УКРАЇНСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ
ЗАЛІЗНИЧНОГО ТРАНСПОРТУ

ФАКУЛЬТЕТ ІНФОРМАЦІЙНО-КЕРУЮЧИХ СИСТЕМ
ТА ТЕХНОЛОГІЙ

Кафедра транспортного зв'язку

ТЕОРІЯ ІНФОРМАЦІЇ ТА КОДУВАННЯ

МЕТОДИЧНІ ВКАЗІВКИ

до виконання практичних занять і самостійних робіт

із дисципліни

«ТЕЛЕКОМУНІКАЦІЙНІ СИСТЕМИ ПЕРЕДАЧІ»

Харків – 2025

Методичні вказівки розглянуто і рекомендовано до друку на засіданні кафедри транспортного зв’язку 13 січня 2025 р., протокол № 7.

Методичні вказівки містять короткі теоретичні відомості, методику розв’язування типових задач і контрольні задачі для самостійного розв’язання для закріплення теоретичного матеріалу, тематика яких охоплює питання, розглянуті в рамках навчальної дисципліни «Телекомунікаційні системи передачі». Задачі, наведені в методичних вказівках, можна використати для проведення поточного контролю знань здобувачів, модульного контролю та на іспиті.

Методичні вказівки також можуть бути використані під час самостійної підготовки і для викладання певних розділів інших дисциплін відповідно до навчальних програм.

Рекомендовано для здобувачів вищої освіти першого (бакалаврського) рівня всіх форм навчання за освітньою програмою «Телекомунікації та радіотехніка».

Укладачі:

проф. К. А. Трубчанінова,

доц. І. В. Ковтун

Рецензент

проф. С. І. Доценко

ЗМІСТ

Завдання 1. ІНФОРМАЦІЙНІ ХАРАКТЕРИСТИКИ ДЖЕРЕЛ ПОВІДОМЛЕНЬ

1.1 Короткі теоретичні відомості

Під ***інформацією*** розуміють сукупність відомостей про якусь подію, об'єкт або процес. Інформація надходить від джерела до отримувача у вигляді повідомлень, які з передаванням дискретної інформації формуються із сукупності обмеженої кількості вихідних символів. Сукупність символів, із яких сформовані повідомлення, називають ***алфавітом***. Якщо алфавіт складається з m символів, а кожен символ позначимо x_i, то кількість інформації в i-му символі визначають як

$$I\left(x_i\right) = -\log_2 p\left(x_i\right). \tag{1.1}$$

Якщо символи алфавіту, з яких складаються повідомлення, статистично незалежні, середня кількість інформації в одному символі дорівнює

$$\overline{I\left(x_i\right)} = -\sum_{i=1}^{m} p\left(x_i\right)\log_2 p\left(x_i\right). \tag{1.2}$$

Оскільки одиницею вимірювання кількості інформації є біт, то в подальшому позначку «2» для log опускатимемо.

Важливою характеристикою джерела інформації є невизначеність його стану, так звана ***ентропія джерела***

$$H\left(X\right) = \overline{I\left(x_i\right)} = -\sum_{i=1}^{m} p\left(x_i\right)\log p\left(x_i\right). \tag{1.3}$$

Максимальну ентропію має джерело рівноймовірних незалежних повідомлень:

$$H(X) = -\sum_{i=1}^{m} p(x_i)\log p(x_i) = \log m \text{ за } p(x_i) = \frac{1}{m}. \tag{1.4}$$

Ентропію джерела безперервних повідомлень *x*, описаних стаціонарним ергодичним процесом, визначають за формулою

$$h(x) = -\int_X w(x)\log w(x)dx - \log \Delta x, \tag{1.5}$$

де $w(x)$ – одномірна густина ймовірності безперервного повідомлення;

Δx – роздільна здатність джерела.

Важливою характеристикою джерела інформації є його продуктивність

$$R_Д = \lim_{T \to \infty} \frac{1}{T} I(x_T), \tag{1.6}$$

де $I(x_T)$ – кількість інформації, продукована джерелом за проміжок часу T.

Для джерел дискретної інформації продуктивність дорівнює

$$R_Д = \frac{H(X)}{\bar{\tau}}, \tag{1.7}$$

де $\bar{\tau}$ – середня тривалість елементарних повідомлень (символів) джерела.

Для використання символів однакової тривалості $R_Д = \frac{H(X)}{\tau}$.

Для забезпечення максимальної швидкості передавання інформації каналом зв'язку необхідно узгоджувати характеристики джерела з характеристиками каналу. Це вирішують шляхом оптимального статистичного кодування повідомлень джерела двійковими кодами. Математична ознака оптимальності має вигляд

$$n_i = I(x_i) = -\log p(x_i), \tag{1.8}$$

де n_i – кількість двійкових символів у кодовій комбінації.

Найбільш відомими є схеми оптимального статистичного кодування Шеннона-Фано та Хаффмена. Розв'язанню цих задач присвячено завдання 2.

1.2 Приклади розв'язання основних типів задач

Приклад 1.1. Визначити ентропію повідомлення з шести літер, якщо загальна кількість літер в алфавіті дорівнює 32, усі повідомлення рівноймовірні.

Розв'язання. Загальна кількість шестилітерних повідомлень $m = 32^6$.

Використовуючи формулу для визначення ентропії рівноймовірних повідомлень, отримаємо

$$H(X) = \log m = \log 32^6 = 6\log 32 = 30 \text{ біт/повідомлення.}$$

Приклад 1.2. Визначити диференціальну ентропію безперервного повідомлення, розподіленого за нормальним законом

$$w(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right),$$

якщо його середня потужність, виражена в нормованих одиницях, дорівнює $\sigma^2=25$.

Розв'язання. Використовуючи вираз для визначення диференціальної ентропії, отримаємо

$$h(x)=-\int_{-\infty}^{\infty}w(x)\log w(x)dx=-\int_{-\infty}^{\infty}w(x)\log\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{x^2}{2\sigma^2}\right)\right]dx=$$

$$=-\log\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}w(x)dx-\int_{-\infty}^{\infty}w(x)\left(-\frac{x^2}{2\sigma^2}\right)\log e dx=$$

$$=\log\sqrt{2\pi\sigma^2}+\frac{\log e}{2\sigma^2}\int_{-\infty}^{\infty}x^2w(x)dx=\log\sqrt{2\pi\sigma^2}+\frac{\sigma^2\log e}{2\sigma^2}=\log\sqrt{2\pi e\sigma^2}=$$

$$=\log\sqrt{2\pi e25}=\log 5\sqrt{2\pi e}=4{,}37,\text{ біт/повідомл.}$$

Приклад 1.3. Вимірювана величина *x* змінюється в межах від x_0 до (x_0+b) і розподілена за законом рівної ймовірності. Знайти диференціальну ентропію величини *x*, якщо *b* дорівнює 32 нормованим одиницям.

Розв'язання. Закон рівної ймовірності можна аналітично подати як

$$w(x)=\begin{cases}\frac{1}{b}\,при\,x_0\le x\le x_0+b;\\ 0\,при\,x\le x_0\,та\,x>x+b.\end{cases}$$

Ентропія дорівнює

$$h(x) = -\int_{-\infty}^{\infty} w(x)\log w(x)\,dx = -\frac{1}{b}\int_{x_0}^{x_0+b}\log\frac{1}{b}dx =$$
$$= \log b = \log 32 = 5 \text{ біт/повідомл.}$$

Приклад 1.4. Джерело інформації генерує повідомлення з символів алфавіту $\{x_i\}\, i = \overline{1,m}$. Обчислити ентропію джерела і його надмірність за умови взаємної незалежності символів, якщо

$$i = \overline{1,4};\ p(x_1) = 0{,}1;\ p(x_2) = 0{,}2;\ p(x_3) = 0{,}3;\ p(x_4) = 0{,}4.$$

Розв'язання. Ентропія джерела дорівнює

$$H(X) = -\sum_{i=1}^{4} p(x_i)\cdot\log p(x_i) = -0{,}1\log 0{,}1 + 0{,}2\log 0{,}2 +$$
$$+0{,}3\log 0{,}3 + 0{,}4\log 0{,}4 = 1{,}85 \text{ біт/повідомл.}$$

Надмірність джерела

$$\gamma = 1 - \frac{H(X)}{\log m} = 1 - \frac{1{,}85}{\log 4} = 0{,}08.$$

Приклад 1.5. Закодувати оптимальним статистичним кодом за схемою Шеннона-Фано ансамбль повідомлень джерела $\{x_i\}$, якщо повідомлення статистично незалежні та задані апріорні ймовірності їх появи на виході джерела $p(x_i)$:

$$p(x_1)=0{,}1;\ p(x_2)=0{,}2;\ p(x_3)=0{,}05;\ p(x_4)=0{,}25;\ p(x_5)=0{,}1;$$
$$p(x_6)=0{,}01;\ p(x_7)=0{,}15;\ p(x_8)=0{,}1.$$

Розв'язання. Кодування за методом Шеннона-Фано здійснюється в такий спосіб. Усі повідомлення записують у таблицю в порядку зменшення їхньої ймовірності. Потім усю сукупність повідомлень розбивають на дві приблизно рівні групи. Усім повідомленням верхньої групи приписують перший кодовий символ «1», а повідомленням нижньої групи — символ «0». Потім кожну групу аналогічно розбивають на підгрупи за можливістю з однаковими ймовірностями, при цьому верхнім підгрупам в обох групах приписують символ «1» (другий символ кодової комбінації), а нижнім — символ «0». Ця процедура здійснюється доти, доки в кожній підгрупі не залишиться по одному повідомленню.

Процес кодування наведений у таблиці 1.1.

Жодна коротка кодова комбінація не має бути початком більш довгої. Середня довжина кодової комбінації

$$\langle n\rangle=\sum_{i=1}^{8} n_i\,p(x_i)=2{,}85\ \text{дв.симв.}$$

Для оптимального двійкового кодування ентропія дорівнює значенням, наведеним у таблиці 1.1.

Таблиця 1.1 – Процес кодування за методом Шеннона-Фано

Повідомлення	$p(x_i)$	*Кодування*	*Кодова комбінація*	*Кількість знаків* n_i	$I(x_i)$
x_4	0,25	11	11	2	2,0
x_2	0,2	10	10	2	2,33
x_7	0,15	011	011	3	2,745
x_1	0,1	010	010	3	3,33
x_5	0,1	0011	0011	4	3,33
x_8	0,1	0010	0010	4	3,33
x_3	0,05	0001	0001	4	4,33
x_6	0,05	0000	0000	4	4,33

У процесі розв’язання задачі завжди має бути виконана умова

$$H(X) \le \langle n \rangle.$$

Приклад 1.6. Сигнал формується у вигляді двійкового коду з імовірностями появи символів 1 і 0, які дорівнюють відповідно $p(x_1)=0,6$ і $p(x_0)=0,4$. Появи символів взаємопов’язані умовними ймовірностями:

$p(x_0/x_0)=0,1$ – імовірність того, що після 0 йтиме 0;

$p(x_1/x_0)=0,9$ – імовірність того, що після 0 йтиме 1;

$p(x_1/x_1)=0,1$ – імовірність того, що після 1 йтиме 1;

$p(x_0/x_1)=0,9$ – імовірність того, що після 1 йтиме 0.

Знайти ентропію сигналів.

Розв'язання. Знаходимо шукану ентропію за формулою

$$H(X) = -\sum_{i=0}^{1} p(x_i) \sum_{j=0}^{1} p(x_j/x_i) \log p(x_j/x_i) =$$

$$= -(0,1\log 0,1 + 0,9\log 0,9) \cdot (0,6+0,4) = 0,467 \text{ біт/повідомл.}$$

Приклад 1.7. Джерело формує дискретні повідомлення x_1, x_2, x_3, x_4, імовірності появи яких дорівнюють $p(x_1)=\frac{1}{2}$, $p(x_2)=\frac{1}{4}$, $p(x_3)=\frac{1}{8}$, $p(x_4)=\frac{1}{8}$.

Знайти ентропію повідомлень для випадків, коли між повідомленнями відсутні статистичні зв’язки і послідовність їх появи в довгому тексті має статистичну залежність, описану двомірними $p(x_i, x_j)$ і умовними ймовірностями $p(x_i/x_j)$, конкретні значення яких наведені в таблиці 1.2.

Таблиця 1.2 - Значення двомірних і умовних імовірностей

$x_i x_j$	$p(x_i, x_j)$	$p(x_i, x_j)$	$x_i x_j$	$p(x_i, x_j)$	$p(x_i, x_j)$
$x_1 x_1$	13/32	13/16	$x_3 x_1$	0	0
$x_1 x_2$	3/32	3/16	$x_3 x_2$	0	0
$x_1 x_3$	0	0	$x_3 x_3$	0	0
$x_1 x_4$	0	0	$x_3 x_4$	1/8	1
$x_2 x_1$	1/32	1/8	$x_4 x_1$	1/16	1/2
$x_2 x_2$	1/8	1/2	$x_4 x_2$	1/32	1/4
$x_2 x_3$	3/32	3/8	$x_4 x_3$	1/32	1/4
$x_2 x_4$	0	0	$x_4 x_4$	0	0

Розв'язання. Ентропія статистично незалежних повідомлень

$$H(X)=-\sum_{i=1}^{4} p(x_i)\log p(x_i)=$$

$$=-\left(\frac{1}{2}\log\frac{1}{2}+\frac{1}{4}\log\frac{1}{4}+\frac{1}{8}\log\frac{1}{8}+\frac{1}{8}\log\frac{1}{8}\right)=1,75 \text{ біт/повідомл.}$$

Ентропія попарно статистично пов'язаних повідомлень

$$H(X) = -\sum_{i=1}^{4}\sum_{j=1}^{4} p(x_i, x_j)\log p(x_i / x_j) =$$

$$= -(13/32 \cdot \log 13/16 + 3/32 \cdot \log 3/16 + 1/32 \cdot \log 1/8 + 1/8 \cdot \log 1/2 +$$

$$+3/32 \cdot \log 3/8 + 1/8 \cdot \log 1 + 1/16 \cdot \log 1/2 + 1/32 \cdot \log 1/4 +$$

$$+1/32 \cdot \log 1/4) = 0{,}886 \text{ біт/повідомл.}$$

Порівняння отриманих результатів свідчить, що наявність статистичних зв'язків майже вдвічі зменшує ентропію повідомлень.

Приклад 1.8. Джерелом інформації є вимірювальний датчик випадкового процесу x, рівномірно розподіленого в межах від 0 до 256 нормованих одиниць. Визначити кількість інформації, яку отримують у результаті одного заміру значення цього випадкового процесу, якщо похибка вимірювання розподілена за нормальним законом і середнє квадратичне значення похибки $\sigma = 4$.

Розв'язання. Диференціальна ентропія випадкової величини x

$$h(x) = -\int_0^{256} w(x)\log w(x)\,dx = -\int_0^{256} \frac{1}{256}\log\frac{1}{256}\,dx = 8 \text{ біт/повідомл.}$$

Диференціальна ентропія похибки вимірювання

$$h(\sigma) = \log\sqrt{2\pi e}\sigma = \log\sqrt{2\pi e}4 = \log 4\sqrt{2\pi e} = 4{,}04 \text{ біт/повідомл.}$$

Кількість інформації, яку отримують у результаті одного вимірювання, визначають як різницю між ентропією самої величини і ентропією похибки:

$$I(x) = h(x) - h(\delta) = 8 - 4{,}04 = 3{,}96 \text{ біт.}$$

Література [1, 2].

1.3 Задачі для самостійного розв’язання

Задача 1.1. Визначити ентропію повідомлень із k літер, якщо кількість літер в алфавіті дорівнює m, усі повідомлення рівноймовірні (таблиця 1.3).

Таблиця 1.3 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
k	4	5	6	7	8	8	7	6	5	4
m	16	32	64	128	256	256	128	64	32	16

Задача 1.2. Джерело інформації видає символи з ансамблю $X = \{x_i\}\left(i = \overline{1,4}\right)$ з імовірностями $p(x_i)$. Знайти кількість інформації, що міститься в кожному з символів джерела в разі їх незалежного вибору (джерело без пам’яті). Обчислити ентропію і надмірність заданого джерела (таблиця 1.4).

Таблиця 1.4 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
$p(x_1)$	0,2	0,15	0,05	0,1	0,55	0,45	0,4	0,5	0,5	0,45
$p(x_2)$	0,3	0,35	0,1	0,2	0,3	0,3	0,25	0,2	0,4	0,25
$p(x_3)$	0,4	0,45	0,15	0,25	0,1	0,15	0,2	0,1	0,05	0,2
$p(x_4)$	0,1	0,05	0,7	0,45	0,05	0,1	0,15	0,2	0,05	0,1

Задача 1.3. Джерелом інформації є вимірювальний датчик випадкового процесу x, рівноймовірно розподіленого в межах від 0 до m нормованих одиниць. Визначити кількість інформації, яку отримують у результаті одного заміру значення цього випадкового процесу, якщо похибка вимірювання розподілена за нормальним законом, середнє квадратичне значення похибки σ (таблиця 1.5).

Таблиця 1.5 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
m	128	256	512	1024	2048	2048	1024	512	256	128
σ	2	3	4	5	6	7	6	5	4	3

Задача 1.4. Визначити диференціальну ентропію неперервного повідомлення, розподіленого за нормальним законом, якщо його середня потужність, виражена в нормованих одиницях, дорівнює σ^2 (таблиця 1.6).

Таблиця 1.6 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
σ^2	10	20	30	40	50	60	70	80	90	100

Задача 1.5. Вимірювана величина x змінюється в межах від x_0 до $(x_0 + b)$ і розподілена за законом рівної ймовірності. Знайти диференціальну ентропію величини x (таблиця 1.7).

Таблиця 1.7 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
b	10	20	30	40	50	60	70	80	90	100

Контрольні питання

1 Дайте визначення середньої кількості інформації в одному символі. Яку одиницю використовують для визначення середньої кількості інформації в одному символі?

2 Що розуміють під ентропією джерела дискретних повідомлень і які властивості ентропії?

3 Як визначають і обчислюють диференціальну ентропію (ентропію джерела безперервних повідомлень)?

4 Дайте визначення продуктивності джерела повідомлень.

5 Наведіть математичну ознаку оптимальності джерела повідомлень.

Завдання 2. ЕНТРОПІЙНЕ КОДУВАННЯ

2.1 Короткі теоретичні відомості

Для зменшення надмірності джерела використовують ***ентропійне кодування*** (оптимальне або економне кодування) – кодування словами (кодами) змінної довжини, за якої довжина коду символу має зворотну залежність від імовірності появи символу в переданому повідомленні. Зазвичай ентропійні кодери використовують для стиснення даних коди, довжини яких пропорційні від'ємному логарифму ймовірності символу. Отже, найбільш імовірні символи використовують найкоротші коди.

Код Шеннона-Фано

Код Шеннона-Фано використовує коди змінної довжини: символ, що часто зустрічається, закодований кодом меншої довжини; той, що рідко зустрічається – кодом більшої довжини. Коди Шеннона-Фано префіксні, тобто жодне кодове слово не є префіксом (групою перших символів) будь-якого іншого слова. Ця властивість дає змогу однозначно декодувати будь-яку послідовність кодових слів. Для побудови шуканих кодів застосовують традиційний табличний спосіб кодування та побудову «кодового дерева».

Основні етапи створення ***табличного коду***:

1 Символи первинного алфавіту M виписують у порядку убування ймовірностей.

2 Символи отриманого алфавіту ділять на дві частини, сумарні ймовірності символів яких максимально близькі один одному.

3 У префіксному коді першій частині алфавіту присвоюють двійкову цифру «1», другій частині – «0».

4 Отримані частини рекурсивно діляться, а їхнім частинам призначають відповідні двійкові цифри в префіксному коді.

Код Шеннона-Фано будується також за допомогою ***дерева***.

Основні етапи побудови дерева:

1 Побудова дерева починається від кореня.

2 Усій множині кодованих елементів відповідають корені дерева (вершини першого рівня). Вона розбита на дві підмножини з приблизно однаковими сумарними ймовірностями. Ці підмножини відповідають двом вершинам другого рівня, з'єднаним із коренем.

3 Далі кожна з цих підмножин розбита на дві підмножини з приблизно однаковими сумарними ймовірностями. Їм відповідають вершини третього рівня.

4 Якщо підмножина містить єдиний елемент, то йому відповідає кінцева вершина кодового дерева; така підмножина розбиттю не підлягає.

5 Подібно поступаємо до тих пір, поки не отримаємо всі кінцеві вершини. Гілки кодового дерева розмічаємо символами 1 і 0.

Для побудови коду Шеннона-Фано розбиття множини елементів може бути вироблено декілька способів. Вибір розбиття на рівні n може погіршити варіанти розбиття на наступному рівні $(n+1)$ і призвести до неоптимальності коду в цілому. Іншими словами, оптимальна поведінка на кожному кроці шляху ще не гарантує оптимальності всієї сукупності дій. Тому код Шеннона-Фано не є оптимальним в загальному сенсі, хоча і дає оптимальні результати за деяких розподілів імовірностей. Для одного і того самого розподілу ймовірностей можна побудувати кілька кодів Шеннона-Фано, і всі вони можуть дати різні результати. Якщо побудувати всі можливі коди Шеннона-Фано для певного розподілу ймовірностей, то серед них будуть знаходитися і всі коди Хаффмана, тобто оптимальні коди. На деяких послідовностях можуть бути сформовані неоптимальні коди Шеннона-Фано, тому більш ефективним вважають стиск методом Хаффмана.

Код Хаффмана

Метод кодування Хаффмана складається з двох основних етапів:

- побудова оптимального кодового дерева;
- побудова відображення «код – символ» на основі побудованого дерева.

Алгоритм побудови коду Хаффмана:

1 Символи вихідного алфавіту виписують у порядку убування ймовірностей.

2 Дві останні (найменші) імовірності (символи) складаються, а отримана сума стає новим елементом таблиці, яка займає відповідне місце в списку ймовірностей, що убувають за величиною.

3 Попередній крок повторюють до тих пір, поки в списку не залишаться лише два елементи.

Бінарне дерево, відповідне коду Хаффмана, називають деревом Хаффмана. Завдання побудови коду Хаффмана рівносильне завданню побудови відповідного йому дерева.

Загальна схема побудови дерева Хаффмана:

1 Складають список кодованих символів.

2 Зі списку вибирають два вузли з найменшою вагою (під вагою слід розуміти частоту використання символу або ймовірність його появи – чим частіше використаний, тим більше важить).

3 Формують новий вузол і приєднують до нього як дочірні два вузли, вибрані зі списку. При цьому вага сформованого вузла дорівнює сумі ваг дочірніх вузлів.

4 Видаляють вибрані вузли зі списку.

5 Додають знову сформований вузол до списку.

6 Якщо в списку більше одного вузла, повторюють пункти 2-6.

2.2 Приклади розв’язання основних типів задач

Приклад 2.1. Закодувати оптимальним статистичним кодом за схемою Шеннона-Фано ансамбль повідомлень джерела $\{x_i\}$, якщо повідомлення статистично незалежні та задані апріорні ймовірності їх появи на виході джерела $p(x_i)$. Провести кодування ***однолітерних кодових комбінацій табличним методом*** (таблиця 2.1).

Таблиця 2.1 – Вихідні дані

Номер літери	1	2	3	4	5	6	7
Імовірність	0,2	0,1	0,1	0,4	0,1	0,05	0,05

Розв’язання

1 Символи первинного алфавіту виписуємо в порядку убування ймовірностей:

$$0{,}4 \rightarrow 0{,}2 \rightarrow 0{,}1 \rightarrow 0{,}1 \rightarrow 0{,}1 \rightarrow 0{,}05 \rightarrow 0{,}05.$$

Розмістимо ймовірності в новому порядку в таблицю Шеннона-Фано (таблиця 2.2).

Таблиця 2.2 – Кодування за методом Шеннона-Фано (перший етап)

Номер літери	Імовірність	Розбиття на підгрупи (римські цифри означають номери груп і підгруп)				Кодове позначення
1	0,4					
2	0,2					
3	0,1					
4	0,1					
5	0,1					
6	0,05					
7	0,05					

2 Символи отриманого алфавіту ділять на дві частини, сумарні ймовірності символів яких максимально близькі один одному (таблиця 2.3). Так, на першому етапі розподілу літер на групи відокремлюємо лише одну першу літеру (І група), залишивши у другій групі всі інші.

Таблиця 2.3 – Кодування за методом Шеннона-Фано (другий етап)

Номер літери	Імовірність	Розбиття на підгрупи (римські цифри означають номери груп і підгруп)				Кодове позначення
1	0,4	} I (0.4)				
2	0,2	} II (0.6)				
3	0,1					
4	0,1					
5	0,1					
6	0,05					
7	0,05					

3 У префіксному коді першій частині алфавіту присвоюють двійкову цифру «1», другій частині – «0» (таблиця 2.4).

4 Отримані частини рекурсивно діляться, а їхнім частинам призначають відповідні двійкові цифри в префіксному коді (таблиця 2.5). Так, друга і третя літери складуть І підгрупу ІІ групи; ІІ підгрупа тієї самої групи складається з решти чотирьох літер (4, 5, 6 і 7 літер).

Таблиця 2.4 – Кодування за методом Шеннона-Фано (третій етап)

Номер літери	Імовірність	Розбиття на підгрупи (римські цифри означають номери груп і підгруп)				Кодове позначення
1	0,4	} I				1
2	0,2	} II				0
3	0,1					0
4	0,1					0
5	0,1					0
6	0,05					0
7	0,05					0

Таблиця 2.5 – Кодування за методом Шеннона-Фано (четвертий етап)

Номер літери	Імовірність	Розбиття на підгрупи (римські цифри означають номери груп і підгруп)				Кодове позначення
1	0,4	} I				1
2	0,2	} II	} I			01
3	0,1					01
4	0,1		} II			00
5	0,1					00
6	0,05					00
7	0,05					00

5 Процес ділення продовжується доти, поки не залишиться пара ймовірностей для ділення, тобто I частина складатиметься лише з однієї літери (таблиця 2.6).

Таблиця 2.6 – Кодування за методом Шеннона-Фано (п’ятий етап)

Номер літери	Імовірність	Розбиття на підгрупи (римські цифри означають номери груп і підгруп)				Кодове позначення
1	0,4	} I				1
2	0,2	} II	} I	} I		011
3	0,1			} II		010
4	0,1		} II	} I	} I	0011
5	0,1				} II	0010
6	0,05			} II	} I	0001
7	0,05				} II	0000

Аналогічно попередньому прикладу розберемо випадок «алфавіту», що включає 18 літер, що мають такі ймовірності: 0.3; 0.2; 0.1 (дві літери); 0.05; 0.03 (п’ять літер); 0.02 (дві літери); 0.01 (шість літер) (таблиця 2.7).

Таблиця 2.7 – Кодування за методом Шеннона-Фано

<table>
<tr><th>Номер літери</th><th>Імовір-ність</th><th colspan="7">Розбиття на підгрупи (римські цифри означають номери груп і підгруп)</th><th>Кодове позначення</th></tr>
<tr><td>1</td><td>0,3</td><td rowspan="2">} I</td><td>} I</td><td colspan="5"></td><td>11</td></tr>
<tr><td>2</td><td>0,2</td><td>} II</td><td colspan="5"></td><td>10</td></tr>
<tr><td>3</td><td>0,1</td><td rowspan="16">}II</td><td rowspan="3">} I</td><td>} I</td><td colspan="4"></td><td>011</td></tr>
<tr><td>4</td><td>0,1</td><td rowspan="2">} II</td><td>} I</td><td colspan="3"></td><td>0101</td></tr>
<tr><td>5</td><td>0,05</td><td>} II</td><td colspan="3"></td><td>0100</td></tr>
<tr><td>6</td><td>0,03</td><td rowspan="13">} II</td><td rowspan="4">} I</td><td rowspan="2">} I</td><td>} I</td><td colspan="2"></td><td>00111</td></tr>
<tr><td>7</td><td>0,03</td><td>} II</td><td colspan="2"></td><td>00110</td></tr>
<tr><td>8</td><td>0,03</td><td rowspan="2">} II</td><td>} I</td><td colspan="2"></td><td>00101</td></tr>
<tr><td>9</td><td>0,03</td><td>} II</td><td colspan="2"></td><td>00100</td></tr>
<tr><td>10</td><td>0,03</td><td rowspan="9">} II</td><td rowspan="3">} I</td><td>} I</td><td colspan="2"></td><td>00011</td></tr>
<tr><td>11</td><td>0,02</td><td rowspan="2">} II</td><td>} I</td><td></td><td>000101</td></tr>
<tr><td>12</td><td>0,02</td><td>} II</td><td></td><td>000100</td></tr>
<tr><td>13</td><td>0,01</td><td rowspan="6">} II</td><td rowspan="3">} I</td><td>} I</td><td></td><td>000011</td></tr>
<tr><td>14</td><td>0,01</td><td rowspan="2">} II</td><td>} I</td><td>0000101</td></tr>
<tr><td>15</td><td>0,01</td><td>} II</td><td>0000100</td></tr>
<tr><td>16</td><td>0,01</td><td rowspan="3">} II</td><td>} I</td><td></td><td>000001</td></tr>
<tr><td>17</td><td>0,01</td><td rowspan="2">} II</td><td>} I</td><td>0000001</td></tr>
<tr><td>18</td><td>0,01</td><td>} II</td><td>0000000</td></tr>
</table>

Приклад 2.2. Закодувати оптимальним статистичним кодом за схемою Шеннона-Фано ансамбль повідомлень джерела $\{x_i\}$, якщо повідомлення статистично незалежні та задані апріорні ймовірності їх появи на виході джерела $p(x_i)$. Провести кодування ***однолітерних кодових комбінацій,*** використовуючи ***метод «дерева».*** Вихідні дані аналогічні прикладу 2.1.

Розв'язання

1 Побудова дерева починається від кореня (рисунок 2.1).

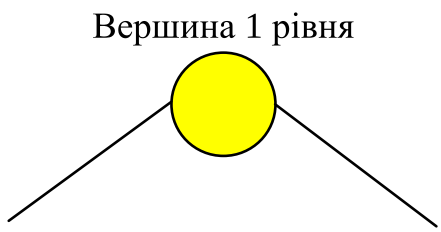

Рисунок 2.1 – Побудова дерева (перший етап)

2 Усій множині кодованих елементів відповідають корені дерева (вершини першого рівня). Вона розбита на дві підмножини з приблизно однаковими сумарними ймовірностями. Ці підмножини відповідають двом вершинам другого рівня, які з’єднані з коренем. Гілки кодового дерева розмічаємо символами 1 і 0 (рисунок 2.2).

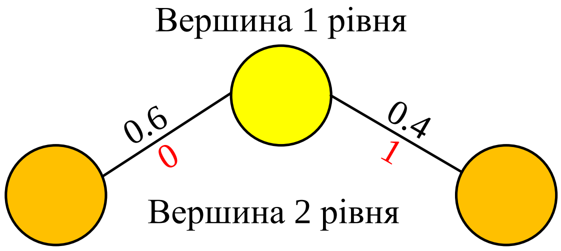

Рисунок 2.2 – Побудова дерева (другий етап)

3 Далі кожна з цих підмножин розбита на дві підмножини з приблизно однаковими сумарними ймовірностями. Їм відповідають вершини третього рівня (рисунок 2.3).

4 Якщо підмножина містить єдиний елемент, то йому відповідає кінцева вершина кодового дерева; така підмножина розбиттю не підлягає. Так, на рисунку 2.3 такому випадку відповідає літера **«1».**

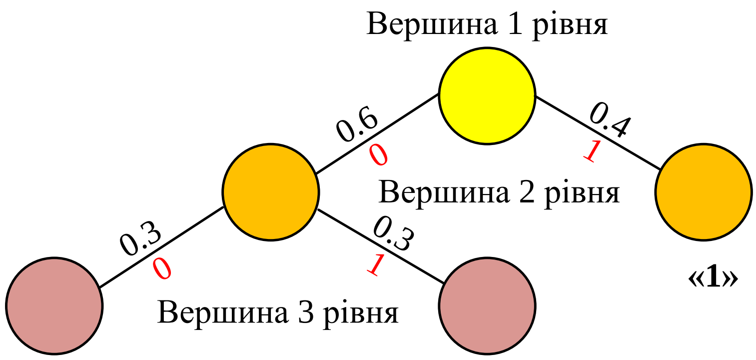

Рисунок 2.3 – Побудова дерева (третій і четвертий етапи)

5 Так само поступаємо до тих пір, поки не отримаємо всі кінцеві вершини (рисунок 2.4).

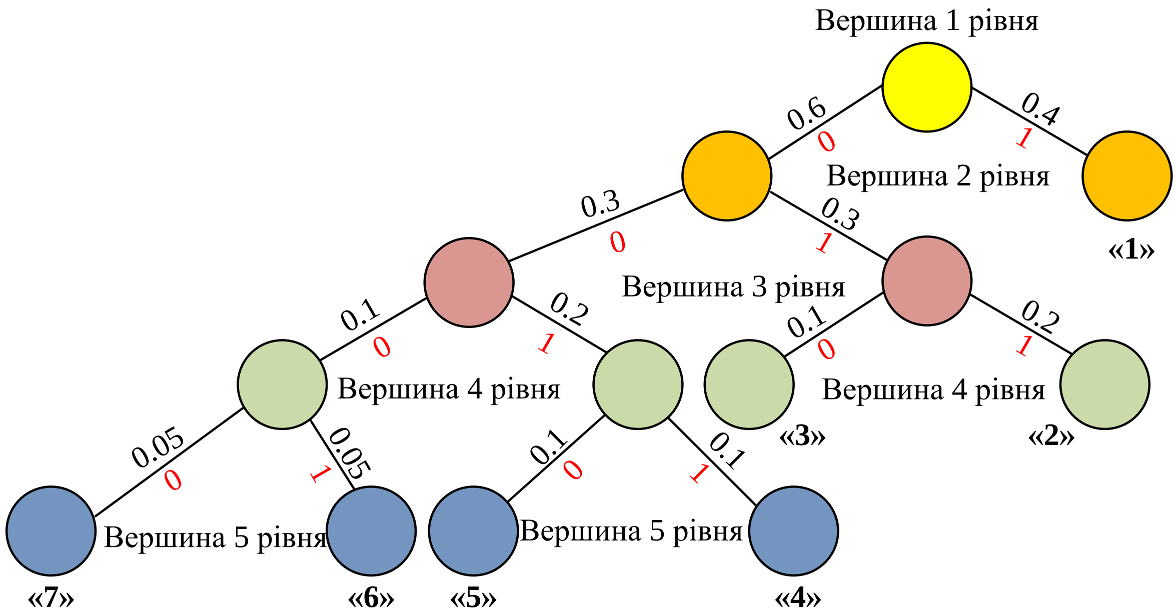

Рисунок 2.4 – Побудова дерева (п’ятий етап)

Визначимо ***середнє значення довжини*** l_c такого кодового позначення. Усе ж таки середнє значення довжини виявляється лише трохи більшим за мінімальне значення ентропії *H(X)*, допущене міркуваннями збереження кількості інформації для кодування. Так, для розглянутого вище прикладу семилітерного алфавіту найкращий рівномірний код складається з тризначних кодових позначень ($2^2 < 7 < 2^3$), тому в ньому на кожну літеру вихідного повідомлення припадає рівно три елементарних сигнали. З використанням же коду Шеннона-Фано середня кількість елементарних сигналів l_c, що припадають на одну літеру повідомлення, дорівнює

$$l_{cp} = \sum_{i=1}^{6} n_i \cdot p_i =$$

$$= 1 \cdot 0{,}4 + 3 \cdot 0{,}2 + 3 \cdot 0{,}1 + 4 \cdot 0{,}1 + 4 \cdot 0{,}1 + 4 \cdot 0{,}05 + 4 \cdot 0{,}05 = 2{,}5,\ біт,$$

де n_i – довжина кодового слова для певної літери;

p_i – імовірність появи цієї літери на виході джерела повідомлення.

Це значення помітно менше, ніж 3, і не дуже далеке від ентропії

$$H_c(X) = -0{,}4\log 0{,}4 - 0{,}2\log 0{,}2 - 3\cdot 0{,}1\log 0{,}1 - 2\cdot 0{,}05\log 0{,}05 \approx$$
$$\approx 2{,}391,\text{біт/літера}.$$

Для розв’язання задачі кодування завжди має бути виконана умова

$$H_c(X) \le l_c.$$

При цьому за будь-якого методу кодування, що використовує m-ічний код, середня кількість елементарних сигналів, що припадають на одну літеру повідомлення, ніколи не може бути менше відношення $\dfrac{H(X)}{\log m}$ (де $H(X)$ – ентропія однієї літери повідомлення); проте воно завжди може бути зроблено як завгодно близьким до цієї величини, якщо кодувати відразу досить довгі «блоки» з N літер. Звідси випливає, що якщо по лінії зв’язку за одиницю часу можна передати L елементарних сигналів (що набувають m різних значень), то швидкість передавання повідомлень по такій лінії не може бути більшою, ніж

$$v = \frac{L\log m}{H} \text{ літер/од. часу.}$$

Однак передавання зі швидкістю, як завгодно близькою до v (але меншою за v), вже є можливим.

Величина $C = L\log m$, що стоїть у чисельнику виразу для v, залежить лише від самої лінії зв’язку (у той час як знаменник $H(X)$ характеризує передане повідомлення). Ця величина C вказує на найбільшу кількість одиниць інформації, яку можна передати по лінії за одиницю часу (бо один елементарний сигнал, як ми знаємо, може містити щонайбільше $logm$ одиниць інформації); її називають ***пропускною здатністю лінії зв'язку***.

Розглянемо приклад 18-літерного алфавіту, для якого найкращий рівномірний код складається з п'ятизначних кодових позначень (оскільки $2^4 < 18 < 2^5$), у разі ж коду Шеннона-Фано є літери, кодовані навіть сьома двійковими сигналами, але середня кількість елементарних сигналів, що припадає на одну букву, тут дорівнює

$$l_C = 2 \cdot 0{,}5 + 3 \cdot 0{,}1 + 4 \cdot 0{,}15 + 5 \cdot 0{,}15 + 6 \cdot 0{,}06 + 7 \cdot 0{,}04 = 3{,}29, \text{ біт.}$$

Останнє значення помітно менше, ніж 5, і вже не набагато відрізняється від величини ентропії

$$H_C(X) = -0{,}3 \log 0{,}3 - 0{,}2 \log 0{,}2 - \ldots - 6 \cdot 0{,}01 \log 0{,}01 \approx 3{,}25, \textit{ біт /літера}.$$

Приклад 2.3. Закодувати оптимальним статистичним кодом за схемою Шеннона-Фано ансамбль повідомлень джерела $\{x_i\}$, якщо повідомлення статистично незалежні та задані апріорні ймовірності їх появи на виході джерела $p(x_i)$. Провести кодування ***дво- і трилітерних кодових комбінацій табличним методом.***

Розв'язання

Особливо вигідно кодувати за методом Шеннона-Фано не окремі літери, а відразу цілі блоки з декількох літер. Правда, при цьому все одно неможливо перевершити граничне значення $H_C(X)$ двійкових знаків на одну літеру повідомлення. Розглянемо, наприклад, випадок, коли є лише дві різні літери **А** і **Б**, що мають імовірності ***p*(*А*) = 0,7** та ***p*(*Б*) = 0,3** (таблиця 2.8), тоді

$$H_c(X) = -0{,}7\,log\,0{,}7 - 0{,}3\,log\,0{,}3 = 0{,}881,\text{біт/літера}.$$

Таблиця 2.8 – Кодування дволітерного джерела

Літера	Імовірність	Кодове позначення
А	0,7	1
Б	0,3	0

Застосування методу Шеннона-Фано для вихідного дволітерного алфавіту тут є безцільним: воно призводить лише до найпростішого рівномірного коду, який потребує для передавання кожної літери одного двійкового знака на 12 % більше мінімального досяжного значення 0,881 біт/літера.

Застосовуючи же метод Шеннона-Фано для кодування ***дволітерних*** комбінацій, імовірність яких визначена правилом множення ймовірностей для незалежних подій, отримаємо кодування, наведене в таблиці 2.9.

Середнє значення довжини кодового позначення

$$l_c = 1 \cdot 0{,}49 + 2 \cdot 0{,}21 + 3 \cdot 0{,}3 = 1{,}81,\ біт,$$

так що на одну літеру алфавіту припадає в середньому $\frac{1{,}81}{2} = 0{,}905$ двійкових знаків, і це лише на 3 % більше значення 0,881 біт/літера.

Таблиця 2.9 – Кодування дволітерних комбінацій джерела

Комбінація літер	Імовірність	Кодове позначення
АА	0,49	1
АБ	0,21	01
БА	0,21	001
ББ	0,09	000

Ще кращі результати отримаємо, застосувавши метод Шеннона-Фано для кодування ***трилітерних*** комбінації. При цьому отримуємо кодування, наведене в таблиці 2.10.

Таблиця 2.10 – Кодування трилітерних комбінацій джерела

Комбінація літер	Імовірність	Кодове позначення
ААА	0,343	11
ААБ	0,147	10
АБА	0,147	011
БАА	0,147	010
АББ	0,063	0010
БАБ	0,063	0011
ББА	0,063	0001
БББ	0,027	0000

Середнє значення довжини кодового позначення тут дорівнює 2,686, тобто на одну літеру тексту припадає в середньому 0,895 двійкових знаків, що лише на 1,5 % більше значення $H_c(X) \approx 0{,}881$ біт/літера.

У разі ще більшої різниці в імовірностях літер А і Б наближення до мінімально можливого значення $H_c(X)$ може бути дещо менш швидким, але воно проявляється не менш наочно. Так, за $p(A)$ = 0,89 і $p(Б)$ = 0,11 значення середньої ентропії дорівнює

$$H_c(X) = -0{,}89 \log 0{,}89 - 0{,}11 \log 0{,}11 = 0{,}5, біт/літера,$$

а рівномірний код $A \to 1, Б \to 0$ потребує витрат одного двійкового знака на кожну літеру у два рази більше.

При цьому застосування коду Шеннона-Фано для дволітерних комбінацій тут призводить до коду, у якому на кожну букву припадає в середньому 0,66 біт, застосування для трилітерних комбінацій – 0,55 біт, для чотирилітерних блоків – у середньому 0,52 біт, а це лише на 4 % більше за мінімальне значення 0,5 біт/літера.

Приклад 2.4. Закодувати оптимальним статистичним кодом за схемою табличного методу Хаффмана ансамбль повідомлень джерела $\{x_i\}$, якщо повідомлення статистично незалежні та задані апріорні ймовірності їх появи на виході джерела $p(x_i)$. Вихідні дані аналогічні прикладу 2.1.

Розв'язання

1 Символи вихідного алфавіту виписують у порядку убування ймовірностей (таблиця 2.11).

Таблиця 2.11 – Кодування за методом Хаффмана (перший етап)

<table>
<tr><th rowspan="3">Номер літери</th><th colspan="6">Імовірність</th></tr>
<tr><th rowspan="2">Вихідний алфавіт A</th><th colspan="5">Стислі алфавіти</th></tr>
<tr><th>A_1</th><th>A_2</th><th>A_3</th><th>A_4</th><th>A_5</th></tr>
<tr><td>1</td><td>0,4</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>2</td><td>0,2</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>3</td><td>0,1</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>4</td><td>0,1</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>5</td><td>0,1</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>6</td><td>0,05</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>7</td><td>0,05</td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

2 Дві останні (найменші) імовірності (символи) складаються, а отримана сума стає новим елементом таблиці, яка займає відповідне місце в списку ймовірностей, що убувають за величиною (таблиця 2.12).

Таблиця 2.12 – Кодування за методом Хаффмана (другий етап)

Номер літери	Імовірність					
	Вихідний алфавіт A	Стислі алфавіти				
		A_1	A_2	A_3	A_4	A_5
1	0,4	0,4				
2	0,2	0,2				
3	0,1	0,1				
4	0,1	0,1				
5	0,1	0,1				
6	0,05	0,1				
7	0,05					

3 Попередній крок повторюють до тих пір, поки в списку не залишаться лише два елементи (таблиця 2.13).

Таблиця 2.13 – Кодування за методом Хаффмана (третій етап)

Номер літери	Імовірність					
	Вихідний алфавіт A	Стислі алфавіти				
		A_1	A_2	A_3	A_4	A_5
1	0,4	0,4	0,4	0,4	0,4	0,6
2	0,2	0,2	0,2	0,2	0,4	0,4
3	0,1	0,1	0,2	0,2	0,2	
4	0,1	0,1	0,1	0,2		
5	0,1	0,1	0,1			
6	0,05	0,1				
7	0,05					

Кодування літер уже сформованої таблиці відбувається «рухом» по таблиці справа наліво. У результаті кожна літера отримує свою кодову комбінацію (таблиця 2.14).

Таблиця 2.14 – Кодування за методом Хаффмана (табличний метод)

Номер літери	Імовірність					
	Вихідний алфавіт A	Стислі алфавіти				
		A_1	A_2	A_3	A_4	A_5
1	0,4 **1**	0,4 **1**	0,4 **1**	0,4 **1**	0,4 **1**	0,6 **0**
2	0,2 **01**	0,2 **01**	0,2 **01**	0,2 **01**	0,4 **00**	0,4 **1**
3	0,1 **0001**	0,1 **0001**	0,2 **001**	0,2 **001**	0,2 **01**	
4	0,1 **0000**	0,1 **0000**	0,1 **0001**	0,2 **000**		
5	0,1 **0011**	0,1 **0011**	0,1 **0000**			
6	0,05 **00101**	0,1 **0010**				
7	0,05 **00100**					

Ціна кодування дорівнюватиме

$$l_c = \sum_{i=1}^{6} n_i \cdot p_i = 1 \cdot 0{,}4 + 2 \cdot 0{,}2 + 4 \cdot (0{,}1 \cdot 3) + 5 \cdot (0{,}05 \cdot 2) = 2{,}5, \text{ біт.}$$

Для передавання даних повідомлень можна перейти від політерного (поцифрового) кодування до кодування «блоків» (як і у випадку методу Шеннона-Фано), що складаються з фіксованої кількості послідовних «літер».

Приклад 2.5. Закодувати оптимальним статистичним кодом за схемою методу «дерева» Хаффмана ансамбль повідомлень джерела $\{x_i\}$, якщо повідомлення статистично незалежні та задані апріорні ймовірності їх появи на виході джерела $p(x_i)$. Вихідні дані аналогічні прикладу 2.4.

Розв'язання

Побудову кодового дерева починають з кореня. Двом ребрам, що походять із кореня, приписують як ваги ймовірності 0,6 і 0,4, які стоять в останньому стовпці. Вершинам дерева, що утворилися при цьому, приписано кодові символи 0 і 1. Далі «рухаємося» по таблиці справа наліво. Оскільки ймовірність 0,6 є результатом додавання двох імовірностей 0,4 і 0,2, з вершини 0 виходять два ребра з вагами 0,4 і 0,2 відповідно, що призводить до утворення двох нових вершин із кодовими символами 00 і 01. Процедура продовжується до тих пір, поки в таблиці залишаються ймовірності, отримані в результаті підсумовування. Дерево, отримане в результаті кодування за Хаффманом, зображено на рисунку 2.5.

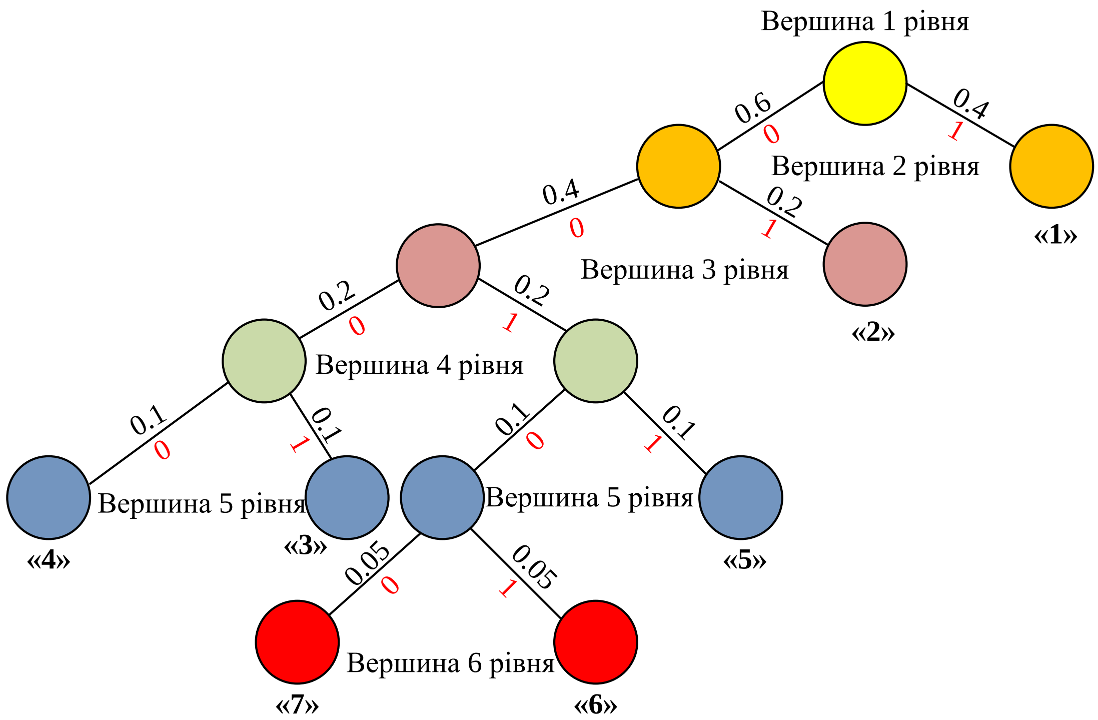

Рисунок 2.5 – Дерево Хаффмана

Література [3, 4].

2.3 Задачі для самостійного розв’язання

Задача 2.1. Закодувати оптимальним двійковим кодом за двома схемами методу Шеннона-Фано (табличний спосіб і «дерево») ансамбль повідомлень $X = \{x_i\}\left(i = \overline{1,8}\right)$, заданих апріорними ймовірностями $p\left(x_i\right)$.

Знайти ентропію ансамблю і середню кількість знаків у кодових комбінаціях. Оцінити виграш оптимального коду порівняно з рівномірним.

Таблиця 2.15 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
$p(x_1)$	0,25	0,2	0,15	0,3	0,4	0,2	0,12	0,04	0,05	0,17
$p(x_2)$	0,2	0,15	0,1	0,2	0,15	0,1	0,16	0,16	0,10	0,13
$p(x_3)$	0,15	0,1	0,05	0,1	0,1	0,2	0,22	0,03	0,15	0,08
$p(x_4)$	0,1	0,05	0,25	0,05	0,15	0,1	0,08	0,17	0,20	0,02
$p(x_5)$	0,05	0,1	0,15	0,1	0,05	0,15	0,1	0,06	0,20	0,1
$p(x_6)$	0,10	0,3	0,2	0,1	0,05	0,15	0,2	0,14	0,15	0,2
$p(x_7)$	0,10	0,05	0,05	0,1	0,05	0,05	0,06	0,15	0,10	0,15
$p(x_8)$	0,05	0,05	0,05	0,05	0,05	0,05	0,06	0,25	0,05	0,15
τ_i, мс	0,8	1,5	0,03	2,4	1,8	2,5	0,01	0,05	2,4	0,8

Задача 2.2. Закодувати оптимальним двійковим кодом за схемою дво- і трилітерного колування методом Шеннона-Фано ансамбль повідомлень $X = \{x_i\}\left(i = \overline{1,8}\right)$, заданих апріорними ймовірностями $p(x_i)$. Знайти ентропію ансамблю і середню кількість знаків у кодових комбінаціях. Оцінити виграш оптимального коду порівняно з рівномірним. Вихідні дані аналогічні задачі 2.1. Для дво- і трилітерного кодування вірогідності повідомлень А і Б мають вигляд

$$p\,(A) = p\,(x_1 + x_2 + x_3 + x_4\,);\ p\,(Б) = p\,(x_5 + x_6 + x_7 + x_8).$$

Задача 2.3. Закодувати оптимальним двійковим кодом за двома схемами методу Хаффмана (табличний спосіб і «дерево») ансамбль повідомлень $X = \{x_i\}\left(i = \overline{1,8}\right)$, заданих апріорними ймовірностями $p(x_i)$.

Знайти ентропію ансамблю і середню кількість знаків у кодових комбінаціях. Оцінити виграш оптимального коду порівняно з рівномірним. Вихідні дані аналогічні задачі 2.1.

Контрольні питання

1 Дайте визначення нерівномірним префіксним кодам.

2 Алгоритм кодування Хаффмана.

3 Алгоритм кодування Шеннона-Фано.

4 За рахунок чого здійснюється стиснення, за Хаффманом?

5 Дво- і трилітерне кодування.

Завдання 3. ДИСКРЕТНІ ТА БЕЗПЕРЕРВНІ КАНАЛИ ЗВ'ЯЗКУ

3.1 Короткі теоретичні відомості

Під час проєктування систем передавання інформації важливо визначити їхні основні інформаційні характеристики: швидкість передавання інформації і пропускну здатність.

У реальних каналах зв'язку завжди присутні завади різного походження, які зменшують швидкість передавання інформації і пропускну здатність каналів, що визначена як максимально можлива швидкість передавання.

Швидкість передавання у двійкових каналах зв'язку

$$R = \frac{1}{\tau}\left[H(X) - H\left(X / \hat{X}\right)\right], \tag{3.1}$$

де τ – тривалість одного двійкового символу;

$H(X)$ – ентропія джерела;

$H\left(X / \hat{X}\right)$ – ентропія втрат інформації в каналі, обумовлена помилками, пов'язаними з наявністю завад.

Пропускна здатність каналу зв'язку

$$C = \max R = \max \frac{1}{\tau}\left[H(X) - H\left(X / \hat{X}\right)\right]. \tag{3.2}$$

Максимізації C досягають максимізацією $H(X)$.

Пропускна здатність двійкового симетричного каналу

$$C = \frac{1}{\tau}\left[1 - p_e \log\frac{1}{p_e} - \left(1 - p_e\right)\log\frac{1}{1 - p_e}\right]. \quad (3.3)$$

Пропускну здатність безперервного каналу зв'язку обчислюють за формулою Шеннона [5]

$$C = \Delta F \log\left(\alpha \frac{P_c}{P_{ш}} + 1\right), \quad (3.4)$$

де ΔF – ширина смуги пропускання каналу;

P_c – середня потужність сигналу;

$P_{ш}$ – середня потужність шуму;

α – коефіцієнт форми сигналу; $\alpha = 1$ для сигналу у вигляді нормального білого шуму; $\alpha = 0,3$ для синусоїдального сигналу.

Важливою характеристикою систем передавання інформації є ефективність використання пропускної здатності:

$$\beta_c = \frac{R}{C}, \quad (3.5)$$

де R – реальна швидкість передавання інформації.

3.2 Приклади розв'язання основних типів задач

Приклад 3.1. Повідомлення дискретного джерела кодують рівномірним двійковим кодом і передають симетричним каналом зв'язку з завадами. Визначити пропускну здатність каналу зв'язку за умови, що тривалість двійкових сигналів $\tau = 1$ мкс, середня ймовірність помилки на один двійковий символ $p_e = 10^{-3}$.

Розв'язання. Пропускна здатність двійкового симетричного каналу

$$C = \frac{1}{\tau}\left[1 - p_e \log\frac{1}{p_e} - (1 - p_e)\log\frac{1}{1 - p_e}\right] =$$

$$= \frac{1}{10^{-6}}\left[1 - 10^{-3}\log\frac{1}{10^{-3}} - \left(1 - 10^{-3}\right)\log\frac{1}{10^{-3}}\right] \approx 9{,}9 \cdot 10^{5} \text{ біт/с}.$$

Приклад 3.2. Повідомлення дискретного джерела кодують m-значним кодом і передають симетричним каналом зв'язку за умови, що тривалість одного сигналу $\tau = 2$ мкс, середня ймовірність помилки на один сигнал $p_e = 10^{-3}$, $m = 8$.

Розв'язання. Пропускна здатність багатопозиційного симетричного каналу

$$C = \frac{1}{\tau}\left[\log m - p_e \log\frac{m-1}{p_e} - (1 - p_e)\log\frac{1}{1 - p_e}\right] =$$

$$= \frac{1}{2 \cdot 10^{-6}}\left[\log 8 - 10^{-3}\log\frac{8-1}{10^{-3}} - \left(1 - 10^{-3}\right)\log\frac{1}{1 - 10^{-3}}\right] =$$

$$= 1{,}4936 \cdot 10^{6} \text{ біт/с.}$$

Приклад 3.3. Обчислити пропускну здатність неперервного каналу радіозв'язку, якщо середня потужність сигналу на вході радіоприймача $P_c = 1$ мкВт, а завадою є тепловий шум приймального пристрою зі смугою $\Delta F = 10$ кГц. Приймач працює за температури 20° С.

Розв'язання. Пропускна здатність неперервного каналу зв'язку, за формулою Шеннона,

$$C = \Delta F \log\left(\frac{P_c}{P_ш} + 1\right).$$

Потужність теплового шуму може бути визначена за формулою $P_ш = 4kTF$, де T – абсолютна температура приймального пристрою; k – стала Больцмана, яка дорівнює $1{,}38 \cdot 10^{-23}$ Вт/°К·Гц. У цьому випадку $\Delta F = 10$ кГц, $T = 273 + t°C = 293°$К, тобто

$$C = 10^4 \log\left(1 + \frac{10^{-6}}{4 \cdot 1{,}38 \cdot 10^{-23} \cdot 293 \cdot 10^4}\right) \approx 3{,}26 \cdot 10^5 \text{ біт/с.}$$

Література [5, 6].

3.3 Задачі для самостійного розв’язання

Задача 3.1. Повідомлення дискретного джерела кодують рівномірним двійковим кодом і передають симетричним каналом зв’язку з завадами. Визначити пропускну здатність каналу зв’язку за умови, що тривалість двійкових символів τ, середня ймовірність помилки p_e (таблиця 3.1).

Таблиця 3.1 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
τ, мкс	1	2	3	4	5	6	7	8	9	10
p_e	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-7}	10^{-3}	10^{-4}	10^{-5}	10^{-6}

Задача 3.2. Обчислити пропускну здатність безперервного радіоканалу, якщо задана середня потужність сигналу P_c на виході каналу,

а завадою є внутрішній тепловий шум радіоприймального пристрою з ефективною смугою пропускання ΔF. Приймач працює за температури $T,°C$. Коефіцієнти форми сигналу $\alpha_1 = 1, \alpha_2 = 0,3$ (таблиця 3.2).

Таблиця 3.2 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
P_c, мкВт	1,5	2	2,5	3	3,5	4	4,5	5	5,5	1
ΔF, кГц	12	13	14	15	16	20	10	25	30	15
$T,°C$	20	25	30	35	20	25	30	35	20	30

Задача 3.3. Визначити необхідне відношення сигнал/шум на вході приймача, необхідне для забезпечення заданої пропускної здатності каналу зв’язку C за заданої ширини смуги пропускання ΔF (таблиця 3.3).

Таблиця 3.3 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
C, кбіт/с	30	40	50	60	70	80	90	100	110	120
ΔF, кГц	2	3	4	5	6	7	8	9	10	11

Задача 3.4. Обчислити збільшення пропускної здатності каналу зв’язку, якщо замість синусоїдального сигналу використовувати сигнал типу нормального білого шуму. Задано смугу пропускання каналу ΔF і відношення середніх потужностей сигналу і шуму $P_c / P_ш$ (таблиця 3.4).

Таблиця 3.4 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
ΔF, кГц	2	3	4	5	6	7	8	9	10	11
$P_c / P_ш$, дБ	20	21	22	23	24	25	26	27	28	29

Контрольні питання

1 Перелічіть моделі дискретного каналу зв’язку.

2 Перелічіть моделі безперервного каналу зв’язку.

3 Дайте визначення пропускної здатності дискретного каналу. Теорема Шеннона для каналу з завадами.

4 Дайте визначення пропускної здатності безперервного каналу. Теорема Шеннона для каналу з завадами.

5 Як і навіщо визначають ефективність використання пропускної здатності?

Завдання 4. ЗАВАДОСТІЙКЕ КОДУВАННЯ

4.1 Короткі теоретичні відомості

Для забезпечення заданої завадостійкості систем зв'язку застосовують завадостійке кодування, що економічно значно ефективніше, ніж підвищення енергетичного потенціалу систем. Завадостійкі коди поділяються на коди, які виявляють помилки, і коди, які виявляють і виправляють помилки. Кодові комбінації складаються з інформаційних символів a і контрольних (перевірочних) символів b. Кількість інформаційних символів у кодовій комбінації блочного коду $k = \log_2 M$, де M — кількість повідомлень алфавіту джерела, а кількість контрольних символів r, яка визначається, виходячи з кількості g помилок у кодовій комбінації, що необхідно виявляти (виправляти).

Отже, довжина кодової комбінації завадостійкого блочного коду (кількість розрядів у кодовій комбінації) $n = k + r$. Для завадостійких блочних кодів, які виправляють помилки, кількість контрольних символів у кодовій комбінації визначається з нерівності:

$$2^r - 1 \geq \sum_{i=1}^{g} C_n^i \, ; \; C_n^i = \frac{n!}{i!(n-1)!}, \qquad (4.1)$$

де g – максимальна кратність помилок, яка має виправлятись.

Якщо кількість помилок у кодовій комбінації перевищує g, то частина кодових комбінацій не буде виправлена. Імовірність цього визначається формулою:

$$p_{пом} = 1 - \left(1 - p_e\right)^n - \sum_{i=1}^{g} C_n^i p_e^i \left(1 - p_e\right)^{n-i}. \qquad (4.2)$$

У лінійних систематичних кодах контрольні символи формуються з інформаційних за правилом:

$$b_j = \sum_{i=1}^{k} \oplus \alpha_{ji} \alpha_i ,$$

де α_{ji} – коефіцієнти, які мають значення 1 або 0 залежно від виду коду.

При декодуванні з прийнятих інформаційних символів $\hat{a}$ за таким самим правилом, що й у кодері, формується допоміжна сукупність контрольних символів яка використовується для виявлення помилок шляхом порівняння їх з прийнятими контрольними символами $\hat{b}_j$:

$$\hat{b}_j \oplus b_j^* = S_j . \qquad (4.3)$$

Сукупність двійкових символів $S_j \left(j = \overline{1, r} \right)$ складає контрольне число $S = S_1 S_2 ... S_j ... S_r$ – синдром.

Якщо контрольне число (синдром) дорівнює нулю $S = 0$, то це означає, що в прийнятій кодовій комбінації помилки відсутні або не виявляються, коли їх кількість перевищує g. Якщо $S \neq 0$, то це означає, що кодова комбінація прийнята помилково, і для її виправлення необхідно ідентифікувати контрольне число з можливим конкретним типом помилки. Для цього в різних кодах існують свої правила.

На цей час найбільш ефективними є завадостійкі циклічні коди, які забезпечують досить просту апаратну реалізацію, оскільки всі операції кодування і декодування, виявлення і виправлення помилок здійснюються в регістрах зсуву шляхом циклічних перестановок символів кодових комбінацій.

Математична теорія циклічних кодів базується на поданні кодових комбінацій у вигляді двійкових поліномів:

$$G(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + ... + a_2x^2 + a_1x + a_0x^0, \qquad (4.4)$$

де коефіцієнти a приймають значення «1» або «0» відповідно до символів кодових комбінацій. Над двійковими поліномами виконують операції підсумовування, віднімання, множення та ділення.

Операція циклічного зсуву – це множення полінома $G(x)$ на x у відповідному степені та приведення результату до стандартної форми. Якщо вихідна кодова комбінація джерела $G(x)$ складається з k розрядів, то процес формування кодової комбінації $F(x)$ циклічного коду, що виправляє одиничні помилки, має вигляд [2]:

$$\left[G(x)\cdot x^r\right] : P(x) = Q(x) + R(x); \qquad (4.5)$$

$$F(x) = G(x)\cdot x^r \oplus R(x), \qquad (4.6)$$

де $P(x)$ – породжуючий поліном степеня r;

$R(x)$ – залишок.

Породжуючі поліноми вибирають із таблиці, фрагмент якої наведений в таблиці 4.1.

Таблиця 4.1 – Фрагмент таблиці породжуючих поліномів

Код	Поліном	Код	Поліном
11	$x+1$	10001	x^4+1
101	x^2+1	10011	x^4+x+1
111	x^2+x+1	10101	x^4+x^2+1
1001	x^3+1	10111	x^4+x^2+x+1
1011	x^3+x+1	11001	x^4+x^3+1
1101	x^3+x^2+1	11011	x^4+x^3+x+1
1111	x^3+x^2+x+1	11111	$x^4+x^3+x^2+x+1$

На приймальному кінці лінії зв'язку в декодері виконуються операції перевірки прийнятих комбінацій $\hat{F}(x)$ за правилом:

$$\hat{F}(x)/P(x)=Q^*(x)+R^*(x). \tag{4.7}$$

Якщо залишок $R^*(x)=0$, то це означає, що кодова комбінація прийнята правильно або помилка не виявляється, коли число спотворених символів більше $g=1$.

При $R^*(x)\neq 0$ необхідно виправляти помилку, визначивши номер спотвореної позиції в кодовій комбінації.

4.2 Приклади розв'язання основних типів задач

Приклад 4.1. Для підвищення завадостійкості радіосистеми передавання дискретної інформації застосоване завадостійке блочне кодування кодовими комбінаціями довжиною n, які забезпечують виправлення g незалежних помилок у кожній комбінації за заданої імовірності помилкового приймання одного розряду p_e. Обчислити ймовірність помилкового приймання кодованої комбінації при $n = 11$, $g = 1$, $p_e = 10^{-4}$.

Розв'язання. Імовірність помилкового приймання кодованої комбінації обчислюється за формулою:

$$p_{пом} = 1 - p_{прав} - p_{некор.п} = 1 - \left(1 - p_e\right)^n - \sum_{i=1}^{g} C_n^i p_e^i \left(1 - p_e\right)^{n-i}, \quad (4.8)$$

де $C_n^i = \dfrac{n!}{i!\left(n-1\right)!}$.

При $g = 1$ імовірність виникнення некоректованих помилок дорівнює:

$$p_{некор.п} = 1 - \left(1 - p_e\right)^n - np_e\left(1 - p_e\right)^{n-1}. \quad (4.9)$$

Оскільки $p_e << 1$, то

$$p_{пом} =\approx 1 - 1 + np_e - np_e + n(n-1)\cdot p_e^2 = n(n-1)\cdot p_e^2 =$$

$$= 11\cdot 10\cdot\left(10^{-4}\right)^2 = 1{,}1\cdot 10^{-6}.$$

Приклад 4.2. Для підвищення завадостійкості радіосистеми передавання дискретної інформації необхідно застосувати завадостійке блочне кодування з виправленням незалежних помилок від 1 до g у кодовій комбінації. Розмір алфавіту джерела повідомлень *М*. Обчислити кількість інформаційних k і контрольних символів у кодовій комбінації, якщо $M = 16$; $g = 1,2$.

Розв'язання. Кількість інформаційних символів обчислюється за формулою:

$$k = \log M = \log 16 = 4.$$

Кількість контрольних символів визначається з нерівності:

$$2^r - 1 \geq \sum_{i=1}^{g} C_{k+r}^{i}$$

$$2^r - 1 \geq 4 + r + \frac{(4+r)!}{2!(4+r-2)!}$$

$$2^r - 1 \geq 4 + r + \frac{(4+r-1)(4+r)}{2},\ r = 6.$$

Приклад 4.3. У радіосистемі передавання дискретної інформації застосоване завадостійке кодування з виправленням однократних помилок. Система працює з джерелом *M* повідомлень, задане максимальне значення ймовірності помилкового приймання одного повідомлення $p_{пом}$. Визначити допустиме значення ймовірності помилкового приймання одного символу p_e, якщо *M* = 64, $p_{пом} \leq 10^{-6}$.

Розв'язання. Кількість інформаційних символів k у кодовій комбінації $k = \log 64 = 6$. Кількість контрольних символів r визначається з нерівності $2^r - 1 \geq 6 + r$, $r = 4$.

Допустима ймовірність помилкового приймання одного символу визначається з рівняння:

$$p_{пом} = 1 - (1 - p_e)^{k+r} - (k+r)\, p_e (1 - p_e)^{k+r-1}.$$

Оскільки $p_e << 1$, аналогічно прикладу 5.1:

$$p_{пом} \approx (k+r)(k+r-1)\,p_e^2 .$$

Звідси

$$p_e = \sqrt{\frac{p_{пом}}{(k+r)(k+r-1)}} = \sqrt{\frac{10^{-6}}{10 \cdot 9}} = 1{,}05 \cdot 10^{-4} .$$

Приклад 4.4. Радіосистема призначена для передавання дискретних повідомлень алфавіту *M*. Визначити значення допустимих імовірностей помилок на один символ при роботі натуральним двійковим кодом і завадостійким блочним кодом, який виправляє однократні помилки, якщо задана ймовірність помилкового приймання кодової комбінації $p_{пом}$. Обчислити еквівалентний виграш у завадостійкості в перерахунку на один символ за таких числових значень: M = 128, $p_{пом} = 10^{-7}$.

Розв'язання. Кількість інформаційних символів $k = \log M = \log 128 = 7$. Кількість контрольних символів $r = 4$ (з нерівності $2^r - 1 \geq 7 + r$).

При натуральному двійковому кодуванні:

$$p_{пом} = 1 - (1 - p_e)^k \approx k p_e . \qquad (4.10)$$

Звідси

$$p_{e1} = \frac{p_{пом}}{k} = \frac{10^{-7}}{7} = 1{,}4 \cdot 10^{-8} .$$

Користуючись формулою, що одержана у процесі розв’язання прикладу 4.3, визначаємо:

$$p_{e2} = \sqrt{\frac{p_{пом}}{(k+r)(k+r-1)}} = \sqrt{\frac{10^{-7}}{11 \cdot 10}} = 3 \cdot 10^{-5}.$$

Еквівалентний виграш у завадостійкості дорівнює:

$$B = \frac{p_{e2}}{p_{e1}} = \frac{3 \cdot 10^{-5}}{1,4 \cdot 10^{-8}} = 2,16 \cdot 10^{3}.$$

Приклад 4.5. У радіосистемі передавання дискретних повідомлень використовується завадостійке кодування кодом Хеммінга. Закодувати повідомлення з номером *N*, якщо задана загальна кількість елементарних повідомлень *M*; відомо, що *N*= 3, *M* = 100.

Розв'язання. Визначимо кількість інформаційних і контрольних символів у кодовій комбінації:

$$k = \log M = \log 100 \approx 7,$$
$$2^{r} - 1 \geq 7 + r, \quad r = 4.$$

Структура комбінації коду Хеммінга є такою:

$$b_1 b_2 a_3 b_4 a_5 a_6 a_7 b_8 a_9 a_{10} a_{11},$$

де a – інформаційні, b – контрольні символи.

Інформаційні символи визначаються переведенням номера повідомлення з десяткового коду в двійковий, тобто:

$$a_3 = 0,\ a_5 = 0,\ a_6 = 0,\ a_7 = 0,\ a_9 = 0,\ a_{10} = 1,\ a_{11} = 1.$$

Контрольні символи визначаються зі співвідношень:

$$\begin{aligned}
&b_1 = a_3 \oplus a_5 \oplus a_7 \oplus a_9 \oplus a_{11} = 0 \oplus 0 \oplus 0 \oplus 0 \oplus 1 = 1,\\
&b_2 = a_3 \oplus a_6 \oplus a_7 \oplus a_{10} \oplus a_{11} = 0 \oplus 0 \oplus 0 \oplus 1 \oplus 1 = 0,\\
&b_4 = a_5 \oplus a_6 \oplus a_7 = 0 \oplus 0 \oplus 0 = 0,\\
&b_8 = a_9 \oplus a_{10} \oplus a_{11} = 0 \oplus 1 \oplus 1 = 0.
\end{aligned}$$

Шукана кодова комбінація є такою: 10000000011.

Приклад 4.6. Дискретна система зв’язку використовує завадостійке кодування повідомлень (кодом Хеммінга). Передана кодова комбінація з умов прикладу 4.5. Система прийняла кодову комбінацію 10001000011. Перевірте, чи правильно прийнята комбінація і за необхідності виправте її.

Розв'язання. Прийнята кодова комбінація має вигляд:

$$\hat{b}_1\hat{b}_2\hat{a}_3\hat{b}_4\hat{a}_5\hat{a}_6\hat{a}_7\hat{b}_8\hat{a}_9\hat{a}_{10}\hat{a}_{11}.$$

Для перевірки прийнятої кодової комбінації необхідно визначити контрольне число $S = s_4 s_3 s_2 s_1$, елементи якого обчислюються зі співвідношень:

$$\begin{aligned}
&s_1 = \hat{b}_1 \oplus \hat{a}_3 \oplus \hat{a}_5 \oplus \hat{a}_7 \oplus \hat{a}_9 \oplus \hat{a}_{11} = 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 \oplus 1 = 1,\\
&s_2 = \hat{b}_2 \oplus \hat{a}_3 \oplus \hat{a}_6 \oplus \hat{a}_7 \oplus \hat{a}_{10} \oplus \hat{a}_{11} = 0 \oplus 0 \oplus 0 \oplus 0 \oplus 1 \oplus 1 = 0,\\
&s_3 = \hat{b}_4 \oplus \hat{a}_5 \oplus \hat{a}_6 \oplus \hat{a}_7 = 0 \oplus 1 \oplus 0 \oplus 0 = 1,\\
&s_4 = \hat{b}_8 \oplus \hat{a}_9 \oplus \hat{a}_{10} \oplus \hat{a}_{11} = 0 \oplus 0 \oplus 1 \oplus 1 = 0.\\
&S = 0101.
\end{aligned}$$

Відмінність контрольного числа від нуля свідчить про наявність помилки в кодовій комбінації. Якщо помилка одна, вона може бути виправлена інвертуванням, оскільки контрольне число, переведене з двійкової системи в десяткову, вказує номер спотвореної позиції, тобто:

$$S = 0101 \Rightarrow 5.$$

Приклад 4.7. У радіосистемі передавання дискретних повідомлень застосовується завадостійке кодування циклічним кодом, здатним виправляти одиничні помилки. Загальна кількість елементарних повідомлень, що передаються, дорівнює *M*. Визначити кодову комбінацію для повідомлення за номером 10, якщо *M* = 16.

Розв'язання. Визначимо кількість інформаційних і контрольних символів у кодовій комбінації:

$$k = \log M = \log 16 = 4, \quad 2^r - 1 \geq 4 + r, \quad r = 3.$$

Кодова комбінація для повідомлення за номером 10 має вигляд 1010 або подана у вигляді двійкового полінома:

$$G(x) = x^3 + x\,.$$

Для r = 3 обираємо згідно з таблицею у Додатку Б породжуючий поліном:

$$P(x) = x^3 + x^2 + 1.$$

Визначаємо комбінацію циклічного коду за формулою:

$$F(x) = G(x) \cdot x^r \oplus R(x),$$

де $R(x)$ – залишок від ділення добутку $G(x)\cdot x^r$ на породжуючий поліном $P(x)$.

$$G(x)\cdot x^r = \left(x^3 + x\right)\cdot x^3 = x^6 + x^4,$$

$$\begin{array}{l|l}
\quad x^6 + x^4 & x^3 + x^2 + 1 \\ \cline{2-2}
+ & \quad x^3 + x^2 + 1 \\
\quad x^6 + x^5 + x^3 & \\ \cline{1-1}
\qquad x^5 + x^4 + x^3 & \\
\quad + & \\
\qquad x^5 + x^4 + x^2 & \\ \cline{1-1}
\qquad\quad x^3 + x^2 & \\
\qquad + & \\
\qquad\quad x^3 + x^2 + 1 & \\ \cline{1-1}
\qquad\qquad R(x) = 1 &
\end{array}$$

$$F(x) = x^6 + x^4 + 1, \text{ або } 1010001.$$

Приклад 4.8. Цифрова система радіозв'язку використовує завадостійке кодування циклічним кодом, що виправляє одиничні помилки. Довжина кодової комбінації за умов $n = 7$. Породжуючий поліном $P(x) = x^3 + x^2 + 1$. Прийнята кодова комбінація 1110001.

Перевірте, чи правильно прийнята комбінація і за необхідності виправте її.

Розв'язання. Запишемо прийняту кодову комбінацію у вигляді двійкового полінома:

$$\hat{F}(x) = x^6 + x^5 + x^4 + 1.$$

Перевіримо комбінацію шляхом ділення її на породжуючий поліном $\hat{F}(x)/P(x)$:

$$
\begin{array}{l|l}
\;x^6 + x^5 + x^4 + 1 & x^3 + x^2 + 1 \\ \cline{2-2}
+ & x^3 + x \\
\quad x^6 + x^5 + x^3 & \\ \cline{1-1}
\qquad x^4 + x^3 + 1 & \\
\quad + & \\
\qquad x^4 + x^3 + x & \\ \cline{1-1}
\qquad R^*(x) = x + 1 &
\end{array}
.
$$

Наявність ненульового залишку $R^*(x)$ свідчить про помилку в прийнятій кодовій комбінації. Для виправлення помилки необхідно знайти вектор помилки $E(x) = x^i$, $i = \overline{0, n-1}$ і виконати операцію $\hat{F}(x) \oplus E(x) = F(x)$, де $F(x)$ — передана комбінація. Запропоновані різні варіанти розв'язання цієї задачі. Розглянемо найпростіший. Знайдемо еталонний залишок, який відповідає помилці в першому (старшому) розряді кодової комбінації. Для цього поділимо вектор помилки в першому розряді на породжуючий поліном $E(x)/P(x)$:

$$
\begin{array}{l|l}
\;x^6 & x^3 + x^2 + 1 \\ \cline{2-2}
+ & x^3 + x^2 + x \\
\quad x^6 + x^5 + x^3 & \\ \cline{1-1}
\qquad x^5 + x^3 & \\
\quad + & \\
\qquad x^5 + x^4 + x^2 & \\ \cline{1-1}
\qquad\quad x^4 + x^3 + x^2 & \\
\qquad + & \\
\qquad\quad x^4 + x^3 + x & \\ \cline{1-1}
\qquad\quad R_{em}(x) = x^2 + x &
\end{array}
.
$$

Зафіксуємо значення еталонного залишку.

Залишок, одержаний з діленням прийнятої кодової комбінації на породжуючий поліном, не збігається з еталонним залишком:

$$R_i^*\left(x\right) \oplus R_{em}\left(x\right) = \left(x+1\right) \oplus \left(x^2+x\right) = x^2+1.$$

Отже, помилка не в першому розряді.

Здійснимо циклічний зсув прийнятої кодової комбінації на один розряд справа наліво і знову поділимо на породжуючий поліном $\left[\hat{F}\left(x\right)\cdot x\right]/P\left(x\right)$:

$$\hat{F}\left(x\right)\cdot x = \left(x^6+x^5+x^4+1\right)\cdot x = x^6+x^5+x+1.$$

$$
\begin{array}{l|l}
\quad x^6+x^5+x+1 & x^3+x^2+1 \\ \cline{2-2}
+ & \quad x^3+1 \\
\quad x^6+x^5+x^3 & \\ \cline{1-1}
\qquad\quad x^3+x+1 & \\
\qquad + & \\
\qquad\quad x^3+x^2+1 & \\ \cline{1-1}
\qquad R_2^*\left(x\right)=x^2+x &
\end{array}
.
$$

$$R_2^*\left(x\right) \oplus R_{em}\left(x\right) = \left(x^2+x\right) \oplus \left(x^2+x\right) = 0.$$

Рівність еталонного та одержаного залишків свідчить про помилку в другому розряді. Якщо рівність не досягнута, описаний процес циклічного зсуву і ділення необхідно продовжувати до досягнення результату $R_i^*\left(x\right) \oplus R_{em}\left(x\right) = 0$. Номер спотвореного розряду визначають як кількість кроків зсуву плюс одиниця.

Література [2, 6].

4.3 Задачі для самостійного розв’язання

Задача 4.1. Для підвищення завадостійкості системи передавання дискретної інформації застосовано завадостійке блочне кодування з виправленням одиничних помилок у кожній кодовій комбінації. Обчислити ймовірність помилкового приймання кодових комбінацій, якщо задані довжина кодової комбінації n та ймовірність помилкового приймання одного символу p_e (таблиця 4.2).

Таблиця 4.2 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
n	7	8	9	10	11	12	13	14	15	16
p_e	10^{-7}	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{-6}

Задача 4.2. У системі передавання дискретних повідомлень використано завадостійке кодування (кодом Хеммінга). Закодувати повідомлення з порядковим номером N, якщо задана загальна кількість елементарних повідомлень M (таблиця 4.3).

Таблиця 4.3 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
M	50	60	70	80	90	100	110	120	130	140
N	5	10	15	20	25	30	35	40	45	50

Задача 4.3. У системі передавання дискретних повідомлень використано завадостійке кодування (кодом Хеммінга). Закодувати повідомлення з порядковим номером N, якщо задана загальна кількість

елементарних повідомлень M (таблиця 4.4). Внести помилку в i-й розряд і показати процес виправлення комбінації.

Таблиця 4.4 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
M	50	60	70	80	90	100	110	120	130	140
N	5	10	15	20	25	30	35	40	45	50
i	1	2	3	4	5	6	7	8	9	10

Задача 4.4. У системі передавання дискретних повідомлень використано завадостійке кодування циклічним кодом, що виправляє одиничні помилки. Закодувати повідомлення з порядковим номером N, якщо задана загальна кількість елементарних повідомлень M (таблиця 4.5).

Таблиця 4.5 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
M	16	16	16	16	16	16	16	16	16	16
N	6	15	14	13	12	11	10	9	8	7

Задача 4.5. У системі передавання дискретних повідомлень використано завадостійке кодування циклічним кодом, що виправляє одиничні помилки. Закодувати повідомлення з порядковим номером N, якщо задана загальна кількість елементарних повідомлень M (таблиця 4.6). Внести помилку в i-й розряд і показати процес виправлення комбінації.

Таблиця 4.6 – Вихідні дані

Вихідні дані	Варіант									
	1	2	3	4	5	6	7	8	9	10
M	16	16	16	16	16	16	16	16	16	16
N	1	2	3	4	5	6	7	8	9	10
i	4	3	2	4	3	2	4	3	2	4

Контрольні питання

1 Поясніть відмінність між рівномірним і нерівномірним кодуванням.

2 Що таке надмірність завадостійкого коду?

3 Що таке відносна швидкість завадостійкого коду?

4 Що таке відстань, за Хеммінгом, і вага кодової комбінації?

5 Що таке мінімальна відстань коду?

6 Як пов'язані мінімальна відстань коду, кількість виправлених і кількість виявлених помилок?

СПИСОК ЛІТЕРАТУРИ

1 Батаєв О. П., Корольова Н. А., Ковтун І. В. Теорія електричного зв'язку: навч. посіб. Харків: УкрДАЗТ, 2010. 630 с.

2 Основи теорії телекомунікацій: підручник /за заг. ред. М. Ю. Ільченка. Київ: ІССЗІ НТУУ «КПІ», 2010. 786 с.

3 Беркман Л. Н., Варфоломеєва О. Г., Коршун Н. В., Макаренко А. О. Сигнали в системах телекомунікацій та методи їх обробки: навч. посіб. Київ: ДУТ ННІТІ, 2017. 92 с.

4 Приходько С. І., Трубчанінова К. А., Батаєв О. П. Основи теорії інформації та кодування: навч. посіб. для техн. спец. ВНЗ. Харків: УкрДУЗТ, 2017. 110 с.

5 Основи теорії інформації та кодування: навч. посіб. / Л. С. Сорока, О. В. Сєвєрінов, О. С. Жученко та ін. Харків: ХНУ імені В. Н. Каразіна, 2008. 264 с.

6 Заполовський М. Й., Порошин С. М., Мезенцев М. В. Теорія інформації та кодування: навч. посіб. Харків : НТУ «ХПІ», 2020. 257 с.

ДОДАТОК А

Поліноми неприведені

Степінь твірного полінома r	Вид полінома неприведеного *P(x)*	Значність коду $n = l = 2^r - 1$
1	$x+1$	1
2	x^2+x+1	3
3	x^3+x^2+1 x^3+x+1	7
4	x^4+x+1 x^4+x^3+1 $x^4+x^3+x^2+x+1$	15
5	x^5+x^2+1 x^5+x^3+1 $x^5+x^4+x^2+1$ $x^5+x^3+x^2+x+1$ $x^5+x^4+x^2+x+1$ $x^5+x^4+x^3+x+1$ $x^5+x^4+x^3+x^2+1$	31
6	x^6+x+1 x^6+x^3+1 x^6+x^5+1 $x^6+x^4+x^2+x+1$ $x^6+x^4+x^3+x+1$ $x^6+x^5+x^2+x+1$ $x^6+x^5+x^3+x^2+1$ $x^6+x^5+x^4+x+1$ $x^6+x^5+x^4+x^2+1$	63
7	x^7+x+1 x^7+x^6+1	127
8	$x^8+x^4+x^3+x^2+1$ $x^8+x^6+x^5+x^4+1$	255
9	x^9+x^4+1 x^9+x^5+1	511
10	$x^{10}+x^3+1$ $x^{10}+x^7+1$	1023

ДОДАТОК Б

Таблиця значень допоміжної функції $H\left(p\right) = -p\log_2 p$

p	*H(p)*	*p*	*H(p)*	*p*	*H(p)*	*p*	*H(p)*
0,00 0,01	0,0000 0,0664	0,26	0,5053	0,51	0,4954	0,76	0,3009
0,02	0,1129	0,27	0,5100	0,52	0,4906	0,77	0,2903
0,03	0,1518	0,28	0,5142	0,53	0,4854	0,78	0,2796
0,04	0,1858	0,29	0,5179	0,54	0,4800	0,79	0,2687
0,05	0,2161	0,30	0,5211	0,55	0,4744	0,80	0,2575
0,06	0,2435	0,31	0,5238	0,56	0,4684	0,81	0,2462
0,07	0,2686	0,32	0,5260	0,57	0,4623	0,82	0,2348
0,08	0,2915	0,33	0,5278	0,58	0,4558	0,83	0,2231
0,09	0,3127	0,34	0,5292	0,59	0,4491	0,84	0,2113
0,10	0,3322	0,35	0,5301	0,60	0,4422	0,85	0,1993
0,11	0,3503	0,36	0,5306	0,61	0,4350	0,86	0,1871
0,12	0,3671	0,37	0,5307	0,62	0,4276	0,87	0,1748
0,13	0,3826	0,38	0,5305	0,63	0,4199	0,88	0,1623
0,14	0,3971	0,39	0,5298	0,64	0,4121	0,89	0,1496
0,15	0,4105	0,40	0,5288	0,65	0,4040	0,90	0,1368
0,16	0,4230	0,41	0,5274	0,66	0,3956	0,91	0,1238
0,17	0,4346	0,42	0,5256	0,67	0,3871	0,92	0,1107
0,18	0,4453	0,43	0,5236	0,68	0,3783	0,93	0,0974
0,19	0,4552	0,44	0,5211	0,69	0,3694	0,94	0,0839
0,20	0,4644	0,45	0,5184	0,70	0,3602	0,95	0,0703
0,21	0,4728	0,46	0,5153	0,71	0,3508	0,96	0,0565
0,22	0,4806	0,47	0,5120	0,72	0,3412	0,97	0,0426
0,23	0,4877	0,48	0,5083	0,73	0,3314	0,98	0,0286
0,24	0,4941	0,49	0,5043	0,74	0,3215	0,99	0,0144
0,25	0,5000	0,50	0,5000	0,75	0,3113	1,00	0,0000

ТЕОРІЯ ІНФОРМАЦІЇ ТА КОДУВАННЯ

МЕТОДИЧНІ ВКАЗІВКИ

до виконання практичних занять і самостійних робіт

із дисципліни

«ТЕЛЕКОМУНІКАЦІЙНІ СИСТЕМИ ПЕРЕДАЧІ»

Відповідальний за випуск Трубчанінова К. А.

Редактор Ібрагімова Н. В.

Підписано до друку 16.01.2025 р.
Умовн. друк. арк. 3,75. Тираж . Замовлення № .
Видавець та виготовлювач Український державний університет залізничного транспорту,
61050, Харків-50, майдан Фейєрбаха,7.
Свідоцтво суб’єкта видавничої справи ДК № 6100 від 21.03.2018