:

$p(x_1)$	$p(x_2)$	$p(x_3)$	$\boldsymbol{P(Y|X)}$
0.5	0.2	0.3	$\begin{pmatrix} 0.95 & 0.025 & 0.025 \\ 0.025 & 0.95 & 0.025 \\ 0.025 & 0.025 & 0.95 \end{pmatrix}$

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.

Скориставшись формулою $p(x,y) = p(y|x) * p(x)$ побудуємо матрицю сумісних ймовірностей

$$P(X,Y) = \begin{pmatrix} 0.95 * 0.5 & 0.025 * 0.5 & 0.025 * 0.5 \\ 0.025 * 0.2 & 0.95 * 0.2 & 0.025 * 0.2 \\ 0.025 * 0.3 & 0.025 * 0.3 & 0.95 * 0.3 \end{pmatrix} = \begin{pmatrix} 0.475 & 0.0125 & 0.0125 \\ 0.005 & 0.19 & 0.005 \\ 0.0075 & 0.0075 & 0.285 \end{pmatrix}$$

Проведемо згортку по i для того, щоб знайти $p(y)$,

$$p(y_1) = 0.475 + 0.005 + 0.0075 = 0.4875$$

$$p(y_2) = 0.0125 + 0.19 + 0.0075 = 0.21$$

$$p(y_3) = 0.0125 + 0.005 + 0.285 = 0.3025$$

Обчислимо ентропію джерел X та Y, а також надлишковість джерел X та Y.

$$H(X) = -(0{,}5\, log_2 0{,}5 + 0{,}2\, log_2 0{,}2 + 0{,}3\, log_2 0{,}3) = 0{,}5 + 0{,}464385618977472 + 0{,}521089678249862 = 1{,}48547529722733$$

$H(Y) = -(0{,}4875\, log_2 0{,}4875 + 0{,}21\, log_2 0{,}21 + 0{,}3025\, log_2 0{,}3025) = 0{,}505306364562243 + 0{,}472823141069153 + 0{,}521810368131289 = 1{,}49993987376269$

$$\rho_X = 1 - \frac{1{,}48547529722733}{1{,}58496250072116} = 0{,}0627694367838705$$

$$\rho_Y = 1 - \frac{1{,}49993987376269}{1{,}58496250072116} = 0{,}0536433050749061$$

Обчислимо ентропію об'єднання двох джерел, використовуючи формулу

$$H(X,Y) = -\sum_{i=1}^{k}\sum_{j=1}^{k} p(x_i, y_j) log_2 p(x_i, y_j))$$

$H(X,Y) = -(0{,}475\, log_2 0{,}475 + 0{,}0125\, log_2 0{,}0125 + 0{,}0125\, log_2 0{,}0125 + 0{,}005\, log_2 0{,}005 + 0{,}19\, log_2 0{,}19 + 0{,}005\, log_2 0{,}005 + 0{,}0075\, log_2 0{,}0075 + 0{,}0075\, log_2 0{,}0075 + 0{,}285\, log_2 0{,}285 =$

$0{,}510150276185794 + 0{,}079024101186092 + 0{,}079024101186092 + 0{,}0382192809488736 + 0{,}455226448502916 + 0{,}0382192809488736 + 0{,}0529417026679018 + 0{,}0529417026679018 + 0{,}516125360048845 = 1{,}82187225434329$

Обчислимо умовну ентропію джерела Y стосовно джерела X за формулою

$$H(Y|X) = \sum_{i=1}^{k} p(x_i) H(Y|x_i)$$

Для цього спершу обчислимо часткові умовні ентропії

$$H(Y|x_i) = -\sum_{j=1}^{k} p(y_j|x_i) log_2(p(y_j|x_i))$$

$$H(Y|x_1) = -(0{,}95\, log_2 0{,}95 + 0{,}025\, log_2 0{,}025 + 0{,}025\, log_2 0{,}025) = 0{,}0703005523715881 + 0{,}133048202372184 + 0{,}133048202372184 = 0{,}336396957115956$$

$$H(Y|x_2) = -(0{,}025\ log_2 0{,}025 + 0{,}95\ log_2 0{,}95 + 0{,}025\ log_2 0{,}025)$$
$$= 0{,}133048202372184 + 0{,}0703005523715881$$
$$+ 0{,}133048202372184 = 0{,}336396957115956$$

$$H(Y|x_3) = -(0{,}025\ log_2 0{,}025 + 0{,}025\ log_2 0{,}025 + 0{,}95\ log_2 0{,}95)$$
$$= 0{,}133048202372184 + 0{,}133048202372184$$
$$+ 0{,}0703005523715881 = 0{,}336396957115956$$

Тепер підставимо отримані значення у формулу для знаходження загальної умовної ентропії

$H(Y|X) = 0{,}5 * 0{,}336396957115956 + 0{,}2 * 0{,}336396957115956 + 0{,}3 * 0{,}336396957115956 =$

$0{,}168198478557978 + 0{,}0672793914231912 + 0{,}100919087134787 = 0{,}336396957115956$

Обчислимо середню взаємну кількість інформації за формулою

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

$I(X;Y) = 1{,}48547529722733 + 1{,}49993987376269 - 1{,}82187225434329 = 1{,}16354291664673$

Оскільки $I(X;Y) \neq 0$, то джерела є статистично залежними.

Обчислимо умовну ентропію джерела X стосовно джерела Y за формулою

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

$$H(X|Y) = 1{,}48547529722733 - 1{,}16354291664673 = 0{,}3219323805806$$

Відповідь:

$\boldsymbol{H(X,Y)}$	$\boldsymbol{H(X\|Y)}$	$\boldsymbol{H(Y\|X)}$	$\boldsymbol{I(X;Y)}$
1,82187225434329	0,3219323805806	0,336396957115956	1,16354291664673

$\boldsymbol{H(X)}$	$\boldsymbol{H(Y)}$	$\boldsymbol{\rho_X}$	$\boldsymbol{\rho_Y}$
1,48547529722733	1,49993987376269	0,0627694367838705	0,0536433050749061