Практичні заняття з теорії інформації та кодування

Практичне заняття 1. Ансамблі та джерела повідомлень

Мета: Отримати загальні відомості про склад дисципліни «Теорія інформації та кодування», математичні моделі каналів зв’язку.

Питання 1

Повідомлення - це сукупність даних або інформації, яка передається від одного об'єкта до іншого з метою передачі знань, інструкцій, або інших даних. Повідомлення можуть бути представлені в різних формах: текст, зображення, звук, відео тощо. Вони можуть бути закодовані та передані через різноманітні канали зв'язку, наприклад, радіохвилі, оптоволокно, інтернет тощо.

Питання 2

Дані представляють собою сировину або необроблені факти та цифри без конкретного значення. Повідомлення - це структуровані дані, які передаються або зберігаються. Інформація виникає, коли дані або повідомлення інтерпретуються та аналізуються в конкретному контексті, надаючи їм значення та корисність.

Питання 3

Неперервні повідомлення виникають з неперервних джерел і можуть приймати будь-яке значення з нескінченної множини, наприклад, аналоговий аудіосигнал. Дискретні повідомлення формуються з дискретних джерел і можуть приймати значення лише з обмеженої, зазвичай скінченної, множини, наприклад, текстові повідомлення.

Питання 4

Операції перетворення сигналу на повідомлення включають кілька етапів: кодування, модуляція, передача, демодуляція та декодування. Ці етапи допомагають адаптувати сигнал до характеристик каналу та відновити оригінальне повідомлення.

Питання 5

Дисципліна «Теорія інформації та кодування» вивчає основні принципи обробки, передачі, зберігання та захисту інформації, методи кодування та декодування, теорію інформації, теорію кодування та їх практичні застосування в комунікаційних системах.

Питання 6

Математичні моделі каналів зв’язку розроблені для моделювання характеристик та обмежень фізичних каналів. Вони включають моделі з адитивним білим гауссівським шумом, канали з вибірковим стиранням та канали з постійним затриманням.

Питання 7

Джерело повідомлень - це елемент комунікаційної системи, який генерує повідомлення для передачі через канал. Джерело може бути фізичним об'єктом, таким як мікрофон, або абстрактним, наприклад, програмою або користувачем.

Питання 8

Ансамбль повідомлень - це сукупність усіх можливих повідомлень, які може генерувати джерело, із вказанням імовірностей кожного повідомлення. Це базове поняття в теорії інформації для аналізу властивостей джерел інформації та способів їх кодування.

Питання 9

Пара множин A = {a1, a2, a3}, P = {1/2; 1/3; 1/5} не визначає ансамбль повідомлень, оскільки сума ймовірностей в множині P
$\frac{1}{2}+\frac{1}{3}+\frac{1}{5}=\frac{15+10+6}{30}=\frac{31}{30}\neq 1$. Для визначення ансамблю повідомлень, сума ймовірностей усіх елементів має бути рівна 1.

Практичне заняття 2. Кількісна міра інформації

Мета: Отримати загальні відомості про методи вимірювання інформації, тим, як вона вимірюється у дисципліні «Теорія інформації та кодування», у математичних моделях каналів зв’язку.

Питання 1

Кількість інформації в повідомленні визначається за формулою Хартлі: $I = log_2 N$, де N - кількість символів в алфавіті, . Ця формула показує, скільки бітів необхідно для кодування одного повідомлення з ансамблю повідомлень.

Для повідомлення із К літер і алфавіту із N літер:

$$I = log_2 N^K = K\, log_2 N$$

Питання 2

Ентропія - це міра невизначеності стану джерела повідомлень. Вона визначається за формулою Шеннона: $H(X) = -\Sigma\,(p(x_i) * \log_2 p(x_i))$, де $p(x_i)$ - ймовірність появи повідомлення x_i. Властивості ентропії:

1) Невід'ємність: Ентропія завжди невід'ємна;
2) Дорівнює нулю, якщо ймовірність одного зі станів джерела інформації дорівнює 1, і тим самим стан джерела повністю визначено.
3) Максимальна, якщо повідомлення рівноймовірні та статистично незалежні: $H(X)_{max} = \log_2 N$
4) Адитивність: Ентропія об'єднаних статистично незалежних джерел інформації дорівнює сумі їх ентропій: $H(X \bullet Y) = H(X) + H(Y)$.

Питання 3

Для визначення кількості інформації в повідомленні a_1 з ймовірністю $p_1 = 0{,}3$, використовується формула: $I(a_1) = -\log_2(p_1)$. Визначено, що кількість інформації, яка міститься в повідомленні a_1, становить $I(a_1) = -\log_2 0{,}3 \approx 1{,}74$ біти.

Питання 4

Якщо ансамбль С містить 16 рівноймовірних повідомлень, кількість інформації, яку містить кожне таке повідомлення, можна знайти за формулою Хартлі: $I = \log_2(16) = 4$ біти.

Питання 5

Джерело А виробляє трилітерне повідомлення a_1 з алфавіту {a, b. c, d}, вибираючи їх рівноймовірно та незалежно одне від одного. Визначити кількість інформації, яку містить кожне таке повідомлення.

Спосіб 1

$$K = 3, \ N = 4$$

Загальна кількість можливих повідомлень в цьому випадку становить $N^K = 4^3$, оскільки кожна літера може бути однією з чотирьох літер алфавіту. Таким чином, кількість інформації, яку містить кожне таке повідомлення, можна знайти за формулою: I = $\log_2(4^3)$ =3 $\log_2 4$ = 6 біт.

Спосіб 2

Для повідомлення із K літер та алфавіту із N літер

$$I = K\, log_2\, N = 3\log_2 4 = \ 6 \text{ біт}$$

Практичне заняття 3. Безумовна ентропія

Мета: Ознайомитись з поняттями ентропія, безумовна ентропія, з тим, як визначається кількість інформації.

Питання 1

Кількість інформації в повідомленні визначається за формулою Хартлі: $I = log_2 N$, де N - кількість символів в алфавіті, . Ця формула показує, скільки бітів необхідно для кодування одного повідомлення з ансамблю повідомлень.

Для повідомлення із К літер і алфавіту із N літер:

$I = log_2 N^K = K\, log_2 N$

Питання 2

Ентропія - це міра невизначеності стану джерела повідомлень. Вона визначається за формулою Шеннона: $H(X) = -\Sigma\, (p\left(x_i\right) * \log_2 p(x_i))$, де $p(x_i)$ - ймовірність появи повідомлення x_i. Властивості ентропії:

1) Невід'ємність: Ентропія завжди невід'ємна;
2) Дорівнює нулю, якщо ймовірність одного зі станів джерела інформації дорівнює 1, і тим самим стан джерела повністю визначено.
3) Максимальна, якщо повідомлення рівноймовірні та статистично незалежні: $H(X)_{max} = \log_2 N$
4) Адитивність: Ентропія об'єднаних статистично незалежних джерел інформації дорівнює сумі їх ентропій: $H(X \bullet Y) = H(X) + H(Y)$.

Питання 3

Безумовна ентропія - це ентропія джерела інформації, яка характеризує середню кількість інформації, що міститься в повідомленнях джерела, не враховуючи будь-яку додаткову інформацію або умови.

Питання 4

Умовна ентропія - це міра невизначеності одного випадкового величини, умовно відносно значення іншої випадкової величини. Вона вимірює середню кількість інформації, що міститься в одній величині, коли значення іншої величини відоме.

Питання 5

Джерела А та В мають розподіли

P_A = {0,1; 0,1; 0,15; 0,125; 0,125; 0,1; 0,15; 0,15},

P_B = {0,5; 0,3; 0,1; 0,025; 0,025; 0,02; 0,15; 0,15}.

Ентропія якого джерела більше?

Визначаємо ентропію кожного джерела за формулою Шеннона

$$H(X) = -\Sigma\left(p\left(x_i\right) * \log_2 p(x_i)\right)$$

Джерело А

									Σ
P	0,1	0,1	0,15	0,125	0,125	0,1	0,15	0,15	1
log2 p	-3,3219	-3,3219	-2,7370	-3,0000	-3,0000	-3,3219	-2,7370	-2,7370	
p*log	-0,3322	-0,3322	-0,4105	-0,3750	-0,3750	-0,3322	-0,4105	-0,4105	-2,9782

Джерело В

									Σ
P	0,5	0,03	0,1	0,025	0,025	0,02	0,15	0,15	1
log2 p	-1,0000	-5,0589	-3,3219	-5,3219	-5,3219	-5,6439	-2,7370	-2,7370	
p*log	-0,5000	-0,1518	-0,3322	-0,1330	-0,1330	-0,1129	-0,4105	-0,4105	-2,1840

Виправлено помилку в умові: при 0,3 сума ймовірностей 1,27, що не може бути вірним.

Ентропія джерел

$$H_A = 2{,}9782\ біт,\quad H_B = 2{,}1840\ біт$$

$$H_A > H_B$$

Практичне заняття 4. Умовна ентропія

Мета: Ознайомитись з поняттям умовної ентропії, з тим, яку роль відіграє вона для вимірювання інформації, втрат інформації.

Питання 1

Умовна ентропія - це міра невизначеності одного випадкового величини, умовно відносно значення іншої випадкової величини. Вона вимірює середню кількість інформації, що міститься в одній величині, коли значення іншої величини відоме.

Питання 2

Різновиди умовної ентропії включають часткову умовну ентропію та загальну умовну ентропію.
Часткова умовна ентропія характеризує середню невизначеність одного символу джерела, умовно відносно іншого символу.
Загальна умовна ентропія характеризує середню невизначеність всього джерела, умовно відносно іншого джерела.

Питання 3

Основні властивості умовної ентропії включають невід'ємність, адитивність та симетрію.
Умовна ентропія завжди невід'ємна, адитивна відносно незалежних величин та симетрична, тобто H(X|Y) = H(Y|X) для незалежних X та Y.

Питання 4

Часткова умовна ентропія визначається як середня невизначеність одного символу джерела, умовно відносно іншого символу.
Для двох дискретних немарківських джерел інформації з алфавітами $A = \left\{a_1, a, \dots, a_M\right\}$ $та$ $B = \left\{b_1, b_2, \dots, b_N\right\}$, якщо вони є статистично залежними, поява символу a_1 на виході першого джерела дає розподіл умовних ймовірностей $p(b_k/a_1)$, який відрізняється від розподілу $p\left(b_k/a_2\right)$ $і$ $т.д.$ Ентропія другого джерела в залежності від символу на виході першого джерела задається наступним виразом

$$H\left(B/a_i\right) = -\sum_{k=1}^{N} p(b_k/a_i) \bullet \log_2 p(b_k/a_i)$$

Питання 5

Загальна умовна ентропія визначається як середня невизначеність всього джерела, умовно відносно іншого джерела.

Якщо $H\left(B/a_i\right)$ усереднити по всіх a_i, то отримаємо загальну умовну ентропію

$$H(B/A) = \sum_{i=1}^{M} p\left(a_i\right)H\left(B/a_i\right) = -\sum_{i=1}^{M}\sum_{k=1}^{N} p\left(a_i\right)p(b_k/a_i)\log_2 p(b_k/a_i) =$$

$$= -\sum_{i=1}^{M}\sum_{k=1}^{N} p\left(a_i, b_k\right)\log_2 p(b_k/a_i)$$

де $p\left(a_i, b_k\right) = p\left(a_i\right)p(b_k/a_i)$ – ймовірність сумісної появи символів b_k *та* a_i на виходах другого та першого джерела.

Питання 6

Ентропія об’єднання двох джерел визначається як H(A, B) = H(A) + H(B|A) = H(B) + H(A|B), де H(A) та H(B) - ентропії джерел A та B відповідно, H(B|A) - умовна ентропія B при умові A, та H(A|B) - умовна ентропія A при умові B.

Завдання 7

Ентропія монітора персонального комп'ютера при виведенні тексту в 28 рядків по 60 рівноймовірних символів у кожному, використовуючи стандартний міжнародний код (128 символів) з двома градаціями яскравості, може бути знайдена за формулою:

$$H = \log_2(128 \bullet 2)^{28 \bullet 60} = 28 \bullet 60 \log_2 2^8 = 1680 \bullet 8 = 13440\ \textit{біт}$$

Завдання 8

Ансамбль повідомлень джерела A визначено, як A = {0; 1} та P_A = {0,75;0,25}. Статистична залежність повідомлень $a_i \in A$ характеризується умовними ймовірностями p(0/1) = 0,12 і p(1/0) = 0,08. Визначити часткову та загальну умовну ентропію цього джерела.

Матриця умовних ймовірностей

$$\begin{pmatrix} p(0/0) & p(0/1) \\ p(1/0) & p(1/1) \end{pmatrix} = \begin{pmatrix} 0{,}75-0{,}12 & 0{,}12 \\ 0{,}08 & 0{,}25-0{,}08 \end{pmatrix} = \begin{pmatrix} 0{,}63 & 0{,}12 \\ 0{,}08 & 0{,}17 \end{pmatrix}$$

	(0/0)	(0/1)	(1/0)	(1/1)	Σ
$p(b_k/a_i)$	0,63	0,12	0,08	0,17	1
$\log_2(b_k/a_i)$	-0,6666	-3,0589	-3,6439	-2,5564	
$p \bullet log$	-0,4199	-0,3671	-0,2915	-0,4346	

$$H\big(B/a_i\big) = -\sum_{k=1}^{N} p(b_k/a_i) \bullet \log_2 p(b_k/a_i)$$

$$H(B/0) = -\big(p(0/0) \bullet \log_2 p(0/0) + p(1/0) \bullet \log_2 p(1/0)\big) =$$

$$= 0{,}4199 + 0{,}2915 \approx 0{,}712 \textit{ біт}$$

$$H(B/1) = -\big(p(0/1) \bullet \log_2 p(0/1) + p(1/1) \bullet \log_2 p(1/1)\big) =$$

$$= 0{,}3671 + 0{,}4346 \approx 0{,}802 \textit{ біт}$$

Загальна умовна ентропія

$$H(B/A) = \sum_{i=1}^{M} p\big(a_i\big) H\big(B/a_i\big)$$

$$H(B/A) = 0{,}75 \bullet H(B/0) + 0{,}25 \bullet H(B/1) \approx$$

$$\approx 0{,}75 \bullet 0{,}712 + 0{,}25 \bullet 0{,}802 \approx 1{,}085$$

Практичне заняття 5. Продуктивність дискретного джерела

Мета: Ознайомити з поняттям продуктивності дискретного джерела, з тим, яку роль відіграє вона для вимірювання інформації, втрат інформації.

Питання 1

Основні властивості ентропії об’єднання двох джерел включають:

1) Адитивність: H(A, B) = H(A) + H(B|A) = H(B) + H(A|B).
2) Невід'ємність: Ентропія об’єднання завжди невід'ємна.
3) Симетрія: H(A, B) = H(B, A).
4) Максимальна для незалежних джерел: Якщо джерела A та B незалежні, то H(A, B) = H(A) + H(B).

Питання 2

Кількість інформації на одне повідомлення двох статистично взаємозв’язаних джерел визначається як середня кількість інформації, що міститься в повідомленнях обох джерел. Математично це виражено як H(A, B), де H(A, B) - ентропія об’єднання джерел A та B.

$$H(A,B) = -\sum_{i=1}^{M}\sum_{k=1}^{N} p\left(a_i, b_k\right)\log_2 p(a_i, b_k)$$

Питання 3

Ентропія двох джерел стає максимальною, коли повідомлення з цих джерел є статистично незалежними. У цьому випадку максимальна ентропія об’єднання рівна сумі ентропій кожного джерела, тобто H(A, B) = H(A) + H(B).

Питання 4

Продуктивність дискретного джерела визначається як середня кількість інформації, яку джерело виробляє за одиницю часу. Математично це виражено як R = H(A)/T, де H(A) - ентропія джерела, а T - середній інтервал часу між послідовними повідомленнями.

Питання 5

Для визначення продуктивності дискретного джерела з різною тривалістю вибору повідомлень, необхідно врахувати середню тривалість вибору кожного повідомлення. Продуктивність можна визначити як $R = H(A)/T_{avg}$, де $H(A)$ - ентропія джерела, а T_{avg} - середня тривалість вибору повідомлення.

Питання 6

При передачі банківської інформації реченнями по 16 рядків на кожні 100 речень цифра 5 зустрічається 90 разів, а цифра 9 - 70 разів. Числа 59 і 95 зустрічаються 12 разів. Визначити умовну ентропію появи в реченні цифри 9, якщо в ньому є цифра 5, та умовну ентропію цифри 5, якщо в ньому з'явилась цифра 9.

Ймовірність сумісної появи 5 і 9: $p(5{,}9) = \frac{12}{100 \bullet 16} = \frac{3}{400}$

Ймовірність появи в реченні цифри 5: $p(5) = \frac{90}{100 \bullet 16} = \frac{9}{160}$

Ймовірність появи в реченні цифри 9: $p(9) = \frac{70}{100 \bullet 16} = \frac{7}{160}$

Умовна ймовірність появи 9, якщо є 5

$$P(9/5) = \frac{p(5{,}9)}{p(5)} = \frac{3}{400} : \frac{9}{160} = \frac{3}{400} \bullet \frac{160}{9} = \frac{1}{5} \bullet \frac{2}{3} = \frac{2}{15}$$

Умовна ймовірність появи 5, якщо є 9

$$P(5/9) = \frac{p(5{,}9)}{p(9)} = \frac{3}{400} : \frac{7}{160} = \frac{3}{400} \bullet \frac{160}{7} = \frac{3}{5} \bullet \frac{2}{7} = \frac{6}{35}$$

Умовна ентропія появи 9, якщо є 5

$$H(9/5) = -p(9/5) \bullet \log_2 p(9/5) = -\frac{2}{15} \bullet \log_2 \frac{2}{15} \approx 0{,}388\ біт$$

Умовна ентропія появи 5, якщо є 9

$$H(5/9) = -p(5/9) \bullet \log_2 p(5/9) = -\frac{6}{41} \bullet \log_2 \frac{6}{41} \approx 0{,}436\ біт$$

Практичне заняття 6. Пропускна здатність каналу

Мета: Ознайомити з поняттям пропускної здатності каналу з тим, яку роль відіграє вона для вимірювання інформації, втрат інформації.

Питання 1

Швидкість передачі інформації по дискретному каналу визначається як кількість біт інформації, яка передається через канал за одиницю часу. Це можна визначити за формулою:

$R = B \log_2(1 + S/N)$

де R - швидкість передачі інформації, B - пропускна здатність каналу, S/N - відношення сигнал/шум.

Питання 2

Інформаційні втрати при передачі інформації по каналу зв’язку дорівнюють різниці між вхідною ентропією джерела інформації та вихідною ентропією, яка досягається при передачі через канал. Це можна визначити як:

Інформаційні втрати = $H(X) - H(X|Y)$

де $H(X)$ - вхідна ентропія, $H(X|Y)$ - умовна ентропія на виході каналу.

Питання 3

Пропускна здатність каналу передачі - це максимальна швидкість передачі даних, яку канал може підтримувати. Для дискретних каналів це можна визначити за формулою Шеннона:

$C = B \log_2(1 + S/N)$

де C - пропускна здатність, B - пропускна здатність каналу, S/N - відношення сигнал/шум.

Питання 4

Пропускна здатність каналу при відсутності завад визначається просто як пропускна здатність каналу, оскільки відсутність завад означає, що відношення сигнал/шум є нескінченно великим. Тобто:

C = B $\log_2(1 + \infty)$ = B log_2(∞) = ∞

Отже, теоретично, при відсутності завад, пропускна здатність каналу є нескінченною.

Завдання 5

Дослідження каналу зв'язку між джерелом A та спостерігачем B виявило такі умовні ймовірності вибору повідомлень $b_j \in B$

$$p(b_j / a_i) = \begin{pmatrix} 0{,}97 & 0{,}02 & 0{,}01 \\ 0{,}1 & 0{,}86 & 0{,}04 \\ 0{,}03 & 0{,}08 & 0{,}89 \end{pmatrix}$$

Визначити часткову та загальну умовну ентропію повідомлень в цьому каналі при рівноймовірному виборі їх джерелом A та при P_A = {0,65; 0,3; 0.05}.

Часткова ентропія

$$H\left(B/a_i\right) = -\sum_{k=1}^{N} p(b_k/a_i) \bullet \log_2 p(b_k/a_i)$$

Загальна умовна ентропія

$$H(B/A) = \sum_{i=1}^{M} p\left(a_i\right) H\left(B/a_i\right)$$

	0,97	0,02	0,01	
p =	0,1	0,86	0,04	
	0,03	0,08	0,89	
	-0,0439	-5,6439	-6,6439	
log p =	-3,3219	-0,2176	-4,6439	
	-5,0589	-3,6439	-0,1681	
				Σ
	-0,0426	-0,1129	-0,0664	-0,2219
p*log =	-0,3322	-0,1871	-0,1858	-0,7051
	-0,1518	-0,2915	-0,1496	-0,5929

Отже, часткові ентропії

$$H\left(B/a_1\right) = 0{,}2219, \quad H\left(B/a_2\right) = 0{,}7051, \quad H\left(B/a_3\right) = 0{,}5929$$

Загальна умовна ентропія при рівномірному виборі джерелом А

$$H(B/A) = \frac{1}{3}\sum_{i=1}^{M} H\left(B/a_i\right) \approx 0{,}507\ біт$$

Загальна умовна ентропія при P_A

$$H(B/A) = 0{,}65 \bullet H\left(B/a_1\right) + 0{,}3 \bullet H\left(B/a_2\right) + 0{,}05 \bullet H\left(B/a_3\right) \approx 0{,}385\ біт$$

Завдання 8

Два статистично незалежних джерела визначаються матрицею сумісних ймовірностей

$$p(a_i, b_j) = \begin{pmatrix} 0{,}25 & 0 & 0{,}1 \\ 0{,}15 & 0{,}3 & 0{,}1 \\ 0 & 0{,}05 & 0{,}05 \end{pmatrix}$$

Визначити часткову та загальну умовну ентропію, ентропію об'єднання, безумовну ентропію цих джерел, а також кількість інформації, що припадає на пару повідомлень ai,bj.
Безумовні ймовірності

$$p(a_i) = \begin{pmatrix} 0,25 + 0 + 0,1 \\ 0,15 + 0,3 + 0,1 \\ 0 + 0,05 + 0,05 \end{pmatrix} = \begin{pmatrix} 0,35 \\ 0,55 \\ 0,1 \end{pmatrix}$$

$$p\left(b_j\right) = (0,25 + 0,15 + 0; \quad 0 + 0,3 + 0,5; \quad 0,1 + 0,1 + 0,05) = (0,4; \quad 0,35; \quad 0,25)$$

Безумовні ентропії

$$H(A) = -\sum_i p(a_i)\log_2(p(a_i))$$

p(a)	log	p*log
0,35	-1,5146	-0,5301
0,55	-0,8625	-0,4744
0,1	-3,3219	-0,3322
		-1,3367

$$H(A) = 1,337 \text{ } біт$$

$$H(B) = -\sum_j p\left(b_j\right)\log_2(p\left(b_j\right))$$

p(b)	0,4	0,35	0,25	
log	-1,3219	-1,5146	-2,0000	
p*log	-0,5288	-0,5301	-0,5000	-1,5589

$$H(B) = 1,559 \text{ } біт$$

Сумісна ентропія (ентропія об'єднання)

$$H(A, B) = -\sum_i \sum_j p(a_i, b_j)\log_2\left(p(a_i, b_j)\right)$$

	0,25	0	0,1
p =	0,15	0,3	0,1
	0	0,05	0,05

	-2,0000	0,0000	-3,3219
log =	-2,7370	-1,7370	-3,3219

	0,0000	-4,3219	-4,3219	
	-0,5000	0,0000	-0,3322	
p*log=	-0,4105	-0,5211	-0,3322	
	0,0000	-0,2161	-0,2161	
				-2,5282

$$H(A,B) = 2{,}528\ біт$$

$$H(A) + H(B) = 1{,}3367 + 1{,}5589 = 2{,}8955 \approx 2{,}896\ біт \neq H(A,B)$$

Отже, джерела не є статистично незалежними.

Часткові умовні ентропії

$$H(A\big|B) = H(A,B) - H(B) = 2{,}528 - 1{,}559 = 0{,}969\ біт$$

$$H(B\big|A) = H(A,B) - H(A) = 2{,}528 - 1{,}337 = 1{,}191\ біт$$

Кількість інформації розраховується як:

I(A;B) = H(A) + H(B) - H(A,B)

$$I(A,B) = 2{,}896 - 2{,}528 = 0{,}367\ біт$$

Практичне заняття 7. Інформаційні втрати при передачі інформації по каналу зв’язку

Мета: Отримати загальні відомості про інформаційні втрати при передачі інформації по каналу зв’язку, про методи боротьби з перешкодами при передачі інформації

Питання 1

Інформаційні втрати при передачі інформації по каналу зв’язку дорівнюють різниці між вхідною ентропією джерела інформації та вихідною ентропією, яка досягається при передачі через канал. Це можна визначити як:

Інформаційні втрати = $H(X)-H(X|Y)$ де $H(X)$ - вхідна ентропія,

$H(X|Y)$ - умовна ентропія на виході каналу.

Питання 2

Інформаційні втрати в каналі з абсолютною статистичною залежністю між входом і виходом дорівнюють нулю, оскільки вся інформація, що входить в канал, передається на вихід без втрат. В цьому випадку:

Інформаційні втрати = 0

Питання 3

Інформаційні втрати в каналі з статистичною незалежністю входу і виходу дорівнюють вхідній ентропії, оскільки інформація, що входить в канал, не впливає на вихід, і тому вся інформація втрачається. В цьому випадку:

Інформаційні втрати = $H(X)$

Завдання 4

Ансамбль повідомлень джерела А визначено,

як А = {a_1, a_2, a_3} та P_A = {0,65; 0,25; 0,1}.

Матриця умовних ймовірностей каналу має вигляд:

$$p\left(b_j/a_i\right) = \begin{pmatrix} 0{,}99 & 0{,}005 & 0{,}005 \\ 0{,}13 & 0{,}75 & 0{,}12 \\ 0{,}15 & 0{,}35 & 0{,}5 \end{pmatrix}$$

Визначити кількість інформації, що передається в одному та 100 повідомленнях. Чому дорівнюють інформаційні втрати в каналі при передачі 100 повідомлень з алфавіту А.

Безумовна ентропія джерела А

p(a)	log	p*log
0,65	-0,6215	-0,4040
0,25	-2,0000	-0,5000
0,1	-3,3219	-0,3322
		-1,2362

$$H(A) = -\sum_i p\left(a_i\right)\log_2(p\left(a_i\right))$$
$$H(A) = 1{,}2362\ біт$$

Часткова ентропія

$$H\left(B/a_i\right) = -\sum_{k=1}^{N} p(b_k/a_i) \bullet \log_2 p(b_k/a_i)$$

Загальна умовна ентропія

$$H(B/A) = \sum_{i=1}^{M} p\left(a_i\right)H\left(B/a_i\right)$$

	0,99	0,005	0,005				
p =	0,13	0,75	0,12				
	0,15	0,35	0,5				
	-0,0145	-7,6439	-7,6439				
log =	-2,9434	-0,4150	-3,0589				
	-2,7370	-1,5146	-1,0000				
					H(B/ai)	Pa	H*Pa
	-0,0144	-0,0382	-0,0382		0,0908	0,65	0,0590
p*log =	-0,3826	-0,3113	-0,3671		1,0610	0,25	0,2652
	-0,4105	-0,5301	-0,5000		1,4406	0,1	0,1441
							0,4683

$$H(B/A) = 0{,}4683 \ біт$$

Сумісна ентропія (ентропія об'єднання)

$$H(A,B) = H(B/A) + H(A) = 0{,}4683 + 1{,}2362 = 1{,}7045 \ біт$$

Безумовна ентропія джерела B

	0,6435	0,00325	0,00325	
p(ai,bj) =	0,0325	0,1875	0,03	
	0,015	0,035	0,05	
p(B)	0,6910	0,2258	0,0833	
log	-0,5332	-2,1472	-3,5864	
p*log	-0,3685	-0,4847	-0,2986	-1,1518

$$H(B) = 1{,}1518 \ біт$$

Кількість інформації в одному повідомленні

I(A;B) = H(A) + H(B) - H(A,B)

$$I(A,B) = 1{,}2362 + 1{,}1518 - 1{,}7045 = 1{,}0825 \ біт$$

Кількість інформації в 100 повідомленнях

$$100 \bullet I(A,B) = 108{,}25 \;\; біт$$

Інформаційні втрати = H(A) - H(A|B)

де H(A) - вхідна ентропія, H(A|B) - умовна ентропія на виході каналу.

$$H(B|A) - H(A|B) = H(B) - H(A)$$
$$H(A) - H(A|B) = H(B) - H(B|A)$$
$$LOSS(A,B) = 1{,}1518 - 0{,}4683 = 0{,}6835 \;\; біт$$

Інформаційні втрати для 100 повідомлень

$$100 \bullet LOSS(A,B) = 68{,}35 \; біт$$

Завдання 5

Визначити інформаційні втрати в каналі передачі з матрицею умовних ймовірностей

$$p(b_j / a_i) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Оскільки це одинична матриця, канал передачі є ідеальним, і в ньому немає інформаційних втрат. Тобто, інформаційні втрати в каналі дорівнюють 0.

Практичне заняття 8. Теорема Шеннона про кодування дискретного джерела

Мета: Ознайомитись з торемою Шеннона про кодування дискретного джерела та її практичним застосуванням.

Питання 1

Продуктивність дискретного джерела визначається як середня кількість інформації, яка виробляється джерелом за одиницю часу. Математично це може бути виражено як $V=H(X)\cdot R$, де $H(X)$ - ентропія джерела, а R - швидкість видачі символів джерелом.

Питання 2

Теорема Шеннона про кодування дискретного джерела формулюється наступним чином: для будь-якого дискретного джерела з ентропією H, існує код з середньою довжиною слова, яка відрізняється від H менше ніж на довільно задану величину $\epsilon > 0$, за умови, що довжина блоку достатньо велика.

Питання 3

Зміст теореми Шеннона про кодування дискретного джерела полягає в тому, що існує можливість стиснення даних дискретного джерела до його ентропії з довільно високою точністю. Тобто, ентропія джерела визначає нижню межу середньої довжини кодового слова, і існують коди, які можуть досягти цього ліміту з довільно великою точністю.

Завдання 4

Повідомлення передаються взаємонезалежними рівноймовірними символами тривалістю $5 \bullet 10^{-4}$ сек. Визначити швидкість передачі кожного символу та всієї інформації, якщо обсяг алфавіту дорівнює 16, 32, 64.

Швидкість передачі – це відношення ентропії одного символу до часу його передачі

$$V = \frac{H_i(X)}{\tau}$$

Ентропія одного символу для алфавіту із N символів $H_i(X) = \log_2 N$

При $N = 16$

$$V_{16} = \frac{\log_2 16}{5 \bullet 10^{-4}} = \frac{4}{5} \bullet 10^4 = 8000\ біт/c$$

При $N = 32$

$$V_{16} = \frac{\log_2 32}{5 \bullet 10^{-4}} = \frac{5}{5} \bullet 10^4 = 10\,000\ біт/c$$

При $N = 64$

$$V_{16} = \frac{\log_2 64}{5 \bullet 10^{-4}} = \frac{6}{5} \bullet 10^4 = 12\,000\ біт/c$$

Завдання 5

Час передачі повідомлення 0 дорівнює 0,1 с., а повідомлення 1 – 0,6 с. Знайти розподіл ймовірностей p0 та p1 , за яких досягається максимальна швидкість передачі інформації.

Швидкість передачі інформації для двох повідомлень

$$V = V_1 + V_2 = p_0\frac{\log p_0}{t_0} + p_1\frac{\log p_1}{t_1} = \frac{p_0 \log p_0}{0{,}1} + \frac{(1-p_0)\log(1-p_0)}{0{,}6} =$$

$$= \frac{6p_0 \log p_0 + (1-p_0)\log(1-p_0)}{0{,}6} \to max$$

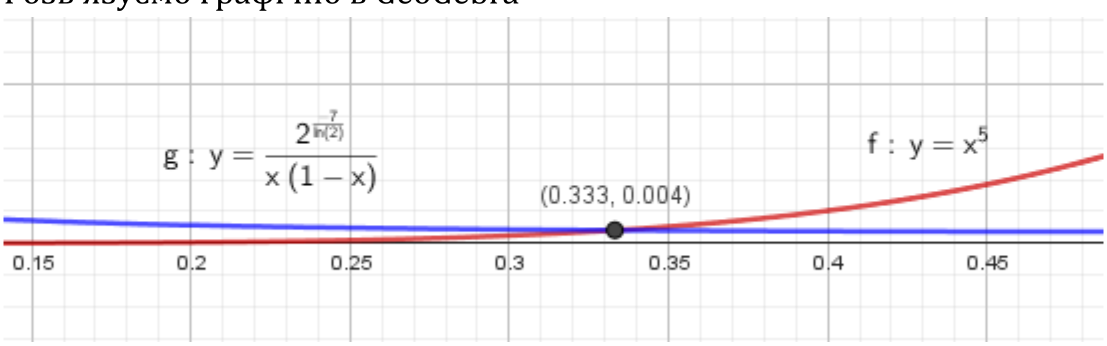
$$(x\log x)' = \log x + \frac{x}{x\ln 2} = \log x + \frac{1}{\ln 2}$$

$$V' = \frac{1}{0{,}6}\left(6\log p_0 + \frac{6}{\ln 2} + \log(1-p_0) + \frac{1}{\ln 2}\right) = 0$$

$$\log p_0^6(1-p_0) = -\frac{7}{\ln 2}$$

$$p_0^6(1-p_0) = 2^{-\frac{7}{\ln 2}}, \;\; y_1 = p_0^5; \quad y_2 = \frac{2^{-\frac{7}{\ln 2}}}{p_0(1-p_0)}$$

Розв’язуємо графічно в GeoGebra



$$p_0 \approx 0{,}333$$

signal	0	1	
p	0,333	0,667	
log	-1,5864	-0,5842	
-p*log	0,5283	0,3897	0,9180
/t	5,2827	0,6495	5,9322

Розподіл

$$p_0 \approx 0{,}333, \quad p_1 \approx 0{,}667$$

Максимальна швидкість

$$V = \frac{p_0 \log p_0}{0{,}1} + \frac{(1-p_0)\log(1-p_0)}{0{,}6} \approx 5{,}93\ біт/с$$

Завдання 6

Визначити пропускну здатність каналу, матриця ймовірностей якого при τ = 10-2 с. має вигляд

$$p(a_i, b_j) = \begin{pmatrix} 0 & 0{,}2 & 0{,}15 \\ 0 & 0{,}2 & 0 \\ 0{,}3 & 0 & 0{,}15 \end{pmatrix}.$$

	0	0,2	0,15
p(a,b) =	0	0,2	0
	0,3	0	0,15

	0	-2,32193	-2,73697
log	0	-2,32193	0
	-1,73697	0	-2,73697

	0	-0,46439	-0,41054	
p*log	0	-0,46439	0	
	-0,52109	0	-0,41054	
				-2,2710

$$H(A,B) = 2{,}2710\ біт$$

p(a)	log	p*log
0,35	-1,5146	-0,5301
0,2	-2,3219	-0,4644
0,45	-1,1520	-0,5184
		-1,5129

$$H(A) = 1{,}5129\ біт$$

p(b)	0,3	0,4	0,3
log	-1,7370	-1,3219	-1,7370

p*log	-0,5211	-0,5288	-0,5211	-1,5710

$$H(B) = 1{,}5710\ біт$$

$$H(B|A) = H(A,B) - H(A) = 2{,}2710 - 1{,}5129 = 0{,}7581\ біт$$

Пропускна здатність каналу

C = H(B) - H(B|A)

$C = 1{,}5710 - 0{,}7581 = 0{,}8129\ біт/символ$

Швидкість передачі інформації по каналу

$$V = \frac{C}{\tau},\quad V = \frac{0{,}8129}{10^{-2}} = 81{,}29\ біт/с$$

Завдання 7

Чи можлива безпомилкова передача інформації по каналу, параметри якого задані в попередній задачі, якщо продуктивність джерела Vдж = 9,6 Кбіт/с?

За теоремою Шеннона для каналу з перешкодами завжди можна знайти таку систему кодування, при якій повідомлення будуть передані з будь-яким великим ступенем вірності, якщо тільки продуктивність джерела не перевищує пропускну здатність каналу.

В даному випадку продуктивність джерела 9600 біт/с більша за швидкість передачі інформації по каналу 81 біт/с.

Отже, безпомилкова передача інформації неможлива.

Список використаної літератури

1. Блейхут Р. Теория и практика кодов, контролирующих ошибки. М.: Мир, 1986. – 576 с.

2. Жураковський Ю. П., Полторак В. П. Теорія інформації та кодування. К.: Вища шк., 2001. – 255с.

3. Жураковський Ю.П., Гніліцький В.В. Теорія інформації та кодування в задачах. Житомир: ЖІТІ, 2002. – 230 с.

4. Кузьмин И.В., Кедрус В.А. Основы теории информации и кодирования. - К.: Вища школа, 1986. — 238 с.