УДК 621.391(075.8)
ББК 32.811я73
Ж91

Гриф надано Міністерством освіти і науки України (лист від 21 травня 2001 р. № 1/11-2367)

Рецензенти: засл. діяч науки і техніки України, д-р техн. наук, проф. *В. К. Стеклов* (Київський філіал Української державної академії зв'язку); д-р техн. наук, проф. *А. М. Лучук* (Інститут кібернетики НАН України); д-р техн. наук, проф. *О. М. Романкевич*; канд. техн. наук, доц. *Л. Ф. Карачун* (Національний технічний університет України «Київський політехнічний інститут»)

Редакція літератури з економіки і фундаментальних наук
Редактор *В. Ф. Хміль*

Жураковський Ю. П., Полторак В. П.
Ж91 Теорія інформації та кодування: Підручник. — К.: Вища шк., 2001. — 255 с.: іл.
ISBN 966-642-031-7

Викладено основні поняття та положення теорії інформації, методи математичного опису та інформаційні характеристики дискретних і неперервних джерел повідомлень, визначення кількості інформації та ентропії в повідомленні. Розглянуто основні принципи кодування повідомлень, способи коректувального кодування двійковими та недвійковими кодами, способи ефективного стиснення повідомлень, методику розрахунку завадостійкості кодованих повідомлень. Наведено контрольні задачі, приклади розв'язання типових задач і контрольні запитання, що сприяє кращому засвоєнню матеріалу.

Для студентів вищих технічних навчальних закладів. Може бути корисним для фахівців з телекомунікацій, обчислювальної техніки, автоматизованих систем.

УДК 621.391(075.8)
ББК 32.811я73

ISBN 966-642-031-7

ЗМІСТ

Передмова 6

Розділ 1. Інформація та інформаційні процеси

1.1. Повідомлення та інформація 8
1.2. Моделі інформаційних систем 10
1.3. Математичні моделі каналу зв'язку 13
1.4. Предмет теорії інформації та кодування 19
Контрольні запитання 20

Розділ 2. Кількісні характеристики інформації

2.1. Ансамблі та джерела повідомлень 21
2.2. Кількісна міра інформації 23
2.3. Ентропія та її властивості 28
2.4. Безумовна ентропія 32
2.5. Умовна ентропія 34
2.6. Ентропія об'єднання двох джерел 45
Контрольні задачі 47
Контрольні запитання 48

Розділ 3. Характеристики дискретних джерел інформації

3.1. Продуктивність дискретного джерела та швидкість передачі інформації 49
3.2. Інформаційні втрати при передачі інформації по дискретному каналу 51
3.3. Пропускна здатність дискретного каналу 53
3.4. Теорема Шеннона про кодування дискретного джерела 54
Контрольні задачі 58
Контрольні запитання 59

Розділ 4. Характеристики неперервних джерел інформації

4.1. Квантування сигналів 60
4.2. Інформаційні втрати при кодуванні неперервних джерел 65
4.3. Продуктивність неперервного джерела та швидкість передачі інформації 68
4.4. Пропускна здатність неперервного каналу 69
Контрольні задачі 70
Контрольні запитання 71

Розділ 5. Кодування в дискретних і неперервних каналах

5.1. Класифікація кодів і характеристики їх 72
5.2. Системи числення 76
5.3. Основні операції над елементами поля 78
5.4. Способи подання кодів 82
5.5. Надмірність повідомлень і кодів 89
5.6. Основні теореми кодування для каналів 92
5.7. Оптимальне кодування 99
Контрольні задачі 105
Контрольні запитання 110

Розділ 6. Кодування повідомлень

6.1. Класифікація первинних кодів 112
6.2. Нерівномірні двійкові первинні коди 113
6.2.1. Код Морзе 113
6.2.2. Число-імпульсні коди 114
6.3. Рівномірні двійкові первинні коди 115
6.3.1. Числові двійкові коди 115
6.3.2. Двійково-десяткові коди 122
6.3.3. Двійково-десяткові коди з самодоповненням 124
6.3.4. Двійково-шістнадцятковий код 125
6.3.5. Рефлексні коди 125
6.4. Недвійкові первинні коди 128
Контрольні задачі 130
Контрольні запитання 133

Розділ 7. Коди, що виявляють помилки

7.1. Двійкові коди, що виявляють помилки 134
7.1.1. Код із перевіркою на парність 134
7.1.2. Код із перевіркою на непарність 135
7.1.3. Код із простим повторенням 135
7.1.4. Інверсний код 136
7.1.5. Кореляційний код 136
7.1.6. Код зі сталою вагою 137
7.1.7. Код із кількістю одиниць у комбінації, кратною трьом 137
7.2. Недвійкові коди, що виявляють помилки 138
7.2.1. Код із перевіркою за модулем q 139
7.2.2. Код із повторенням 139
7.2.3. Незвідні змінно-позиційні коди 140
7.3. Штрихові коди 146
Контрольні задачі 153
Контрольні запитання 157

Розділ 8. Коди, що виправляють помилки

8.1. Двійкові групові коди 158
8.1.1. Лінійний систематичний груповий (блоковий) код 158
8.1.2. Коди Хеммінга 166
8.1.3. Циклічні коди 171
8.1.4. Коди Боуза — Чоудхурі — Хоквінгема 180
8.1.5. Код Файра 185
8.1.6. Код із багатократним повторенням 187
8.1.7. Ітеративні коди 188
8.1.8. Каскадні коди 190
8.2. Рекурентні коди 191
8.3. Недвійкові коди 192
8.3.1. Код із багатократним повторенням 193
8.3.2. Узагальнений код Хеммінга 194
8.3.3. Коди Боуза — Чоудхурі — Хоквінгема 196
8.3.4. Коди Ріда — Соломона 198
8.3.5. Багатовимірні ітеративні коди 200
8.3.6. Недвійковий ланцюговий код 202
Контрольні задачі 203
Контрольні запитання 214

Розділ 9. Ефективність кодування та передачі інформації

9.1. Вірогідність передачі кодованих повідомлень 216
9.2. Стиснення інформації 221
9.2.1. Способи стиснення даних при передачі 223
9.2.2. Способи стиснення даних при архівації 235
9.3. Збільшення основи коду 239
9.4. Використання зворотного зв'язку для підвищення ефективності передачі інформації 243
Контрольні задачі 246
Контрольні запитання 250

Додатки 252

Список використаної та рекомендованої літератури 254

З кожним роком зростають обсяги інформації, які потрібно передавати, обробляти та зберігати для своєчасного використання їх при вирішенні виробничих, наукових, економічних та інших завдань.

Деякі теоретичні знання про інформацію, її кількість, способи кодування повідомлень та захисту їх від завад і спотворень, викладені в цьому підручнику, допоможуть користувачам краще зрозуміти процеси перетворення повідомлень при передачі, а також скористатися цими знаннями для захисту інформації від спотворень.

Елементи теорії інформації розглядалися Р. Хартлі в праці, присвяченій вимірюванню кількості інформації (1928 р.), але першим фундаментальним дослідженням, яке, по суті, започаткувало теорію інформації, є праця К. Шеннона «Математична теорія зв'язку» (1948 р.). У ній обґрунтовано методику вимірювання кількості інформації та запропоновано теореми про оптимальне кодування, що доводять можливість досягнення максимальних швидкостей передачі інформації.

Вагомий внесок у розвиток окремих розділів теорії інформації зробили вчені Х. Найквіст, Н. Вінер, Д. Міддлтон, А. Файнстейн, К. Хелстром, В. Котельников, О. Харкевич, А. Колмогоров, О. Хінчин, В. Фано, У. Пітерсон, Р. Боуз, Д. Рой-Чоудхурі, П. Еліас, Е. Берлекемп, Л. Бородін, Л. Фінк, Р. Стратонович, А. Зюко, Ф. Катков, К. Зігангіров.

Без знання основ теорії інформації та кодування неможливе створення нових сучасних систем передачі інформації. Тому її вивчення є невід'ємною частиною теоретичної підготовки фахівців у галузі комп'ютеризованих систем, автоматики та управління, комп'ютерної інженерії, телекомунікацій тощо.

В основу підручника покладено матеріали лекційних і практичних занять з цієї дисципліни, яка викладається авторами в Національному технічному університеті України («Київський політехнічний інститут»). У ньому також використано матеріали, опубліковані в літературних джерелах, та дані деяких науково-дослідних робіт, виконаних авторами.

Структура підручника дає змогу читачеві послідовно від простого до більш складного вивчити всі основні поняття та положення теорії інформації й кодування, хоча й потребує від нього деяких знань з таких розділів математики, як теорія ймовірностей та математична статистика, теорія матриць, комбінаторика.

Набутими знаннями з цієї дисципліни студенти та фахівці можуть скористатися під час вивчення інших дисциплін, тісно пов'язаних з теорією інформації та кодуванням, а також при курсовому та дипломному проектуванні зі споріднених дисциплін.

Розд. 1, 5—8 (крім п. 8.3.4), 9 (крім п. 9.3) та п. 4.1 написано д-ром техн. наук, проф. *Ю. П. Журавковським*, решту матеріалу — канд. техн. наук, доц. *В. П. Полтораком.*

РОЗДІЛ 1 ІНФОРМАЦІЯ ТА ІНФОРМАЦІЙНІ ПРОЦЕСИ

Як відомо, всі процеси, що відбуваються в природі, в тому числі виробничі, пов'язані з інформацією — її здобуттям, перетворенням, накопиченням, передаванням, зберіганням і відображенням.

У цьому розділі наводяться основні поняття та визначення з теорії інформації, описуються моделі інформаційних систем і рух інформації в них, подаються моделі каналів інформації, а також оцінюється вплив характеристик каналів на якість передачі інформації.

1.1. ПОВІДОМЛЕННЯ ТА ІНФОРМАЦІЯ

Насамперед слід визначитися з такими поняттями, як дані, повідомлення та інформація.

Під *даними* розуміють усі відомості, здобуті від навколишнього світу та подані у формалізованому вигляді (літерами, цифрами, символами тощо). Дані, що підлягають передачі, називаються *повідомленнями.* Повідомлення стають інформацією тільки в момент їх застосування, тобто інформація — це використовувані повідомлення, причому такі, які відзначаються новизною і раніше не були відомі одержувачеві (оператору ЕОМ). Є й інше, більш широке визначення інформації, згідно з яким інформація — це відомості, що є об'єктом зберігання, передавання та перетворення [47].

Для можливості технічного оброблення (передачі, запису та ін.) повідомлення має бути перетворене на *сигнал* — матеріальний носій, що відображує повідомлення [47].

Розрізняють сигнали *звукові* (акустичні), *електричні*, *оптичні*, *гідравлічні* та ін. Один вид сигналу можна перетворювати на інший (електричний на звуковий, оптичний на електричний тощо).

Будь-який сигнал характеризується такими основними параметрами: тривалістю, шириною частотного спектра та динамічним діапазоном.

Під *тривалістю* T_c сигналу розуміють час, протягом якого він знаходиться в каналі зв'язку. *Частотний спектр* F_c сигналу визначає смугу частот, яку він охоплював під час передачі по каналу зв'язку. Залежно від виду сигналу (аналоговий, дискретний) частотний спектр може бути і нескінченним; тому на практиці його обмежують для можливості передачі по каналах з обмеженою смугою частот. Так, телефонні розмови ведуться по каналах зі смугою пропускання 3100 Гц (300 ... 3400 Гц), хоча сам початковий сигнал займає спектр до 15 ... 17 кГц.

Середньою потужністю P_c сигналу є потужність, яка забезпечується апаратурою під час його надходження до каналу зв'язку. На практиці частіше замість P_c користуються поняттям *динамічного діапазону* D_c, що визначається логарифмом відношення найбільшої (максимальної) миттєвої потужності сигналу ($P_{c\,max} = P_c$) до найменшої (мінімальної) $P_{c\,min}$, дозволене значення якої дорівнює потужності завад ($P_{c\,min} = P_з$):

$$D_c = \log (P_c/P_з).$$

Усі ці параметри сигналу є його *обсягом*:

$$V_c = T_c F_c D_c.$$

Аналогічними параметрами характеризується також канал зв'язку. Ними є *тривалість використання* $T_к$ каналу, *смуга його частот* $F_к$ та *динамічний діапазон* $D_к$. У цьому разі під $T_к$ розуміють час використання каналу для передачі сигналів, під $F_к$ — смугу частот, яка забезпечується каналом, а під $D_к$ — динамічний діапазон рівнів сигналів, які можуть бути передані ним. Добуток цих трьох параметрів визначає *ємність* каналу зв'язку

$$V_к = T_к F_к D_к.$$

Для забезпечення передачі сигналів по каналу зв'язку необхідно, щоб $V_к > V_c$; крім того, мають виконуватися такі умови:

$$T_к > T_c;\ F_к > F_c;\ D_к > D_c.$$

Якщо деякі з них не виконуються, треба досягти їх за рахунок інших. Так, якщо $V_к > V_c$, $T_к > T_c$ і $D_к > D_c$, але $F_к < F_c$, то, збільшуючи T_c, можна зменшити частотний спектр F_c сигналу і виконати умову $F_к > F_c$.

Повідомлення та відповідні сигнали можуть бути *неперервними* (аналоговими) та *дискретними* (знаковими). Перші описуються неперервною функцією часу. До них належать такі повідомлення, як музика, телевізійне зображення, радіомовлення. За допомогою спеціальних пристроїв неперервні повідомлення перетворю-

ються на неперервні електричні сигнали, якими передаються повідомлення по каналу зв'язку від передавача до приймача.

Дискретними повідомленнями є скінченна послідовність окремих символів (знаків, літер) з обмеженою тривалістю. Вони характерні для телеграфії, передачі даних, телекомунікацій. Для перетворення дискретного повідомлення на сигнал потрібна операція кодування.

Неперервні повідомлення можна передавати дискретними способами. В цьому разі неперервні сигнали, якими передаються ці повідомлення, перетворюються на дискретні за допомогою операцій квантування за рівнем та дискретизації в часі. На приймальному боці виконується обернене перетворення: за прийнятими дискретними сигналами відновлюються передані неперервні сигнали.

Дискретні сигнали як засіб передачі повідомлень більш поширені, ніж неперервні, завдяки тому що вони меншою мірою зазнають впливу завад і спотворень в каналах зв'язку, а в разі спотворення їх легше регенерувати (відновити) і, крім того, вони досить легко обробляються в ЕОМ.

1.2. МОДЕЛІ ІНФОРМАЦІЙНИХ СИСТЕМ

Під *інформаційною* розумітимемо будь-яку систему, яка за допомогою технічних засобів виконує одну або кілька таких функцій, як збирання, передавання, перетворення, накопичення, зберігання та оброблення інформації.

За функціональною ознакою інформаційні системи можна поділити на: системи електрозв'язку; системи передачі даних; інформаційно-вимірювальні системи; системи перетворення інформації; інформаційно-пошукові системи; системи зберігання інформації; автоматизовані системи управління; системи експериментальних досліджень.

Найпоширенішими в повсякденному житті є системи електрозв'язку та передачі даних, які можна об'єднати назвою *систем передачі інформації* (СПІ). Як приклад розглянемо роботу одноканальної системи передачі даних, структурну схему якої для передачі інформації в одному напрямку зображено на рис. 1.1. Тут *ДП* і *ОП* — відповідно джерело і одержувач повідомлення; $П_{вх}$ і $П_{вих}$ — відповідно вхідний та вихідний перетворювачі; *К1*, *К2* — кодери; *ДК1*, *ДК2* — декодери; *М* — модулятор; *ДМ* — демодулятор.

Лінія зв'язку — це фізичне середовище, в якому поширюються сигнали.

Каналом зв'язку називається сукупність технічних засобів, що забезпечує передачу повідомлень від джерела до одержувача

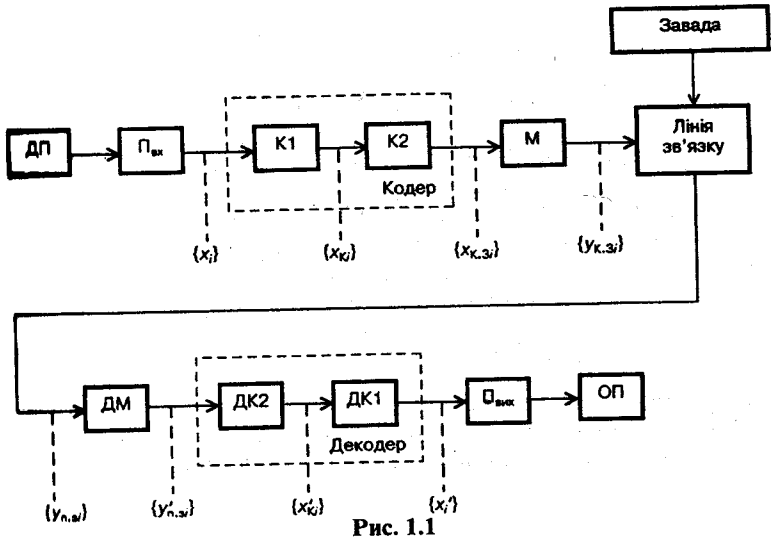


Рис. 1.1

по одній лінії зв'язку незалежно від передачі повідомлення від інших джерел до інших одержувачів.

Сукупність лінії зв'язку, модулятора та демодулятора (пристроїв перетворення сигналів) утворює *неперервний* канал передачі інформації, а якщо до цієї сукупності додати ще кодер і декодер, то дістанемо *дискретний* канал цієї передачі.

Модулятор і демодулятор, конструктивно об'єднані в одному блоці, називаються *модемом*, а конструктивне об'єднання кодера та декодера — *кодеком*.

Нехай джерело повідомлень *ДП* за допомогою вхідного перетворювача $П_{вх}$ створює випадкову дискретну послідовність сигналів $\{x_i\}$, яка подається на кодувальний пристрій кодера *К1*, призначеного для перетворення первинної послідовності сигналів $\{x_i\}$ на послідовність $\{x_{ki}\}$ (див. рис. 1.1). Це перетворення необхідне для більш ефективного використання лінії зв'язку, а також для перетворення дискретної послідовності сигналів одного алфавіту на дискретну послідовність іншого алфавіту.

Далі послідовність сигналів $\{x_{ki}\}$ кодером *К2* перетворюється на дискретну випадкову послідовність сигналів $\{x_{к.зі}\}$. Кодер *К2* виконує функції пристрою захисту інформації від помилок, кодуючи комбінації первинного коду коректувальним кодом, який виявляє та виправляє помилки, що дає змогу зменшити вплив завад і спотворень на інформацію, яка передається. У свою чергу, послідовність сигналів $\{x_{к.зі}\}$ після їх модуляції в модуляторі *М* перетворюється на випадкову послідовність сигналів $\{y_{к.зі}\}$, що однозначно відповідає послідовності $\{x_{к.зі}\}$.

Під *способом передачі сигналів* розуміють сукупність операцій перетворення повідомлення на сигнал [5], яку можна подати так:

$$\{y_{к.зі}(t)\} = \Lambda_{прд}\{x_i\} = \Lambda_м\Lambda_к\{x_i\}, \quad (1.1)$$

де $\Lambda_{прд}$ — оператор способу передачі сигналів; $\Lambda_м$ — оператор їх модуляції; $\Lambda_к$ — оператор кодування сигналів.

Сигнали при передачі по лінії зв'язку загасають, піддаючись дії завад і спотворень, що спричинює значні відхилення послідовності сигналів на вході приймача $\{y_{п.зі}(t)\}$ від переданої в лінію послідовності сигналів $\{y_{к.зі}(t)\}$, тобто

$$\{y_{п.зі}(t)\} = \Lambda_л\{y_{к.зі}(t)\} = \Lambda_л\Lambda_{прд}\{x_i\}, \quad (1.2)$$

де $\Lambda_л$ — оператор лінії зв'язку.

Якщо в лінії є адитивна завада у вигляді випадкового процесу $\omega(t)$, то на вході приймача діятиме неперервний випадковий процес

$$y_i(t) = \{y_{п.зі}(t)\} + \omega(t). \quad (1.3)$$

У приймачі після підсилення сигналів, яке необхідне для компенсації загасання їх у лінії зв'язку, сигнали демодулюються в демодуляторі *ДМ*. На виході останнього утворюється дискретна послідовність сигналів $\{y'_{п.зі}(t)\}$, яка має відповідати послідовності сигналів $\{y_{к.зі}(t)\}$ на виході модулятора *М*. Цього, однак, може й не бути через дію завад і спотворень у лінії та похибки перетворень сигналів у модуляторі та демодуляторі. Після декодування сигналів декодером *ДК2* послідовність $\{y'_{п.зі}\}$ перетворюється на послідовність кодових комбінацій сигналів $\{x'_{кі}\}$, яка має відповідати переданій послідовності сигналів $\{x_{кі}\}$. Ця послідовність залежатиме від властивостей лінії зв'язку, способу приймання сигналів і коректувального коду, що використовується для їх передачі.

Після декодера *ДК1*, який перетворює послідовність сигналів $\{x'_{кі}\}$ на послідовність $\{x'_i\}$, дискретна послідовність сигналів подається (в разі необхідності перетворення дискретних сигналів на неперервну форму) на вихідний перетворювач $П_{вих}$, з виходу якого вона спрямовується до одержувача повідомлення *ОП*.

Сукупність операцій перетворення сигналів на повідомлення називається *способом їх приймання*, який можна відобразити так:

$$\{x'_i\} = \Lambda_{прм}[y_i(t)] = \Lambda_{прм}[\{y_{п.зі}(t)\} + \omega(t)], \quad (1.4)$$

де $\Lambda_{прм} = \Lambda_{дм}\Lambda_{дк}$ — оператор способу приймання сигналів; $\Lambda_{дм}$ — оператор їх демодуляції; $\Lambda_{дк}$ — оператор декодування сигналів.

З урахуванням (1.1) і (1.4) процес передавання дискретної інформації можна подати у вигляді

$$\{x'_i\} = \Lambda_{дм}\Lambda_{дк}[\Lambda_л\Lambda_м\Lambda_к\{x_i\} + \omega(t)]. \quad (1.5)$$

Завдання, яке необхідно вирішити при побудові СПІ, полягає в тому, щоб дістати послідовність сигналів $\{x'_i\}$, яка найменше відрізняється від переданої послідовності сигналів $\{x_i\}$, і забезпечити при цьому високі техніко-економічні показники системи — швидкість передачі інформації, її вірогідність, прийнятну вартість тощо. При побудові СПІ, як правило, задаються ансамблем повідомлень джерела та параметрами лінії (каналу) зв'язку.

Одним з основних завдань при проектуванні СПІ є вибір способу передавання інформації, від якого значною мірою залежатимуть рішення щодо вибору окремих вузлів і блоків системи.

У багатоканальних СПІ на відміну від одноканальної забезпечується одночасна та взаємно незалежна передача по одній загальній лінії повідомлень від багатьох джерел (відправників). Спрощену структурну схему СПІ з частотним поділом сигналів показано на рис. 1.2, де $УП_{прд}$ і $УП_{прм}$ — відповідно ущільнювальні пристрої передавача та приймача. Решта блоків багатоканальної СПІ за призначенням не відрізняються від одноканальної. За допомогою ущільнювальних передавальних і приймальних пристроїв розносять спектри сигналів, які передаються різними каналами, в діапазоні частот, що відводиться в лінії зв'язку для організації *n* каналів передачі інформації. В цілому процес математичного опису такої системи не відрізняється від аналогічного опису одноканальної СПІ.

1.3. МАТЕМАТИЧНІ МОДЕЛІ КАНАЛУ ЗВ'ЯЗКУ

Від вибору каналу зв'язку залежить не тільки кількість інформації, яку можна передати від передавача до одержувача повідомлень, а й швидкість передачі інформації та її вірогідність.

У той же час точний математичний опис будь-якого реального каналу зв'язку досить складний. Тому на практиці, як правило, користуються більш спрощеними математичними моделями, які дають змогу визначити основні закономірності каналів. При побудові таких моделей враховуються найсуттєвіші особливості каналу зв'язку і відкидаються другорядні чинники, що майже не впливають на якість зв'язку.

Розрізняють математичні моделі неперервних і дискретних каналів зв'язку [10, 11, 33]. Розглянемо більш детально першу

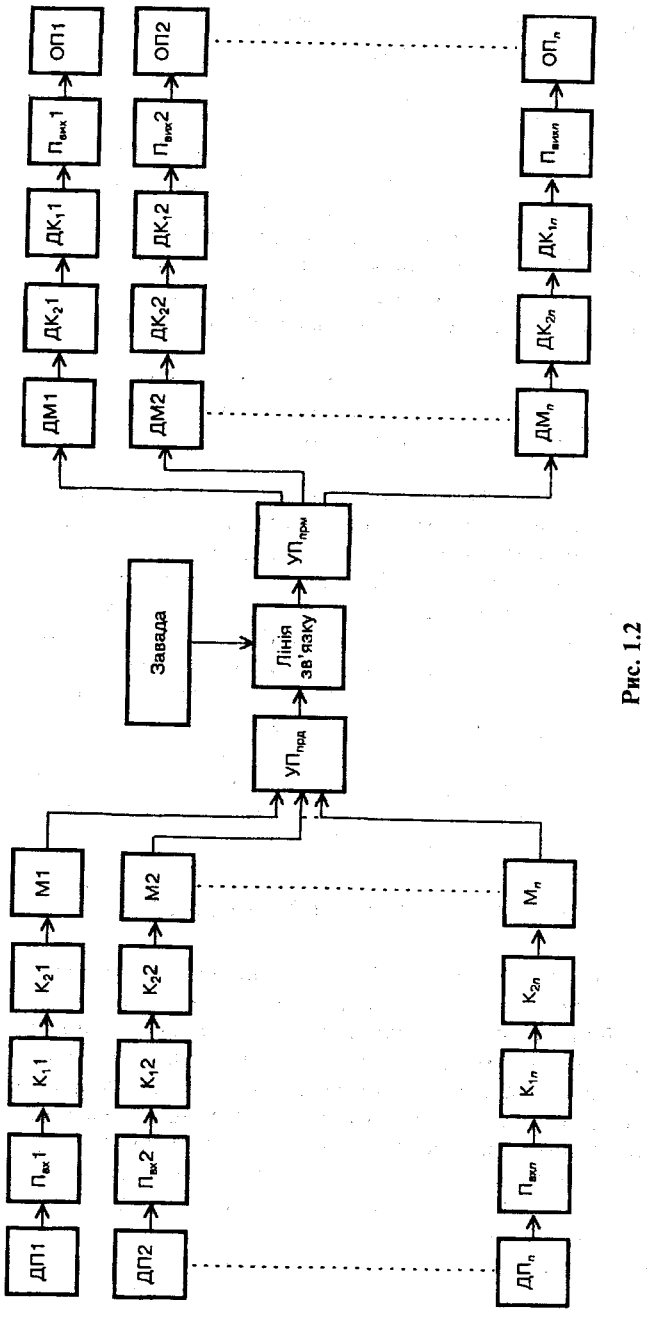


Рис. 1.2

групу цих моделей, зазначивши, що на вхід неперервного каналу можуть надходити будь-які сигнали зі спектром, який лежить у повній смузі частот і характеризується обмеженою середньою або піковою потужністю.

Ідеальний канал зв'язку без завад — це лінійне коло зі сталою функцією передачі, що звичайно зосереджена в обмеженій смузі частот, де завади будь-якого виду відсутні. Вихідний сигнал при заданому вхідному буде детермінованим.

Ця модель каналу зв'язку частіше застосовується для опису кабельних каналів дуже короткої довжини, де наявність адитивних завад ще не відчувається.

Канал зв'язку з адитивним гауссовим шумом — це канал з «білим» або «квазібілим» (з рівномірною спектральною щільністю в смузі спектра сигналу) шумом, у якому коефіцієнт передачі та час затримки сигналів не залежать від часу й є відомими детермінованими величинами. Вихідний сигнал такого каналу визначається виразом

$$z(t) = k_{\text{п}}\, s(t-\tau) + w(t), \qquad (1.6)$$

де $s(t)$ — вхідний сигнал; $k_{\text{п}}$ — коефіцієнт передачі; τ — час затримки сигналу в каналі; $w(t)$ — гауссовий адитивний шум з нульовим математичним сподіванням і заданою кореляційною функцією.

Якщо змінити початок відліку часу на виході каналу, то запізнення τ сигналу можна не враховувати.

Модель (1.6) застосовується для опису реальних проводових каналів зв'язку й однопроменевих радіоканалів без завмирань на метрових хвилях для зв'язку в межах прямої видимості, а також радіоканалів з повільними завмираннями, для яких можна надійно передбачити значення $k_{\text{п}}$ та τ.

Канал з невизначеною фазою сигналу — це канал, який відрізняється від попереднього тим, що в ньому запізнення τ є випадковою величиною. Вихідний сигнал такого каналу визначається виразом

$$z(t) = k_{\text{п}}\,[s(t)\cos\theta_{\text{к}} + \tilde{s}(t)\sin\theta_{\text{к}}] + w(t), \qquad (1.7)$$

де $\tilde{s}(t)$ — перетворення Гільберта від $s(t)$; $\theta_{\text{к}} = \omega_0\tau$ — випадкова початкова фаза, причому розподіл імовірностей $\theta_{\text{к}}$ найчастіше задається рівномірним в інтервалі від 0 до 2π.

Модель (1.7) використовується для опису тих самих каналів зв'язку, що й попередня, якщо фаза сигналу в них змінюється. Ця флуктуація фази пояснюється незначними змінами довжини каналу, фазовою нестабільністю генераторів носійної час-

тоти, а також незначними змінами властивостей середовища, через яке передається сигнал.

Однопроменевий гауссовий канал із загальними завмираннями (флуктуаціями амплітуд і фаз сигналу) також описується виразом (1.7), але тут, крім запізнення τ, випадковими вважаються також коефіцієнт передачі $k_п$ та фаза $\theta_к$, тобто випадковими є квадратурні компоненти

$$X = k_п \cos \theta_к;\ Y = k_п \sin \theta_к.$$

Тоді, якщо врахувати зміну квадратурних компонентів $X(t)$ і $Y(t)$ в часі, вихідний сигнал каналу визначиться виразом

$$z(t) = X(t)\, s(t) + Y(t)\, \tilde{s}(t) + w(t) = \\ = k_п(t)\,[s(t) \cos \theta_к(t) + s(t) \sin \theta_к(t)] + w(t). \qquad (1.8)$$

Розподіл коефіцієнта передачі $k_п(t)$ може бути релеєвським або узагальненим релеєвським [10, 33]. Звідси походять і відповідні назви каналів зв'язку: з *релеєвськими* або з *узагальненими релеєвськими* завмираннями.

Модель (1.8) досить добре описує більшість радіоканалів різних хвильових діапазонів і проводових каналів з випадковими та змінними параметрами.

Гауссовий багатопроменевий канал з завмираннями та адитивним шумом (канал з міжсимвольною інтерференцією та адитивним шумом) є каналом, у якому сигнали від передавача до приймача поширюються кількома шляхами (каналами), причому тривалість проходження їх і коефіцієнти передачі каналів є неоднаковими та випадковими. Крім того, в таких каналах випадкова імпульсна реакція $G(t, \tau)$ від часу t не залежить (або змінюється дуже повільно), так що розсіяння за частотою практично не спостерігається.

Причиною міжсимвольної інтерференції є розсіяння сигналу в часі при проходженні по каналу, що спричинюється головним чином нелінійністю фазочастотної характеристики каналу або поширенням сигналу кількома шляхами. В разі міжсимвольної інтерференції на виході каналу сигнал, який описується виразом (1.8), деформується так, що одночасно присутніми є відгуки каналу на частини вхідного сигналу, які припадають на досить віддалені моменти часу. При передачі дискретних повідомлень це зводиться до того, що під час приймання одного елемента (символу) на вході приймального пристрою діють також відгуки на попередні, а іноді й на більш пізні елементи (символи), які є завадами і призводять до спотворення сигналу, що приймається.

Вихідний сигнал такого каналу визначається виразом

$$z(t) = \sum_{r=-Q}^{D} s_r(t - rT) + w(t) = s_0(t) + g_{м.і}(t) + w(t), \qquad (1.9)$$

де $\sum_{\substack{r=-Q \\ r\neq 0}}^{D} s_r(t - rT) = g_{м.і}(t)$ — сигнал міжсимвольної інтерференції, зумовлений елементами (символами), які передані до та після сигналу, що аналізується; $w(t)$ — адитивний шум у каналі; $s_0(t)$ — сигнал, обумовлений елементом (символом), який аналізується.

Модель (1.9) застосовується для опису радіоканалів з багатопроменевим поширенням сигналів, а також проводових каналів значної довжини, де відчувається вплив фазочастотних спотворень їх.

Моделі дискретного каналу. При побудові моделей дискретних каналів треба враховувати, по-перше, те, що для них вхідними та вихідними сигналами є послідовності кодових елементів (символів), і, по-друге, те, що розподіл умовних імовірностей вихідного сигналу залежить від заданого вхідного розподілу. Тому для визначення вхідних сигналів досить указати кількість q різних символів (алфавіт або основу коду), а також довжину T кожного символу. Якщо значення T вважати однаковими для всіх символів, то величина $v = 1/T$ визначатиме кількість символів, які передаються за одиницю часу. Ця величина називається *технічною швидкістю передачі* (швидкістю модуляції сигналів) [33] і виражається в бодах.

Нехай синхронізація в каналі ідеальна і кожний символ, що подається на його вхід, зумовлює появу одного вихідного символу, тобто технічна швидкість передачі на вході та виході каналу однакова. Отже, всі вхідні та вихідні послідовності (вектори) завдовжки n, кількість яких дорівнює q^n, утворюють кінцевий q^n-вимірний векторний простір.

Для оцінки правильності прийнятої послідовності символів визначають різницю між прийнятою та переданою послідовностями порозрядним відніманням їх за модулем q. При цьому дістають *вектор помилки*. З цього випливає, що проходження дискретного сигналу через канал можна розглядати як додавання вхідного вектора з вектором помилки, тобто останній в дискретному каналі відіграє таку саму роль, як завада в неперервному каналі.

Таким чином, для будь-якої моделі дискретного каналу можна записати

$$\hat{A}^{(n)} = A^{(n)} + E^{(n)}, \qquad (1.10)$$

де $A^{(n)}$ і $\hat{A}^{(n)}$ — випадкові послідовності n символів на вході та виході каналу; $E^{(n)}$ — випадковий вектор помилки, який у загальному випадку залежить від $A^{(n)}$.

Відомі моделі дискретних каналів різняться розподілом імовірностей вектора $E^{(n)}$. Розглянемо деякі з них.

Симетричний канал без пам'яті — це дискретний канал, у якому ймовірність помилкового приймання символу не залежить від передісторії, тобто від того, які символи передавалися раніше та як вони були прийняті. В такому каналі кожний переданий кодовий символ може бути прийнятий помилково з фіксованою ймовірністю p і правильно з імовірністю $1-p$, причому в разі помилки замість переданого символу a може бути з однаковою ймовірністю прийнятий будь-який інший символ з алфавіту q.

Таким чином, імовірність того, що був прийнятий символ $\hat{a}_j$ замість переданого символу a_i, можна визначити так:

$$p(\hat{a}_j / a_i) = \begin{cases} p/q-1 & \text{при} \quad i \neq j; \\ 1-p & \text{при} \quad i = j. \end{cases} \tag{1.11}$$

Імовірність будь-якого n-вимірного вектора помилки в такому каналі визначається виразом

$$p(E^{(n)}) = [p/(q-1)]^{\nu}(1-p)^{n-\nu}, \tag{1.12}$$

де ν — кількість ненульових символів у векторі помилки (вага помилки).

Імовірність виникнення ν будь-яких помилок (і в будь-якому порядку) у векторі завдовжки n можна знайти за формулою Бернуллі

$$p(\nu) = C_n^{\nu}\left(\frac{p}{q-1}\right)^{\nu}(1-p)^{n-\nu}, \tag{1.13}$$

де $C_n^{\nu} = n!/[\nu!\,(n-\nu)!]$ — біномний коефіцієнт, який дорівнює кількості різних сполучень ν помилок у послідовності з n елементів.

Модель (1.13) ще називається *біномним каналом*. Вона використовується тоді, коли в неперервному каналі, що входить до складу дискретного, відсутні завмирання й є тільки адитивний «білий шум» (у крайньому разі — «квазібілий»).

Симетричний канал без пам'яті зі стираннями відрізняється від попереднього тільки введенням у початкову послідовність символів додаткового символу, що позначається знаком «?», який використовується тоді, коли розв'язувальна схема демодулятора не може надійно розпізнати переданий символ. Завдяки введенню цього додаткового елемента стирання досягається значне зменшення ймовірності помилки.

Несиметричний канал без пам'яті відрізняється від симетричного тим, що помилки виникають у ньому незалежно одна від одної, але ймовірність їх залежить від того, який символ передається. Так, для двійкового (бінарного) несиметричного каналу ($q = 2$) ймовірність $P(1/0)$ приймання символу «1» при передачі символу «0» не дорівнює ймовірності $P(0/1)$ приймання символу «0» при передачі символу «1». У цій моделі ймовірність вектора помилки залежить від послідовності символів, які передаються.

Марківська модель є найпростішою моделлю дискретного каналу з пам'яттю. Ймовірність виникнення помилки в цій моделі утворює просте коло Маркова, тобто залежить від того, правильно чи помилково прийнято попередній символ, але в той же час не залежить від того, який символ передається. Ця модель застосовується тоді, коли в неперервному каналі з гауссовим шумом використовується відносна фазова модуляція [33].

Канал з адитивним дискретним шумом є узагальненою моделлю дискретних каналів, де ймовірність виникнення вектора помилки $E^{(n)}$ не залежить від послідовності символів, які передаються. При цьому ймовірність виникнення кожного вектора помилки вважається заданою і взагалі не визначається його вагою. У багатьох двовекторних каналах з однаковою вагою більш імовірним є той, в якому присутні пачки (пакети) помилок або помилки розташовані ближче одна від одної.

1.4. ПРЕДМЕТ ТЕОРІЇ ІНФОРМАЦІЇ ТА КОДУВАННЯ

Теорія інформації — це розділ кібернетики, в якому за допомогою математичних методів вивчаються способи вимірювання кількості інформації, що міститься в будь-яких повідомленнях, способи кодування для економічного подання повідомлень і надійної передачі їх по каналах зв'язку з завадами.

Курс теорії інформації об'єднує такі теоретичні напрями, як кількісна оцінка інформації, кодування повідомлень, їх стиснення, оцінка ефективності та завадостійкості передачі кодованих повідомлень.

Одним з головних завдань теорії інформації є максимальне використання потенційних можливостей каналів зв'язку на основі оптимального кодування джерела повідомлення та його дальшого завадостійкого кодування. Це збігається з завданням *теорії кодування* — здобуттям ефективних алгоритмів кодування для джерел повідомлень і передачі даних по каналах зв'язку.

Теорія інформації та кодування за своєю природою дуже близька до математичних дисциплін; тому як апарат дослідження в ній застосовуються теорія скінченних полів, лінійна алгебра, комбінаторика, теорія матриць, теорія ймовірностей та математична статистика.

Без розвитку теорії інформації та кодування і впровадження її в життя практично неможливо створення складних систем керування супутниками Землі та ракетами, систем і мереж зв'язку та передачі даних, складних ЕОМ і комплексів тощо.

Засвоєння матеріалу курсу теорії інформації та кодування є необхідним для дальшого оволодіння знаннями в галузі збирання, передавання, перетворення, оброблення, накопичення та зберігання інформації, проектування та розроблення інформаційних систем і засобів автоматики та інформатики будь-якого призначення.

Інтерес до теорії інформації та кодування, а також до технічних засобів реалізації її положень зростатиме зі збільшенням обсягів потоків обміну інформацією. Ця тенденція прослідковується в усіх галузях науки та техніки. Провадиться багато досліджень щодо розробки нових класів кодів, способів захисту інформації від несанкціонованих втручань у потоки інформації, вдосконалення способів і засобів її кодування та декодування.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що таке повідомлення?
2. У чому полягає різниця між даними, повідомленнями та інформацією?
3. Що таке сигнал?
4. Які бувають сигнали?
5. Якими параметрами характеризуються сигнали?
6. Як узгодити ємність каналу та обсяг сигналу?
7. Чим характеризуються неперервні та дискретні повідомлення?
8. Які основні вузли систем передачі інформації та яке призначення їх?
9. Що таке лінія та канал зв'язку?
10. У чому полягає різниця між лінією та каналом зв'язку?
11. У чому полягають операції перетворення сигналу на повідомлення при передачі інформації?
12. Які відмінності багатоканальної системи передачі інформації від одноканальної?
13. Які бувають математичні моделі каналів зв'язку?
14. Що таке ідеальний канал зв'язку без завад?
15. У чому полягає особливість каналу зв'язку з адитивним гауссовим шумом?
16. Яка різниця між одно- та багатопроменевим каналами зв'язку?
17. Які особливості каналу зв'язку з невизначеною фазою сигналу?
18. Які відмінності між симетричним і несиметричним каналами зв'язку?
19. Що таке симетричний канал зв'язку без пам'яті?
20. Які математичні моделі дискретних каналів зв'язку застосовуються при моделюванні реальних каналів?
21. Що вивчається в дисципліні «Теорія інформації та кодування»?

РОЗДІЛ 2 КІЛЬКІСНІ ХАРАКТЕРИСТИКИ ІНФОРМАЦІЇ

Обговорюючи кількісні характеристики будь-якого досліджуваного об'єкта, мають на увазі насамперед можливість і вміння вимірювати ту чи іншу характеристику. Ця можливість визначається доступністю об'єкта та наявністю вимірювального пристрою, а вміння — наявністю та знанням кількісної міри і процедур чи методик її застосування.

Звісно, вибір міри є актом довільним. Можна вибрати загальноприйняту міру, скажімо, метр для вимірювання довжини деякого об'єкта. А можна вибрати й щось інше, як у відомому мультиплікаційному фільмі, де герої вимірювали довжину удава в папугах, доставши 38 папуг.

У теорії інформації поряд з іншими вирішуються питання вимірювання кількості інформації. При цьому потребують свого визначення об'єкти дослідження — джерела повідомлень і моделі їх, зокрема ансамблі повідомлень. Нижче описуються особливості джерел повідомлень, дається оцінка кількості інформації в повідомленнях, наводяться властивості ентропії та її різновиди.

2.1. АНСАМБЛІ ТА ДЖЕРЕЛА ПОВІДОМЛЕНЬ

Матеріальному світові, що оточує людину, притаманна безліч фізичних явищ, багато з яких змінюються в часі, маючи форму фізичних процесів, тобто таких явищ, фізичні показники яких не є миттєвими, а розподіленими в часі, які можна спостерігати кожної миті.

Будь-який матеріальний об'єкт разом із спостерігачем утворює систему, яка називається *джерелом повідомлень* [15, 44]. На рис. 2.1 зображено схему системи взаємозв'язаних об'єктів і спостерігачів, вкладених одне в одне, стосовно передачі відомостей про певний фізичний об'єкт певному одержувачеві. Для кожної стрілки на рис. 2.1 частина системи, розміщена ліворуч, може розглядатися, як спостережуваний об'єкт, а розташована праворуч — як спостерігач. При цьому не має значення природа спостерігача: чи це людина, чи це якийсь прилад. Його головне завдання полягає в перетворенні відомостей про стан

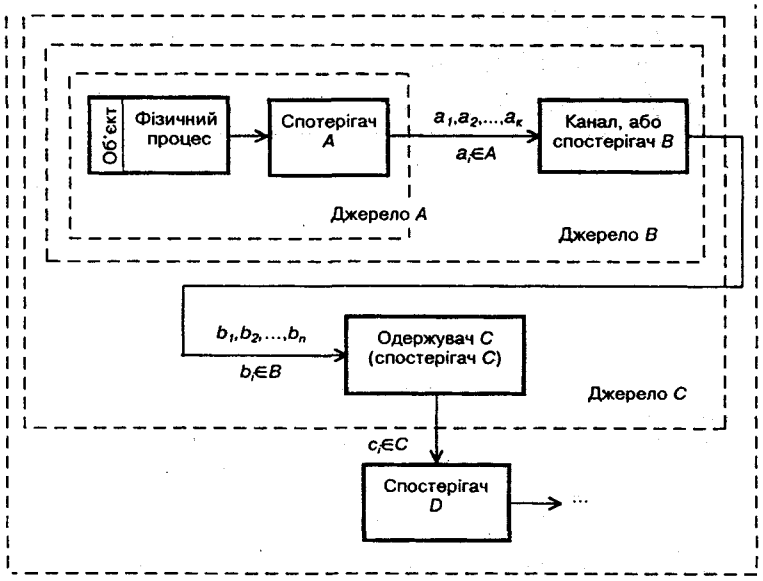


Рис. 2.1

спостережуваного об'єкта на форму, зручну для прийняття іншими людиною або приладом.

Стан матеріального об'єкта, а отже, і його фізичні показники можуть набувати значення з певного дискретного набору значень. Джерело повідомлень з таким об'єктом є *дискретним*. Якщо стан матеріального об'єкта, відбитий у його фізичних показниках, набуває значення з нескінченної множини можливих значень, то таке джерело повідомлень є *неперервним*. Принципово воно може бути зведене до дискретного, якщо прийняти допустимий рівень похибки та за її допомогою з нескінченної множини можливих значень повідомлень вибрати певний дискретний набір. Саме тут у наявності похибки та її допустимому рівні криється принципова різниця між дискретним і неперервним джерелами повідомлень. Докладніше це питання розглядається в розд. 4.

Якщо під час деякого часового проміжку дискретним джерелом вибрано деяке повідомлення a_i, яке ніяк не зумовлене повідомленням a_{i-1}, вибраним у попередній проміжок часу, то таке джерело є *дискретним джерелом без пам'яті*.

Якщо в деякому часовому проміжку дискретним джерелом вибрано повідомлення a_i, пов'язане з попереднім повідомленням a_{i-1} і статистично зумовлене ним, то таке джерело називається *дискретним джерелом із пам'яттю*.

Крім дискретних, можуть бути також *неперервні джерела повідомлень із пам'яттю* та *без пам'яті* (див. п. 4.4). Переліченими тут типами не вичерпуються всі відомі джерела повідомлень. Специфічні джерела повідомлень, в яких враховано статистичні якості фізичних параметрів, розглядаються в [3, 35, 44].

Якщо кожного проміжку часу дискретне джерело повідомлень вибирає одне з k можливих повідомлень $a_1, a_2, ..., a_k$, то кажуть, що $A = \{a_1, ..., a_k\}$ є дискретною множиною повідомлень, або просто множиною повідомлень A. Як правило, для повнішого опису джерела повідомлень на множині A визначають її ймовірнісну міру, тобто з кожним дискретним повідомленням a_i пов'язують ймовірність p_i його вибору джерелом. Таким чином, множині $A = \{a_1, ..., a_k\}$ зіставляється ймовірнісна міра у вигляді множини $P_k = \{p_1, ..., p_k\}$, на яку накладено обмеження у вигляді $\sum_{i=1}^{k} p_i = 1$.

Дві множини A та P дають достатньо повний опис дискретного джерела повідомлень у вигляді його ймовірнісної моделі, а тому разом вони утворюють ансамбль повідомлень дискретного джерела. Це означає, що кожного проміжку часу дискретним джерелом вибирається певне повідомлення $a_i \in A$ з імовірністю $p(a_i) = p_i \in P$. Наведене вище обмеження є природною умовою включення до складу множини A повної групи подій, якими виступають дискретні повідомлення. Ця вимога з'являється тут тому, що надалі треба скористатися апаратом математичної статистики, звідки й походить цей термін. Це дає змогу врахувати при розгляді всі обговорювані події-повідомлення. За своїм розсудом можна змінити склад можливих повідомлень в A, але слід пронормувати їх ймовірності так, щоб сума ймовірностей дорівнювала одиниці.

2.2. КІЛЬКІСНА МІРА ІНФОРМАЦІЇ

Розглянемо докладніше деякі властивості дискретного джерела повідомлень та моделі його ансамблю.

Нехай дискретне джерело повідомлень складається з людини-спостерігача та такого фізичного процесу, як колір неба. Якщо яскравого безхмарного дня в певну мить таким джерелом вибрано повідомлення «синій», то ніяких нових відомостей до наших знань про цей об'єкт воно не дасть, тобто можна сказати, що є повідомлення, але немає інформації, або її кількість дорівнює нулю.

Якщо ж за цих самих обставин джерелом вибрано якесь інше повідомлення про колір неба (наприклад, «зелений»), то воно

дасть нові відомості про спостережуваний об'єкт, які додадуть до наших знань щось нове. Тоді можна сказати, що є повідомлення і в ньому є якась кількість інформації.

Зараз інформацію на якісному рівні інтуїтивно можна визначити як нове знання про стан спостережуваного об'єкта, а її кількість — як кількість нового знання про нього. Звичайно, якщо нове знання збільшує загальний рівень знання про стан об'єкта, то кількість інформації має накопичуватися додаванням і повинна мати адитивний характер. З іншого боку, спостерігач у певний момент часу нічого не знає про новий стан об'єкта. Він тільки спостерігає його і, лише вибравши нове повідомлення, дозволяє дістати якесь нове знання про об'єкт. Можна сказати, що до вибору повідомлення джерелом невизначеність стану об'єкта з боку спостерігача має певний рівень.

Після вибору повідомлення джерелом утворюється деяка кількість інформації про стан спостережуваного об'єкта, яка певною мірою зменшує цю невизначеність.

Розглянемо, дещо ускладнивши, відомий приклад, пов'язаний з грою в шахи. Нехай дискретне джерело повідомлень складається з шахівниці з фігурами та спостерігача, який у певні моменти часу повинен фіксувати стан справ на шахівниці та передавати його у вигляді повідомлень. Дискретність такого джерела визначається дискретністю та черговістю дій гравців, внаслідок чого змінюється стан спостережуваного об'єкта.

Припустимо, що спостерігач починає фіксувати стан гри на шахівниці з певного моменту після чергового ходу (тобто після чергової зміни позицій). Він має вибрати одне з множини A можливих повідомлень. Які ж відомості він повинен включити до повідомлення? Звичайно, перелік усіх наявних фігур на шахівниці та положення їх на полі (тобто координати кожної позиції) на даний момент часу. Закони побудови повідомлень з множини A можуть бути різними: за мовою, знаками чи символами, складом елементів повідомлення тощо. Визначимося з елементами повідомлення. Як уже з'ясовано вище, елементами повідомлення з множини A мають бути: а) зазначення гравця або кольору фігури; б) зазначення самої фігури (яка саме фігура) або її відсутності; в) відомості про її позицію або клітинку шахівниці, про яку йдеться.

Спостерігач може закодувати елементи повідомлення звичайною мовою або якимись літерами, символами, ієрогліфами тощо. Наприклад, елемент *а* (білий або чорний) можна закодувати (позначити) цифрами 0 і 1 відповідно; елемент *б* (відсутність або назву відповідної фігури) — звичайними цифрами. Скажімо, відсутність будь-якої шахової фігури можна позначити цифрою 0, а наявність пішака — 1, слона — 2, коня — 3, тури — 4, ферзя — 5, короля — 6 (відомо, що ці фігури в шаховій літературі теж позначаються символами, але літерними).

Позицію шахової фігури можна позначити координатами клітинки, як прийнято в грі. При цьому знадобляться вісім позначок для визначення рядка клітинки та стільки ж позначок для визначення самої клітинки в рядку. Цими позначками можуть бути цифри 0, 1, 2, ..., хоча й не обов'язково. Тоді, наприклад, таке повідомлення спостерігача, як 0368 означає, що йдеться про білу (позначка 0 на першій позиції) фігуру під назвою «кінь» (позначка 3 на другій позиції), розміщену на восьмій клітинці шостого рядка.

Множина A містить стільки повідомлень, скільки взагалі може бути подібних цифрових записів у цьому форматі. При такому форматі всього може бути $2 \times 7 \times 8 \times 8 = 896$ повідомлень, тобто в множині A_1 є два повідомлення на першій позиції, сім — на другій та по вісім — на третій і четвертій позиціях.

Розглянемо трохи інший формат повідомлень спостерігача. Нехай він містить два, а не чотири елементи, як у попередньому випадку: перший — це факт відсутності шахової фігури (символ 0) або зазначення однієї з 12 фігур (шість чорних і стільки ж білих), відображеної цифрами від 1 до 12; другий — це просто номер однієї з 64 клітинок шахівниці, що розглядається. Таким чином, перший елемент повідомлення може мати 13 значень, а другий — 64, тобто в множині A_2 є $13 \times 64 = 832$ повідомлення. Наприклад, повідомлення 037 означає, що на 37-й клітинці шахівниці немає ніякої фігури, а повідомлення 936 свідчить про те, що на 36-й клітинці розміщується, скажімо, білий кінь (цифра 9 позначає саме білого коня).

Така помітна різниця між кількістю можливих повідомлень (896 у першому та 832 в другому форматах) пояснюється дуже просто: в першому форматі до множини A_1 включено абсурдні повідомлення із взаємовиключними елементами. Це повідомлення, в яких зазначено колір фігури, факт її відсутності та координати шахової клітинки. Якщо за відсутності фігури на кожній з 64 клітинок перший формат дає змогу вказати колір фігури, то таких повідомлень має бути 128 (по 64 на кожний колір). Однак половина цих 128 повідомлень є зайвими. Звісно, немає значення, якого кольору фігура відсутня на даній клітинці. Має значення сам факт відсутності фігури, тому достатньо лише 64 таких повідомлень.

Отже, бачимо, що із 896 повідомлень 64 є зайвими, які потрібно виключити з множини A_1. Тоді $A_1 = 896 - 64 = 832 = A_2$. Таким чином, якщо спостерігач не повинен вибирати абсурдні повідомлення, то кількість можливих повідомлень дискрет-

ного джерела за обома форматами має бути однаковою, тобто $A_1 = A_2$.

Можна сказати, що кількість інформації в повідомленні про стан такого спостережуваного об'єкта, як шахівниця, незалежно від побудови цього повідомлення (чи то з множини A_1, чи то з множини A_2) залишається однаковою. Якщо перекодувати повідомлення про певний стан об'єкта в інших символах, скажімо, двійкових, то для повідомлення з множини A_1 потрібно 10 цифр: одна — колір шахової фігури (це 0 або 1) та по три — її назва, номер рядка і номер клітинки в рядку. Для повідомлення з множини A_2 також потрібно 10 цифр: чотири — назва фігури чи її відсутність і шість — номер клітинки. З іншого боку, кількість повідомлень $A_1 = A_2 = 832$ можна подати в двійковій формі, що потребує теж 10 двійкових цифр, оскільки задовольняється нерівність $2^{10} > 832 > 2^9$.

Таким чином, на інтуїтивному, якісному рівні можна дійти висновку, що кількість інформації може бути нульовою або мати якесь ненульове значення, повідомлення можна порівнювати за кількістю інформації в них, кількість інформації має накопичуватись додаванням, якщо складність повідомлення зростає.

Отже, розгляд дискретного джерела та множини його повідомлень приводить до формулювання таких природних вимог для визначення кількості інформації:

1) у повідомленні про вірогідний випадок вона має дорівнювати нулю;

2) у двох незалежних повідомленнях вона має дорівнювати сумі кількостей інформації в кожному з них;

3) вона не повинна залежати від якісного змісту повідомлення (ступеня його важливості, відомостей тощо).

Про зміст і значення першої вимоги йшлося вище. Друга вимога має на увазі статистично не зумовлені повідомлення. Скажімо, в наведених прикладах з шахами на формування елементарних повідомлень із множини A_1 потрібно мати таку кількість двійкових цифр: 1; 3; 3; 3, а повідомлень із множини A_2 — 4 та 6. Однак складне повідомлення з обох множин потребує 10 двійкових цифр, що дорівнює сумі кількостей цифр в обох випадках.

Третя вимога зумовлена необхідністю абстрагуватися від конкретного змісту повідомлення заради досягнення найзагальнішого характеру визначення кількості інформації.

Таким чином, для визначення кількості інформації в повідомленні треба виходити з найзагальнішої його характеристики. Таку характеристику, очевидно, дає модель джерела — ансамбль повідомлень.

Вище ми користувалися лише множиною повідомлень з ансамблю. Звернімо тепер увагу на множину ймовірностей цих повідомлень $p(a_i) \in P$. Якщо йдеться про вірогідний випадок, імовірність якого $p(a_i) = p_i = 0$ або $p_i = 1$ (ніколи не відбудеться неможливий випадок або точно відбудеться), то кількість інформації в ньому $I(a_i) = 0$, оскільки рівень невизначеності щодо стану об'єкта після вибору джерелом такого повідомлення не змінився. Якщо ж джерело вибере не таке вже наперед визначене повідомлення, то рівень невизначеності щодо стану об'єкта знизиться. Отже, кількість інформації в повідомленні має бути функцією ймовірності цього повідомлення, тобто $I(a_i) = f(p_i)$, $a_i \in A$, $p_i \in P$.

Для задоволення другої вимоги можна врахувати, що ймовірність двох незалежних повідомлень a_1 та a_2 за законом множення ймовірностей $p(a_1, a_2) = p(a_1)\, p(a_2)$. Проте кількість інформації у цих двох повідомленнях

$$i(a_1, a_2) = i(a_1) + i(a_2). \tag{2.1}$$

Звідси випливає необхідність вибору такої функції $f(p_i)$, значення якої при перемноженні її аргументів p_i додавалися б і щоб $f(0) = f(1) = 0$. Єдиною функцією з такими властивостями є логарифмічна функція

$$I(a_i) = k \log p(a_i), \tag{2.2}$$

де k — коефіцієнт, який узгоджує розмірності (згадаємо обговорення розмірів і мір для кожного з них), а логарифм береться за будь-якою зручною основою.

При такому визначенні кількості інформації задовольняються всі три наведені вище вимоги. Вибір основи логарифма не принциповий, тому що від неї залежить лише одиниця фізичної величини (змінюється значення k):

$$k_1 \log_m p(a) = k_1 \log_m n \log_n p(a) = k_2 \log_n p(a). \tag{2.3}$$

Для того щоб кількість інформації I виражалась додатним числом, покладемо $k = -1$, оскільки $p(a) \le 1$ та $\log p(a) \le 0$, якщо основа логарифма більша від одиниці. Тоді

$$I(a) = -\log p(a) = \log(1 / p(a)). \tag{2.4}$$

За основу логарифма найчастіше вибирають двійку. При цьому одиниця кількості інформації називається *двійковою*, або *бітом* (*Binary digit*). Вона дорівнює кількості інформації в повідомленні про такий випадок, який з однаковою ймовірністю може як відбутися, так і не відбутися, тобто коли моделлю

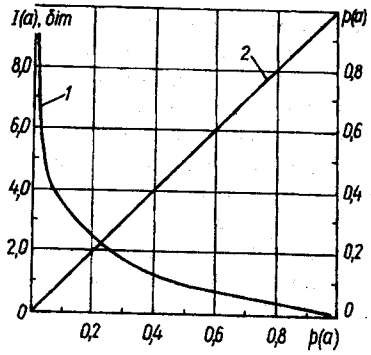


Рис. 2.2

еталонного джерела є ансамбль $A = \{a_1, a_2\}$ та $P = \{0{,}5; 0{,}5\}$.

Якщо за основу логарифма вибрано число e, то така одиниця інформації називається *натуральною*.

Таким чином, кількість інформації в повідомленні тим більша, чим воно менш ймовірне, тобто більш неочікуване. Цю залежність відображено на рис. 2.2 кривою *1*. Лінія *2*, тобто графік $p\,(a) = p\,(a)$, на цьому рисунку знадобиться далі.

Коли джерело інформації породжує послідовність взаємозалежних повідомлень, то здобуття кожного з них змінює ймовірність наступних, а отже, кількість інформації I в них. Остання має вже визначатися умовною ймовірністю вибору джерелом цього повідомлення a_n, якщо до нього вибрано повідомлення $a_{n-1}, a_{n-2}, \ldots$, тобто

$$I\,(a_n / a_{n-1}, a_{n-2}, \ldots) = \log [1 / p\,(a_n / a_{n-1}, a_{n-2}, \ldots)]. \quad (2.5)$$

Визначена таким чином кількість інформації є величиною випадковою тому, що самі повідомлення випадкові. Розподіл її ймовірностей визначається розподілом повідомлень у цьому ансамблі.

Якщо припустити статистичну незалежність та однакову ймовірність повідомлень у наведеному вище прикладі стосовно шахів і визначити ймовірність кожного повідомлення як $p\,(a_i) = p_i = 1/832$, то згідно з (2.4) кількість інформації в одному такому повідомленні в бітах становитиме

$$I\,(a_i) = -\log_2 p_i = \log_2 (1 / p_i) = \log_2 832 \approx 9{,}70044.$$

Як бачимо, це відповідає майже 10 двійковим символам, потрібним для запису такого повідомлення або формулювання його спостерігачем.

Фрагмент таблиці значень двійкових логарифмів цілих чисел уміщено в дод. 1.

2.3. ЕНТРОПІЯ ТА ЇЇ ВЛАСТИВОСТІ

Уже йшлося про те, що здобуття інформації від джерела знімає певною мірою невизначеність стану спостережуваного об'єкта. Якщо за час формування джерелом нового повідомлення

об'єкт не змінює свій стан (тобто джерелом вибирається попереднє повідомлення з множини A), можна уточнити відомості про попередній стан об'єкта, включивши до цієї множини нові можливі повідомлення та перенормувавши ймовірності з множини P.

Взагалі з самого початку до складу множини A слід включати такі повідомлення та таку кількість їх, щоб одним повідомленням можна було б визначити стан об'єкта з потрібною точністю. Це означає, що, формуючи модель джерела повідомлень (його ансамбль), треба заздалегідь передбачити всі необхідні повідомлення.

Інша справа, що кожне таке повідомлення може бути відображене певною кількістю символів, знаків тощо, переносячи певну кількість інформації. При цьому множина повідомлень A є *алфавітом повідомлень*, а множина символів, знаків тощо, за допомогою яких спостерігач подає кожне повідомлення у формі, зручній для одержувача, — *алфавітом джерела*. В літературі перше іноді називають *первинним*, а друге — *вторинним* алфавітами [42].

Ми вже бачили вище, що немає ніякого значення, в якому алфавіті подаються повідомлення. Модель джерела (ансамбль) враховує лише склад їх і розподіл імовірностей (поки що йдеться про статистично незалежні повідомлення).

Розглянемо, наприклад, дискретне джерело повідомлень з ансамблем, наведеним у табл. 2.1. Обчислимо, яку кількість інформації несе кожне таке повідомлення, й занесемо ці дані в табл. 2.1. Третій рядок її підтверджує, що кількість інформації

Таблиця 2.1

$a_i \in A$	a_1	a_2	a_3	a_4	a_5	a_6
$p_i \in P$	0,4	0,3	0,15	0,1	0,03	0,02
$I\,(a_i)$, біт	1,322	1,737	2,737	3,322	5,059	5,644

при прийнятому її визначенні відображує міру неочікуваності кожного повідомлення.

Розглянемо, яку кількість інформації несуть більш-менш довгі послідовності таких повідомлень:

- перша послідовність

$$a_3, a_1, a_4, a_4, a_2, a_2, a_1, a_3, a_1, a_2,$$
$$a_5, a_1, a_1, a_2, a_3, a_2, a_1, a_2, a_1, a_1,$$
$$a_1, a_2, a_1, a_1, a_1, a_2, a_2, a_3, a_4, a_3;$$

• друга послідовність

$$a_2, a_1, a_3, a_2, a_2, a_1, a_2, a_1, a_3, a_1, a_4,$$
$$a_2, a_4, a_1, a_1, a_3, a_1, a_2, a_1, a_5.$$

У першій послідовності є 30 повідомлень, а в другій — 20. Розподіл імовірностей з ансамблю (див. табл. 2.1) настільки нерівномірний, що ні до першої, ні до другої послідовностей не ввійшло повідомлення a_6. Справа в тому, що його за законами статистики можна було б помітити в послідовності повідомлень при довжині останньої, значно більшій від $n = 1 / p_6 = 50$. Отже, кількість інформації в першій послідовності

$$i_1 = \sum_{i=1}^{30} I(a_i) = 12I(a_1) + 9I(a_2) + 5I(a_3) + 3I(a_4) + I(a_5) =$$
$$= 15{,}864 + 15{,}633 + 13{,}685 + 9{,}966 + 5{,}059 = 60{,}207 \text{ біт},$$

а в другій

$$i_2 = \sum_{i=1}^{20} I(a_i) = 8I(a_1) + 6I(a_2) + 3I(a_3) + 2I(a_4) + I(a_5) =$$
$$= 10{,}576 + 10{,}422 + 8{,}211 + 6{,}644 + 5{,}059 = 40{,}912 \text{ біт}.$$

Як бачимо, ці послідовності різняться не тільки кількістю повідомлень, а й кількістю інформації в кожній з них. Однак, якщо обчислити кількість інформації, яка припадає на одне повідомлення в одній послідовності та в іншій, то виявиться, що $i_1/30 = 2{,}0069$ біт/повідомлення та $i_2/20 = 2{,}0456$ біт/повідомлення. Це означає, що середня кількість інформації, яка припадає на одну літеру алфавіту повідомлень (це те саме, що й на одне повідомлення), не залежить від конкретних повідомлень і довжини послідовності їх.

Деяка різниця тут між $i_1/30$ та $i_2/20$ пояснюється лише недостатньою довжиною послідовностей повідомлень. Відповідно до статистичного закону великих чисел ці відношення збігатимуться краще, чим більшими будуть довжини порівнюваних послідовностей.

Можна сказати, що це відношення (тобто кількість інформації, яка припадає на одне повідомлення) характеризує дискретне джерело повідомлень в цілому. Інше джерело з іншим ансамблем повідомлень матиме зовсім іншу питому кількість інформації. Ця загальна характеристика джерела повідомлень називається його *ентропією* $H(A)$. Вона має фізичний зміст середньостатистичної міри невизначеності відомостей спостерігача A (див. рис. 2.1) відносно стану спостережуваного об'єкта.

Точно ентропію можна визначити як математичне сподівання питомої кількості інформації

$$H(A) = \sum_i p(a_i) I(a_i) = -\sum p(a_i) \log p(a_i). \qquad (2.6)$$

Згідно з даними табл. 2.1 маємо $H(A) = 0{,}4 \cdot 1{,}322 + 0{,}3 \times \times 1{,}737 + 0{,}15 \cdot 2{,}737 + 0{,}1 \cdot 3{,}322 + 0{,}03 \cdot 5{,}059 + 0{,}02 \cdot 5{,}644 = = 2{,}0573$ біт/повідомлення.

Бачимо, що точне значення ентропії $H(A)$ не дуже відрізняється від значень, здобутих у наведених вище прикладах послідовностей повідомлень.

Для полегшення розрахунків ентропії за (2.6) у дод. 2 вміщено фрагмент таблиці значень функції $-p \log_2 p$.

У виразі (2.6) усереднення (як обчислення математичного сподівання) виконується по всьому ансамблю повідомлень. При цьому потрібно враховувати всі ймовірнісні зв'язки між різними повідомленнями. З цього виразу випливає, що чим вища ентропія, тим більшу кількість інформації в середньому закладено в кожне повідомлення даного джерела, тим важче запам'ятати (записати) або передати таке повідомлення по каналу зв'язку.

Необхідні витрати енергії на передачу повідомлення пропорційні його ентропії (середній кількості інформації на одне повідомлення). Виходить, що кількість інформації в послідовностях визначається кількістю повідомлень N у послідовності та ентропією $H(A)$ джерела, тобто

$$i(N) = N H(A). \qquad (2.7)$$

Наприклад,

$$i_1 = 30 \cdot 2{,}0573 = 61{,}719 \text{ біт}; \quad i_2 = 20 \cdot 2{,}0573 = 41{,}146 \text{ біт}.$$

Ці точні дані можна порівняти з наведеними вище розрахунками i_1 та i_2 стосовно двох послідовностей з $N = 30$ і 20 й ансамблю повідомлень із табл. 2.1. Розбіжність тут пояснюється невеликими значеннями N, адже ймовірності p_i обчислюються, як правило, за умови $N \to \infty$.

Розглянемо вироджене дискретне джерело з єдиним повідомленням $a \in A$ з $p(a) = 1$. Тоді $H(A) = 0$ згідно з (2.6). Якщо $p(a) = 0$, то $H(A)$ теж дорівнюватиме нулю. Таким чином, *ентропія завжди додатна або дорівнює нулю*, тобто невід'ємна. Це перша її властивість.

Друга властивість ентропії випливає з виразу (2.6), згідно з яким вона є величиною адитивною. Якщо N-вимірні послідовності повідомлень $a_1, a_2, \ldots, a_N$ розглядати як збільшені повідомлення нового джерела, то його ентропія буде в N разів більшою від початкової.

Якщо алфавіт $A = \{a_1, a_2, \ldots, a_k\}$ має k різних повідомлень, то $H(A) \le \log k$. Тут рівність стосується тільки рівноймовірних і статистично незалежних повідомлень $a_i \in A$. Число k називається *обсягом алфавіту повідомлень*.

У розглядуваному прикладі $A = \{a_1, \ldots, a_6\}$. Вважаючи повідомлення статистично незалежними за умови рівноймовірності їх із $p_i = p = 1/6$ для $i = 1, \ldots, 6$, матимемо

$$H(A) = -\sum_{i=1}^{6} \frac{1}{6} \log_2 \frac{1}{6} = \log_2 6 = 2{,}585 \quad \text{біт/повідомлення.}$$

У дійсності нерівноймовірність повідомлень призводить до зменшення деяких складових у виразі (2.6). Тому ми й дістали для джерела повідомлень з ансамблем, наведеним у табл. 2.1, значення $H(A) = 2{,}0573$ біт/повідомлення як розплату за нерівноймовірність повідомлень.

2.4. БЕЗУМОВНА ЕНТРОПІЯ

Термін «безумовна ентропія» запозичений з математичної статистики за аналогією з безумовною ймовірністю, що стосується статистично незалежних подій, тут — повідомлень. Отже, *безумовна ентропія* — це кількість інформації, яка припадає на одне повідомлення джерела із статистично незалежними повідомленнями.

За цим визначенням розглянута в п. 2.3 ентропія є безумовною. Зупинимося докладніше на безумовній ентропії та її властивостях.

Якщо є дискретне джерело статистично незалежних повідомлень з ансамблем $A = \{a_1, a_2, \ldots, a_i, \ldots, a_k\}$ та $p = \{p_1, p_2, \ldots, p_i, \ldots, p_k\}$, то кількість інформації (середня), що припадає на одне повідомлення $a_i \in A$ й визначається формулою Шеннона

$$H(A) = -\sum_{i=1}^{k} p_i \log p_i, \qquad (2.8)$$

є характеристикою цього джерела в цілому. Вона має фізичний зміст середньої за ансамблем невизначеності вибору джерелом повідомлення з A, причому байдуже, якого саме повідомлення, оскільки обчислення ентропії (2.8) «поглинає» індекс i. Наприклад, джерело з $k = 8$ незалежними та рівноймовірними повідомленнями має ентропію

$$H(A) = -\sum_{i=1}^{8} \frac{1}{8} \log_2 \frac{1}{8} = 3 \quad \text{біт/повідомлення.}$$

Тут ураховано, що $p_i = p = 1/8$. Для нерівноймовірних повідомлень у цьому разі $H(A) < 3$ біт/повідомлення.

Наведені в п. 2.3 властивості ентропії при цьому зберігаються, тобто якщо $p = 1$ або 0, то до ансамблю A не може входити більш як одне повідомлення. Таким чином,

$$H_1(A) = \sum_{i=1}^{1} 1 \cdot \log_2 \frac{1}{1} = 1 \log_2 1 = 0 \qquad (2.9)$$

або

$$H_0(A) = \sum_{i=1}^{1} 0 \cdot \log_2 \frac{1}{0} = 0 \cdot \infty,$$

де невизначеність $0 \cdot \infty$, якщо її розкрити за правилом Лопіталя через граничний перехід, дає

$$H_0(A) = 0. \qquad (2.10)$$

Останній випадок потребує пояснення. Якщо є ансамбль з алфавітом A, в якому певне повідомлення a_k має ймовірність $p_k = 0$, то таке повідомлення a_k можна просто виключити з ансамблю та далі не розглядати, оскільки $\sum_i p_i = 1$ для всіх $i \ne k$, а складова частка $H(A)$ із (2.8) дорівнює нулю за (2.10). Якщо при цьому в алфавіті A залишиться лише одне повідомлення (припустимо, їх було два), то очевидно, що $p = 1$ для залишеного повідомлення і маємо випадок (2.9).

Якщо в алфавіті A буде більше, ніж одне повідомлення, то виключення a_k з A ($p(a_k) = 0$) лише спростить модель джерела — його ансамбль і не змінить результат обчислення ентропії $H(A)$ за (2.8).

Безумовна ентропія K рівноймовірних повідомлень завжди максимальна й визначається виразом

$$H(A) = \log_2 K, \qquad (2.11)$$

який називається *формулою Хартлі*. Її легко дістати з формули Шеннона (2.8), поклавши $p_i = 1/K$ для $i = 1 \ldots K$, хоча хронологічно першою була запропонована формула (2.11).

Корисно дослідити вплив імовірності $p(a_i)$ на складові $I_i = -p(a_i) \log_2 p(a_i)$ формули (2.8). Наочне подання цього впливу дає графік $I_i = f(p_i)$ — крива 1 на рис. 2.3, яку здобуто з рис. 2.2 домно-

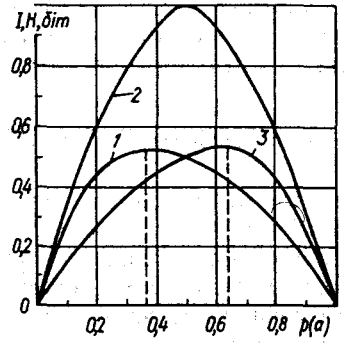


Рис.2.3

женням ординат кривої *1* на множник p (a), відображений на рис. 2.2 лінією *2*. Як випливає з рис. 2.3 (крива *1*), малі значення p (a) значно сильніше впливають на рівень складових формули Шеннона, ніж великі.

Крім того, за графіком прослідковується екстремум цих складових: якщо вони мають його, то й сума H (A) їх теж матиме екстремум.

Для прикладу розглянемо безумовну ентропію двійкового джерела, коли $A = \{a_1, a_2\}$ та $P = \{p, 1-p\}$. Тоді

$$H(A) = -\sum_{i=1}^{2} p_i \log_2 p_i = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p}, \qquad (2.12)$$

оскільки $p_1 + p_2 = p + 1 - p = 1$.

Графік першої складової відображено на рис. 2.3 кривою *1*, графік другої складової — кривою *3*, а графік ентропії H (A)= $= f(p)$ — кривою *2*. Бачимо, що остання є симетричною і має максимум $H(A)_{max} = 1$ при $p = 0{,}5$, тобто максимальна невизначеність повідомлень джерела при $p = 0{,}5$ спричинює його максимальну безумовну ентропію.

Дослідження недвійкових джерел ($K > 2$) дають результат з тією самою тенденцією: безумовна ентропія їх максимальна при рівноймовірності повідомлень, коли $p_i = 1/K$ для всіх $i = 1, 2, \ldots, K$ [3, 42, 44].

Таким чином, основними властивостями безумовної ентропії дискретних повідомлень є такі:

- ентропія — величина дійсна, обмежена та невід'ємна;
- ентропія вірогідних повідомлень дорівнює нулю;
- ентропія максимальна, якщо повідомлення рівноймовірні та статистично незалежні;
- ентропія джерела з двома альтернативними подіями може змінюватися від 0 до 1;
- ентропія складеного джерела, повідомлення якого складаються з часткових повідомлень кількох статистично незалежних джерел, дорівнює сумі ентропії цих джерел.

2.5. УМОВНА ЕНТРОПІЯ

Припустимо, що повідомлення a зустрічалося у довгому ланцюзі з $N = 1000$ повідомлень $l_1 = 200$ разів, повідомлення b у цьому самому ланцюзі — $l_2 = 200$ разів, а разом вони зустрілися лише $l_3 = 50$ разів. Скориставшись теорією математичної статистики, можна встановити, що ймовірність появи повідомлення a в цьому ланцюзі $p(a) = l_1/N = 0{,}25$, а повідомлення b – $p(b) = l_2/N = 0{,}2$. Крім того, $p(ab) = l_3/N = 0{,}05$. При цьому $p(ab) = p(a)\,p(b) = 0{,}25 \cdot 0{,}2 = 0{,}05$, що є явною ознакою статистичної незалежності повідомлень a та b. Саме для таких повідомлень існує безумовна ентропія, про яку йшлося вище.

Коли б у цьому прикладі пара ab зустрілася $l_3 = 30$ разів, то виявилося б, що $p(ab) = l_3/N = 0{,}03 < p(a)\,p(b)$. Це є ознакою порушення статистичної незалежності повідомлень a та b, яка відбиває той факт, що вони не «прагнуть» зустрічатися разом у послідовностях повідомлень, тобто поява одного з них дає підставу з більшою впевненістю не очікувати появи іншого, ніж це було б до появи першого повідомлення. Це означає, що ймовірність появи, скажімо, повідомлення b в послідовності відразу після появи повідомлення a трохи зменшується, і навпаки, хоча взагалі безумовна ймовірність $p(a)$ чи $p(b)$ по всій послідовності в цілому є сталою.

З іншого боку, при $l_3 = 100$ маємо $p(ab) = l_3/N = 0{,}1 > p(a)\,p(b)$, що дає підставу підозрювати взаємне «тяжіння» a до b, і навпаки. Тут теж проглядається порушення статистичної незалежності, тобто відношення «байдужості» між повідомленнями a та b. Поява в послідовності одного з них, наприклад a, трохи збільшує ймовірність появи повідомлення b відразу за повідомленням a, і навпаки. Проте безумовні ймовірності $p(a)$ та $p(b)$ по послідовності в цілому теж є сталими.

Мірою порушення статистичної незалежності (стану «байдужості») між повідомленнями a та b є умовна ймовірність появи повідомлення a за умови, що вже з'явилося повідомлення b і $p(a/b)$, або умовна ймовірність появи повідомлення b, коли вже з'явилося повідомлення a: $p(b/a)$, причому взагалі $p(a/b) \neq p(b/a)$.

Теорія математичної статистики визначає умовну ймовірність через безумовні ймовірності $p(a)$, $p(b)$ та сумісну безумовну ймовірність $p(ab)$ за законом множення ймовірностей:

$$p(ab) = p(a)\,p(b/a) = p(b)\,p(a/b). \qquad (2.13)$$

Звідси випливає, що

$$p(b/a) = p(ab) / p(a);\ p(a/b) = p(ab) / p(b). \qquad (2.14)$$

Зокрема, для статистично незалежних повідомлень a та b (див. вище) маємо

$$p(b/a) = 0{,}05 / 0{,}25 = 0{,}2 = p(b);$$

$$p(a/b) = 0{,}05 / 0{,}2 = 0{,}25 = p(a),$$

тобто умовні ймовірності появи повідомлень вироджуються в безумовні.

Тоді для $l_3 = 30$ згідно з (2.14) знаходимо

$$p\,(b/a) = 0{,}03/0{,}25 = 0{,}12;\ p\,(a/b) = 0{,}03/0{,}2 = 0{,}15, \qquad (2.15)$$

тобто встановлений факт появи повідомлення a зменшує безумовну ймовірність $p\,(b) = 0{,}2$ до умовної ймовірності $p\,(b/a) = 0{,}12$ появи повідомлення b за умови наявного вже повідомлення a. І навпаки, факт появи повідомлення b зменшує безумовну ймовірність $p\,(a) = 0{,}25$ до умовної ймовірності $p\,(a/b) = 0{,}15$ появи повідомлення a за умови наявності повідомлення b. Як бачимо, ймовірність появи повідомлення b зменшується на 0,08, а повідомлення a — на 0,1 кожного разу, як інше з них з'явиться в послідовності.

Аналогічно стосовно наведеного вище прикладу з $l_3 = 100$ відповідно до (2.14) маємо

$$p\,(b/a) = 0{,}1/0{,}25 = 0{,}4;\ p\,(a/b) = 0{,}1/0{,}2 = 0{,}5, \qquad (2.16)$$

що підкреслює згадане вище «тяжіння» a та b одне до одного через підсилення ймовірностей появи їх.

Насправді, при безумовній імовірності, скажімо, $p\,(b) = 0{,}2$ (саме з такою ймовірністю джерело вибирає b серед інших можливих повідомлень у довгому їх ланцюзі), як тільки буде вибрано a, ймовірність вибору b слідом різко (вдвічі в розглядуваному прикладі) підсилюється до значення умовної ймовірності $p\,(b/a) = 0{,}4$ появи повідомлення b за умови наявності повідомлення a. І навпаки, при $p\,(a) = 0{,}25$ взагалі, як тільки буде виявлено повідомлення b , ймовірність появи повідомлення a слідом або поряд з b різко підсилиться до значення умовної ймовірності $p\,(a/b) = 0{,}5$ згідно з (2.16). Отже, при виявленні одного повідомлення ймовірність появи слідом або поряд із ним статистично зумовленого, зв'язаного повідомлення збільшується [для умов (2.16)] або зменшується [для умов (2.15)].

Ці локальні порушення ймовірностей при статистичній залежності повідомлень не можуть бути непоміченими джерелом. І воно на них реагує відповідним зменшенням або збільшенням кількості інформації в кожному такому повідомленні згідно з (2.5). Звісно й ентропія такого джерела має змінюватися належним чином, причому називається вона *умовною*, визначається виразом (2.6), але з урахуванням умовних імовірностей повідомлень.

Розрізняють два різновиди умовної ентропії: *часткову* та *загальну*. Першу знаходять так:

$$H(A/b_j) = -\sum_i p(a_i/b_j)\log p(a_i/b_j),\ j = 1\ldots l; \qquad (2.17)$$

$$H(B/a_i) = -\sum_j p(b_j/a_i)\log p(b_j/a_i),\ i = 1\ldots k, \qquad (2.18)$$

де $A = \{a_1, a_2, \ldots, a_i, \ldots, a_k\}$, $B = \{b_1, b_2, \ldots, b_j, \ldots, b_l\}$ — алфавіти повідомлень; a_i — конкретне повідомлення, відносно якого визначається часткова умовна ентропія $H\,(B/a_i)$ алфавіту B за умови вибору попереднього повідомлення a_i; b_j — конкретне повідомлення, відносно якого обчислюється часткова умовна ентропія $H\,(A/b_j)$ алфавіту A за умови вибору попереднього повідомлення b_j; i — номер повідомлення з алфавіту A; j — номер повідомлення з алфавіту B; $p\,(a/b)$, $p\,(b/a)$ — умовні ймовірності.

Термін «вибір попереднього» досить умовний, оскільки повідомлення a та b можуть бути рознесені в часі, але знаходитися разом у просторі (як записано вище a/b чи b/a), або бути одночасними чи майже одночасними та рознесеними в просторі (як на рис. 2.1). Ніякі обмеження на алфавіти A та B не накладаються. Вони можуть навіть збігатися ($A = B$). Тому можна аналізувати і враховувати взаємозв'язок між повідомленнями одного й того самого джерела в одному й тому самому алфавіті повідомлень, рознесеними в часі, але не в просторі (наприклад, на виході джерела A чи B згідно з рис. 2.1). Такі послідовності зумовлених повідомлень називаються *ланцюгами Маркова* [38, 46, 49].

З іншого боку, алфавіти A та B можуть не збігатися ($A \neq B$), хоча між елементами їх може бути й відповідність. Цю ситуацію відбито на рис. 2.1, де джерело A має свою модель — ансамбль A та P_A. Оскільки воно може мати k дискретних станів, джерело A виступає як об'єкт спостерігача B і разом із ним утворює нове джерело B, яке має свою модель — ансамбль B та P_B. Між джерелом A та спостерігачем B існує канал зв'язку, в якому діють, умовно кажучи, завади, що порушують процес вибору спостерігачем B повідомлень $b_j \in B$. Це, в свою чергу, порушує відповідність між повідомленнями $a_i \in A$ та $b_j \in B$.

Алфавіти A та B можуть бути однакового ($k = l$) і неоднакового ($k \neq l$) обсягів. Звичайно розглядають ситуації, коли $k = l$ або $k < l$. Система спостереження з $k = l$ має природне пояснення. Тут спостерігач B повинен реагувати своїм повідомленням $b_j\,(j = 1 \ldots l)$ на кожний стан джерела A, висловлений повідомленням $a_i\,(i = 1 \ldots k)$. При цьому діє принцип перетворення $a_1 \to b_1$, $a_2 \to b_2$, …, $a_i \to b_i$, …, $a_k \to b_k$, коли кожному повідомленню a_i джерела A відповідає повідомлення b_i джерела B. Цю модель зображено на рис. 2.4, за винятком елемента b_l із B, де $l = k + 1$. На рисунку виділено напрямки взаємооднозначної відповідності $A <=> B$.

Як згадувалося вище, завади порушують вибір повідомлень b_j джерелом B. За цих обставин, якщо джерелом A вибрано певне повідомлення a_i, якому має відповідати повідомлення b_j при

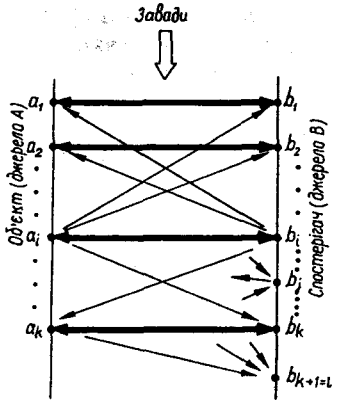


Рис. 2.4

$j = i$, то джерело B може вибрати будь-яке повідомлення b_j при $j = 1 \dots k$ з імовірністю $p(b_j/a_i)$, причому умовна ймовірність правильного вибору (тут воно має смисл перетворення $a_i \to b_i$) дорівнює $p(b_i/a_i)$. Решта виборів будуть помилковими з умовними ймовірностями $p(b_j/a_i)$ при $j \neq i$. Це дає змогу дослідити систему спостереження, передачі, перетворення, збирання, збереження інформації з боку спостерігача B.

Водночас таку систему можна дослідити з позицій спостерігача B. Для цього потрібно знати факт вибору джерелом B певного повідомлення b_j. Після встановлення такого факту можна з умовною ймовірністю $p(a_i/b_j)$ сказати, що об'єкт A знаходиться в стані a_i. Це твердження може бути правильним при $i = j$ або неправильним при $i \neq j$.

Існують певні ситуації, за яких систему доводиться ускладнювати, вибираючи $l > k$ (частіше $l = k + 1$, див. рис. 2.4). При цьому повідомлення b_l не відповідає ніякому повідомленню a_i ($i \neq l$), а є ознакою та повідомленням про особливий стан спостерігача B, який відмовляється від вибору якогось певного повідомлення b_j ($j = 1 \dots k$) у відповідь на повідомлення a_i ($i = 1 \dots k$). Цей стан спостерігача називається *стиранням повідомлення*. Відбувається воно з умовною ймовірністю $p(b_l/a_i)$, якщо стан a_i джерела A відомо. Якщо ж, навпаки, відомим є стан стирання повідомлення b_l джерела B, то з умовною ймовірністю $p(a_i/b_l)$ можна вважати, що джерело A знаходилося в стані a_i. Коли ж $i = 1 \dots k$ та $j = 1 \dots k$, все відбувається аналогічно вищерозглянутій моделі без $b_l = b_{k+1}$ у складі джерела B, тобто коли $l = k$.

Згадаємо ще також про те, що крім статистичного зв'язку між парами повідомлень у довгих послідовностях (рознесення в часі) або в складених системах (рознесення в просторі) часто спостерігається статистичний зв'язок між трьома, чотирма й більше повідомленнями, що зумовлює наявність умовної ентропії більш високого порядку [3, 42].

Тепер повернемося знову до часткової умовної ентропії і будемо дотримуватися моделі системи спостереження, показаної на рис. 2.4 без стирання, тобто коли $l = k$. Звернемося до фізичного змісту умовної ентропії. Як і при визначенні безумовної ентропії, часткова умовна ентропія є математичним сподіванням значення $p(a_i/b_j)$ або $p(b_j/a_i)$. Це поняття відбиває середню за повним алфа-

вітом кількість інформації, що припадає на одне повідомлення цього алфавіту (джерела). Тому $H(A/b_j)$ в (2.17) є питомою кількістю інформації джерела A за умови, що вже встановлено, факт вибору джерелом B певного повідомлення b_j; $H(B/a_i)$ в (2.18) — питомою кількістю інформації джерела B за умови, що вже відомо стан джерела A. Іншими словами, $H(A/b_j)$ — це середня кількість інформації, яка містилася в будь-якому повідомленні джерела A, якщо джерело B вибрало повідомлення b_j ($j = 1 \dots k$), а $H(B/a_i)$ — середня кількість інформації, здобутої після вибору джерелом B будь-якого свого повідомлення, коли відомо, що джерело A знаходилося в стані a_i ($i = 1 \dots k$). Часткова умовна ентропія визначається як статистичне усереднення за методом зваженої суми (2.17), (2.18) по індексу $i = 1 \dots k$ або $j = 1 \dots k$ відповідно.

Ураховуючи викладене вище, загальну умовну ентропію можна визначити так: якщо часткова умовна ентропія джерела A відносно конкретного повідомлення b_j дорівнює $H(A/b_j)$, а розподіл імовірностей P_B джерела B задано ансамблем B із $P_B = \{p(b_1), \dots, p(b_j), \dots, p(b_k)\}$, то цілком природно обчислити середнє по j значення $H(A/b_j)$ за всіма j як статистичне усереднення методом зваженої суми, тобто

$$H(A/B) = \sum_j p(b_j) H(A/b_j), \tag{2.19}$$

де $H(A/B)$ — загальна умовна ентропія джерела A відносно джерела B. Це питома (середньостатична) кількість інформації, що припадає на будь-яке повідомлення джерела A, якщо відомо його статистичну взаємозалежність із джерелом B.

Аналогічно (2.19) загальна умовна ентропія джерела B відносно джерела A визначається виразом

$$H(B/A) = \sum_i p(a_i) H(B/a_i), \tag{2.20}$$

що є питомою (середньостатичною) кількістю інформації, яка припадає на будь-яке повідомлення джерела B, якщо відомо його статистичну взаємозалежність із джерелом A.

Вирази (2.19) і (2.20) є операціями згортки за індексом, внаслідок якої операція (2.19) «поглинає» індекс j, а операція (2.20) — індекс i.

З урахуванням (2.13), (2.17), (2.18) вирази (2.19) і (2.20) набувають вигляду

$$\begin{aligned} H(A/B) &= -\sum_j p(b_j) \sum_i p(a_i/b_j) \log p(a_i/b_j) = \\ &= -\sum_j \sum_i p(b_j) p(a_i/b_j) \log p(a_i/b_j) = \\ &= -\sum_j \sum_i p(a_i/b_j) \log p(a_i/b_j); \end{aligned} \tag{2.21}$$

$$H(B/A) = -\sum_i p(a_i) \sum_j p(b_j/a_i) \log p(b_j/a_i) =$$
$$= -\sum_i \sum_j p(a_i) p(b_j/a_i) \log p(b_j/a_i) = \qquad (2.22)$$
$$= -\sum_i \sum_j p(a_i/bj) \log p(b_j/a_i).$$

Загалом статистична залежність джерела B від джерела A відбивається матрицею прямих переходів повідомлень a_i ($i = 1 \dots k$) джерела A в повідомлення b_j ($j = 1 \dots k$) джерела B:

$P(B/A) =$

A	B					
	b_1	b_2	…	b_j	…	b_k
a_1	$\underline{p(b_1/a_1)}$	$p(b_2/a_1)$	…	$p(b_j/a_1)$	…	$p(b_k/a_1)$
a_2	$p(b_1/a_2)$	$\underline{p(b_2/a_2)}$	…	$p(b_j/a_2)$	…	$p(b_k/a_2)$
…	…	…	…	…	…	…
a_i	$p(b_1/a_i)$	$p(b_2/a_i)$	…	$\underline{p(b_j/a_i)}$	…	$p(b_k/a_i)$
…	…	…	…	…	…	…
a_k	$p(b_1/a_k)$	$p(b_2/a_k)$	…	$p(b_j/a_k)$	…	$\underline{p(b_k/a_k)}$

(2.23)

На головній діагоналі цієї матриці розміщено умовні ймовірності прямої відповідності типу $a_1 \to b_1, \dots, a_k \to b_k$, які характеризують правильний вибір джерелом B своїх повідомлень (тобто відповідно до повідомлень джерела A). Решта ймовірностей відповідають неправильному вибору повідомлень джерелом B.

Матриця (2.23) відбиває вплив завад у каналі між джерелом A та спостерігачем B (див. рис. 2.1). Якщо завади непомітні або зовсім відсутні, то маємо однозначну відповідність $a_i \to b_i$ з умовної ймовірності $p(b_i/a_i) = 1$ для $i = 1 \dots k$. Решта ймовірностей $p(b_j/a_i) = 0$ для всіх $j \neq i$.

Кожний рядок у (2.23) є спотвореним розподілом імовірностей P_B появи повідомлення $b_j \in B$. Джерело B має розподіл безумовних імовірностей P_B. Врахування статистичного впливу повідомлення $a_i \in A$ спотворює цей розподіл (або уточнює його) і дає новий розподіл імовірностей $P(B/a_i)$ для i-го рядка матриці. Саме тому виконується закон нормування

$$\sum_j p(b_j/a_i) = 1, \quad i = 1 \dots k. \qquad (2.24)$$

Статистична залежність джерела A від джерела B відбивається матрицею зворотних переходів типу $a_i \leftarrow b_j$, складеною з умовних імовірностей $p(a_i/b_j)$:

$$P(A/B) = \begin{bmatrix} p(a_1/b_1) & p(a_1/b_2) & \dots & p(a_1/b_j) & \dots & p(a_1/b_k) \\ p(a_2/b_1) & p(a_2/b_2) & \dots & p(a_2/b_j) & \dots & p(a_2/b_k) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p(a_i/b_1) & p(a_i/b_2) & \dots & p(a_i/b_j) & \dots & p(a_i/b_k) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p(a_k/b_1) & p(a_k/b_2) & \dots & p(a_k/b_j) & \dots & p(a_k/b_k) \end{bmatrix} =$$

$= P(A/b_j)$, $i = 1 \dots k$. (2.25)

Вона відображає k розміщених стовпцями варіантів спотворених первинних розподілів імовірностей ансамблю A, які відчувають на собі статистичний вплив повідомлень b_j ($j = 1 \dots k$), вибраних джерелом B. Саме тому для кожного такого розподілу

$$\sum_i p(a_i/b_j) = 1, j = 1 \dots k. \qquad (2.26)$$

Таким чином, якщо задано ансамбль A та матрицю прямих переходів (2.23), то, враховуючи (2.13), можна знайти сумісні умовні ймовірності $p(a_i, b_j) = p(b_j, a_i)$, взявши з ансамблю A безумовні ймовірності $P_A = \{p(a_i)\}$, $i = 1 \dots k$. Виконавши згортку по i, дістанемо такий розподіл безумовних імовірностей $\{p(b_j)\}$, $j = 1 \dots k$:

$$p(b_j) = \sum_{i=1} p(a_i b_j), \ j = 1 \dots k. \qquad (2.27)$$

Відси за формулою, подібною до (2.8), визначаємо безумовну ентропію

$$H(B) = -\sum_{j=1}^{k} p(b_j) \log p(b_j). \qquad (2.28)$$

Звичайно, цю ентропію можна знайти безпосередньо за (2.8), скориставшись значенням $p(a_i)$ з ансамблю A. Умовні ентропії як часткові, так і загальні можна визначити за наявними даними $\{p(a_i)\}$, $\{p(b_j/a_i)\}$ та $p(a_i b_j)$. Взагалі треба виділити розподіл сумісних імовірностей $P(A, B) = \{p(a_i, b_j)\}$, за яким можна дістати розподіл P_B згорткою по i (2.27), або розподіл P_A згорткою за j:

$$p(a_i) = \sum_{j=1}^{k} p(a_i b_j), \ i = 1 \dots k, \qquad (2.29)$$

після чого легко знайти розподіли умовних імовірностей для матриць (2.23) та (2.25) на підставі (2.14). Маючи ці дані, можна визначити всі розглянуті різновиди ентропії: $H(A)$ за (2.8), $H(B)$ за (2.28), $H(A/b_j)$ за (2.17), $H(B/a_i)$ за (2.18), $H(A/B)$ за (2.21) і $H(B/A)$ за (2.22).

На завершення розглянемо конкретний приклад.

Нехай задано таку матрицю $P(a_i, b_j)$ сумісних імовірностей появи повідомлень двох джерел A та B (див. рис. 2.1) з алфавітами a_i та b_j, де $i = 1, 2, 3$ та $j = 1, 2, 3$ ($k = l = 3$):

$$P(a_i,b_j)=\begin{matrix}\begin{bmatrix}0{,}2 & 0 & 0{,}1\\ 0 & 0{,}2 & 0{,}1\\ 0{,}1 & 0{,}1 & 0{,}2\end{bmatrix} & \begin{matrix}j\\1\\2.\\3\end{matrix}\\ \underbrace{\begin{matrix}1 & 2 & 3\end{matrix}}_{i} & \end{matrix} \qquad (2.30)$$

Оскільки $p(ab)$ — безумовні ймовірності, для них виконується закон нормування

$$\sum_{i=1}^{3}\sum_{j=1}^{3} p(a_i b_j) = 1.$$

Окрема розгортка по рядках (по i) без зміни j «поглинає» алфавіт a_i і дає безумовну ймовірність $p(b_j)$, тобто ймовірнісну міру для джерела B:

$$\begin{aligned} &0{,}2 + 0 + 0{,}1 = 0{,}3 = p(b_1) \text{ при } j = 1;\\ &0 + 0{,}2 + 0{,}1 = 0{,}3 = p(b_2) \text{ при } j = 2;\\ &0{,}1 + 0{,}1 + 0{,}2 = 0{,}4 = p(b_3) \text{ при } j = 3. \end{aligned} \qquad (2.31)$$

Перевіримо нормування:

$$p(b_1) + p(b_2) + p(b_3) = 0{,}3 + 0{,}3 + 0{,}4 = 1.$$

Окрема згортка по стовпцях (за j) без зміни i «поглинає» алфавіт b_j і дає безумовну ймовірність $p(a_i)$, тобто ймовірнісну міру для джерела A:

$$\begin{aligned} &0{,}2 + 0 + 0{,}1 = 0{,}3 = p(a_1) \text{ при } i = 1;\\ &0 + 0{,}2 + 0{,}1 = 0{,}3 = p(a_2) \text{ при } i = 2;\\ &0{,}1 + 0{,}1 + 0{,}2 = 0{,}4 = p(a_3) \text{ при } i = 3. \end{aligned} \qquad (2.32)$$

Перевіряємо нормування:

$$p(a_1) + p(a_2) + p(a_3) = 0{,}3 + 0{,}3 + 0{,}4 = 1.$$

З урахуванням (2.14) на підставі (2.30) дістаємо таку матрицю умовних імовірностей:

$$p(a_i/b_j)=\begin{matrix}\begin{bmatrix}\frac{0{,}2}{0{,}3} & \frac{0}{0{,}3} & \frac{0{,}1}{0{,}3}\\ \frac{0}{0{,}3} & \frac{0{,}2}{0{,}3} & \frac{0{,}1}{0{,}3}\\ \frac{0{,}1}{0{,}4} & \frac{0{,}1}{0{,}4} & \frac{0{,}2}{0{,}4}\end{bmatrix} & =\begin{bmatrix}2/3 & 0 & 1/3\\ 0 & 2/3 & 1/3\\ 0{,}25 & 0{,}25 & 0{,}5\end{bmatrix} & \begin{matrix}j\\1\\2.\\3\end{matrix}\\ \underbrace{\begin{matrix}1 & 2 & 3\end{matrix}}_{i} & & \end{matrix} \qquad (2.33)$$

Перевірка на виконання закону нормування (2.26) дає позитивну відповідь, тобто

$$\frac{2}{3}+0+\frac{1}{3}=\frac{3}{3}=1=p(a_1/b_1)+p(a_2/b_1)+p(a_3/b_1) \text{ при } j=1;$$

$$0+\frac{2}{3}+\frac{1}{3}=\frac{3}{3}=1=p(a_1/b_2)+p(a_2/b_2)+p(a_3/b_2) \text{ при } j=2;$$

$$0{,}25+0{,}25+0{,}5=\frac{3}{3}=1=p(a_1/b_3)+p(a_2/b_3)+p(a_3/b_3) \text{ при } j=3.$$

Отже, кожний рядок (2.33) є розподілом повідомлень a_i, спотвореним через статистичний вплив повідомлень $b_j \in B$ [пор. із (2.32)]:

$$p(b_j/a_i)=\begin{matrix}\begin{bmatrix}\frac{0{,}2}{0{,}3} & \frac{0}{0{,}3} & \frac{0{,}1}{0{,}4}\\ \frac{0}{0{,}3} & \frac{0{,}2}{0{,}3} & \frac{0{,}1}{0{,}4}\\ \frac{0{,}1}{0{,}3} & \frac{0{,}1}{0{,}3} & \frac{0{,}2}{0{,}4}\end{bmatrix} & \begin{matrix}j\\1\\2\\3\end{matrix} & =\begin{bmatrix}2/3 & 0 & 0{,}25\\ 0 & 2/3 & 0{,}25\\ 1/3 & 1/3 & 0{,}5\end{bmatrix}.\\ \underbrace{\begin{matrix}1 & 2 & 3\end{matrix}}_{i} & & \end{matrix} \qquad (2.34)$$

Перевірка виконання закону нормування (2.24) дає позитивну відповідь, тобто

$$\begin{aligned} &2/3 + 0 + 1/3 = 3/3 = 1 = p(b_1/a_1) + p(b_2/a_1) + p(b_3/a_1) \text{ при } i = 1;\\ &0 + 2/3 + 1/3 = 3/3 = 1 = p(b_1/a_2) + p(b_2/a_2) + p(b_3/a_2) \text{ при } i = 2;\\ &0{,}25 + 0{,}25 + 0{,}5 = 1 = p(b_1/a_3) + p(b_2/a_3) + p(b_3/a_3) \text{ при } i = 3. \end{aligned}$$

Таким чином, кожний стовпець (2.34) є розподілом повідомлень b_j, спотвореним через статистичний вплив повідомлень $a_i \in A$ [пор. із (2.31)].

Тепер, маючи дані (2.32), за (2.8) розрахуємо безумовну ентропію джерела A:

$$H(A) = -(0{,}3 \log 0{,}3 + 0{,}3 \log 0{,}3 + 0{,}4 \log 0{,}4) = 1{,}57095 \text{ біт/повідомлення.}$$

Аналогічно, маючи дані (2.31), за (2.28) обчислимо безумовну ентропію джерела B:

$$H(B) = -(0{,}3 \log 0{,}3 + 0{,}3 \log 0{,}3 + 0{,}4 \log 0{,}4) = 1{,}57095 \text{ біт/повідомлення.}$$

У розглядуваному прикладі $H(B) = H(A)$ тому, що збігаються розподіли безумовних імовірностей. До речі, максимального значення $H(A)$ та $H(B)$ досягли б при рівноймовірному розподілі ймовірностей, тобто при $p(a_i) = 1/3$ та $p(b_j) = 1/3$ для $i, j \in \{1, 2, 3\}$.

Тоді було б

$$H_{max}(A) = H_{max}(B) = -\log(1/k) = \log k = \log 3 = 1{,}58496 \text{ біт/повідомлення.}$$

Як бачимо, максимальна ентропія перевищує значення безумовної ентропії при нерівномірному розподілі ймовірностей.

Часткова умовна ентропія джерела A за (2.17) з урахуванням (2.33) становитиме

$$H(A/b_1) = -[(2/3) \log (2/3) + 0 \log 0 + (1/3) \log (1/3)] = 0{,}91829 \text{ біт/повідомлення;}$$

$$H(A/b_2) = -[0 \log 0 + (2/3) \log (2/3) + (1/3) \log (1/3)] = \\ = 0{,}91829 \text{ біт/повідомлення};$$

$$H(A/b_3) = -(0{,}25 \log 0{,}25 + 0{,}25 \log 0{,}25 + 0{,}5 \log 0{,}5) = \\ = 1{,}5 \text{ біт/повідомлення}.$$

Загальна умовна ентропія джерела A відносно джерела B згідно з **(2.19)** або (2.21)

$$H(A/B) = 0{,}3\, H(A/b_1) + 0{,}3\, H(A/b_2) + 0{,}4\, H(A/b_3) = \\ = 2 \cdot 0{,}3 \cdot 0{,}91829 + 0{,}4 \cdot 1{,}5 = 1{,}15097 \text{ біт/повідомлення}.$$

Часткова умовна ентропія джерела B за (2.18) з урахуванням **(2.34)** становитиме

$$H(B/a_1) = -[(2/3) \log (2/3) + 0 \log 0 + (1/3) \log (1/3)] = \\ = 0{,}91829 \text{ біт/повідомлення};$$

$$H(B/a_2) = -[0 \log 0 + (2/3) \log (2/3) + (1/3) \log (1/3)] = \\ = 0{,}91829 \text{ біт/повідомлення};$$

$$H(B/a_3) = -(0{,}25 \log 0{,}25 + 0{,}25 \log 0{,}25 + 0{,}5 \log 0{,}5) = \\ = 1{,}5 \text{ біт/ повідомлення}.$$

Загальна умовна ентропія джерела B відносно джерела A згідно з (2.20) або (2.22)

$$H(B/A) = 0{,}3 \cdot 2 \cdot 0{,}91829 + 0{,}4 \cdot 1{,}5 = 1{,}15097 \text{ біт/повідомлення}.$$

Відзначимо окремо такі властивості умовної ентропії:

1. Якщо джерела повідомлень A та B статистично **незалежні,** то умовна ентропія джерела A відносно джерела B дорівнює безумовній ентропії джерела A й навпаки:

$$H(A/B) = H(A);\ H(B/A) = H(B).$$

2. Якщо джерела повідомлень A та B настільки статистично взаємозв'язані, що виникнення одного з повідомлень спричинює безумовну появу іншого, то умовні ентропії їх дорівнюють нулю:

$$H(A/B) = H(B/A) = 0.$$

3. Ентропія джерела взаємозалежних повідомлень (умовна ентропія) менша від ентропії джерела незалежних повідомлень (безумовна ентропія).

4. Максимальну ентропію мають джерела взаємонезалежних рівноймовірних повідомлень, умовна ентропія яких дорівнює нулю, а ймовірність появи символів алфавіту $p = 1/k$, де k — кількість повідомлень в алфавіті.

2.6. ЕНТРОПІЯ ОБ'ЄДНАННЯ ДВОХ ДЖЕРЕЛ

Ентропію $H(A, B)$ об'єднання двох джерел A та B знаходимо через імовірність сумісної появи пар повідомлень a_i, b_j для всіх $i = 1 \ldots k$ та $j = 1 \ldots l$, яку позначимо $p(a_i, b_j)$. Для цього складемо матрицю типу (2.23), що визначає розподіл сумісної безумовної ймовірності двох джерел:

$$p(a_i, b_j) = \begin{bmatrix} p(a_1,b_1) & p(a_2,b_1) & p(a_3,b_1) & \ldots & p(a_k,b_1) \\ p(a_1,b_2) & p(a_2,b_2) & p(a_3,b_2) & \ldots & p(a_k,b_2) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ p(a_1,b_l) & p(a_2,b_l) & p(a_3,b_l) & \ldots & p(a_k,b_l) \end{bmatrix} \quad (2.35)$$

Оскільки $p(a_i, b_j)$ — це ймовірність сумісної появи двох повідомлень, ентропія $H(A, B)$ є середньою кількістю інформації, що припадає на два довільних повідомлення джерел A та B й визначається так:

$$H(A,B) = -\sum_i \sum_j p(a_i, b_j) \log p(a_i, b_j). \quad (2.36)$$

Зрозуміло, що

$$H(A, B) = H(B, A), \quad (2.37)$$

оскільки $p(a_i, b_j) = p(b_j, a_i)$. Для сумісної появи повідомлень a_i та b_j послідовність запису їх не має значення.

Розглянемо докладніше вираз (2.36). Якщо врахувати (2.13), то можна записати

$$H(A,B) = -\sum_i \sum_j p(a_i) p(b_j/a_i) \log[p(a_i) p(b_j/a_i)]. \quad (2.38)$$

Згадавши, що логарифм добутку дорівнює сумі логарифмів **(це** властивість адитивності кількості інформації), матимемо

$$H(A,B) = -\sum_i \sum_j \{p(a_i) p(b_j/a_i)[\log p(a_i) + \log p(b_j/a_i)]\} = \\ = -\{\sum_i \sum_j p(a_i) p(b_j/a_i) \log p(a_i) + \sum_i \sum_j p(a_i) p(b_j/a_i) \log p(b_j/a_i)\} = \\ = -\{\sum_i p(a_i) \log p(a_i) \sum_j p(b_j/a_i) + \sum_i p(a_i) \sum_j p(b_j/a_i) \log p(b_j/a_i)\}. \quad (2.39)$$

Тепер потрібно врахувати, що згідно з (2.24) $\sum_j p(b_j/a_i) = 1$.

Тому перша складова в (2.39) відповідно до (2.8) визначає безумовну ентропію джерела A, тобто

$$-\sum_i p(a_i) \log p(a_i) = H(A),$$

а друга складова відповідно до (2.22) є загальною умовною ентропією $H(B/A)$. Отже,

$$H(A, B) = H(A) + H(B/A). \qquad (2.40)$$

Властивість симетрії (2.37) дає змогу провести подібні перетворення ще раз і дістати інший вираз ентропії $H(A, B)$, а саме:

$$H(A, B) = H(B) + H(A/B). \qquad (2.41)$$

З урахуванням (2.40) і (2.41) загальну умовну ентропію можна визначити порівнянням безумовної ентропії об'єднання джерел та ентропії одного з них — A чи B:

$$\begin{aligned} H(B/A) &= H(A, B) - H(A); \\ H(A/B) &= H(A, B) - H(B). \end{aligned} \qquad (2.42)$$

Кількість інформації, що припадає на одне повідомлення, передане по каналу зв'язку в системі спостереження (див. рис. 2.1) від джерела A спостерігачеві B за наявності завад у каналі та статистичній зумовленості ансамблів A та B [3, 42], визначається виразом

$$\begin{aligned} I(A, B) &= H(A) - H(A/B) = H(B) - H(B/A) = \\ &= H(A) + H(B) - H(A, B). \end{aligned} \qquad (2.43)$$

Наведений вище приклад (див. п. 2.5) дає уяву про ентропію об'єднання. Так, згідно з (2.36) і (2.30) маємо

$$\begin{aligned} H(A, B) &= -(3 \cdot 0{,}2 \log 0{,}2 + 4 \cdot 0{,}1 \log 0{,}1) = \\ &= 2{,}72193 \text{ біт/ два повідомлення.} \end{aligned}$$

З іншого боку, враховуючи (2.40), дістаємо

$$\begin{aligned} H(A, B) &= H(A) + H(B/A) = 1{,}57095 + 1{,}15097 = \\ &= 2{,}72192 \text{ біт/два повідомлення,} \end{aligned}$$

а користуючись (2.41), маємо

$$\begin{aligned} H(A, B) &= H(B) + H(A/B) = \\ &= 2{,}72192 \text{ біт/два повідомлення.} \end{aligned}$$

Відзначимо такі основні властивості ентропії об'єднання двох джерел:

1. При статистично незалежних повідомленнях джерел A та B ентропія об'єднання їх дорівнює сумі ентропії окремих джерел, тобто

$$H(A, B) = H(A) + H(B).$$

2. При повній статистичній залежності джерел A та B ентропія об'єднання їх дорівнює безумовній ентропії одного із джерел, тобто

$$H(A, B) = H(A) = H(B).$$

3. Ентропія об'єднання двох джерел відповідає нерівності

$$H(A, B) \le H(A) + H(B).$$

КОНТРОЛЬНІ ЗАДАЧІ

1. Відомо, що повідомлення $a_i \in A$ з'являється з імовірністю $p_i = 0{,}03$. Визначити кількість інформації, що міститься в цьому повідомленні.

2. Ансамбль C містить 16 рівноймовірних повідомлень. Визначити кількість інформації, яку містить кожне таке повідомлення.

3. Джерело A виробляє трилітерне повідомлення $a_i \in A$ з алфавіту $\{a, b, c, d\}$, вибираючи їх рівноймовірно й незалежно одне від одного за типом $a_1 = abc$, $a_2 = abd$ і т. д. Визначити ентропію цього джерела.

4. Джерела A та B мають розподіли ймовірностей повідомлень $P_A = \{0{,}1; 0{,}1; 0{,}15; 0{,}125; 0{,}125; 0{,}1; 0{,}15; 0{,}15\}$ і $P_B = \{0{,}5; 0{,}3; 0{,}1; 0{,}025; 0{,}025; 0{,}02; 0{,}015; 0{,}015\}$. Ентропія якого джерела більша? Яка максимальна ентропія цього джерела та за якої умови?

5. Визначити ентропію монітора персональної ЕОМ при виведенні тексту в 28 рядків по 60 рівноймовірних символів у кожному, якщо використовується стандартний міжнародний код (128 символів) із двома градаціями яскравості.

6. При передачі банківської інформації реченнями по 16 рядків на кожні 100 речень цифра 5 зустрічається 90, 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 появи цифри 5, якщо в реченні з'явилася цифра 9.

7. 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.

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

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

Визначити часткову та загальну умовну ентропію повідомлень в цьому каналі при рівноймовірному виборі їх джерелом A та при $P_A = \{0{,}65; 0{,}1; 0{,}05\}$.

9. Два статистично незалежних джерела A та B визначаються матрицею сумісних імовірностей

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

Визначити часткову та загальну умовну ентропію, ентропію об'єднання, безумовну ентропію цих джерел, а також кількість інформації, що припадає на пару повідомлень a_i, b_j.

10. Розв'язати попередню задачу, якщо матриця сумісних імовірностей джерел має вигляд

$$p(a_i, b_j) = \begin{bmatrix} 0{,}2 & 0{,}01 & 0{,}02 & 0{,}03 \\ 0{,}02 & 0{,}16 & 0{,}03 & 0{,}01 \\ 0{,}01 & 0{,}04 & 0{,}17 & 0{,}02 \\ 0{,}03 & 0{,}05 & 0{,}1 & 0{,}1 \end{bmatrix}.$$

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що таке джерело повідомлень?
2. Що таке ансамбль повідомлень?
3. Як визначається кількість інформації в одному повідомленні?
4. Що таке ентропія та які її властивості?
5. Що таке безумовна ентропія?
6. Що таке умовна ентропія?
7. Які основні властивості безумовної ентропії?
8. Які є різновиди умовної ентропії?
9. Які основні властивості умовної ентропії?
10. Як визначається часткова умовна ентропія?
11. Як визначається загальна умовна ентропія?
12. Як визначається ентропія об'єднання двох джерел?
13. Які основні властивості ентропії об'єднання двох джерел?
14. Як визначається кількість інформації на одне повідомлення двох статистично взаємозв'язаних джерел?
15. За яких умов ентропія джерела стає максимальною?

РОЗДІЛ 3 ХАРАКТЕРИСТИКИ ДИСКРЕТНИХ ДЖЕРЕЛ ІНФОРМАЦІЇ

Будь-яку інформацію потрібно доставити одержувачеві повідомлень, інакше вона не матиме для нього ніякої цінності. Проте при передачі інформації від дискретних джерел по каналах зв'язку можуть виникати її втрати, якщо в каналах діють завади та є спотворення сигналів, якими передаються повідомлення.

Від інтенсивності завад у каналах залежать швидкість передачі інформації та пропускна здатність каналів. Тому в цьому розділі йтиметься про швидкість передачі інформації від дискретних джерел, оцінювання втрат її в каналах, пропускну здатність дискретного каналу, теореми Шеннона про кодування дискретних джерел.

3.1. ПРОДУКТИВНІСТЬ ДИСКРЕТНОГО ДЖЕРЕЛА ТА ШВИДКІСТЬ ПЕРЕДАЧІ ІНФОРМАЦІЇ

Якщо джерело A вибирає повідомлення a_i, то можна говорити, що воно виробляє певну кількість інформації $I_i = -\log p_i$. Її означення за різних умов розглядалося в попередньому розділі. Там же згадувалось про можливість залучення часового вирівнювання до моделі джерела. Розглянемо трохи складнішу його модель, а саме: ансамбль A повідомлень, урахувавши розподіл імовірностей p_i та розподіл проміжків часу τ_i, протягом яких джерело вибирає різні повідомлення a_i.

Продуктивність джерела щодо певного повідомлення a_i можна визначити як

$$V_{джі} = I_i / \tau_i. \qquad (3.1)$$

Її одиниця залежить від вибору одиниці кількості інформації I_i. Наприклад, це може бути біт за секунду.

Як правило, джерело вибирає досить велику кількість повідомлень протягом певного часу. Тому природно як загальну характеристику джерела прийняти середню за ансамблем його

продуктивність, користуючись відомим з попереднього розділу методом статистичного усереднення:

$$V_{\text{дж}} = \sum_{i=1}^{k} p_i V_{\text{дж}\,i}. \qquad (3.2)$$

Загалом, коли $\tau_i \neq \tau_j$ при $i \neq j$ вираз (3.1) після усереднення за часом можна перетворити до такого вигляду:

$$V_{\text{дж}\,i} = I_i / (\sum_{i=1}^{k} p_i \tau_i) = I_i / \tau_{\text{сер}}, \qquad (3.3)$$

де $\tau_{\text{сер}}$ — середній час вибору джерелом одного повідомлення.

Вираз (3.3) дійсний також для більш поширеного випадку, коли джерело вибирає всі свої повідомлення за один і той самий проміжок часу $\tau = \tau_{\text{сер}}$.

З урахуванням (3.3) вираз (3.2) набуває вигляду:

$$V_{\text{дж}} = \frac{1}{\tau_{\text{сер}}} \sum_{i=1}^{k} p_i I_i = -\frac{1}{\tau_{\text{сер}}} \sum_{i=1}^{k} p_i \log p_i = \frac{H(A)}{\tau_{\text{сер}}}. \qquad (3.4)$$

Бачимо, що така загальна характеристика, як продуктивність дискретного джерела, визначається його середніми показниками: ентропією (середньою кількістю інформації в одному повідомленні) та часом утворення останнього.

Розглянемо узагальнену модель каналу передачі інформації (рис. 3.1). Зазначимо, що передача можлива як у просторі, так і в часі. Це залежить від мети передачі та вибраних носіїв інформації в лінії зв'язку. Виходячи з наведеного в п. 2.1 визначення маємо джерело A з ентропією $H(A)$, утворене спостерігачем A та спостережуваним фізичним процесом A.

Повідомлення $a_i \in A$ певної форми передаються лінією зв'язку, де на них впливають завади, що можуть спотворювати їх. Внаслідок цього утворюється новий фізичний процес B, який спостерігається на виході лінії. Сама лінія при цьому розглядається як фізичний об'єкт B. Роль спостерігача B тут відіграє приймальний пристрій, що вибирає повідомлення b_j з алфавіту B відповідно до фізичного процесу B. Таким чином, утворюється джерело B зі своєю безумовною ентропією $H(B)$.

Вибір повідомлень $b_j \in B$ характеризує при цьому процес передавання інформації по каналу зв'язку від джерела A на вихід джерела B. Позначимо через $I(A, B)$ середню кількість інформації про стан джерела A, яка міститься в одному повідомленні джерела B. Якщо на вибір кожного повідомлення витрачається

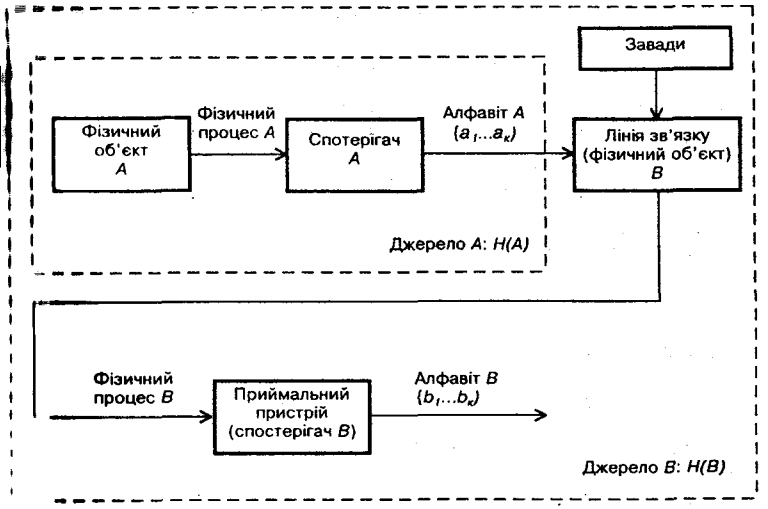


Рис. 3.1

час τ, то питома кількість переданої інформації є швидкістю її передачі по каналу, тобто

$$V = I'(A, B) = I(A, B) / \tau. \qquad (3.5)$$

3.2. ІНФОРМАЦІЙНІ ВТРАТИ ПРИ ПЕРЕДАЧІ ІНФОРМАЦІЇ ПО ДИСКРЕТНОМУ КАНАЛУ

Задача дискретного каналу полягає в тому, щоб повідомлення b_i однозначно відповідало повідомленню a_i. Це було б можливе тоді, коли $p(a_i/b_i) = p(b_i/a_i) = 1$ для всіх $i = 1 \ldots K$. А це означає, що мають виконуватися рівності $p(a_i, b_i) = p(a_i) = p(b_i)$ та $p(a_i/b_j) = p(b_j/a_i) = 0$ або $p(a_i, b_j) = 0$ для всіх $j \neq i$. Цей випадок стосується каналу з малими завадами, які не можуть спотворити повідомлення a_i джерела A так, щоб спостерігач B помилився під час розпізнавання стану джерела B та вибору повідомлення b_i.

Наведені вище умови означають повний збіг ансамблів A та B (нагадаємо, що до ансамблю входять множина повідомлень і множина імовірностей їх). А якщо це так, то середня кількість переданої інформації на одне повідомлення $H(A)$ при повній відсутності інформаційних втрат дорівнює такій самій кількості прийнятої інформації $H(B)$, тобто

$$I(A, B) = H(A) = H(B) = H(A, B). \qquad (3.6)$$

Остання рівність випливає з того, що за наведених умов

$$H(A/B) = H(B/A) = 0. \qquad (3.7)$$

Таким чином, кількість переданої інформації за відсутності завад дорівнює ентропії об'єднання джерел A та B або безумовній ентропії одного з них.

У разі повної статистичної незалежності джерел A та B, що характеризує високий рівень завад, коли повідомлення b_i ніяк статистично незумовлене повідомленням a_i, маємо (див. п. 2.5)

$$H(A/B) = H(A);\ H(B/A) = H(B). \qquad (3.8)$$

При цьому ентропія об'єднання двох джерел становитиме

$$H(A, B) = H(B, A) = H(A) + H(B). \qquad (3.9)$$

У разі статистичної незалежності джерел A та B ніяка інформація від A до B через її повне спотворення не передається. Інформаційною мірою цього спотворення є умовна ентропія одного джерела відносно іншого, яка збільшується від нуля згідно з (3.7) до максимуму згідно з (3.8) у міру зростання статистичної зумовленості джерел A та B.

У проміжному випадку неабсолютної статистичної залежності джерел A та B завади деякою мірою спотворюють повідомлення, що статистичне відбивається у вигляді матриць імовірностей перехресних переходів (див. п. 2.5). При цьому умовна ентропія має певні рівні:

$$0 \le H(A/B) \le H(A);\ 0 \le H(B/A) \le H(B). \qquad (3.10)$$

Ураховуючи викладене, кількість інформації, що передається в каналі, можна визначити так. Якщо джерело A вибрало певне повідомлення, то воно виробляє кількість інформації, що дорівнює $H(A)$. Джерело B, вибравши повідомлення $b_j \in B$, за умови порушення повної статистичної залежності джерел A та B виробляє певну кількість інформації про джерело A, що міститься в джерелі B й дорівнює $H(A/B)$.

Спостерігач B (приймальний пристрій), вибравши повідомлення $b_j \in B$, приймає також рішення про повідомлення $a_i \in A$, яке передавалося, що є одним із його завдань. Прийнявши таке рішення, він виробляє кількість інформації про джерело A, яка дорівнює $H(A)$. Проте перед цим він уже одержав $H(A/B)$ бітів інформації про це джерело, тому кількість переданої в каналі інформації про джерело A як кількість нового відсутнього знання визначається різницею

$$I(A, B) = H(A) - H(A/B). \qquad (3.11)$$

Вираз (3.11) збігається повністю з (3.6) за умови (3.7) при малих завадах і з урахуванням того, що за умови статистичної незалежності (3.8) джерел (великі завади в каналі)

$$I(A, B) = 0. \qquad (3.12)$$

Враховано, що $H(A, B) = H(B, A)$; тому

$$H(A) + H(B/A) = H(B) + H(A/B). \qquad (3.13)$$

Знак рівності тут не зміниться, якщо від обох частин (3.13) відняти суму $H(B/A) + H(A/B)$, тобто

$$H(A) - H(A/B) = H(B) - H(B/A), \qquad (3.14)$$

звідки

$$I(A, B) = I(B, A). \qquad (3.15)$$

Таким чином, інформаційні втрати при передачі інформації в каналі визначаються умовною ентропією одного джерела відносно іншого, а кількість переданої інформації — безумовною ентропією джерела та інформаційними втратами за виразом (3.11) або (3.15).

З урахуванням (3.11), (3.13) і (3.15) можна записати

$$I(A, B) = H(A) - H(A/B) = H(A) - [H(A, B) - H(B)] = H(A) + H(B) - H(A, B); \qquad (3.16)$$
$$I(B, A) = H(B) - H(B/A) = H(B) - [H(B, A) - H(A)] = H(B) + H(A) - H(B, A).$$

Спостерігається повна симетрія цих виразів.

3.3. ПРОПУСКНА ЗДАТНІСТЬ ДИСКРЕТНОГО КАНАЛУ

Максимально можлива швидкість передачі інформації по каналу зв'язку називається його *пропускною здатністю* C. Виходячи з виразу (3.5) маємо

$$C = \frac{1}{\tau}[I(A,B)]_{\max} = \frac{1}{\tau}[H(A) - H(A/B)]_{\max}. \qquad (3.17)$$

Очевидно, вираз (3.17) досягає максимуму при абсолютному статистичному зв'язку джерел A та B, коли виконуються рівності (3.6) і (3.7). Це випадок малого рівня або відсутності спотворюючих завад. Тоді

$$C = \frac{1}{\tau}H(A)_{\max}. \qquad (3.18)$$

Із п. 2.4 відомо, що безумовна ентропія джерела досягає максимуму при рівноймовірних і статистично незалежних повідомленнях a_i, $i = 1 \ldots k$. При цьому $p(a_i) = 1/k$ для всіх i та $H(A)_{\max} = \log_2 k$. Отже,

$$C = \frac{1}{\tau}\log_2 k. \qquad (3.19)$$

Цей вираз і визначає пропускну здатність каналу без завад.

Якщо в каналі є відчутні завади, а умовна ентропія на його вході та виході лежить у діапазоні (3.10), то пропускна здатність каналу із завадами визначається виразом

$$C = \frac{1}{\tau}[\log_2 k - H(A/B)]. \qquad (3.20)$$

При зменшенні завад цей вираз прямує до (3.19), а при їх збільшенні – до нуля.

3.4. ТЕОРЕМА ШЕННОНА ПРО КОДУВАННЯ ДИСКРЕТНОГО ДЖЕРЕЛА

Головна теорема інформації, сформульована та доведена К. Шенноном [44], називається *теоремою кодування дискретного джерела*. Пізніше з'явилося багато її модифікацій для різних обумовленостей та ситуацій [3, 13, 28, 29, 35, 37, 38]. Нижче цю теорему та її модифікацію подамо без доведення, яке можна знайти в [3, 42], і коротко прокоментуємо їх.

Нехай джерело повідомлень має ансамбль $A = \{a_i\}$, $i = \overline{1, l}$ та $P = \{p_i\}$, $i = \overline{1, l}$. Його безумовна ентропія дорівнює $H(A)$, а продуктивність — $V_{дж} H(A)$, де $V_{дж} = T_{дж}^{-1}$ — кількість повідомлень джерела за одиницю часу.

Нехай канал без завад має алфавіт $B = \{b_j\}$, $j = \overline{1, k}$ та ентропію $H(B)$. Тоді швидкість передачі інформації по каналу $R_к = V_к H(B)$. Використаємо безнадмірні вхідні повідомлення каналу, які мають максимальну ентропію, що забезпечує максимально можливу швидкість передачі інформації в каналі $C_к = V_к \log_2 k$ (згадаємо, що $H(B) = \log_2 k$). Величина $C_к$ називається *пропускною здатністю каналу* з параметрами k та $V_к = T_к^{-1}$, де $T_к$ — тривалість передачі одного символу каналу.

Якщо взагалі надмірність повідомлень джерела не дорівнює нулю, то $R_к < C_к$. Як довів К. Шеннон, відповідним добором способу кодування при будь-якій надмірності джерела A можна забезпечити швидкість передачі інформації по каналу без завад як завгодно близьку до його пропускної здатності $C_к$.

ним чином, умовою узгодження джерела та каналу є відповід-
ь продуктивності першого пропускній здатності другого.
(ля доведення цього твердження К. Шеннон використав
яття типових і нетипових послідовностей повідомлень. Не-
джерело A виробляє статистично незалежні повідомлення
$\overline{1, l}$ (це не має особливого значення, але спрощує роз-
). Розглянемо деякі властивості довгих послідовностей по-
омлень джерела A. В бінарному випадку (алфавіт складає-
я з двох повідомлень) імовірність того, що в послідовності
буде t повідомлень a_1 і $n - t$ повідомлень a_2, визначається бі-
номним законом розподілу імовірностей

$$p_t = C_n^t p_1^t (1 - p_1)^{n-t},$$

де $C_n^t = n!/[t!(n-t)!]$ — кількість різних повідомлень із t елементами a_1 і $n - t$ елементами a_2; p_1 — імовірність появи повідомлення a_1; $1 - p_1 = p_2$ — ймовірність появи повідомлення a_2.

Зі збільшенням n значення t і $n - t$ в кожній реалізації довгої
підовності прямуватимуть до своїх математичних сподівань,
орівнюють np_1 і $n(1 - p_1)$ відповідно. Саме такі послідов-
и були названі *типовими*, оскільки незалежно від взаємного
ташування повідомлень a_1 та a_2 в довгій послідовності кіль-
кість їх буде однакова. Послідовність з іншим співвідношен-
ням кількості повідомлень a_1 та a_2 називаються *нетиповими*,
причому ймовірність появи їх із збільшенням n прямує до нуля.

Усі типові послідовності повідомлень мають однакову ймо-
вірність

$$P_т = p_1^{p_1 n} p_2^{p_2 n},$$

якщо n прямує до нескінченності.

Аналогічно для k різних символів алфавіту B на виході кана-
лу (де k збігається з кількістю символів каналу) маємо ансамбль
$B = \{b_j\}$, $j = \overline{1, k}$ та $P = \{p_j\}$, $j = \overline{1, k}$. Тут із збільшенням n кіль-
кість появ кожного символу b_j прямує до свого математичного
сподівання, й саме така довга послідовність символів назива-
ється типовою. В ній незалежно від конкретної реалізації кіль-
кість кожного символу b_j буде однаковою й дорівнює np_j. Ймо-
вірність такої послідовності символів визначається виразом

$$p_т = p_1^{p_1 n} p_2^{p_2 n} \ldots p_k^{p_k n}.$$

Як бачимо, ймовірність $p_т$ не залежить від конкретної реалі-
зації типової послідовності символів, а залежить тільки від ан-
самблю (тобто від обсягу алфавіту та розподілу ймовірностей).
Таким чином, імовірність $p_т$ для всіх типових послідовностей
джерела є величиною сталою.

Кількість різних типових послідовностей символів завдовжки n дорівнює $n!/[(np_1)!\,(np_2)!\ldots(np_k)!]$, що випливає із законів комбінаторики. Всі типові послідовності символів рівноймовірні, а сума ймовірностей появи їх прямує до одиниці (це повна група випадків) при $n \to \infty$. При цьому ймовірність нетипової послідовності прямує до нуля.

Таким чином, імовірність типової послідовності символів може бути подана як $p_т = 1/N_т(A)$, де $N_т(A)$ — кількість типових послідовностей символів джерела.

Розглянуті властивості послідовностей формулюються у вигляді такої теореми [3, 42]: *кожна реалізація довгої послідовності повідомлень стаціонарного джерела A за умови достатньої довжини їх з імовірністю, як завгодно близькою до одиниці, збігається з однією з рівноймовірних типових послідовностей повідомлень.*

Їх рівноймовірність зумовлює максимальне значення ентропії джерела типових послідовностей повідомлень так, що на кожну з них припадає найбільша з можливих кількостей інформації $I_т = \log_2 N_т(A)$. З іншого боку, ентропія джерела A визначається як

$$H(A) = (1 - R_{над}) \log_2 k,$$

де $R_{над}$ — надмірність джерела A.

Це дає змогу визначити $I_т$ як

$$I_т = n\,H(A) = n(1 - R_{над}) \log_2 k,$$

звідки

$$\log_2 N_т(A) = n(1 - R_{над}) \log_2 k.$$

Виходячи з визначення логарифма знаходимо

$$N_т(A) = 2^{n(1 - R_{над}) \log_2 k}.$$

Кількість усіх можливих послідовностей повідомлень завдовжки n, складених з елементів-повідомлень джерела A, становить

$$N(A) = k^n = 2^{n \log_2 k}.$$

Тоді частка типових послідовностей повідомлень серед усіх можливих визначиться як

$$N_т(A)/N(A) = 2^{-nR_{над} \log_2 k}.$$

При надмірності $R_{над} \neq 0$ та $n \to 0$ ця частка прямує до нуля й лише при $R_{над} = 0$ маємо $N_т(A)/N(A) = 1$, тобто в цьому разі кількість типових послідовностей повідомлень збігається з загальною кількістю послідовностей завдовжки n і всі вони рівноймовірні та використовуються для передачі інформації. Таким чином, імовірність появи нетипових послідовностей повідомлень прямує до нуля.

На цьому ґрунтується теорема Шеннона про кодування дискретного джерела, або, як її інакше називають, *теорема кодування дискретного каналу без завад*. Формулюється вона так: *якщо пропускна здатність дискретного каналу без завад перевищує продуктивність джерела повідомлень, тобто*

$$V_к \log_2 k > V_{дж} H(A), \qquad (3.21)$$

то існує спосіб кодування та декодування повідомлень джерела з ентропією $H(A)$, що гарантує як завгодно високу надійність зіставлення прийнятих комбінацій повідомлень з переданими; якщо $V_к \log_2 k < V_{дж} H(A)$, то такого способу немає.

Для доведення теореми всі типові послідовності повідомлень тривалістю T кодують в алфавіті B обсягом k у вигляді кодових комбінацій повідомлень тієї самої тривалості T. Кількість символів на одну комбінацію нового алфавіту, що відображає типову послідовність повідомлень джерела A, становить $n_B = TV_к$. Тоді кількість різних кодових комбінацій повідомлень у новому алфавіті B становитиме

$$N(B) = 2^{TV_к \log 2k}.$$

Кількість же типових послідовностей повідомлень тривалістю T з розрядністю $n = TV_{дж}$ визначається як

$$N_т(A) = 2^{TV_{дж} H(A)}.$$

Умову теореми можна записати у вигляді $TV_к \log_2 k > TV_{дж} H(A)$, або

$$N(B) > N_т(A).$$

Перепишемо цю умову так: $N(B) = N_т(A) + \varepsilon$ або $N(B)/N_т(A) > 2^{T\varepsilon}$, де $\varepsilon > 0$ — як завгодно мала величина. Виберемо $\varepsilon > (\log_2 e)/[TN_т(A)]$; тоді $N(B)/N_т(A) > e^{1/N_т(A)}$ і, врахувавши розвинення в ряд $e^x = \sum_{n=0}^{\infty} x^n/n!$, дістанемо

$$N(B) > N_т(A) + 1.$$

Таким чином, при виконанні умов (3.21) кількість різних кодових комбінацій повідомлень в алфавіті B принаймні на одиницю перевищує кількість типових послідовностей повідомлень джерела A, що забезпечує безпомилкове декодування їх. При невиконанні умови (3.21), коли $V_к \log_2 k < V_{дж} H(A)$, маємо $2^{TV_к \log_2 k} < 2^{TV_{дж} H(A)}$, або

$$N_т(A)/N(B) > 2^{T\varepsilon}.$$

Вибравши $\varepsilon < (\log_2 e)/[TN(B)]$, запишемо

$$N_т(A)/N(B) > e^{1/N(B)} = 1 + 1/N(B) + 1/[2!N^2(B)] + \ldots,$$

звідки випливає, що $N_т(A) > N(B) + 1$, тобто кількість типових послідовностей повідомлень принаймні на одиницю перевищує кількість різних комбінацій коду з максимальною ентропією. Це означає, що навіть при найкращому кодуванні, яке забезпечує однакову ймовірність використання всіх кодових комбінацій на вході каналу, а також максимальну швидкість передачі інформації, неможливо закодувати та передати всі типові послідовності повідомлень $N_т(A)$ джерела. В цьому й полягає доведення розглядуваної теореми.

Питання, пов'язані з точністю передачі повідомлень, оцінкою вірогідності безпомилкової передачі їх, імовірністю помилок при відтворенні повідомлень тут не розглядалися. Про це йтиметься в розд. 9, де ці відомості знадобляться для оцінювання якості конкретних кодів і методів передачі інформації.

КОНТРОЛЬНІ ЗАДАЧІ

1. Ансамбль повідомлень джерела A визначено як $A = \{a_1, a_2, a_3\}$ та $p(a_1) = 0{,}65$; $p(a_2) = 0{,}25$; $p(a_3) = 0{,}1$. Матриця умовних імовірностей каналу має вигляд

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

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

2. Визначити інформаційні втрати в каналі передачі з матрицею умовних імовірностей

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

3. Середня кількість інформації в будь-якому повідомленні $b_j \in B$ дорівнює 2,312 біт. Умовна ентропія на виході B каналу передачі відносно його входу A становить $H(B/A) = 0{,}312$ біт/повідомлення. Визначити кількість інформації, що передається в 10 000 повідомленнях, а також середню швидкість її передачі, якщо на передачу зазначеної кількості повідомлень витрачається 1/3 хв.

4. Матриця сумісних імовірностей каналу передачі має вигляд

$$p(a_i, b_j) = \begin{bmatrix} 0{,}15 & 0{,}15 & 0 \\ 0 & 0{,}25 & 0{,}1 \\ 0 & 0{,}2 & 0{,}15 \end{bmatrix}.$$

Визначити інформаційні втрати в каналі та швидкість передачі інформації, якщо на передачу одного повідомлення витрачається 10^{-3} с.

5. Повідомлення передаються взаємонезалежними рівноймовірними символами тривалістю $5 \cdot 10^{-4}$ с. Визначити швидкість передачі кожного символу та всієї інформації, якщо обсяг алфавіту дорівнює 8, 16, 32.

6. Час передачі повідомлення 0 дорівнює 0,1 с, а повідомлення 1 – 0,6 с. Знайти розподіл імовірностей p_0 та p_1, за яких досягається максимальна швидкість передачі інформації.

7. Визначити продуктивність джерела з ансамблем $A = \{a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8\}$ та $p_i \in \{0{,}1; 0{,}2; 0{,}1; 0{,}15; 0{,}05; 0{,}1; 0{,}2; 0{,}1\}$; $\tau_i \in \{0{,}01; 0{,}001; 0{,}01; 0{,}005; 0{,}008; 0{,}006; 0{,}003; 0{,}001\}$. За яких умов ця продуктивність буде максимальною? Визначити її значення для того самого розподілу τ_i.

8. Канал передачі задано ансамблем $A = \{a_1, a_2, a_3\}$ та $p_i \in \{0{,}3; 0{,}2; 0{,}5\}$. Матриця умовних імовірностей каналу має вигляд

$$p(b_j/a_i) = \begin{bmatrix} 0{,}97 & 0{,}015 & 0{,}015 \\ 0{,}015 & 0{,}97 & 0{,}015 \\ 0{,}015 & 0{,}015 & 0{,}97 \end{bmatrix}.$$

Визначити пропускну здатність каналу при $\tau = 10^{-3}$ с і швидкість передачі інформації.

9. Визначити пропускну здатність каналу, матриця ймовірностей якого при $\tau = 10^{-2}$ с має вигляд

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

10. Чи можлива безпомилкова передача інформації по каналу, параметри якого задано в попередній задачі, якщо продуктивність джерела I'_{a6} 9,6 Кбіт/с?

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Чим визначається продуктивність дискретного джерела?
2. Як можна визначити продуктивність дискретного джерела з різною тривалістю вибору повідомлень?
3. Як визначається швидкість передачі інформації по дискретному каналу?
4. Чому дорівнюють інформаційні втрати при передачі інформації по каналу зв'язку?
5. Чому дорівнюють втрати інформації в каналі з абсолютною статистичною залежністю його виходу та входу?
6. Чому дорівнюють втрати інформації в каналі із статистично незалежними його виходом і входом?
7. Як визначається кількість інформації, що передається в одному повідомленні?
8. Як проілюструвати принцип симетрії для кількості інформації, що передається по каналу?
9. Як визначається пропускна здатність каналу передачі?
10. Як визначається пропускна здатність каналу за відсутності завад?
11. Як формулюється теорема Шеннона про кодування дискретного джерела?
12. У чому полягає зміст теореми Шеннона про кодування дискретного джерела?

РОЗДІЛ 4 ХАРАКТЕРИСТИКИ НЕПЕРЕРВНИХ ДЖЕРЕЛ ІНФОРМАЦІЇ

Не всі повідомлення мають дискретний характер. Є й такі (мовні, телевізійні, факсимільні тощо), характер яких неперервний, на що треба звернути увагу. В цьому плані слід ураховувати втрати при перетворенні неперервних повідомлень на дискретні та інформаційні втрати при кодуванні неперервних джерел, уміючи оцінити швидкість передачі неперервних повідомлень і пропускну здатність неперервного каналу. Про все це йтиметься в цьому розділі.

4.1. КВАНТУВАННЯ СИГНАЛІВ

При передачі повідомлень за допомогою неперервних сигналів стикаються з труднощами, пов'язаними з виникненням апаратурних похибок, а також похибок від нестабільності параметрів ліній і каналів зв'язку та ін. В той же час передача неперервних сигналів з використанням дискретних значень їх дає змогу усунути ці похибки повною регенерацією імпульсів у проміжних пунктах і на приймальному боці системи передачі. Тому при передачі неперервних сигналів їх, як правило, перетворюють на дискретні. Крім зазначеного вище, дискретна форма подання сигналів дає також значні переваги при зберіганні та обробленні інформації.

Процес перетворення неперервних сигналів на дискретні, який називається *квантуванням*, має кілька різновидів. При цьому розрізняють: *дискретизацію* (квантування за часом), *квантування* (квантування за рівнем) і *квантизацію* (квантування за часом і рівнем, або комбіноване квантування).

Дискретизація. Для того щоб краще зрозуміти, як виконується дискретизація неперервних сигналів, розглянемо теорему відліків, яку сформулював В. О. Котельников. Це основна теорема, що доводить можливість передачі неперервних сигналів за допомогою дискретних.

Теорема стверджує: *якщо неперервна в часі функція має обмежений частотний спектр, який не містить складових з частотами, що перевищують F_m, то вона повністю визначається сукупністю своїх миттєвих значень (дискрет), які відлічуються через інтервали часу $\Delta t = (1/2)F_m$, де F_m — максимальна частота спектра неперервного сигналу.*

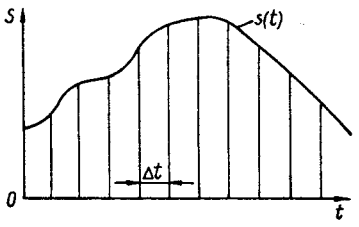


Рис. 4.1

Дискрети ще визначаються *відліками* (визначальними ординатами), а моменти відліку — *тактовими точками*. Інтервал між відліками (дискретами) Δt називається *інтервалом* (або *кроком*) *дискретизації* (кроком квантування за часом). Доведення теореми наведено в [31, 38].

Таким чином, якщо є потреба передачі неперервного сигналу з обмеженим спектром, то досить передати його окремі дискретні значення, відлічені через проміжки часу $\Delta t \le F_m/2$ (рис. 4.1). Чим менші інтервали Δt, тим точніше буде передана функція $s(t)$. Тривалість $\tau_д$ дискрет теоретично може бути дуже малою, але практично вона вибирається з урахуванням ширини смуги пропускання Δf каналу зв'язку, оскільки $\Delta f = k_д/\tau_д$, де $k_д$ — стала величина, близька до 1.

Відповідно до теореми відліків неперервна детермінована функція часу $s(t)$, що має обмежений спектр, може бути подана рядом [28, 31]

$$s(t) = \sum_{k=-\infty}^{\infty} s(k\Delta t)\frac{\sin\omega_m(t-k\Delta t)}{\omega_m(t-k\Delta t)}, \qquad (4.1)$$

де $k = \ldots,-1, 0, +1,\ldots$ — відліки миттєвих значень функції $s(t)$; $\omega_m = 2\pi F_m$ — максимальна частота спектра неперервного сигналу.

З (4.1) випливає, що неперервна функція з обмеженим спектром може бути подана у вигляді суми нескінченного числа членів, кожний з яких є добутком функції відліку виду $\frac{\sin y}{y}$ і коефіцієнта $s(k\Delta t)$, який визначає значення функції $s(t)$ в точках відліку. Графічно функцію відліків зображено на рис. 4.2, де $\tau = t - k\Delta t$. Максимального значення, що дорівнює одиниці, ця функція досягає в момент часу $t = k\Delta t$, а в моменти часу $t = (k + i)\Delta t$, де $i = 1,2,3,\ldots,\infty$, вона дорівнює нулю.

Функція виду $\frac{\sin\omega_m\tau}{\omega_m\tau}$ є не що інше, як реакція ідеального фільтра нижніх частот з граничною частотою ω_m на дельта-функцію. Якщо через такий фільтр пропустити дискретизований сигнал з частотою дискретизації $f_д = 2F_m = \omega_m/\pi$, то підсу-

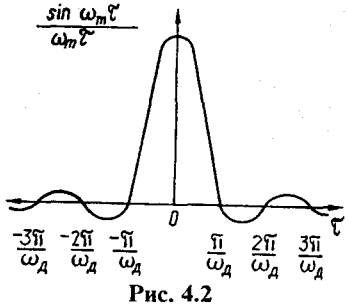


Рис. 4.2

мовуючи сигнали на виході фільтра, можна дістати неперервний первинний сигнал $s(t)$.

Теорема відліків визначає теоретичний підхід до перетворення неперервних функцій на дискретні. Так, для точного відтворення неперервної первинної функції $s(t)$ необхідно підсумувати реакції фільтра на вхідні імпульси в часі від $-\infty$ до $+\infty$. Проте на практиці використовуються обмежені в часі сигнали, що, як наслідок, мають нескінченно широкий спектр. Останній чинник суперечить головній умові теореми відліків.

Тому з деякими втратами точності відтворення, які допускаються на практиці при дискретизації, реальний спектр сигналу, що простягається від нуля до нескінченності, умовно обмежують діапазоном частот, в якому зосереджено основну частину енергії спектра сигналу з максимальною частотою ω_m. При цьому енергія частини спектра сигналу, що залишилася поза обмеженим частотним діапазоном, визначає похибку при відтворенні форми первинної функції $s(t)$. Для зменшення цієї похибки в разі високих вимог до точності відтворення форми первинного повідомлення частоту дискретизації $f_д$ вибирають значно більшою, ніж $2F_m$.

Квантування за рівнем. При квантуванні за рівнем неперервна функція $s(t)$ замінюється множиною дискретних значень (дискрет) шкали рівнів, які виникають у моменти часу, коли значення неперервного сигналу досягає певного рівня, що віддалений від попереднього на певну величину Δs, яка називається *інтервалом* (*кроком*) *квантування за рівнем* (рис. 4.3).

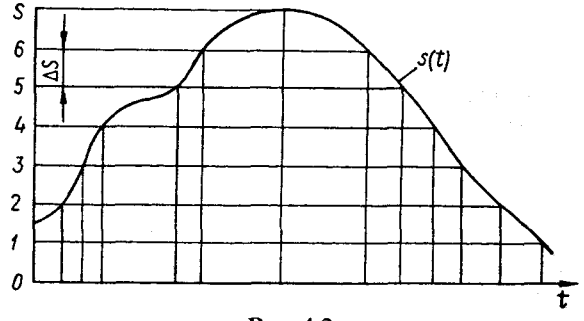


Рис. 4.3

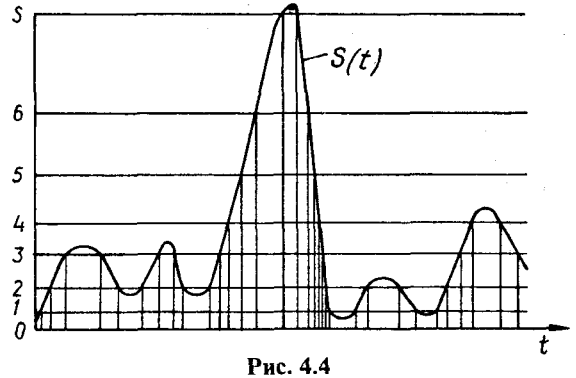


Рис. 4.4

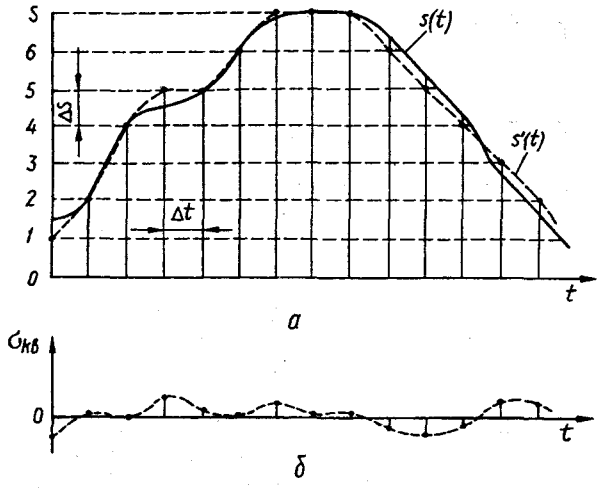
Таким чином, при квантуванні за рівнем частота проходження дискретних сигналів визначається крутістю функції неперервного сигналу $s(t)$ та інтервалом квантування Δs, а кількість дозволених рівнів квантування m — максимальним рівнем сигналу $s_{max}(U_{max}, I_{max}, P_{max})$ та інтервалом квантування Δs, тобто $m = s_{max}/\Delta s$.

Розрізняють *рівномірне* (Δs = const) і *нерівномірне* ($\Delta s \neq$ const) квантування за рівнем. На практиці найпоширенішим є рівномірне (лінійне) квантування (див. рис. 4.3) завдяки більш простій його технічній реалізації.

При неоднаковій імовірності розподілу значень функції $s(t)$ по шкалою рівнів ефективніше застосовувати нерівномірне квантування, оскільки його основною метою є зменшення усередненої за параметром дисперсії похибки квантування. При такому квантуванні значення неперервної функції $s(t)$, які мають велику ймовірність виникнення, передаються з меншою похибкою квантування, а менш імовірні — з більшою.

Для нерівномірного квантування можна застосовувати різні принципи побудови шкали рівнів. Так, на рис. 4.4 показано перетворення неперервного сигналу $s(t)$ з логарифмічною шкалою квантування за рівнем, де малі значення сигналу $s(t)$ перетворюються на дискретні з меншою похибкою, а великі сплески, що мають малу ймовірність, — з більшою похибкою.

Квантизація. При квантизації (комбінованому квантуванні, тобто квантуванні за часом і рівнем) неперервна функція $s(t)$ квантується за часом і рівнем, тобто замінюється множиною дискретних значень (дискрет) шкали рівнів, які виникають у моменти часу, що відстоять один від одного на інтервал (крок) дискретизації Δt, причому значення дискрет визначається найближчим до значення неперервної функції $s(t)$ рівнем квантування в момент відліку (рис. 4.5, *а*).



а

б

Рис. 4.5

Квантизований сигнал не може бути відтворений на приймальному боці повною мірою навіть за відсутності завад. Це пояснюється тим, що під час квантизації вносяться спотворення, усунути які неможливо. Ці спотворення називаються *шумом квантування*, а визначаються вони різницею між значеннями квантованих дискретних сигналів $s'(t)$ та функцією неперервного сигналу $s(t)$ у відповідних точках відліку (тактових точках). Похибка (або шум) квантування (рис. 4.5, *б*) приблизно оцінюється потужністю шуму квантування

$$\sigma_{\text{кв}} = \frac{1}{\Delta s}\int_{-\Delta s/2}^{\Delta s/2} s^2 ds = \frac{\Delta s}{2\sqrt{3}}, \quad \text{або} \quad \sigma^2_{\text{кв}} = \frac{\Delta s^2}{12},$$

де Δs — інтервал квантування за рівнем.

Для зниження $\sigma_{\text{кв}}$ необхідно зменшувати інтервал квантування Δs, але при цьому знижується завадостійкість передачі, оскільки під впливом завад рівень сумарного сигналу (сигнал + завада) може опинитися ближче до іншого рівня квантування, ніж до рівня, що передається. Тому вибір Δs залежить як від потрібної точності відтворення функції $s(t)$ на приймальному боці, так і від рівня завад у каналі зв'язку.

При квантизації, як і при квантуванні за рівнем, також використовують нерівномірне квантування, при якому зі збільшенням номера амплітудної градації зростає також інтервал квантування за рівнем. Це дає змогу при певній похибці (шумі) квантування суттєво зменшити кількість рівнів квантування, необхідних для дискретного подання неперервної функції з заданою точністю.

Квантизація широко застосовується в системах телевимірювання. До недоліків таких систем можна віднести складність декодування квантизованих сигналів на приймальному боці. Тому квантизація частіше використовується як проміжний етап при імпульсно-кодовій модуляції [7, 8], коли замість дискрет у канал зв'язку надсилаються кодові комбінації, яким однозначно відповідають ці дискрети. Це значно поліпшує завадостійкість систем передачі, але потребує розширення частотної смуги пропускання каналу при збереженні первинної швидкості передачі повідомлень.

4.2. ІНФОРМАЦІЙНІ ВТРАТИ ПРИ КОДУВАННІ НЕПЕРЕРВНИХ ДЖЕРЕЛ

Стосовно неперервного джерела на відміну від дискретного можна говорити про нескінченний алфавіт повідомлень, кожне з яких відрізняється від сусідніх на нескінченно малу величину, та нескінченний ансамбль повідомлень. Однак у цьому разі замість імовірностей окремих повідомлень з алфавіту прийнято говорити про диференціальний закон розподілу ймовірностей $w(x)$ випадкової величини x. Інакше $w(x)$ називається *функцією розподілу щільності ймовірностей неперервного повідомлення*. При цьому кількість інформації, наявна в прийнятому неперервному повідомленні, як і раніше визначається різницею значень ентропії (невизначеності) джерела повідомлень до та після одержання повідомлення.

Нехай щільність імовірності $w(x)$ має вигляд, показаний на рис. 4.6. Проквантуємо за рівнем випадкову величину x із дискретою Δx. Імовірність того, що $x_n \le x \le x_{n+1}$, становить $p(x_i) = w(x_i)\Delta x$, тобто визначається площею S_i прямокутника:

$$p(x_i) = S_i \int_{x_n}^{x_{n+1}} w(x)dx = w(x_i)\Delta x_i . \qquad (4.2)$$

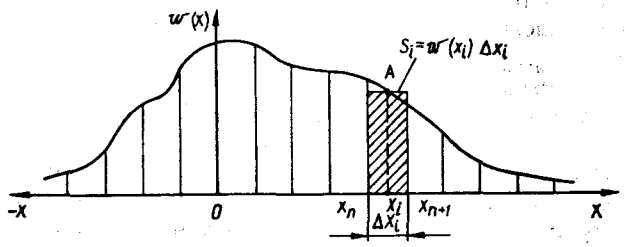


Рис. 4.6

Координату точки x_i визначає теорема про середнє [3]. При цьому має виконуватися умова нормування

$$\int_{-\infty}^{\infty} w(x)dx = 1. \tag{4.3}$$

Ентропія такого штучно утвореного дискретного джерела визначається так, як викладено в п.2.3, а саме [18]:

$$\begin{aligned} H_д(x) &= -\sum_{i=1}^{k} p(x_i)\log p(x_i) = -\sum_{i=1}^{k} w(x_i)\Delta x \log[w(x_i)\Delta x] = \\ &= -\sum_{i=1}^{k} w(x_i)\Delta x \log w(x_i) - \sum_{i=1}^{k} w(x_i)\Delta x \log \Delta x = \\ &= -\sum_{i=1}^{k} [w(x_i)\log w(x_i)]\Delta x - \log \Delta x \sum_{i=1}^{k} w(x_i)\Delta x. \end{aligned} \tag{4.4}$$

Зробимо зворотний перехід до неперервного джерела через спрямування Δx до 0 та граничний перехід [18]:

$$\begin{aligned} H(x) &= \lim_{\Delta x \to 0} \{H_д(x)\} = \lim_{\Delta x \to 0} \left\{ -\sum_{i=1}^{k} [w(x_i)\log w(x_i)]\Delta x + \right. \\ &\quad \left. + \lim_{\Delta x \to 0} \left\{ -\log \Delta x \sum_{i=1}^{k} w(x_i)\Delta x \right\} = \right. \\ &= -\int_{-\infty}^{\infty} w(x)\log w(x)dx - \lim_{\Delta x \to 0} \log \Delta x, \end{aligned} \tag{4.5}$$

оскільки згідно з (4.3)

$$\lim_{\Delta x \to 0} \sum_{i=1}^{k} w(x_i)\Delta x = \int_{-\infty}^{\infty} w(x)dx = 1.$$

Друга складова в (4.5) прямує до нескінченності. Отже, ентропія $H(x)$ неперервного джерела має дорівнювати нескінченності, тобто точне подання випадкового відліку неперервного джерела (одного його повідомлення) потребує нескінченної кількості, скажімо, двійкових розрядів, тому що несе нескінченну кількість інформації. Проте в реальних умовах відлік неперервних повідомлень на приймальному боці виконують у дискретних точках x_n, x_{n+1}, ... (див. рис. 4.6). Це зумовлено скінченною точністю та роздільною здатністю технічних засобів (здатністю їх розрізняти x_n і x_{n+1} при $\Delta x \to 0$). За цих обставин величина Δx є малою, але має скінченне значення.

Таким чином, вираз (4.5) ентропії неперервного джерела має дві складові, одна з яких визначається законом неперервного розподілу ймовірностей, а інша — допустимою точністю (похибкою) Δx кодування неперервного джерела.

Перша складова

$$h(x) = -\int_{-\infty}^{\infty} w(x)\log w(x)dx \tag{4.6}$$

називається *диференціальною ентропією*, що залежить від статистичних властивостей неперервного джерела. Якщо пронормувати x, щоб зробити цю величину безрозмірною, то можна визначити $h(x)$ у двійкових одиницях (при цьому основа логарифма має дорівнювати двом).

Друга складова зовсім не залежить від статистики джерела, а визначається лише дискретністю квантування Δx неперервного повідомлення.

Можливий стан джерела повідомлень до одержання повідомлення y на виході каналу визначається розподілом $w(x)$, а після одержання відліку з неперервного ансамблю y на виході — неперервним законом розподілу умовної імовірності $w(x/y)$, за яким можна знайти умовну ймовірність, користуючись поняттям елементарної площини $\Delta x \Delta y$ та роблячи, як і раніше, граничний перехід $\Delta x \to 0$, $\Delta y \to 0$. Як і в (4.5), загальна умовна ентропія $H(x/y)$ дорівнює нескінченності, але кількість інформації в цьому відліку, як і раніше, дорівнює різниці безумовної та умовної ентропії джерела відносно виходу каналу, тобто

$$\begin{aligned} I(x,y) &= H(x) - H(x/y) = h(x) - h(x/y) = \\ &= -\int_{-\infty}^{\infty} w(x)\log w(x)dx + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(x,y)\log w(x/y)dxdy, \end{aligned} \tag{4.7}$$

де $w(x, y)$ — сумісна щільність імовірності значення відліку y та фактичного значення неперервного повідомлення на виході каналу. Тут $h(x)$, $h(x/y)$ — безумовна й умовна диференціальна ентропія джерела.

Друга складова $\lim_{\Delta x \to 0} \log \Delta x$ в $H(x)$ та $H(x/y)$ виявляється однаковою і при відніманні в (4.7) зникає.

Таким чином, кількість інформації в одному відліку, що передається неперервним каналом, визначається різницею безумовної та умовної диференціальної ентропії неперервного джерела відносно виходу каналу, причому вираз $h(x/y)$ характеризує інформаційні втрати на один відлік при кодуванні неперервного джерела.

У теорії вимірювання випадкових величин користуються іншою кількісною мірою інформації — епсилон-ентропією. Припустимо, що випадкова величина Y містить інформацію про іншу випадкову величину X — дійсне значення вимірюваної величини, тобто Y є відліком X у певних одиницях. За цієї умо-

ви мінімальна кількість інформації про X, яка міститься в Y і потрібна для того, щоб за Y відтворити X із середньоквадратичною похибкою, що не перевищує ε^2 (де $\varepsilon > 0$ — наперед задана величина), називається *епсилон-ентропією*:

$$\overline{(Y-X)}^2 \le \varepsilon^2.$$

При цьому мінімум кількості інформації відшукується за всіма законами розподілу ймовірностей $w(X/Y)$, тобто

$$H_\varepsilon(X) = \min_{w(X/Y)} [H(X) - H(X/Y)].$$

Цю величину, як правило, подають у вигляді

$$H_\varepsilon(X) = \min_{w(X/Y)} \iint w(X,Y) \log \frac{w(X,Y)}{w(X)w(Y)} dXdY,$$

де $w(X, Y)$ — сумісна щільність розподілу ймовірностей X і Y; $w(X)$, $w(Y)$ — одновимірні щільності розподілу цих імовірностей.

4.3. ПРОДУКТИВНІСТЬ НЕПЕРЕРВНОГО ДЖЕРЕЛА ТА ШВИДКІСТЬ ПЕРЕДАЧІ ІНФОРМАЦІЇ

Виходячи з того, що ентропія $H(x)$ неперервного джерела за абсолютним значенням є нескінченною, продуктивність такого джерела також нескінченна. Про продуктивність неперервного джерела доцільно говорити лише в диференціальному відношенні, допускаючи похибку або порівнюючи диференціальну ентропію $h_1(x)$ і $h_2(x)$ сусідніх відліків повідомлення, взятих з інтервалом часу $\Delta t = t_2 - t_1$ [18].

Відповідно до теореми відліків неперервні повідомлення x, y можуть бути подані сукупностями відліків їх x_i та y_i в дискретні моменти часу з кроком Δt.

Розподіл сукупності випадкових величин описується багатовимірною щільністю розподілу ймовірностей $w(x_1, x_2, \ldots, x_m)$, $w(y_1, y_2, \ldots, y_m)$. Якщо вважати випадкові величини незалежними та врахувати, що ентропія сукупності незалежних випадкових величин дорівнює сумі ентропії окремих таких величин, то диференціальна ентропія повідомлення визначиться як

$$h_T(x) = \sum_{i=1}^{m} h(x_i), \qquad (4.8)$$

де $h(x_i) = -\int_{-\infty}^{\infty} w(x_i) \log w(x_i) dx_i$ — диференціальна ентропія i-го відліку повідомлення у формі (4.6); $m = T/\Delta t$ — кількість

відліків повідомлення тривалістю T, зроблених з інтервалом часу Δt.

Обмежившись стаціонарними випадковими процесами, дістанемо

$$h(x_1) = h(x_2) = \ldots = h(x_m) = h(x),$$

звідки

$$h_T(x) = mh(x), \qquad (4.9)$$

де $h(x)$ — диференціальна ентропія одного відліку повідомлення у формі (4.6).

Аналогічно можна показати, що умовна диференціальна ентропія

$$h_T(x/y) = mh(x/y), \qquad (4.10)$$

де $h(x/y)$ — умовна диференціальна ентропія одного відліку повідомлення.

Тоді вираз кількості інформації в неперервному повідомленні тривалістю T матиме вигляд

$$I_T(x,y) = m[h(x) - h(x/y)]. \qquad (4.11)$$

Назвемо *середньою швидкістю* передачі інформації неперервним джерелом кількість інформації, що передається за одиницю часу, тобто

$$R_T(x,y) = \frac{I_T(x,y)}{T} = \frac{m}{T}[h(x) - h(x/y)] = F_д[h(x) - h(x/y)], \qquad (4.12)$$

де $F_д = m/T$ — частота дискретизації повідомлення.

4.4. ПРОПУСКНА ЗДАТНІСТЬ НЕПЕРЕРВНОГО КАНАЛУ

Пропускною здатністю неперервного каналу називається максимально можлива швидкість передачі інформації в ньому [44]:

$$C = \max[R_T(x,y)] = F_д \max[h(x) - h(x/y)]. \qquad (4.13)$$

Як і для дискретного каналу, вираз (4.13) досягає максимуму при максимальному ступені статистичної зумовленості неперервних повідомлень на виході та вході каналу. При цьому умовна диференціальна ентропія прямує до нуля завдяки низькому рівню завад і все меншому спотворенню повідомлень у каналі. Проте функція $h(x)$ матиме максимум лише при певних

законах розподілу $w(x)$ імовірностей (наприклад, якщо як фізичний процес X використовується стаціонарний випадковий процес у вигляді «білого шуму»).

Вираз (4.13) за наявності завади у вигляді «білого шуму» набуває вигляду

$$C = F_m \log_2(1 + P_c/P_з), \qquad (4.14)$$

де $F_m = F_д/2$ — максимальна частота смуги прозорості каналу; P_c, $P_з$ — середні потужності сигналу неперервного повідомлення та завади у вигляді «білого шуму».

Пропускну здатність неперервного каналу можна регулювати, змінюючи F_m, P_c і $P_з$. Суть виразу (4.14) полягає у тому, що сума $(1 + P_c/P_з)$ визначає кількість рівнів (квантів) неперервного повідомлення, які надійно розпізнаються на фоні завади при заданому відношенні сигнал/завада. Тому кількість інформації тут, що припадає на один відлік повідомлення, буде такою самою, як і для дискретного джерела з кількістю станів $k = 1 + P_c/P_з$, коли $I(x,y) = \log_2(1 + P_c/P_з)$.

КОНТРОЛЬНІ ЗАДАЧІ

1. Неперервний канал характеризується відношенням середніх потужностей сигналу та шуму $P_c/P_ш$, де $P_ш = N_0\Delta f$ (N_0 — спектральна щільність потужності завад; Δf — ширина смуги частот каналу). Сигнал, який передається, має тривалість τ. Визначити потрібне відношення $P_c/P_ш$ для випадку, коли тривалість сигналу зменшується до τ_1, а ширина смуги частот каналу змінюється до Δf_1.

2. Довести, що $h(z) = h(y)$ при $z = y \pm k$ або $z = -y$, де k = const.

3. Визначити диференціальну ентропію випадкового відліку повідомлення X, якщо розподіл його ймовірностей рівномірний у проміжку x_1, x_2.

4. Неперервна випадкова величина змінюється за рівноймовірним законом розподілу в проміжку $x = 0\ldots1{,}8$ та $x = 0\ldots0{,}3$. Визначити диференціальну ентропію джерела для цих випадків і порівняти здобуті результати.

5. Визначити швидкість передачі інформації в неперервному каналі з повідомленням, розподіленим за законом

$$w(y) = \begin{cases} 0 \text{ при } y \le 0; \\ y^2 \text{ при } 0 \le y \le 1; \\ 1 \text{ при } y > 1 \end{cases}$$

та з періодом відліків 10^{-6} с.

6. Телефонний канал зв'язку характеризується такими даними: Δf = = 3100 Гц; $P_c/P_ш = 10$. Текст передається з ентропією 3,5 біт/символ, а середня швидкість його читання дорівнює 200 символів/хв. Визначити, наскільки ефективно використовується при цьому пропускна здатність каналу.

7. Напруга в електричному колі вимірюється в межах 150...180 мВ. Як зміниться ентропія випадкової величини напруги при вимірюванні її в мікровольтах?

8. При частотній модуляції носійної неперервним сигналом з рівномірним розподілом частота змінюється в межах 15...60 МГц. Визначити ентропію сигналу при вимірюванні частоти з похибкою 5 кГц.

9. Пропускна здатність неперервного каналу C = 9600 біт/с при відношенні $P_c/P_з = 10$. Як зміниться C при зменшенні цього відношення до 1?

10. Неперервний процес має нормальний розподіл імовірностей із щільністю $w(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/(2\sigma^2)}$. Визначити ентропію цього процесу при похибці вимірювання Δx.

11. Значення сигналу рівномірно лежать у діапазоні 0...10 В. Визначити диференціальну ентропію цього джерела. Якою буде ця ентропія, якщо значення сигналу виразити в мілівольтах?

12. Телеметрична станція за 10 с передає покази 20 датчиків. Спектр частот неперервних повідомлень лежить у межах 0...30 Гц. Рівень сигналів становить 0...10 В, а допустима відносна похибка дорівнює 0,5 % його максимуму. Визначити потрібну швидкість передачі інформації в каналі.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що таке ентропія неперервного джерела?
2. Що таке диференціальна ентропія неперервного джерела?
3. Як обчислюється диференціальна ентропія неперервного джерела?
4. Як визначається кількість інформації на одне повідомлення неперервного джерела?
5. За яких умов диференціальна ентропія неперервного джерела буде від'ємною?
6. Що означає відносність диференціальної ентропії неперервного джерела?
7. Як визначається пропускна здатність неперервного каналу?
8. Яке граничне значення пропускної здатності $C = \Delta f \log[1 + P_c/(N_0\Delta f)]$ (біт/с) неперервного каналу з нормальним шумом, якщо ширина Δf смуги його частот прямує до ∞ при P_c = const і N_0 = const?
9. Як визначається швидкість передачі інформації про неперервний процес?
10. Як визначається період відліків неперервної величини за теоремою відліків Котельникова?
11. Що таке епсилон-ентропія випадкової величини?

РОЗДІЛ 5 КОДУВАННЯ В ДИСКРЕТНИХ І НЕПЕРЕРВНИХ КАНАЛАХ

Передача інформації по лініях і каналах зв'язку пов'язана з певними труднощами. Головним чином — це дія завад на сигнали, що несуть інформацію. Тому для усунення помилок, які виникають при передачі сигналів під впливом завад, повідомлення кодують, оскільки кодування є найефективнішим способом їх захисту.

Для кращого розуміння процесу кодування повідомлень треба знати системи числення, деякі основні операції над елементами поля GF(q), способи подання кодів, основні теореми кодування для каналів. Про все це йтиметься в цьому розділі. Крім того, тут викладено класифікацію кодів, наведено характеристики їх, описано основні принципи оптимального кодування.

5.1. КЛАСИФІКАЦІЯ КОДІВ І ХАРАКТЕРИСТИКИ ЇХ

Кодування — це процес перетворення повідомлення на впорядкований набір символів, елементів, знаків. При кодуванні кожному повідомленню ставиться у відповідність зумовлена кодова комбінація — набір символів (елементів, знаків) з деякої скінченної кількості їх, яка називається *алфавітом*.

Множина кодових комбінацій, побудованих за одним правилом кодування, називається *кодом*. Залежно від алфавіту, який застосовується для побудови кодових комбінацій, коди поділяються на *двійкові*, алфавіт яких складається з двох символів (0 і 1), та *недвійкові* (багатопозиційні, q-коди), алфавіт яких містить більше двох символів.

Розрізняють дві групи кодів: *безнадмірні* (некоректувальні, первинні, прості) та *надмірні* (завадостійкі). Перші не дають змоги виявити та виправити спотворені елементи в своїх комбінаціях, другі — забезпечують можливість виявлення та/або виправлення елементів кодових комбінацій, спотворених унаслідок дії завад.

У надмірних кодах комбінації можуть мати інформаційні та перевірні елементи. Обидві групи кодів поділяються на *рівномірні* та *нерівномірні*, тобто коди зі сталою та змінною кількістю розрядів.

Надмірні коди також бувають *неперервними* (рекурентними) і *блоковими*. В неперервних кодах процес кодування та декодування має неперервний характер, у блокових — кожному повідомленню відповідає кодова комбінація (блок) зі скінченної кількості елементів. Блоки кодуються та декодуються окремо.

Блокові коди, в свою чергу, можуть бути *подільними* та *неподільними*. До перших належать коди, що будуються доповненням інформаційних елементів перевірними; до других — коди, в яких немає чітко зумовлених інформаційних і перевірних елементів.

Подільні блокові коди бувають систематичними та несистематичними. *Систематичним подільним блоковим* кодом називається такий код, у комбінаціях якого перші k позицій (розрядів) зайнято інформаційними елементами, а решту $r = n - k$ позицій, де n — загальна кількість позицій в кодовій комбінації, — перевірними. До *несистематичних подільних блокових* кодів належать коди, в яких інформаційними елементами не зайнято всі k перших позицій.

Різновидом подільних систематичних блокових кодів є *циклічні* коди.

При виборі кодів для передачі інформації керуються вимогами до вірогідності інформації, що передається, та швидкості передачі, які визначаються такими характеристиками кодів:

- кількістю k інформаційних елементів;
- кількістю r перевірних елементів (для коректувальних кодів);
- довжиною (розрядністю) n — кількістю елементів (символів), які входять до складу кодової комбінації ($n = k + r$);
- основою (алфавітом) q;
- потужністю $N_д$ — кількістю дозволених кодових комбінацій, що використовуються для передачі повідомлень;
- повною кількістю N кодових комбінацій, тобто кількістю всіх можливих комбінацій, яка дорівнює q^n (для двійкових кодів $N = 2^n$);
- надмірністю

$$R_{над} = 1 - \frac{\log_q N_д}{\log_q N} \text{ (для неподільних кодів)}$$

або

$R_{над} = 1 - k/n = r/n$ (для подільних кодів при $N_д = 2^k$ і $N = 2^n$);

- відносною швидкістю R, яка характеризує ступінь використання в надмірному коді інформаційних можливостей його потужності, причому

$$R = \frac{\log_q N_д}{\log_q N} \quad \text{або} \quad R = \frac{k}{n} = 1 - R_{над};$$

вагою кодової комбінації (для двійкового коду визначається кількістю одиниць у ній);

• мінімальною кодовою відстанню $d = \min d_{ij}$, тобто мінімальною відстанню між парами кодових комбінацій

$$d_{ij} = \sum_{l=1}^{n} \left| a_{li} - a_{lj} \right|,$$

де a_{li}, a_{lj} — елементи, що знаходяться в l-му місці в i- та j-й комбінаціях. Це значить, що d_{ij} визначається кількістю однойменних розрядів з різними значеннями;

• імовірністю $P_{\text{нв.п}}$ невиявленої помилки, тобто ймовірністю такої події, за якої прийнята кодова комбінація відрізняється від переданої, а властивості коду не дають змоги визначити факт наявності помилки;

• імовірністю $P_{\text{в.п}}$ виявленої помилки, тобто ймовірністю, за якої прийнята кодова комбінація відрізняється від переданої і завдяки властивостям коду встановлюється факт наявності помилки в кодовій комбінації;

• імовірністю $P_{\text{вп.п}}$ виправленої помилки, тобто ймовірністю такої події, за якої прийнята кодова комбінація відрізняється від переданої і завдяки властивостям коду виправляється помилка в кодовій комбінації;

• імовірністю $P_{\text{п}}$ виникнення помилки, тобто ймовірністю такої події, за якої прийнята кодова комбінація відрізняється від переданої ($P_{\text{п}} = P_{\text{нв.п}} + P_{\text{в.п}}$ — для кодів, які виявляють помилки, та $P_{\text{п}} = P_{\text{нв.п}} + P_{\text{в.п}} + P_{\text{нв.п}}$ — для кодів, які виправляють помилки, де $P_{\text{нв.п}}$ — імовірність невиправленої виявленої помилки);

• кратністю ν помилки, що визначається кратністю $\nu_{\text{в}}$ виявлених та $\nu_{\text{вп}}$ виправлених помилок;

• ефективністю

$$r_{\text{еф}} = \frac{N_{\text{д}}}{N} \frac{P_{\text{п}}}{P_{\text{п}} - \sum_{i=1}^{\nu} P_i},$$

де P_i – ймовірність виявленої або виправленої помилки залежно від властивостей коду.

Ступінь захисту інформації від помилок відповідним способом кодування залежить головним чином від мінімальної кодової відстані $d_{\min}$ даного коду [12].

Розрізняють три види кодової відстані: Хеммінга, Лі та матричну. Перша знайшла найбільше поширення в теорії кодування. Кодова відстань Хеммінга нероздільно пов'язана з поняттям *ваги w кодової комбінації* — кількістю її елементів, які не дорівнюють нулю.

Кодова відстань Хеммінга d між двома комбінаціями однієї довжини n визначається як кількість однойменних розрядів (позицій), які мають неоднакові елементи. Так, для двійкових кодів, оскільки в двійковій арифметиці додавання однакових елементів дає 0, а неоднакових — 1, відстань Хеммінга між двома кодовими комбінаціями можна визначити порозрядним додаванням їх за модулем 2 та подальшим підрахунком кількості ненульових елементів, тобто визначенням ваги w такої суми.

Загальна кількість кодових комбінацій завдовжки n дорівнює 2^n, а кількість тих з них, які віддалені від заданої на відстань d, — кількості сполучень з n по d:

$$C_n^d = n!/[d!(n-d)!].$$

Щоб визначити кодову комбінацію, яка віддалена від заданої на відстань d, до цієї комбінації можна додати будь-яку комбінацію вагою $w = d$ (з d одиницями та $n - d$ нулями). Додавання — порозрядне за модулем 2.

Для виявлення всіх помилок кратністю $\nu_{\text{в}}$ кодова відстань має становити $d \geq \nu_{\text{в}} + 1$, а виправлення помилок кратністю $\nu_{\text{вп}}$ – $d \geq 2\nu_{\text{вп}} + 1$. Щоб виправити та виявити всі помилки, має виконуватися умова

$$d \geq \nu_{\text{вп}} + \nu_{\text{в}} + 1.$$

Через те, що загалом кожний елемент (розряд) комбінації недвійкового (багатопозиційного) коду може мати на відміну від двійкового й понад однієї позиції ($m \geq 1$) з алфавіту q, кодова відстань при цьому визначається виразом

$$d = \sum_{i=1}^{m} d_i,$$

де m — кількість позицій в кожному розряді (поодиноковому часовому інтервалі, що відповідає тривалості одного елемента) кодової комбінації.

У метриці Хеммінга кодова відстань, як і для двійкового коду, визначається кількістю однойменних розрядів з різними позиціями (символами):

$$d_i(x_k, x_l) = \begin{cases} 0, x_k = x_l; \\ 1, x_k \neq x_l. \end{cases}$$

У метриці Лі

$$d_i(x_k, x_l) = \min\{|x_k - x_l|, q - |x_k - x_l|\} \equiv \min\{d_{j\,\text{mod}}, q - d_{j\,\text{mod}}\},$$

де $d_{j\text{mod}} = |x_k - x_l|$.

У модульній метриці $d_i(x_k,x_l) = |x_k - x_l|$, тобто слід виконувати віднімання за модулем q.

Відзначимо, що коли значення кодової відстані для двійкового коду в різних матрицях збігаються, для недвійкового коду при $q = 3$ значення d в метриках Хеммінга та Лі також збігаються. При $q > 3$ значення d у різних метриках різняться.

При конкретній реалізації недвійкового коду з використанням позицій тих або інших ознак сигналу кодова відстань визначається відповідною метрикою.

5.2. СИСТЕМИ ЧИСЛЕННЯ

У теорії інформації, кодування, передачі даних і системах обміну інформацією найпоширенішими є двійкова, вісімкова та шістнадцяткова системи числення. Проте це ні в якому разі не означає, що на практиці не користуються іншими системами числення, такими як трійкова, четвіркова, шісткова тощо.

Узагалі ціле число N у будь-якій системі числення можна записати у вигляді ряду

$$N = \sum_{i=0}^{n-1} \alpha_i q^i, \qquad (5.1)$$

де α_i — розрядні коефіцієнти, значення яких змінюються від 0 до $q - 1$; q – основа (алфавіт) системи числення; i — номер розряду; n — кількість їх.

Назва системи числення походить від основи (алфавіту) q : $q = 2$ — двійкова, $q = 3$ — трійкова, $q = 8$ — вісімкова системи числення тощо.

Для запису чисел (табл. 5.1) у дев'ятковій системі використовують 10 цифр (0, 1, 2, 3, 4, 5, 6, 7, 8, 9); у двійковій — дві (0 і 1); у трійковій — три (0,1,2); у вісімковій — вісім (0, 1, 2, 3, 4, 5, 6, 7), а в шістнадцятковій — 16 знаків, з них — 10 цифр (0...9) і шість літер (A, B, C, D, E, F).

Так, користуючись виразом (5.1), десяткове число 375 можна записати у вигляді ряду

$$(375)_{10} = 3 \cdot 10^2 + 7 \cdot 10^1 + 5 \cdot 10^0.$$

Те саме число в двійковій системі числення запишеться як

$$(375)_2 = 1 \cdot 2^8 + 0 \cdot 2^7 + 1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 \rightarrow 101110111;$$

у трійковій — як

$$(375)_3 = 1 \cdot 3^5 + 1 \cdot 3^4 + 1 \cdot 3^3 + 2 \cdot 3^2 + 2 \cdot 3^1 + 0 \cdot 3^0 \rightarrow 111220,$$

Таблиця 5.1

Числа			
десяткові	двійкові	вісімкові	шістнадцяткові
1	1	1	1
2	10	2	2
3	11	3	3
4	100	4	4
5	101	5	5
6	110	6	6
7	111	7	7
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F
16	10000	20	10
17	10001	21	11
18	10010	22	12
32	100000	40	20
64	1000000	100	40
100	1100100	144	64
128	10000000	200	80
256	100000000	400	100

у вісімковій — як

$$(375)_8 = 5 \cdot 8^2 + 6 \cdot 8^1 + 7 \cdot 8^0 \rightarrow 567,$$

а в шістнадцятковій — як

$$(375)_{16} = 1 \cdot 16^2 + 7 \cdot 16^1 + 7 \cdot 16^0 \rightarrow 177.$$

Кількість комбінацій простого коду в будь-якій системі числення залежить від алфавіту коду та довжини кодової комбінації ($N = q^n$) і, звичайно, чим більший алфавіт коду, тим менша кількість розрядів у комбінації.

Для запису десяткового числа в будь-якій системі числення треба поділити його на основу вибраної системи. Після першого ділення дістанемо цілу частку й остачу. Продовживши ділення цілої частки, матимемо нові цілу частку та остачу. Ділення цілих часток продовжуємо доти, доки частка не стане меншою від основи q системи числення. Ця остання частка й буде старшим розрядом числа у вибраній системі числення. Інші розряди відповідатимуть остачам від ділення. Молодший розряд — це остача від першого ділення.

Нехай число 217, записане в десятковій системі числення, треба перевести у двійкову систему. Виконуємо послідовне ділення числа 217 на 2 (в дужках зазначено остачі від ділення):

217 : 2 = 108 + (1);	13 : 2 = 6 + (1);
108 : 2 = 54 + (0);	6 : 2 = 3 + (0);
54 : 2 = 27 + (0);	3 : 2 = 1 + (1);
27 : 2 = 13 + (1);	1 : 2 = 0 + (1).

Відповідно до викладеного вище остання частка від ділення, значення якої менше від основи системи числення (в даному разі це значення частки в передостанньому діленні, коли $1 < 2$), є коефіцієнтом при основі системи числення у найвищому степені (в даному разі це 1) або остання остача, що рівнозначно. Решта остач будуть коефіцієнтами при основі системи числення менших степенів. Таким чином, число 217 у двійковій системі числення записується у вигляді 11011001.

5.3. ОСНОВНІ ОПЕРАЦІЇ НАД ЕЛЕМЕНТАМИ ПОЛЯ

Основа (алфавіт) q коду може мати різні значення ($q \geq 2$). Методика побудови багатьох кодів ґрунтується на використанні властивостей послідовностей двійкових чисел. Розглянемо деякі операції над елементами двійкових кодів ($q = 2$).

Правила додавання за модулем 2 визначаються такими операціями:

$$0 \oplus 0 = 0;\ 1 \oplus 1 = 0;\ 0 \oplus 1 = 1;\ 1 \oplus 0 = 1.$$

Наприклад, при додаванні за модулем 2 двійкових послідовностей чисел 0111000 і 10010 матимемо

$$\begin{array}{r} \oplus\ 0111000 \\ \underline{10010} \\ 0101010\ . \end{array}$$

На відміну від звичайного арифметичного додавання операція додавання двійкових чисел за модулем 2 полягає в тому, що в даному разі розглядають конкретну пару двійкових знаків незалежно від інших чисел двійкової послідовності. Тому результат попередніх операцій при додаванні чергової пари двійкових знаків не враховують, тоді як при звичайних арифметичних операціях цей результат треба враховувати обов'язково, коли при додаванні двох двійкових чисел (наприклад, одиниць) записується 0, а 1 переноситься в старший розряд. Для прикладу, що розглядається, при арифметичній операції було б:

$$\begin{array}{r} \oplus\ 0111000 \\ \underline{10010} \\ 1001010\ . \end{array}$$

Операція віднімання за модулем 2 нічим не відрізняється від операції додавання.

Множення та ділення двійкових чисел за модулем 2 виконують за допомогою операції додавання за модулем 2. Так, при множенні за модулем 2 множене зсувають у бік старшого розряду стільки разів, скільки розрядів є у множнику, а потім додають їх за модулем 2. Множене виписують тільки в тому разі, коли в множнику є 1. Якщо ж у множнику є 0, то черговий зсув виконують без виписування множеного:

$$\begin{array}{r} \times\ 1111 \\ \underline{1101} \\ 1111 \\ \oplus\ 1111 \\ \underline{1111} \\ 1001011. \end{array}$$

При діленні за модулем 2 дільник підписують під діленим так, щоб збігалися старші розряди. Якщо кількість розрядів діленого перевищує або дорівнює кількості розрядів дільника, то в частку переносять 1, після чого виконують додавання за модулем 2 й до здобутого числа дописують праворуч наступну цифру діленого. Якщо ж число остачі разом з дописаною цифрою дорівнює кількості розрядів дільника, то до частки дописують ще одну 1, а якщо ні — то 0 доти, доки кількість розрядів остачі не дорівнюватиме кількості розрядів дільника. Після цього виконують додавання за модулем 2. Операцію повторюють стільки разів, поки всі розряди діленого не перенесуться до остачі. Наприклад:

$$\begin{array}{r|l} \oplus\ 1001011 & 1101 \\ \underline{1101} & \hline 1111 \\ \oplus\ 1000 & \\ 1101 & \\ \underline{1011} & \\ \oplus\ 110 & \\ 1101 & \\ \underline{1101} & \\ \oplus\ 000 & \\ 000 & \end{array} \qquad \begin{array}{r|l} \oplus\ 10101 & 1011 \\ \underline{1011} & \hline 1 + \dfrac{11}{1011} \\ 11 & \end{array}$$

Дуже зручно операції додавання, віднімання, множення та ділення за модулем 2 виконувати з двійковими числами, записаними у вигляді поліномів. Так, двійкові комбінації 110010 і 100001 можна записати поліномами

$$V_1(x) = x^5 + x^4 + x;\ V_2(x) = x^5 + 1.$$

Тоді додавання $V_1(x) \oplus V_2(x)$ за модулем 2 дасть

$$V_1(x) \oplus V_2(x) = x^5 + x^4 + x + x^5 + 1 = x^4 + x + 1 \rightarrow \mathbf{010011}.$$

Результатом множення буде

$$V_1(x)V_2(x) = (x^5 + x^4 + x)(x^5 + 1) = x^{10} + x^9 + x^6 + x^5 + x^4 + x \rightarrow$$
$$\rightarrow 11001110010.$$

Після ділення цих поліномів дістанемо

$$V_1(x):V_2(x) = (x^5 + x^4 + x):(x^5 + 1) = 1 + \frac{x^4 + x + 1}{x^5 + 1}.$$

Операції додавання, віднімання, множення та ділення при основі $q > 2$ коду мають свої особливості. Всі недвійкові коди (q-коди) поділяються на дві великі групи: коди з простою основою $q = p$, де $p \in \{3,5,7,11,13,...\}$, тобто з простими числами, що діляться тільки на самих себе; коди з розкладною основою q. Найбільший практичний інтерес становить підклас кодів з розкладною основою $q = 2^l$, елементи якого мають інформаційну ємність l бітів і можуть бути порівняні з усіма l-розрядними двійковими числами, що дає змогу використовувати двійкову техніку для виконання операцій додавання, ділення та множення.

До підкласу кодів з основою $q = 2^l$ належать недвійкові коди, в яких $q = \{4,8,16,...\}$. Для таких кодів необхідно задати операції над елементами поля $GF(q) = GF(2^l)$. Якщо взяти за основу поліномне подання елементів поля $G = \{0,1,2,..., q-1\} = \{0_l...0_20_1, 0_l...0_21_1,..., 1_l...1_21_1\}_2 = \{0,1,..., x^{l-1} + ... + x + 1\}$, то операція додавання двох елементів виконується як порозрядне додавання за модулем 2 двійкових елементів їх. Наприклад, при $q = 8 = 2^3$ матимемо $G = \{0,1,2,...,7\}_8 = \{000,001,010,...,111\}_2$. Результати додавання за модулем 8 наведено в табл. 5.2.

Операція множення двох елементів алгоритмічно виконується як множення двох поліномів, що відповідають елементам поля, які множать за модулем незвідного полінома степеня l. Так, при $q = 8 = 2^3$, $l = 3$ незвідний поліном степеня $l = 3$ згідно з [25,32] вибираємо виду $P(x) = x^3 + x + 1$. Множимо два елементи поля $G = \{0,1,2,...,7\}_8 = \{000,001,010,...,111\}_2$. Здобуті значення ділимо на незвідний поліном і результати заносимо в табл. 5.3.

Таблиця 5.2

⊕	0	1	2	3	4	5	6	7
0	0	1	2	3	4	5	6	7
1	1	0	3	2	5	4	7	6
2	2	3	0	1	6	7	4	5
3	3	2	1	0	7	6	5	4
4	4	5	6	7	0	1	2	3
5	5	4	7	6	1	0	3	2
6	6	7	4	5	2	3	0	1
7	7	6	5	4	3	2	1	0

Таблиця 5.3

⊗	0	1	2	3	4	5	6	7
0	0	0	0	0	0	0	0	0
1	0	1	2	3	4	5	6	7
2	0	2	4	6	3	1	7	5
3	0	3	6	5	7	4	1	2
4	0	4	3	7	6	2	5	1
5	0	5	1	4	2	7	3	6
6	0	6	7	1	5	3	2	4
7	0	7	5	2	1	6	4	3

Операція віднімання збігається з операцією додавання, а операції ділення та обернення (піднесення до степеня –1) елементів виконуються як множення, оскільки обернення елемента a рівнозначне діленню: $a^{-1} = 1/a$.

Нехай, наприклад, потрібно знайти результат від ділення $5/6 = \beta$. Це рівнозначне запису $6\beta = 5$. З табл. 5.3 знаходимо, що необхідно помножити 6 на 4 ($\beta = 4$), щоб дістати 5. Таким чином, $\beta = 4$ й є результатом ділення 5/6.

Аналогічно можуть бути побудовані таблиці для операцій додавання та множення при інших основах кодів ($q = 4,16,32,...$). Так, для виконання розрахунків над полем $GF(q)$ при $q = 16$ зручніше користуватися табл. 5.4, де наведено деякі форми по-

Таблиця 5.4

Десяткове подання елементів поля $GF(16)$	Подання елементів поля $GF(16)$ у формі двійкового вектора				Мультиплікативна форма у вигляді степеня примітивного елемента $\beta \in GF(16)$	Адитивна форма $a\beta^3 + b\beta^2 + c\beta^1 + d\beta^0$, $a,b,c,d \in \{0,1\}$ у базисах			
	2^3	2^2	2^1	2^0		β^3	β^2	β^1	β^0
0	0	0	0	0	$0 = \beta^{-\infty}$				0
1	0	0	0	1	β^0				1
2	0	0	1	0	β^1			β	
3	0	0	1	1	β^4			β	+1
4	0	1	0	0	β^2		β^2		
5	0	1	0	1	β^8		β^2		+1
6	0	1	1	0	β^5		β^2	+β	
7	0	1	1	1	β^{10}		β^2	+β	+1
8	1	0	0	0	β^3	β^3			
9	1	0	0	1	β^{14}	β^3			+1
10	1	0	1	0	β^9	β^3		+β	
11	1	0	1	1	β^7	β^3		+β	+1
12	1	1	0	0	β^6	β^3	β^2		
13	1	1	0	1	β^{13}	β^3	β^2		+1
14	1	1	1	0	β^{11}	β^3	β^2	+β	
15	1	1	1	1	β^{12}	β^3	β^2	+β	+1

дання елементів поля $GF(16)$ на основі модульного полінома $P(x) = x^4 + x + 1$ ($q = 16 = 2^4 \rightarrow l = 4$), який задає розширювальне рівняння $x^4 = x + 1$ і за аналогією рівняння $\beta^4 = \beta + 1$.

При додаванні за допомогою табл. 5.4 примітивних елементів, наприклад $\beta^4 + \beta^5 + \beta^9 + \beta^{14}$, треба перевести їх з мультиплікативної форми в адитивну, скоротити повторення, а здобутий результат перенести знову в мультиплікативну (або десяткову) форму, тобто

$$\beta^4 + \beta^5 + \beta^9 + \beta^{14} = \beta + 1 + \beta^2 + \beta + \beta^3 + \beta + \beta^3 + 1 = \\ = \beta^2 + \beta = \beta^5 = 6.$$

5.4. СПОСОБИ ПОДАННЯ КОДІВ

Код кожного виду має свій найраціональніший спосіб подання, що випливає з його властивостей. Проте відомо також кілька загальних способів подання кодів, які є досить універсальними і можуть застосовуватися для опису широких класів кодів. До цих способів належать подання кодів у вигляді: 1) таблиць кодових комбінацій; 2) кодового дерева; 3) геометричної моделі; 4) матриці.

Перший спосіб полягає в поданні коду у вигляді таблиці всіх його комбінацій. Наприклад, п'ятиелементний двійковий блоковий код зі сталою вагою, в кожній комбінації якого містяться три одиниці, задається так:

Номер кодової комбінації	Комбінація двійкового блокового коду з вагою 3
1	00111
2	01011
3	01101
4	01110
5	10011
6	10101
7	10110
8	11001
9	11010
10	11100

Цей спосіб застосовується для подання будь-яких блокових кодів, але не може бути використаний для неперервних кодів.

Другий спосіб подання кодів полягає в зображенні комбінацій коду у вигляді кодового дерева, коли комбінації розміщуються в його вузлах. Під *кодовим деревом* розумітимемо графічний образ, який складається з точок і ліній, що розходяться від них і також закінчуються точками. Останні називатимемо *вузлами*, а лінії, які їх з'єднують, — *ребрами*. Перший вузол, від якого починається розходження ребер, називається *коренем дерева*, а

кількість ребер, які треба пройти від кореня до будь-якого вузла — *рівнем*, або *порядком*, цього вузла.

Максимальна кількість вузлів, які зустрічаються під час руху вздовж кодового дерева в напрямку від кореня до вершини, визначає *висоту h* кодового дерева. Вона дорівнює максимальній довжині комбінації коду, побудованому за допомогою цього дерева.

Вузли кодового дерева розташовуються на різних рівнях. Кожний рівень дерева рівномірного коду може мати q^i вузлів, де q — основа коду, i — номер рівня ($i = 1,2,\ldots,n$, тут n — довжина коду). Для рівномірного двійкового простого коду кількість вузлів на останньому рівні n дорівнює кількості N комбінацій коду, тобто $2^n = N$.

Вузли, що не з'єднуються з наступними рівнями, називаються *кінцевими*; вони відповідають комбінаціям коду.

Основою коду обмежується максимальна кількість ребер, яка може виходити з кожного вузла дерева, а максимальною довжиною кодової комбінації — максимальна кількість рівнів кодового дерева. Кожному вузлу приписується значення розрядів комбінації, що відповідає напрямкам руху вздовж ребер від кореня дерева до вузла. Ребра, що йдуть від кореня до вузлів першого рівня, визначають значення першого зліва розряду кодової комбінації, а ті, що з'єднують вузли першого та другого рівнів, — значення другого зліва розряду і т. д. На рис. 5.1 показано приклади кодових дерев: рівномірних двоелементного

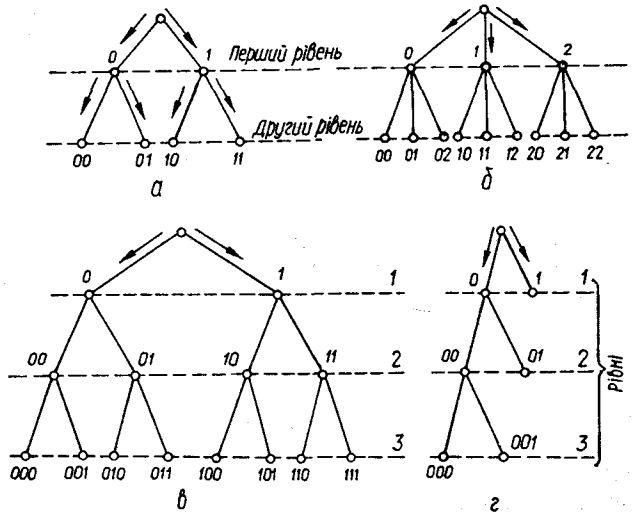


Рис. 5.1

двійкового (рис. 5.1, *а*), двоелементного трійкового (рис. 5.1, *б*), триелементного двійкового (рис. 5.1, в) та нерівномірного двійкового (рис. 5. 1, *г*). Цей спосіб застосовується для зображення як блокових, так і неперервних кодів.

За допомогою кодових дерев легко зобразити *префіксні* коди, що мають властивість префікса й можуть бути утворені послідовним викреслюванням останнього розряду кодової комбінації, причому жодна з комбінацій даного префіксного коду не може бути префіксом його комбінації. Наприклад, префіксами кодової комбінації 10111001 будуть 1, 10, 101, 1011, 10111, 101110, 1011100, 10111001, тобто для однозначного її декодування жодна з комбінацій цього коду не повинна мати перелічені вище комбінації.

Та частина, яка доповнює префіксний код до повної кодової комбінації, утворює суфікс, тобто кожна кодова комбінація складається з префікса та суфікса.

Префіксні коди можна утворити за допомогою кодового дерева, в якого немає вершини і кожний його кінцевий вузол відповідає комбінації префіксного коду.

Третій спосіб подання кодів полягає в зображенні комбінацій коду точками дискретного n-вимірного векторного простору. Так, кожну комбінацію рівномірного блокового коду (з основою q і довжиною n) $V = (V_n, V_{n-1}, \ldots, V_2, V_1)$ можна розглядати як вектор або точку деякого n-вимірного векторного простору з координатами $V_n, V_{n-1}, \ldots, V_2, V_1$. Якщо значення q скінченне, а будь-яка координата вектора є цілим додатним числом від 0 до $q - 1$, то зазначений код можна розглядати як дискретний n-вимірний простір, що складається з $N = q^n$ точок, які відповідають кінцям усіх можливих векторів.

Цей n-вимірний простір дістав назву *кодового*. Кількість просторових вимірювань кодового простору для коду з будь-якою основою дорівнює довжині n коду, а кількість градацій по кожній з осей (напрямків вимірювання) визначається основою коду і становить $q - 1$.

Якщо для дискретного n-вимірного простору, що тут розглядається, ввести поняття кодової відстані d між точками V_i та V_j, то матимемо

$$d(V_i, V_j) = \sum_{k=1}^{n} (V_{ki} - V_{kj}). \qquad (5.2)$$

Цілком природно, що для простору з відстанню (5.2), як і для будь-якого іншого кодового простору, $d(V_i, V_j) = d(V_j, V_i)$.

Одним з основних параметрів коду з довільною основою q, що визначають його завадостійкість, є *мінімальна кодова відстань* $d_{\min}$. На відміну від кодової відстані d, що визначає

кількість станів, які мають пройти якісні ознаки кодової комбінації, щоб опинитися в стані, який відповідає порівнюваній кодовій комбінації, мінімальна кодова відстань характеризує не дві окремо взяті комбінації, а код у цілому, і визначається мінімальною кількістю якісних ознак, за якими відрізняються одна від одної будь-яка пара комбінацій цього коду.

Для визначення кодової відстані між комбінаціями коду з основою q треба виконати їх порозрядне віднімання за модулем q. Кодова відстань дорівнює вазі комбінації, що складається з різниці значень комбінацій, між якими визначається ця відстань.

З'єднавши кожну точку простору, що розглядається, прямими лініями з усіма точками, віддаленими на відстань $d(V_i, V_j) = 1$, дістанемо геометричну фігуру сіткової структури. Цю фігуру називають *геометричною моделлю* n-елементного q-коду.

Точки дискретного простору, які містить ця геометрична фігура, називаються її *вершинами*, а лінії, що їх з'єднують, — *ребрами*.

На рис. 5.2 зображено геометричні моделі деяких кодів. Моделлю будь-якого двозначного набору якісних ознак (двоелементного коду) є фігура двовимірного простору — квадрат (рис. 5.2, *а*) або фігура, що складається з квадратів (рис. 5.2, *б*, *д*); моделлю будь-якого тризначного набору якісних ознак (триелементного коду) — фігура тривимірного простору — куб (рис. 5.2, *в*) або фігура, що складається з кубів (рис. 5.2, *г, е*).

Побудова моделі чотиризначного набору якісних ознак (чотириелементного коду), тобто фігури чотиривимірного простору, можлива, якщо тривимірний куб або кожну з його вершин змістити в новому напрямку. Взагалі в цьому разі n-вимірний куб повинен мати $2n$ вершин, $n \cdot 2^{n-1}$ ребер, $n(n-1) \cdot 2^{n-3}$ граней, а найвіддаленіша від певної його вершини точка має знаходитися на відстані n ребер.

Розглянемо більш докладно властивості, що випливають з геометричної фігури, яка є моделлю n-елементного двійкового коду й дістала назву *n-вимірного куба*. Відстань між будь-якими його вершинами, тобто між двома кодовими комбінаціями, згідно з (5.2) можна визначити як кількість розрядів, якими вони різняться. Так, відстань між комбінаціями 010 і 100 дорівнює двом, оскільки вони різняться елементами в двох розрядах — першому та другому.

Відмінність елементів однойменних розрядів комбінацій двійкового коду легко визначити, застосувавши до них операцію додавання за модулем 2. Як відомо з п. 5.2, результат такого додавання тільки тоді дорівнює 1, коли числа, що додаються, різняться. З урахуванням цього відстань між будь-якими дво-

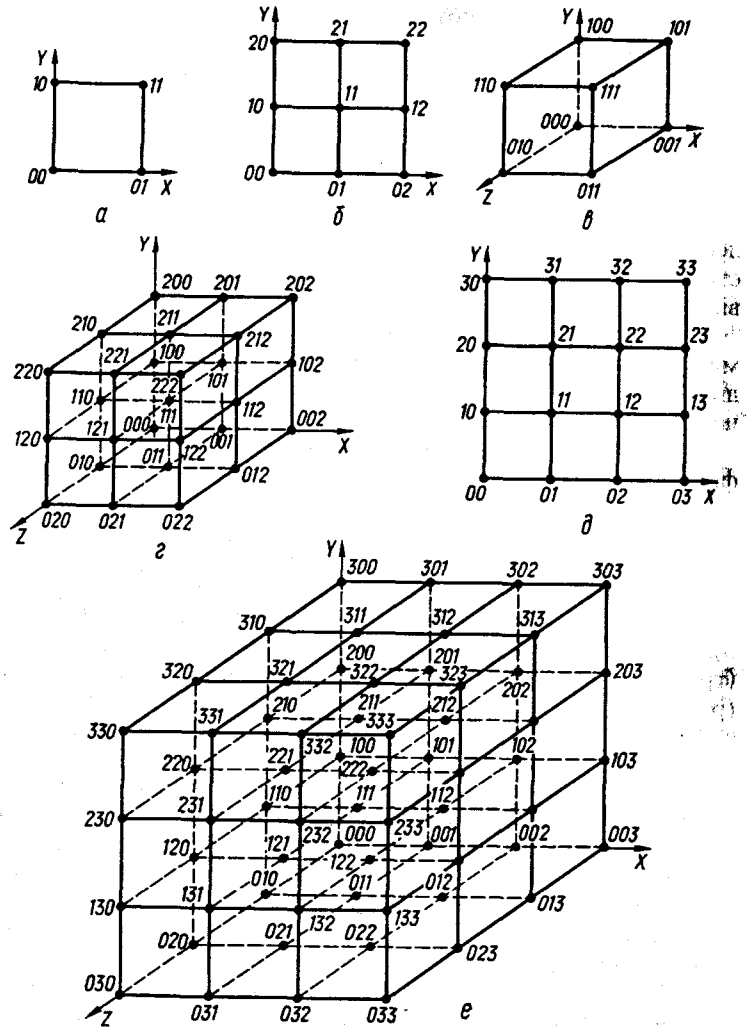


Рис. 5.2

ма комбінаціями n-елементного двійкового коду (вершинами n-вимірного куба) визначається виразом

$$d(V_i,V_j)=\sum_{k=1}^{n}(V_{ki}\oplus V_{kj}). \qquad (5.3)$$

Відстань між комбінаціями V_i та V_j можна знайти також через кількість одиниць у деякій комбінації V_l [або через її вагу

$w(V_l)$], здобуту після додавання за модулем 2 комбінацій V_i та V_j ($V_l = V_i \oplus V_j$):

$$d(V_i,V_j)=w(V_l). \qquad (5.4)$$

Зручність геометричної моделі зображення будь-якого коду полягає в тому, що кожна її вершина відповідає одній комбінації коду, а відстань між комбінаціями V_i та V_j згідно з (5.2) дорівнює кількості ребер, які треба пройти найкоротшим шляхом з вершини V_i до вершини V_j.

Недоліком геометричної моделі є те, що при довжині коду $n > 3$ зобразити її у звичайному тривимірному просторі неможливо. Тому вона застосовується лише для рівномірних блокових кодів з метою наочного зображення та полегшення аналізу їхніх властивостей.

Четвертий спосіб подання кодів у вигляді матриці з 2^n рядками та n стовпцями можливий тільки для рівномірних n-елементних двійкових блокових кодів. Якщо матрицею подається сукупність ненульових комбінацій коду, то кількість рядків дорівнюватиме $2^n - 1$.

З урахуванням того, що матриця n-елементного коду складається з $2^n - 1$ комбінацій, записаних у вигляді рядків, особливість такого запису полягає в тому, що додавання за модулем 2 будь-якої кількості рядків цієї матриці приводить до появи дозволеної комбінації коду, в тому числі й нульової. Якщо останню відкинути, то дістанемо нову матрицю коду, але вже з меншою кількістю рядків. Повторивши аналогічну операцію додавання рядків матриці за модулем 2, можна знову дістати нульову комбінацію коду. Ця операція повторюється доти, поки не буде здобута матриця з лінійно незалежними рядками, додавання яких за модулем 2 вже не приведе до утворення нульової комбінації коду.

Наприклад, щоб побудувати матрицю триелементного двійкового простого коду, треба, записавши у вигляді матриці всі $2^n - 1$ комбінацій простого коду, крім нульової, послідовно додати їх за модулем 2, виключаючи кожного разу ті комбінації, які в сумі з попередніми утворюють нульову комбінацію:

1)	2)	3)	4)	5)	6)	7)
001	001	001	001	001	001	100
010	010	010	010	010	010	010
011	000	100	100	100	100	001
100	100	000	000	000		
101	101	110	111			
110	110	111				
111	111					

Друга колонка тут формується так: якщо до третього рядка першої колонки додати суму її першого та другого рядків, то утвориться нульова комбінація, яку виключаємо при запису третьої колонки. Якщо до п'ятого рядка другої колонки після цього додати суму її першого та четвертого рядків, то дістанемо нульову комбінацію в третій колонці. Цю комбінацію також виключаємо, записуючи четверту колонку, і т. д. Таким чином, відкинувши всі нульові комбінації, матимемо шосту колонку з кодовими комбінаціями, що містять тільки по одній одиниці. Це й буде матриця даного коду (сьома колонка).

Квадратна матриця, діагональ якої складається з одиниць, а решта її елементів — нулі, називається *одиничною*. Якщо рядки такої n-елементної матриці додавати за модулем 2, то підбором відповідної комбінації їх можна дістати всі комбінації n-елементного коду. Тому такі матриці ще називаються *визначальними*.

Якщо напрямок головної діагоналі матриці проходить справа наліво, то матриця називається *транспонованою*. Для розглянутого коду це буде матриця (шоста колонка)

$$\begin{bmatrix} 001 \\ 010 \\ 100 \end{bmatrix}.$$

Загалом визначальна матриця n-елементного коду записується так:

$$\left.\begin{bmatrix} 100\ldots000 \\ 010\ldots000 \\ 001\ldots000 \\ \ldots\ldots\ldots\ldots \\ 000\ldots100 \\ 000\ldots010 \\ 000\ldots001 \end{bmatrix}\right\} n \text{ рядків}$$

$$\underbrace{\qquad\qquad}_{n \text{ стовпців}}$$

Розглянутий приклад утворення визначальної матриці й здобута одинична матриця можуть бути використані тільки для побудови всіх комбінацій двійкового простого коду з мінімальною кодовою відстанню $d_{\min} = 1$.

Матричний спосіб може бути застосований також для побудови коректувальних кодів з $d_{\min} > 1$, здатних виявляти та виправляти помилки. Проте при цьому твірна (породжувальна) матриця складається з двох підматриць — уже відомої одиничної (інформаційної) та додаткової (перевірної).

За допомогою інформаційної одиничної підматриці E_k утворюють інформаційну частину кодової комбінації коректувального коду, яка складається з k інформаційних елементів і визначає розмір підматриці $(k \times k)$, що відповідає розмірам визначальної квадратної матриці двійкового простого коду, оскільки кількість $N_д$ комбінацій коректувального коду дорівнює кількості N комбінацій початкового двійкового простого коду, які треба закодувати цим коректувальним кодом.

За допомогою додаткової перевірної підматриці $C_{r,k}$, правило побудови якої описується в п. 8.1.1, де розглядаються коректувальні коди, утворюють перевірну частину комбінації коректувального коду, що складається з r перевірних елементів. Тому додаткова перевірна підматриця має розмір $r \times k$.

Таким чином, загальний розмір твірної матриці коректувальних кодів дорівнює $n \times k$, оскільки $n = k + r$.

5.5. НАДМІРНІСТЬ ПОВІДОМЛЕНЬ І КОДІВ

Від надмірності повідомлень і кодів, якими вони передаються, залежить максимальна кількість інформації, що може бути передана по каналу за одиницю часу. Якщо повідомлення передаються алфавітом q, то максимальну кількість інформації на один елемент (символ, знак) повідомлення $H = \log q$ можна дістати лише в разі його рівноймовірних і незалежних елементів. Реальні коди, які використовуються для кодування повідомлень, майже ніколи не задовольняють цю умову, тому що інформаційне навантаження кожного елемента їх, як правило, менше від того, яке вони могли б забезпечувати. Це свідчить про те, що повідомлення мають інформаційну надмірність.

Розрізняють два види надмірності: *природну* та *штучну*. Першою описується надмірність первинних алфавітів, а другою — вторинних. Природна надмірність поділяється на семантичну та статистичну.

Семантична надмірність випливає з того, що будь-яку думку, яка міститься в повідомленні, можна висловити коротше. Взагалі вважають, що коли повідомлення можна скоротити без втрат його змісту, а потім поновити останній, воно має семантичну надмірність.

Так, повідомлення: «До вечора наступного дня передати інформацію не зможемо у зв'язку з пошкодженням ліній, що з'єднують пункт збирання інформації з периферійними абонентськими пунктами» без значної втрати цінності інформації можна було б сформулювати коротше: «Передача інформації затримується до вечора наступного дня у зв'язку з пошкодженням

абонентських ліній», тобто перше повідомлення має семантичну надмірність відносно другого.

Є багато способів усунення семантичної надмірності: заміною деяких типових повідомлень, які зустрічаються досить часто, умовними позначеннями; введенням таблиць, куди заносяться характерні елементи повідомлення; застосуванням скорочень тощо. Нагадаємо, що всі ці перетворення стосуються первинного алфавіту.

Систематична надмірність спричинена нерівномірним розподілом якісних ознак первинного алфавіту та взаємозалежністю їх. Це можна побачити на прикладі англійського алфавіту, що містить 26 літер. Максимальне значення ентропії англійського алфавіту $H_{max} = \log_2 q = \log_2 26 = 4{,}7$ біт [42]. Проте у зв'язку з тим, що ймовірність появи літер англійського алфавіту не однакова, ентропія англійської мови значно менша ніж 4,7 біт і без урахування взаємозалежності між словами становить приблизно 2,35 біт [42].

Якщо ж урахувати дійсну частоту появи літер у текстах, різних сполученнях і слів у різних повідомленнях, то інформацію, що передається, можна значно скоротити, стиснути. Коефіцієнт ущільнення інформації визначається виразом

$$K_{ущ} = H/H_{max},$$

а надмірність — виразом

$$R_{над} = 1 - K_{ущ} = 1 - H/H_{max}. \qquad (5.5)$$

Із (5.5) випливає, що для зменшення надмірності повідомлення необхідно збільшити ентропію первинного алфавіту.

Для англійської мови

$$R_{над} = 1 - \frac{2{,}35}{4{,}7} = 1 - 0{,}5 = 0{,}5,$$

тобто можна відновити зміст англійських текстів, складених з 50 % алфавіту.

До видів *статистичної надмірності* алфавітів належать такі поняття, як надмірність $R_{над.зв}$, зумовлена статистичним зв'язком між елементами повідомлення, та надмірність $R_{над.р}$, спричинена нерівноймовірним розподілом елементів у повідомленні.

Надмірність $R_{над.зв}$ вказує на інформаційний резерв повідомлень із взаємозалежними елементами відносно повідомлень, які мають статистичний зв'язок між елементами [42]:

$$R_{над.зв} = 1 - H/H',$$

де $H = -\sum_i \sum_j p(a_i) p(b_j/a_i) \log p(b_j/a_i)$; $H' = -\sum_i p_i \log p_i$.

Тут H' теж має надмірність через нерівномірний розподіл ймовірностей окремих елементів алфавіту.

Надмірність $H_{над.р}$ вказує на інформаційний резерв повідомлень, елементи яких нерівноймовірні [42]:

$$R_{над.р} = 1 - H/H_{max},$$

де $H_{max} = \log q$.

Повна статистична надмірність алфавіту визначається виразом

$$R_{над} = R_{над.зв} + R_{над.р} - R_{над.зв} R_{над.р}.$$

При незначних $R_{над.зв}$ і $R_{над.р}$ цей вираз набуває вигляду

$$R_{над} = R_{над.зв} + R_{над.р},$$

тому що зі зменшенням $R_{над.зв}$ і $R_{над.р}$ добуток їх прямує до нуля.

Для усунення статистичної надмірності алфавітів використовують оптимальні нерівномірні коди (див. п.5.7); при цьому статистична надмірність первинного алфавіту значно зменшується завдяки більш раціональній побудові повідомлень у вторинному алфавіті.

Іноді статистична надмірність випливає з природи самого коду. Так, при передачі десяткових чисел двійковим кодом трьома двійковими розрядами можна передати і цифру 5, і цифру 8, тобто для передачі п'яти та восьми повідомлень треба мати коди однакової довжини.

Довжина комбінації двійкового коду визначається виразом

$$n \geq \frac{\log_2 N}{\log_2 q}, \text{ або } n \geq \frac{\log_2 q_1}{\log_2 q_2},$$

де N — кількість повідомлень, яку необхідно передати; q_1, q_2 — відповідно якісні ознаки первинного та вторинного алфавітів.

Так, для передачі $N = 5$ повідомлень двійковим кодом ($q = 2$) потрібно

$$n \geq \frac{\log_2 5}{\log_2 2} = 2{,}32 \approx 3 \text{ двійкових символів.}$$

Загалом надмірність від округлення визначається виразом

$$R_{над.окр} = \frac{K - K_{над}}{K},$$

де K — округлене до найближчого цілого значення $K_{над} = \frac{\log_2 q_1}{\log_2 q_2}$.

У даному разі

$$R_{\text{над.окр}} = \frac{3-2{,}32}{3} \approx 0{,}227,$$

що характеризує недовантаженість коду.

Вираз

$$n \geq \frac{H}{\log q} = \frac{\log N}{\log q} \qquad (5.6)$$

можна застосувати для визначення довжини кодів з рівноймовірними та взаємонезалежними елементами. Для двійкового коду ($q = 2$) цей вираз дійсний тільки тоді, коли ймовірність появи 0 та 1 однакові. Проте в рівномірних кодах, як правило, нулі зустрічаються частіше, ніж одиниці [42]. Тому надмірність, закладену в природу коду, повністю усунути не можна. Однак надмірність від нерівноймовірності появи елемента та надмірність від округлення зменшуються зі збільшенням довжини кодового блока.

На відміну від природної надмірності, яка характерна для первинних алфавітів і присутня в повідомленні ще до того, як воно перетворюється на код, *штучна надмірність* вводиться в нього у вигляді r додаткових елементів спеціально для підвищення його завадостійкості. Таким чином, з n розрядів коду, з яких k несуть інформаційне навантаження, $r = n - k$ розрядів вводяться як коректувальні. Ця величина характеризує *абсолютну* коректувальну надмірність, а величина

$$R_{\text{над}\,r} = \frac{n-k}{n} = 1 - \frac{k}{n} = \frac{r}{n}$$

— відносну коректувальну надмірність коду (див. п. 5.1).

5.6. ОСНОВНІ ТЕОРЕМИ КОДУВАННЯ ДЛЯ КАНАЛІВ

Повідомлення при передачі по каналах кодуються для того, щоб зменшити вплив завад у каналі та забезпечити надійний зв'язок між джерелом й одержувачем повідомлень.

Імовірність неправильної передачі повідомлення по каналу може бути дуже малою, якщо воно передається за допомогою досить великої кількості повторень одного й того самого вхідного сигналу. Проте це пропорційно збільшує час, який відводиться на передачу; при цьому швидкість передачі (тобто кількість інформації, що передається за одиницю часу) прямує до нуля.

Теореми кодування для каналів допомагають зрозуміти, що існують нетривіальні способи кодування, які дають змогу здійснити передачу повідомлень зі скільки завгодно високою вірогідністю та відносно великою швидкістю. Ці теореми не вказують конкретних шляхів побудови пристроїв кодування та декодування, але показують, що вплив завад може бути зведений до мінімуму завдяки вибраному способу кодування та його реалізації.

Теорема кодування для каналу із завадами (яку ще називають *основною теоремою Шеннона* для дискретного каналу із завадами) доводить, що його пропускна здатність визначає верхню межу швидкості безпомилкової передачі інформації по каналу [42]. Формулюється вона так: *існує такий спосіб кодування для дискретного каналу із завадами, при якому можна забезпечити безпомилкову передачу інформації від джерела, якщо продуктивність останнього менша від пропускної здатності каналу, тобто*

$$V_{\text{дж}} H(A) < V_k[\log_2 k - H(B/B')] = C_{\text{к}}, \qquad (5.7)$$

де $V_{\text{дж}}$ — кількість повідомлень, вироблених джерелом A за одиницю часу; $H(A)$ — ентропія джерела; V_k — кількість символів коду, що подаються на вхід каналу за одиницю часу; $[\log_2 k - H(B/B')]$ — максимальна кількість інформації, яка переноситься одним символом коду; $H(B/B')$ — надійність каналу, що визначається дією завад; B — алфавіт обсягом k символів на вході каналу; B' — алфавіт символів, які з'являються на виході каналу.

Для доведення цієї теореми використаємо поняття типових послідовностей повідомлень джерела A, які кодуються на вході каналу символами з множини B (алфавіт входу каналу) та відображаються символами з множини B на виході каналу (алфавіт виходу каналу). Кількість типових послідовностей джерела A великої довжини T при його продуктивності $V_{\text{дж}} H(A)$ визначається виразом

$$N_{\text{т}}(A) = 2^{TV_{\text{дж}} H(A)},$$

де $TV_{\text{дж}}$ — кількість повідомлень джерела A в типовій послідовності.

Для кодування типових послідовностей на вході каналу застосовуємо дискретні кодові комбінації в алфавіті B обсягом k, що дорівнює обсягу алфавіту символів каналу. Якщо довжина цих комбінацій становить також T, то кількість символів у ній буде TV_k, а кількість можливих кодових комбінацій становитиме

$$N(B) = k^{TV_k} = 2^{TV_k \log_2 k}.$$

Тоді нерівність (5.7) можна записати у вигляді

$$V_{\text{дж}}H(A) < V_k \log_2 k - V_k H(B/B').$$

Оскільки $H(B/B') > 0$ за наявності завад у каналі, виключивши другий член у правій частині нерівності, тільки підсилимо останню:

$$V_{\text{дж}}H(A) < V_k \log_2 k,$$

звідки

$$N_{\text{т}}(A) = 2^{TV_{\text{дж}}H(A)} << N(B) = 2^{TV_k \log_2 k}.$$

Це означає, що код B як множина комбінацій алфавіту B, які можуть бути використані для кодування типових послідовностей джерела A, набагато перевищує множину останніх. Таким чином, при кодуванні типових послідовностей джерела A використовується невелика частина кодових комбінацій в алфавіті B, що забезпечує велику різноманітність можливих способів кодування.

У теорії інформації обчисленням середньої імовірності помилки декодування доведено, що серед множини способів кодування є принаймні один, який дає змогу порівняти будь-яку прийняту кодову комбінацію з алфавіту B з однією з комбінацій алфавіту B, використаних при кодуванні, що є умовою вірогідного декодування, причому середня ймовірність помилки декодування із збільшенням довжини коду прямує до нуля.

Теорема Шеннона не встановлює певний спосіб кодування для каналу із завадами, який забезпечує безпомилкову передачу інформації зі швидкістю, як завгодно близькою до пропускної здатності каналу, а лише доводить наявність такого способу на рівні його існування. В цьому відношенні вона є неконструктивною, але дає змогу зробити важливий висновок: *вірогідність передачі інформації по дискретному каналу із завадами тим вища, чим більша довжина кодової комбінації та менша продуктивність джерела А відносно пропускної здатності каналу.*

Таким чином, стає можливою заміна ефективності використання каналу вірогідністю передачі інформації, що широко застосовується в реальних каналах і системах передачі даних.

Для однозначного декодування прийнятих повідомлень, а також для передачі великих обсягів інформації з якомога мінімальнішими матеріальними та часовими витратами, коди мають задовольняти деякі вимоги, які краще викласти у вигляді теорем [13]. Спочатку доведемо теорему, в якій формулюється необхідна умова однозначного декодування коду.

ТЕОРЕМА 5.1. *Нехай код, що однозначно декодується, складається з N комбінацій завдовжки $n_1, n_2, \ldots, n_N$, а його алфавіт містить q символів. Тоді*

$$\sum_{i=1}^{N} q^{-n_i} \le 1.$$

Доведення. Позначивши через L довільне додатне ціле число, запишемо таку рівність:

$$\left(\sum_{i=1}^{N} q^{-n_i}\right)^L = \sum_{i_1=1}^{N}\sum_{i_2=1}^{N}\cdots\sum_{i_L=1}^{N} q^{-\left(n_{i_1}+n_{i_2}+\ldots+n_{i_L}\right)}, \qquad (5.8)$$

кожна складова правої частини якої відповідає кожній можливій послідовності L кодових комбінацій. Сума $n_{i_1} + n_{i_2} + \ldots + n_{i_L}$ дорівнює сумарній довжині послідовності цих комбінацій. Якщо через A_j позначити кількість послідовностей L кодових комбінацій, що мають сумарну довжину j, то (5.8) можна подати у вигляді

$$\left(\sum_{i=1}^{N} q^{-n_i}\right)^L = \sum_{j=1}^{Ln} A_j q^{-j}, \qquad (5.9)$$

де n — максимальне з чисел $n_1, n_2, \ldots, n_N$.

Код однозначно декодується, якщо при будь-яких L та j існує єдина послідовність кодових символів завдовжки j, утворена L кодовими комбінаціями. Через те що q^j — максимальна кількість різних послідовностей завдовжки j, то $A_j \le q^j$. Підставивши цю нерівність у (5.9), дістанемо

$$\left(\sum_{i=1}^{N} q^{-n_i}\right)^L \le Ln, \qquad (5.10)$$

або

$$\left(\sum_{i=1}^{N} q^{-n_i}\right)^L \le (Ln)^{1/L} = 2^{\frac{1}{L}\log nL}. \qquad (5.11)$$

Нерівність (5.11) дійсна для всіх додатних цілих чисел L. Перейшовши до границі при $L \to \infty$, матимемо твердження теореми

$$\sum_{i=1}^{N} q^{-n_i} \le 1. \qquad (5.12)$$

Нерівність (5.12) називається *нерівністю Крафта* [13, 42] для префіксних кодів, який довів, що для того щоб існував префіксний код в алфавіті обсягом q з комбінаціями завдовжки $n_1, n_2, \ldots, n_N$, необхідно й достатньо, щоб $\sum_{i=1}^{N} q^{-n_i} \le 1$. Ця нерівність

є необхідною та достатньою умовою існування кодового дерева, вершини якого мають порядки $n_1, n_2, \ldots, n_N$.

Нехай $\{X, p(x)\}$, $X = \{x_1, x_2, \ldots, x_N\}$ — довільний дискретний ансамбль повідомлень і $H(x)$ — його ентропія. Позначимо через $\bar{n}(x)$ середню довжину q-коду (q – його алфавіт), тобто

$\bar{n}(x) = \sum_{i=1}^{N} n_i p(x_i)$, де N — кількість кодових комбінацій.

ТЕОРЕМА 5.2. *Для будь-якого коду з властивістю однозначного декодування виконується умова*

$$\bar{n}(x) \geq H(x)/\log q.$$

Доведення. З визначення середньої довжини кодових комбінацій маємо

$$\bar{n}(x)\log q = \sum_{i=1}^{N} p(x_i) n_i \log q = \sum_{i=1}^{N} p(x_i)\log q^{n_i}. \qquad (5.13)$$

Розглянемо різницю

$$H(x) - \bar{n}(x)\log q = -\sum_{i=1}^{N} p(x_i)\log p(x_i) + \sum_{i=1}^{N} p(x_i)\log q^{-n_i} = \sum_{i=1}^{N} p(x_i)\log\frac{q^{-n_i}}{p(x_i)}. \qquad (5.14)$$

Урахувавши те, що $\ln x \leq x - 1$ [49], дістанемо

$$H(x) - \bar{n}(x)\log q \leq \log e \sum_{i=1}^{N} p(x_i)\left(\frac{q^{-n_i}}{p(x_i)} - 1\right) = \log e\left(\sum_{i=1}^{N} q^{-n_i} - 1\right) \leq 0. \qquad (5.15)$$

Остання нерівність є наслідком того, що код має властивість однозначного декодування (див. теорему 5.1). Теорему доведено.

ТЕОРЕМА 5.3. *Існує q-код з властивістю однозначного декодування, для якого виконується нерівність*

$$\bar{n}(x) < H(x)/\log q + 1. \qquad (5.16)$$

Доведення. Нехай n_i' — найменше ціле число, що відповідає умові n_i' і $I(x_i)/\log q$, де $I(x_i) = -\log p(x_i)$ — індивідуальна

інформація повідомлення x_i, $i = 1, 2, \ldots, N$. Тоді

$$I(x_i)/\log q \leq n_i' < I(x_i)/\log q + 1. \qquad (5.17)$$

Оскільки

$$\sum_{i=1}^{N} q^{-n_i'} \leq \sum_{i=1}^{N} q^{-I(x_i)/\log q} = \sum_{i=1}^{N} p(x_i) = 1, \qquad (5.18)$$

згідно з (5.12) існує дерево з кінцевими вершинами порядків $n_1', n_2', \ldots, n_N'$. Відповідний код матиме середню довжину

$$\bar{n}(x) = \sum_{i=1}^{N} n_i' p(x_i) < H(x)/\log q + 1. \qquad (5.19)$$

Теорему доведено.

Теореми 5.2 та 5.3 можна узагальнити в разі кодування послідовностей повідомлень, коли ансамбль їх є n-м степенем ансамблю X.

Для цього розглянемо довільний дискретний ансамбль послідовностей повідомлень $\{X^n, p(x)\}$, де n — довжина послідовностей і $H(X^n)$ — ентропія цього ансамблю. Тоді (див. теореми 5.2 та 5.3) для будь-якого q-коду, що однозначно кодує послідовності повідомлень з ансамблю X^n, виконується умова

$$\bar{n}\Delta\frac{\bar{n}(X^n)}{n} \geq \frac{H(X^n)}{n\log q}, \qquad (5.20)$$

де $\bar{n}$ — середня кількість символів, яка припадає на одне повідомлення.

Крім того, існує код, для якого справджується нерівність

$$\bar{n} < \frac{H(X^n)}{n\log q} + \frac{1}{n}. \qquad (5.21)$$

Нехай є стаціонарне джерело U_x, що вибирає повідомлення з множини X і має ентропію $H(X/X^\infty)$. Нехай також це джерело кодується за допомогою нерівномірного q-коду, тобто кожній послідовності повідомлень $x = (x^{(1)}, x^{(2)}, \ldots, x^{(n)}) \in X^n$ ставиться у відповідність кодова комбінація завдовжки $n(x)$. Середня довжина коду $\bar{n}(X^n) = \sum_{X^n} p(x)n(x)$, а середня кількість кодових символів, яка припадає на одне повідомлення джерела,

$$\bar{n} = \frac{\bar{n}(X^n)}{n} = \frac{1}{n}p(x)n(x). \qquad (5.22)$$

При цьому середня швидкість кодування

$$R \overset{\Delta}{=} \overline{n} \log q \overset{\Delta}{=} \frac{\log q}{n} \sum_{X^n} p(x) n(x). \qquad (5.23)$$

ТЕОРЕМА 5.4 (обернена теорема кодування). *Для будь-якого коду, що однозначно кодує джерело U_x, середня швидкість кодування задовольняє умову*

$$R \le H(X/X^{\infty}). \qquad (5.24)$$

Доведення. З визначення середньої швидкості кодування (5.23) та умови (5.20), а також виходячи з того, що для будь-якого дискретного стаціонарного джерела послідовність $H(X_n/X^{n-1})$ має границю

$$\lim_{n\to\infty} H\left(X_n \; X^{n-1}\right) \overset{\Delta}{=} H\left(X \; X^{\infty}\right),$$

випливає, що для всіх $n = 1,2,\ldots$ виконується умова

$$R \ge H(X^n)/n \ge H(X_n/^{n-1}) \ge H(X/X^{\infty}). \qquad (5.25)$$

Теорему доведено.

ТЕОРЕМА 5.5 (пряма теорема кодування). *Нехай* ε — *довільне ціле додатне число, а q — алфавіт коду. Існує однозначно декодований q-код, що кодує джерело U_x, для якого виконується умова*

$$R < H(X/X^{\infty}) + \varepsilon. \qquad (5.26)$$

Доведення. Для будь-якого додатного цілого числа ε_1 знайдеться таке число $N(\varepsilon_1)$, що для всіх $n > N(\varepsilon_1)$ справджується нерівність

$$H(X^n)/n < H(X/X^{\infty}) + \varepsilon_1. \qquad (5.27)$$

З (5.21) і (5.27) випливає, що для довільного цілого числа $q > 0$ існує однозначно декодований q-код із середньою кількістю елементів на одне повідомлення

$$\overline{n} < \frac{H(X^n)}{n \log q} + \frac{1}{n} < \frac{H(X/X^{\infty})}{\log q} + \frac{\varepsilon_1}{\log q} + \frac{1}{n} \qquad (5.28)$$

і, як наслідок, зі швидкістю кодування

$$R < H(X/X^{\infty}) + \varepsilon_1 + \frac{\log q}{n} \qquad (5.29)$$

Беручи $\varepsilon_1 = \varepsilon/2$, дістаємо, що для всіх $n \ge \frac{\log q}{\varepsilon - \varepsilon_1} = \frac{2\log q}{\varepsilon}$ виконується умова $\varepsilon_1 + \frac{\log q}{n} \le \varepsilon$.

Твердження теореми дійсне тепер для всіх n, більших ніж максимальне з чисел $N(\varepsilon/2)$ та $\frac{2\log q}{\varepsilon}$.

Теорему доведено.

З теорем 5.4 та 5.5 випливає, що середня швидкість створення інформації дискретним стаціонарним джерелом дорівнює $H(X/X^{\infty})$ — ентропії на одне повідомлення, тобто тій самій величині, що й у задачі рівномірного кодування. Звідси випливає, що мінімальна кількість двійкових елементів, яка припадає в середньому на одне повідомлення джерела, може бути дуже близькою до $H(X/X^{\infty})$ як при рівномірному, так і при нерівномірному кодуванні.

5.7. ОПТИМАЛЬНЕ КОДУВАННЯ

Знайти код, який був би оптимальним з усіх точок зору практично неможливо. Тому код може бути оптимальним тільки за певних умов (з точок зору швидкості передачі інформації, здатності виправляти помилки тощо).

У теорії інформації існує кілька методик побудови оптимальних з точки зору швидкості передачі інформації безнадмірних кодів.

До *оптимальних безнадмірних кодів* (з точки зору довжини їх, тобто швидкості передачі інформації) належать нерівномірні коди, які передають повідомлення комбінаціями мінімальної середньої довжини. Це зовсім не означає, що вони дійсно є абсолютно безнадмірними, оскільки такими вважаються коди, які задовольняють умову рівності обсягу та кількості інформації (див. п.2.2). Ці коди все ж таки мають потенціальну надмірність через заборонені кодові комбінації, до яких належать комбінації, що доповнюють вершини неповного кодового дерева, яке відповідає оптимальному нерівномірному коду (ОНК), до повного (див. п.5.4) утворення рівномірного коду.

Відповідно до [42] *оптимальним кодуванням* називається процедура перетворення символів первинного алфавіту q_1 на кодові комбінації вторинного алфавіту q_2, при якій середня довжина повідомлення у вторинному алфавіті мінімальна.

Таким чином, основним завданням оптимального кодування є досягнення рівності між кількістю інформації I, що виробляється джерелом повідомлень, та обсягом інформації Q на вхо-

ді приймача повідомлень. Якщо $I = QI_{сер} = H$, то збільшення швидкості передачі інформації завдяки поліпшенню процедури кодування стає неможливим.

Ґрунтуючись на наведених у п.5.6 теоремах, можна запропонувати кілька методик побудови ОНК для дискретних ансамблів повідомлень $\{X, p(x)\}$ із середньою довжиною кодових комбінацій

$$\bar{n}(X) = H(X)/\log q.$$

Перша універсальна методика побудови ОНК ґрунтується на методиці Шеннона — Фано [36, 44] і передбачає цю побудову в кодовому алфавіті з кількістю якісних значень q. Згідно з цією методикою виконують такі процедури:

1) множину з N повідомлень, які кодуються, розташовують у порядку спадання ймовірностей;

2) впорядковані за ймовірностями повідомлення розбивають, по можливості, на q рівноймовірних груп;

3) кожній з груп завжди в одній і тій самій послідовності присвоюють символи алфавіту q (всім повідомленням першої групи — першу якісну ознаку цього алфавіту, всім повідомленням другої групи — другу якісну його ознаку тощо);

4) створені групи розбивають, по можливості, на рівноймовірні підгрупи, кількість яких дорівнює або менша ніж q (якщо після розбивання в групі залишається одне повідомлення, то подальший поділ стає неможливим);

5) кожній з утворених підгруп присвоюють якісні ознаки з алфавіту q за процедурою п.3;

6) розбивання та присвоєння ознак алфавіту q повторюють доти, поки після чергового поділу в утворених підгрупах залишиться не більш як одне повідомлення.

Для побудови ОНК за викладеною методикою слід ураховувати також відхилення від рівноймовірних значень, що утворюються при поділі на підгрупи. Вони враховуються згідно з правилами заліку остач ділення та середнього відхилення:

1) для того щоб повідомлення первинного джерела можна було поділити по можливості на якомога рівноймовірні підгрупи при побудові ОНК з алфавітом q, остача попереднього ділення додається за абсолютним значенням сумарної ймовірності чергового ділення [*остачею ділення* називається різниця між квантом ділення та реальним значенням сумарної ймовірності в групі (підгрупі), де квант ділення дорівнює $1/q$];

2) середнє відхилення має бути меншим або дорівнювати значенню ймовірності першого символу чергового ділення. Якщо середнє відхилення не дорівнює нулю, то середнє значення сумарної ймовірності в групі (підгрупі) при черговому діленні підраховується з додаванням значення середнього відхилення (*середнім відхиленням* називається абсолютне значення суми остач ділень на проміжних етапах побудови коду).

Розглянемо приклад побудови ОНК для передачі 16 повідомлень за допомогою четвіркового коду з алфавітом $q = 4$, якщо повідомлення на виході джерела з'являються з імовірностями $p(x_i)$, як зазначено в табл. 5.5. Послідовність цієї побудови така:

1. Визначаємо квант поділу $1/q = 1/4 = 0{,}25$.

2. Розбиваємо всі повідомлення, по можливості, на чотири рівномірні групи:

$\sum_{i=1}^{1} p(x_i) = 0{,}22$, остача від ділення дорівнює 0,03, оскільки (0,25 – 0,22) < (0,32 – 0,25);

$\sum_{i=2}^{4} p(x_i) = 0{,}28$, остача дорівнює 0,03, середнє відхилення по двох діагоналях нульове;

$\sum_{i=5}^{8} p(x_i) = 0{,}26$, остача дорівнює 0,01, середнє відхилення по двох діагоналях становить 0,01;

$\sum_{i=9}^{16} p(x_i) = 0{,}24$, остача дорівнює — 0,01, середнє відхилення по двох діагоналях нульове.

Таблиця 5.5

Номер повідомлення	Імовірність повідомлення $p(x_i)$	Поділ на групи (підгрупи)			Кодова комбінація ОНК
		Перша	Друга	Третя	
1	0,22	→			0
2	0,1	→			10
3	0,1	→			11
4	0,08	→			12
5	0,07	→			20
6	0,07	→			21
7	0,06	→			22
8	0,06	→			23
9	0,05	→			30
10	0,05	→			31
11	0,04	→			320
12	0,03	→			321
13	0,02	→			330
14	0,02	→			331
15	0,02	→			332
16	0,01	→			333

3. За результатами першого поділу груп як перший символ кодових комбінацій присвоюємо послідовно якісні ознаки алфавіту $q = 4$ згідно з третьою процедурою першої універсальної методики побудови ОНК.

4. Утворені групи, крім першої, розбиваємо на підгрупи. Друга та третя групи мають до чотирьох повідомлень, тому як другий символ кодових комбінацій їм присвоюємо відповідно три та чотири якісні ознаки алфавіту q.

5. Четверта група має вісім повідомлень, тому розбиваємо її на рівноймовірні підгрупи. Квант поділу (0,24 + 0,01) : 4 = 0,0625 (0,01 додається як остача від попереднього поділу).

Поділ присвоєння ознак алфавіту q виконуємо доти, поки після чергового поділу в утворених підгрупах залишиться не більш як одне повідомлення.

6. Для того щоб перевірити оптимальність коду відносно довжини кодових комбінацій, визначаємо середню довжину $n_{сер}$ кодової комбінації ОНК. У разі оптимальності ця довжина не повинна перевищувати довжину рівномірного четвіркового коду, яким можна закодувати 16 повідомлень, тобто $q^n = 4^2 = 16 (n = 2)$:

$$n_{сер} = \sum_{i=1}^{16} p(x_i) n_i = 0{,}22 \cdot 1 + (0{,}1 + 0{,}1 + 0{,}08 + 0{,}07 + 0{,}07 + 0{,}06 +$$

$$+ 0{,}06 + 0{,}05 + 0{,}05) \cdot 2 + (0{,}04 + 0{,}03 + 0{,}02 + 0{,}02 + 0{,}01 + 0{,}01) \cdot 3 =$$

$$= 0{,}22 + 0{,}64 \cdot 2 + 0{,}14 \cdot 3 = 0{,}22 + 1{,}28 + 0{,}42 = 1{,}92 < 2.$$

Таким чином, утворений код дійсно оптимальний, оскільки $n_{сер} < n$.

Оптимальність кодування можна визначити також порівнянням ентропії, що припадає на одне повідомлення, із середньою довжиною кодової комбінації, які мають бути дуже близькі за значеннями, причому ентропія має бути меншою від середньої довжини кодової комбінації або дорівнювати їй.

Отже,

$$H = -\sum_{i=1}^{16} p(x_i) \log p(x_i) = -(0{,}22 \log 0{,}22 + 0{,}1 \log 0{,}1 + 0{,}1 \log 0{,}1 + 0{,}08 \log 0{,}08 +$$

$$+ 0{,}07 \log 0{,}07 + 0{,}06 \log 0{,}06 + 0{,}06 \log 0{,}06 + \ 0{,}05 \log 0{,}05 + 0{,}05 \log 0{,}05 +$$

$$+ 0{,}04 \log 0{,}04 + 0{,}03 \log 0{,}03 + 0{,}02 \log 0{,}02 + 0{,}02 \log 0{,}02 + 0{,}02 \log 0{,}02 +$$

$$+ 0{,}01 \log 0{,}01) = 0{,}4806 + 2 \cdot 0{,}3322 + 0{,}2915 + 2 \cdot 0{,}2686 + 2 \cdot 0{,}2435 +$$

$$+ 2 \cdot 0{,}2161 + 0{,}1857 + 0{,}1517 + 3 \cdot 0{,}1129 + 0{,}0664 = 3{,}6354 \text{ біт/повідомлення.}$$

При цьому треба звернути увагу на те, що ентропія залежить від алфавіту оптимального коду, яким кодуються повідомлення, тобто слід урахувати кількість інформації, яка міститься в одному елементі кодової комбінації. Для оптимального коду $q = 4$ в одному елементі кодової комбінації буде 2 біти інформації.

Таким чином,

$$H = 3{,}6354 \text{ біт/повідомлення} < 2n_{сер} =$$
$$= 3{,}84 \text{ біт/повідомлення} < 2n = 2 \cdot 2 = 4 \text{ біт/повідомлення,}$$

тобто код є оптимальним.

Друга універсальна методика побудови ОНК ґрунтується на відомій методиці Хаффмена [12, 42]. Вона, як і методика Шеннона — Фано, передбачає побудову ОНК у кодовому алфавіті з кількістю якісних значень q. Згідно з цією методикою виконують такі процедури:

1) множину з N повідомлень, що кодуються, розташовують у порядку спадання ймовірностей;

2) останні N_0 повідомлень ($2 \le N_0 \le q$) об'єднують у нове повідомлення з імовірністю, що дорівнює сумі ймовірностей об'єднуваних повідомлень;

3) утворену множину ($N - N_0 + 1$) повідомлень розташовують у порядку спадання ймовірностей;

4) об'єднують останні q повідомлень і впорядковують множину повідомлень у порядку спадання ймовірностей. Так діють доти, доки ймовірність чергового об'єднаного повідомлення не дорівнюватиме одиниці;

5) будують кодове дерево, починаючи з кореня, і гілкам цього дерева присвоюють якісні ознаки кодового алфавіту q.

Кодові комбінації ОНК — це послідовність якісних ознак, які зустрічаються на шляху від кореня до вершини кодового дерева.

Побудову ОНК за допомогою другої універсальної методики розглянемо стосовно передачі 16 повідомлень комбінаціями четвіркового коду, які задано в попередньому прикладі. Послідовність цієї побудови така:

1. Множину з $N = 16$ повідомлень розташовуємо в порядку спадання ймовірностей.

2. Оскільки $q = 4$, об'єднуємо останні чотири повідомлення й утворюємо нове умовне повідомлення з імовірністю, що дорівнює сумі ймовірностей об'єднаних повідомлень.

3. Утворену множину з 16 – 4 + 1 = 13 повідомлень розташовуємо в порядку спадання ймовірностей.

4. Знову об'єднуємо останні чотири повідомлення та впорядковуємо множину повідомлень у порядку спадання ймовірностей. Цю процедуру повторюємо ще три рази, поки при останньому об'єднанні сумарна ймовірність не досягне значення одиниці (табл. 5.6).

5. Будуємо кодове дерево (рис. 5.3). Його гілкам присвоюємо якісні ознаки кодового алфавіту від 0 до 3.

6. Кодові комбінації ОНК заносимо до табл. 5.6. Вони визначаються послідовністю якісних ознак, які зустрічаються на шляху від кореня до певної вершини кодового дерева.

7. Для того щоб перевірити оптимальність коду відносно довжини кодових комбінацій, визначаємо середню довжину $n_{сер}$ кодової комбінації ОНК та ентропію, що припадає на одне повідомлення:

$$n_{сер} = 0{,}22 \cdot 1 + (0{,}1 + 0{,}1 + 0{,}08 + 0{,}07 + 0{,}07 + 0{,}06 + 0{,}06 + 0{,}05 +$$
$$+ 0{,}05 + 0{,}04 + 0{,}03) \cdot 2 + (0{,}02 + 0{,}02 + 0{,}02 + 0{,}01) \cdot 3 =$$
$$= 0{,}22 + 0{,}71 \cdot 2 + 0{,}07 \cdot 3 = 1{,}85 < 2;$$

$$H = 0{,}6354 \text{ біт/повідомлення} < 2n_{сер} = 3{,}7 \text{ біт/повідомлення} < 2n =$$
$$= 4 \text{ біт/повідомлення.}$$

Таблиця 5.6

Номер повідомлення	Імовірність повідомлення $p(x_i)$	Імовірності повідомлень при об'єднаннях					Кодова комбінація ОНК
		першому	другому	третьому	четвертому	п'ятому	
1	0,22	0,22	0,22	0,26	0,35	1	2
2	0,1	0,1	0,17	0,22	0,26		00
3	0,1	0,1	0,1	0,17	0,22		01
4	0,08	0,08	0,1	0,1	0,17		02
5	0,07	0,07	0,08	0,1			03
6	0,07	0,07	0,07	0,08			10
7	0,06	0,07	0,07	0,07			12
8	0,06	0,06	0,07				13
9	0,05	0,06	0,06				30
10	0,05	0,05	0,06				31
11	0,04	0,05					32
12	0,03	0,04					33
13	0,02	0,03					110
14	0,02						111
15	0,02						112
16	0,01						113

Отже, цей код також є оптимальним і має кращі показники, ніж оптимальний код, побудований за першою універсальною методикою, оскільки $n_{сер2} < n_{сер1}$ (1,85 < 1,92).

Обидві універсальні методики мають неоднозначність, але перша з них дає змогу точніше будувати ОНК. До недоліків другої універсальної методики побудови ОНК слід віднести громіздкість (особливо зі збільшенням кількості повідомлень N та алфавіту q коду), що пояснюється необхідністю побудови кодового дерева.

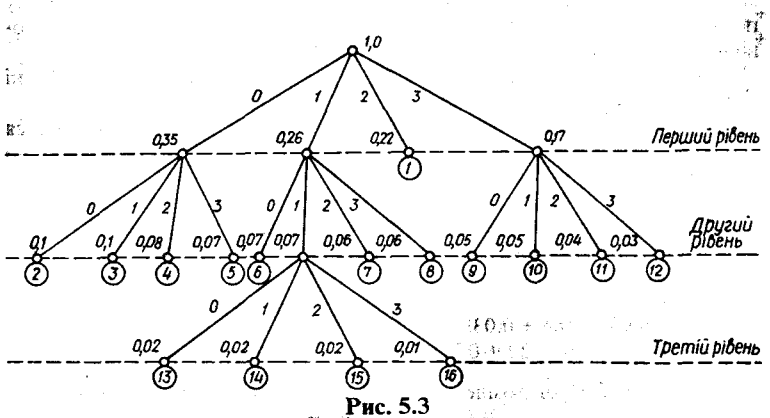


Рис. 5.3

Відзначимо, що переваги другої універсальної методики побудови ОНК з $q > 2$ при $N < q^n$ будуть вагоміші при більш ретельному виборі кількості найменш імовірних повідомлень, що об'єднуються на першому етапі ($2 \le N_0 \le q$). На всіх наступних етапах ця кількість має дорівнювати q.

КОНТРОЛЬНІ ЗАДАЧІ

1. Записати десяткове число 184 у двійковій системі числення.

Розв'язання. Виконуємо послідовне ділення десяткового числа 184 на основу двійкової системи числення $q = 2$:

184 : 2 = 92 + (остача 0); 92:2 = 46 + (0); 46 : 2 = 23 + (0); 23 : 2 = 11 + (1); 11 : 2 = 5 + (1); 5 : 2 = 2 + (1); 2 : 2 = 1 + (0); 1 : 2 = 0 + (1).

Записуємо остачі від ділення так: 10111000. Це й буде двійковим поданням десяткового числа 184.

Перевіряємо правильність запису десяткового числа 184 у двійковій системі числення, для чого виконуємо зворотний перехід від двійкового числа 10111000 до десяткового 184:

$(10111000)_2 = 1 \cdot 2^7 + 0 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 0 \cdot 2^1 + 0 \cdot 2^0 = 128 + 32 + 16 + 8 = (184)_{10}$.

2. Записати десяткове число 382 у вісімковій системі числення.

Розв'язання. Виконуємо послідовне ділення десяткового числа 382 на основу вісімкової системи числення $q = 8$:

382 : 8 = 47 + (остача 6); 47 : 8 = 5 + (7); 5 : 8 = 0 + (5).

Записуємо остачі від ділення так: 576. Це і є записом десяткового числа 382 у вісімковій системі числення.

Перевіряємо правильність запису десяткового числа 382 у вісімковій системі числення, для чого виконуємо зворотний перехід від вісімкового числа 576 до десяткового 382:

$(576)_8 = 5 \cdot 8^2 + 7 \cdot 8^1 + 6 \cdot 8^0 = 320 + 56 + 6 = (382)_{10}$.

3. Записати десяткове число 333 у двійковій системі числення.

4. Записати десяткове число 91 у трійковій системі числення.

5. Записати десяткове число 815 у четвірковій системі числення.

6. Записати десяткове число 4327 у шістнадцятковій системі числення.

7. Визначити кодову відстань між комбінаціями A та B двійкового коду і записати всі комбінації, які знаходяться від комбінації A на кодовій відстані $d = 3$, якщо $A = 01001$, $B = 11101$.

Розв'язання. Виконуємо додавання за модулем 2 комбінацій A та B коду ($A \oplus B = C$):

$$\begin{array}{r} 01001 \\ \oplus\ 11101 \\ \hline 10100, \end{array}$$

тобто $C = 10100$. Вага w комбінації C дорівнює 2, тому що в комбінаціях A та B на трьох однойменних позиціях (першій справа, другій та четвертій) знаходяться однакові елементи, а на двох (третій справа та п'ятій) — різні елементи, сукупність яких і визначає ступінь різниці між комбінаціями A та B. Вага комбінації C і є кодовою відстанню Хеммінга між комбінаціями A та B, тобто $d = 2$.

Будь-яка комбінація вагою $w = 3$, якщо її порозрядно додати до комбінації A (тієї самої довжини), дає нову комбінацію, що знаходитиметься від комбінації A на кодовій відстані $d = 3$. Випишемо всі комбінації вагою $w = 3$ завдовжки $n = 5$ (їх кількість становитиме $C_n^w = C_n^d = C_5^3 = 10$, де $C_n^d = n!/[d!(n-d!)]$ – число сполучень із n по d): 00111, 01011, 01101, 01110, 10011, 10101, 10110, 11001, 11010, 11100.

Додаючи порозрядно кожну з цих комбінацій до комбінації A, дістаємо шукані кодові комбінації:

$$\begin{array}{lllll} \oplus\ \begin{array}{r}00111\\01001\end{array} & \oplus\ \begin{array}{r}01011\\01001\end{array} & \oplus\ \begin{array}{r}01101\\01001\end{array} & \ldots, & \oplus\ \begin{array}{r}11100\\01001\end{array} \\ \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} & & \underline{\quad\quad} \\ 01110, & 00010, & 00100, & \ldots, & 10101; \end{array}$$

8. Побудувати всі комбінації n-розрядного двійкового простого коду, що знаходяться від комбінації A на кодовій відстані $d = 1, 2, 3$, якщо $A = 10101$ і $n = 5$.

Розв'язання. Для побудови шуканих комбінацій потрібно до заданої комбінації A додати комбінацію n-розрядного ($n = 5$) коду з відповідною вагою. Додавання комбінацій виконуємо порозрядно за модулем 2. При цьому дістаємо такі кодові комбінації, які знаходяться від комбінації A на відстанях:

$d = 1$

$$\begin{array}{lllll} \oplus\ \begin{array}{r}10101\\00001\end{array} & \oplus\ \begin{array}{r}10101\\00010\end{array} & \oplus\ \begin{array}{r}10101\\00100\end{array} & \oplus\ \begin{array}{r}10101\\01000\end{array} & \oplus\ \begin{array}{r}10101\\10000\end{array} \\ \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} \\ 10100, & 10111, & 10001, & 11101, & 00101; \end{array}$$

$d = 2$

$$\begin{array}{llll} \oplus\ \begin{array}{r}10101\\00011\end{array} & \oplus\ \begin{array}{r}10101\\00101\end{array} & \oplus\ \begin{array}{r}10101\\00110\end{array} & \oplus\ \begin{array}{r}10101\\11000\end{array} \\ \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} \\ 10110, & 10000, & 10011, & 01101; \end{array}$$

$d = 3$

$$\begin{array}{llll} \oplus\ \begin{array}{r}10101\\00111\end{array} & \oplus\ \begin{array}{r}10101\\01011\end{array} & \oplus\ \begin{array}{r}10101\\01101\end{array} & \oplus\ \begin{array}{r}10101\\11100\end{array} \\ \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} & \underline{\quad\quad} \\ 10010, & 11110, & 11000, & 01001. \end{array}$$

Взагалі кількість комбінацій відповідної ваги визначається як $C_n^w = C_n^d$. Отже, кількість комбінацій, які знаходяться від комбінації A на відстані $d = 1$, буде $C_5^1 = 5$, на відстані $d = 2$ – $C_5^2 = 10$ і на відстані $d = 3$ – $C_5^3 = 10$.

9. Визначити кодову відстань між комбінаціями A та B двійкового коду і записати всі комбінації, які знаходяться від комбінацій A та B на відстані $d = 2$, якщо $A = 110101$, $B = 101100$.

10. Визначити мінімальну та максимальну кодові відстані d Хеммінга між комбінаціями двійкового простого коду, вказавши на пари комбінацій з d_{min} і d_{max} для комбінацій 000011, 110111, 010100, 101001, 011101.

11. Побудувати всі комбінації n-розрядного двійкового простого коду, які знаходяться від комбінації A на кодовій відстані $d = 1,2,3$; згрупувати їх за вагою та підрахувати кількість комбінацій, які знаходяться від комбінації A на відстані $d = 4...n$, якщо $A = 10110$ і $n = 5$.

12. Підрахувати кількість усіх комбінацій n-розрядного двійкового простого коду, які знаходяться від комбінації A на відстані d, якщо $A = 00010$, $d = 2$.

13. Визначити кодову відстань між комбінаціями A, B, C, D рівномірного трійкового коду завдовжки $n = 3$, якщо $A = 021$, $B = 002$, $C = 212$, $D = 120$.

Розв'язання. Виконуємо порозрядне віднімання комбінацій за модулем 3. Кодова відстань дорівнює вазі комбінації, що містить різницю комбінацій, між якими визначається ця відстань:

$$\begin{array}{lll} \ominus\ \left.\begin{array}{r}021\\002\\\hline 021\end{array}\right\} \begin{array}{l} d=3\\ w=3\end{array} & \ominus\ \left.\begin{array}{r}021\\212\\\hline 211\end{array}\right\} \begin{array}{l} d=4\\ w=4\end{array} & \ominus\ \left.\begin{array}{r}021\\120\\\hline 101\end{array}\right\} \begin{array}{l} d=2\\ w=2\end{array} \\ \ominus\ \left.\begin{array}{r}002\\212\\\hline 210\end{array}\right\} \begin{array}{l} d=3\\ w=3\end{array} & \ominus\ \left.\begin{array}{r}002\\120\\\hline 122\end{array}\right\} \begin{array}{l} d=5\\ w=5\end{array} & \ominus\ \left.\begin{array}{r}212\\120\\\hline 112\end{array}\right\} \begin{array}{l} d=4\\ w=4\end{array} \end{array}$$

14. Визначити кодову відстань між такими комбінаціями рівномірного трійкового коду завдовжки $n = 4$: 1020, 0211, 0122, 2012, 2221, 1110.

15. Визначити кодову відстань між такими комбінаціями рівномірного четвіркового коду завдовжки $n = 3$: 031, 123, 303, 210, 022, 111.

16. Задати в табличній формі кодові комбінації двійкового блокового коду постійною вагою $w = 2$ завдовжки $n = 5$.

Розв'язання. Визначаємо кількість комбінацій заданого коду $C_n^w = C_5^2 = 10$.

Записуємо комбінації коду у табличній формі:

Номер комбінації	Комбінація двійкового блокового коду з $w = 2$ та $n = 5$
1	00011
2	00101
3	00110
4	01001
5	01010
6	01100
7	10001
8	10010
9	10100
10	11000

17. Задати в табличній формі кодові комбінації двійкового блокового коду зі сталою вагою $w = 4$ завдовжки $n = 6$.

18. Задати в табличній формі кодові комбінації двійкового простого коду завдовжки $n = 4$.

19. Задати в табличній формі кодові комбінації трійкового блокового простого коду завдовжки $n = 3$.

20. Задати за допомогою кодового дерева комбінації нерівномірного двійкового коду з максимальною довжиною $n = 4$.

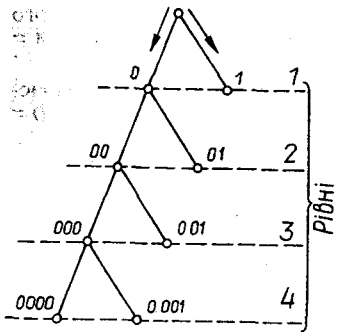


Рис. 5.4

Розв'язання. Максимальна довжина $n = 4$ нерівномірного двійкового коду визначає кількість рівнів кодового дерева; тому воно матиме вигляд, показаний на рис. 5.4.

З рисунка випливає, що кількість комбінацій (1, 01, 001, 0000, 0001) цього коду дорівнює п'яти.

21. Задати за допомогою кодового дерева комбінації рівномірного двійкового коду завдовжки $n = 4$.

22. Задати за допомогою кодового дерева комбінації рівномірного трійкового коду завдовжки $n = 3$.

23. Задати за допомогою кодового дерева комбінації рівномірного вісімкового коду завдовжки $n = 2$.

24. Задати за допомогою кодового дерева комбінації нерівномірного двійкового коду з максимальною довжиною $n = 5$.

25. Побудувати визначальну матрицю чотириелементного двійкового простого коду.

Розв'язання. Записуємо всі $2^n - 1$ комбінацій простого коду, крім нульових, у вигляді матриці та послідовно додаємо їх за модулем 2, виключивши ті комбінації, які в сумі з попередніми утворюють нульову комбінацію:

1)	2)	3)	4)
0001	0001	0001	1000
0010	0010	0010	0100
0011	0000	0100	0010
0100	0100	1000	0001
0101	0000		
0110	0000		
0111	0000		
1000	1000		
1001	0000		
1010	0000		
1011	0000		
1100	0000		
1101	0000		
1110	0000		
1111	0000		

Таким чином, визначальна матриця чотириелементного двійкового простого коду має вигляд

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

26. Побудувати визначальну матрицю п'ятиелементного двійкового простого коду.

27. Побудувати визначальну матрицю восьмиелементного двійкового простого коду.

Таблиця 5.7

Номер повідомлення	Імовірність повідомлення $P(x_i)$	Поділ на групи (підгрупи)					Кодова комбінація ОНК
		першу	другу	третю	четверту	п'яту	
1	0,3	→					00
2	0,2	→					01
3	0,15	→					100
4	0,12	→					101
5	0,1	→					110
6	0,08	→					1110
7	0,03	→					11110
8	0,02	→					11111

28. Побудувати двійковий ОНК Шеннона — Фано для восьми повідомлень джерела з імовірностями $P(x_1) = 0,3$; $P(x_2) = 0,2$; $P(x_3) = 0,15$; $P(x_4) = 0,12$; $P(x_5) = 0,1$; $P(x_6) = 0,08$; $P(x_7) = 0,03$; $P(x_8) = 0,02$.

Розв'язання. Користуючись першою універсальною методикою побудови ОНК (див. п. 5.7), будуємо заданий код (табл. 5.7).

Перевіряємо утворений код на оптимальність, для чого визначаємо середню кількість елементів, яка припадає на одну комбінацію коду Шеннона — Фано:

$$n_{сер} = 2(0,3 + 0,2) + 3(0,15 + 0,12 + 0,1) + 4 \cdot 0,08 + 5(0,03 + 0,02) = \\ = 1 + 1,11 + 0,32 + 0,25 = 2,68 < 3,$$

тобто код оптимальний.

29. Побудувати двійковий ОНК Шеннона — Фано для ансамблю повідомлень з імовірностями 0,16; 0,2; 0,14; 0,4; 0,02; 0,03; 0,05.

30. Розв'язати попередню задачу для ансамблю повідомлень з імовірностями 0,16; 0,11; 0,04; 0,12; 0,07; 0,07; 0,09; 0,03; 0,1; 0,02; 0,02; 0,01; 0,06; 0,04; 0,01; 0,05.

Таблиця 5.8

Номер повідомлення	Імовірність повідомлення $p(x_i)$	Імовірності повідомлень при об'єднаннях							Кодова комбінація ОНК
		першому	другому	третьому	четвертому	п'ятому	шостому	сьомому	
1	0,3	0,3	0,3	0,3	0,3	→ 0,42	→ 0,58	→ 1	00
2	0,2	0,2	0,2	→ 0,22	→ 0,28	0,3	0,42		11
3	0,15	0,15	0,15	0,2	0,22	0,28			010
4	0,12	0,12	→ 0,13	0,12	0,2				100
5	0,1	0,1	0,12	0,13					101
6	0,08	0,08	0,1						0110
7	0,03	→ 0,05							01110
8	0,02								01111

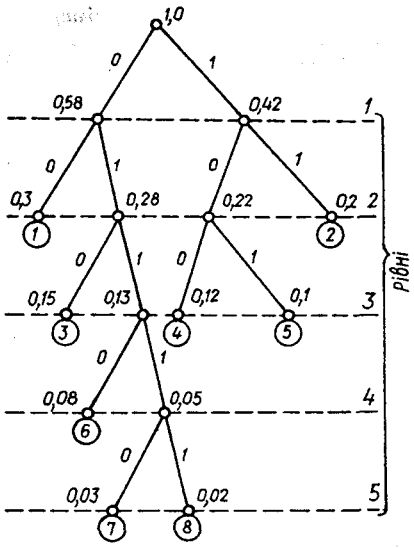


Рис. 5.5

31. Побудувати трійковий ОНК Шеннона — Фано для ансамблю повідомлень з імовірностями 0,15; 0,35; 0,2; 0,03; 0,02; 0,05; 0,1; 0,04; 0,06.

32. Побудувати двійковий ОНК Хаффмена для ансамблю повідомлень з імовірностями, заданими в задачі 28.

Розв'язання. Користуючись другою універсальною методикою побудови ОНК (див. п.5.7), будуємо заданий код (табл. 5.8).

Будуємо кодове дерево, що має вигляд, зображений на рис. 5.5.

Перевіряємо утворений код на оптимальність, для чого визначаємо середню кількість елементів, яка припадає на одну комбінацію коду Хаффмена:

$$n_{сер} = 2(0,3 + 0,2) + 3(0,15 + 0,12 + 0,1) + 4 \cdot 0,08 + 5\,(0,03 + 0,02) =$$
$$= 1 + 1,11 + 0,32 + 0,25 = 2,68 < 3,$$

тобто код оптимальний.

33. Розв'язати попередню задачу для ансамблю повідомлень з імовірностями 0,07; 0,1; 0,03; 0,05; 0,05; 0,16; 0,08; 0,14; 0,1; 0,1; 0,04; 0,01; 0,03; 0,02; 0,02.

34. Побудувати двійковий ОНК Хаффмена для ансамблю повідомлень з імовірностями 0,06; 0,25; 0,1; 0,05; 0,2; 0,04; 0,3.

35. Побудувати трійковий ОНК Хаффмена для ансамблю повідомлень з імовірностями 0,03; 0,08; 0,055; 0,2; 0,04; 0,07; 0,14; 0,36.

36. Для ансамблю повідомлень з імовірностями 0,15; 0,1; 0,05; 0,25; 0,02; 0,03; 0,35 побудувати двійкові ОНК Шеннона — Фано та Хаффмена, порівнявши їх за оптимальністю.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Яка різниця між двійковими та недвійковими кодами?
2. Чим відрізняються рівномірні коди від нерівномірних?
3. Яка різниця між подільними та неподільними блоковими кодами?
4. Яка особливість побудови систематичних подільних блокових кодів?
5. Якими параметрами характеризуються коди?
6. Що таке кодова відстань і що характеризує мінімальна кодова відстань коду?
7. Як перейти з однієї системи числення до іншої?
8. Які основні операції виконуються над елементами поля $G\overline{F}(q)$?
9. Які способи використовуються для подання кодів?
10. У чому полягає спосіб зображення комбінації коду у вигляді кодового дерева?
11. У чому полягає спосіб зображення комбінацій коду у вигляді геометричної моделі?
12. Які особливості матричного способу подання коду?
13. Чим різняться систематична та статистична надмірності повідомлень і кодів?
14. На що спрямовані основні теореми кодування для каналу зв'язку?
15. З чим пов'язана нерівність Крафта?
16. Що доводять пряма та обернена теореми кодування?
17. Що таке оптимальне кодування?
18. На чому ґрунтується перша універсальна методика побудови ОНК?
19. На чому ґрунтується друга універсальна методика побудови ОНК?
20. Чим різняться ОНК Шеннона — Фано та Хаффмена?

РОЗДІЛ 6 КОДУВАННЯ ПОВІДОМЛЕНЬ

Повідомлення, що надходять від первісних джерел, як правило, кодуються. Кодування застосовується як для спрощення оброблення повідомлень, так і для підвищення завадостійкості їх передачі по лініях і каналах зв'язку, де спотворюються сигнали, що спричинює появу помилок у повідомленнях.

Для кодування повідомлень, які надходять з джерела інформації, на першому етапі (первинне оброблення повідомлень) використовуються первинні коди, які мають мінімальну кодову відстань $d_{min} = 1$ і не можуть застосовуватися для виявлення та виправлення помилок. Для підвищення завадостійкості передачі повідомлень використовується вторинне кодування комбінацій первинного коду коректувальними кодами, що виявляють і виправляють помилки.

У цьому розділі розглядаються тільки нерівномірні та рівномірні первинні коди, що знайшли застосування в техніці зв'язку, телемеханіці, передачі даних.

6.1. КЛАСИФІКАЦІЯ ПЕРВИННИХ КОДІВ

Для кодування повідомлень при підготовці та введенні даних у системи передачі й оброблення інформації застосовуються, як правило, первинні коди, до яких належать n-розрядні коди з основою (алфавітом) q, в яких використовуються всі q^n кодових комбінацій з потужністю $q^n \geq N_д > q^{n-1}$.

Розрізняють *нерівномірні* та *рівномірні* первинні коди. З перших найвідомішими є оптимальні двійкові коди Шеннона — Фано та Хаффмена, які розглядалися в п.5.7, а також двійковий код Морзе [12].

До рівномірних первинних кодів, які широко застосовуються на практиці [7, 39, 43], належать рекомендовані МККТТ (Міжнародний консультативний комітет з телеграфії та телефонії — тепер Міжнародний союз електрозв'язку) та Міжнародною організацією із стандартизації (ISO) коди: п'ятирозрядний двійковий, міжнародний стандартний телеграфний код № 2 (МТК-2),

міжнародний семирозрядний стандартний двійковий код № 5 для передачі даних. З метою кодової сумісності Єдиної системи ЕОМ у 70-ті роки XX ст. були розроблені стандарти, прийняті свого часу країнами Ради Економічної Взаємодопомоги, у зв'язку з чим було рекомендовано використовувати двійкові коди міжмашинного обміну інформацією КОІ-7H_0, КОІ-7C_1, КОІ-8, код ДКОІ для внутрішнього обміну інформацією та код КПК-12 для подання даних на перфокартах.

Крім перелічених вище, до первинних кодів належать також коди, що мають специфічне використання. Це рівномірні рефлексні коди, що застосовуються в техніці аналого-цифрового перетворення і телевимірюванні, та двійково-десяткові коди, що поширені в системах відображення інформації або використовуються як проміжні при введенні в ЕОМ даних, поданих у десятковому коді.

6.2. НЕРІВНОМІРНІ ДВІЙКОВІ ПЕРВИННІ КОДИ

6.2.1. КОД МОРЗЕ

Крім двійкових ОНК Шеннона — Фано та Хаффмена, які досить докладно розглядалися в п. 5.7, до цього часу широко застосовується нерівномірний (неповний) код Морзе, комбінації якого передаються елементами різної тривалості (крапки та тире). Цей код в основному використовується для передачі телеграфних повідомлень при радіозв'язку з морськими суднами, геологорозвідувальними та пошуковими партіями, полярними станціями.

Спочатку код Морзе було розроблено для латинського алфавіту з урахуванням частоти появи окремих літер у тексті. При переході до українського алфавіту необхідно збільшити кількість кодових комбінацій, щоб можна було передавати літери, які не мають аналогів у латинському алфавіті (ш, щ, я тощо).

До переваг коду Морзе (табл. 6.1) слід зарахувати його простоту, легкість запам'ятовування, можливість візуального прий-

Таблиця 6.1

Номер комбінації	Літери алфавітів		Комбінація	Номер комбінації	Цифри, знаки	Комбінація
	національного (укр./рос.)	латинського				
1	А	A	·–	32	1	·– – – –
2	Б	B	–···	33	2	··– – –
3	В	W	·––	34	3	···– –

Закінчення табл. 6.1

Номер комбінації	Літери алфавітів: національного (укр./рос.)	Літери алфавітів: латинського	Комбінація	Номер комбінації	Цифри, знаки	Комбінація
4	Г	G	– – ·	35	4	· · · · –
5	Д	D	– · ·	36	5	· · · · ·
6	Е	E	·	37	6	– · · · ·
7	Ж	V	· · · –	38	7	– – · · ·
8	З	Z	– – · ·	39	8	– – – · ·
9	І / И	I	· ·	40	9	– – – – ·
10	Й	J	· – – –	41	0	– – – – –
11	К	K	– · –	42	Крапка	· · · · · ·
12	Л	L	· – · ·	43	Крапка з комою	– · – – ·
13	М	M	– –	44	Кома	· – · – · –
14	Н	N	– ·	45	Лапки	· – · · – ·
15	О	O	– – –	46	Двокрапка	– – – · · ·
16	П	P	· – – ·	47	?	· · – – · ·
17	Р	R	· – ·	48	!	– – · · – –
18	С	S	· · ·	49	Апостроф	· – – – – ·
19	Т	T	–	50	Тире	– · · · · –
20	У	U	· · –	51	Дужки	– · – – · –
21	Ф	F	· · – ·	52	Підкреслення	· · – – · –
22	Х	H	· · · ·	53	№	– · · – ·
23	Ц	C	– · – ·	54	Чекати	· – · · ·
24	Ч		– – – ·	55	Зрозумів	– – – – · –
25	Ш		– – – –	56	Дробова риска	– · · – ·
26	Щ	Q	– – · –	57	Знак поділу	– · · · –
27	Ь / Ъ	X	– · · –	58	Перебій	· · · · · · · ·
28	И / Ы		– · – –	59	Початок передачі	– · · – – ·
29	Ю		· · – –			
30	Я		· – · –			
31	Є / Э	E	· · – · ·			

мання та приймання на слух, до недоліків — необхідність декодування тексту перед врученням споживачеві, а також надмірність. Крім того, цей код не враховує особливостей української мови, тобто частоту появи літер в українському тексті.

6.2.2. ЧИСЛО-ІМПУЛЬСНІ КОДИ

У число-імпульсному коді, який ще має назву *одинично-десяткового* [34], кожний розряд десяткового числа записується у вигляді відповідної кількості одиниць. Для можливості приймання їх приймачем окремі розряди кодових комбінацій відокремлюються інтервалами. Код не є рівномірним, хоча може бути перетворений на рівномірний дописуванням у кожній комбінації зліва нулів для заповнення загальної кількості їх елементів до 10. Так, запис десяткового числа 45 має вигляд 1111, 11111 (у варіанті рівномірного число-імпульсного коду це число записується так: 0000001111, 0000011111).

6.3. РІВНОМІРНІ ДВІЙКОВІ ПЕРВИННІ КОДИ

Ці коди широко застосовуються для передачі телеграфних повідомлень і даних, а різняться вони кількістю елементів, з яких складаються кодові комбінації, та комбінаціями цих елементів.

6.3.1. ЧИСЛОВІ ДВІЙКОВІ КОДИ

У цих кодах, які ще називаються *простими*, всі повідомлення нумеруються порядковою послідовністю в двійковій системі числення, що утворює їхній двійковий код. Кількість комбінацій двійкового коду $N = 2^n$, тобто для запису в двійковому коді N повідомлень треба мати n розрядів: $n = \log_2 N$, де n — ціле число.

У числових двійкових кодах використовуються всі можливі комбінації ($N_д = N$); тому ці коди є безнадмірними та завадонезахищеними, а мінімальна кодова відстань у них $d_{min} = 1$.

Міжнародний телеграфний код. Для використання в телеграфних апаратах МККТТ рекомендується міжнародний телеграф-

Таблиця 6.2

Номер комбінації	Регістр: національний (укр/рос)	Регістр: латинський	Регістр: цифровий	Комбінація
1	Т	T	5	00001
2	Повернення каретки			00010
3	0	0	9	00011
4	Пробіл			00100
5	Х	H	Щ	00101
6	Н	N	,	00110
7	М	M	.	00111
8	Переведення рядка			01000
9	Л	L	)	01001
10	Р	R	4	01010
11	Г	G	Ш	01011
12	І / И	I	8	01100
13	П	P	О	01101
14	Ц	C	:	01110

Закінчення табл. 6.2

Номер комбінації	Регістр			Комбінація
	національний (укр/рос)	латинський	цифровий	
15	Ж	V	=	01111
16	Е	E	3	10000
17	З	Z	+	10001
18	Д	D	Хто там?	10010
19	Б	B	?	10011
20	С	S	Апостроф	10100
21	И / Ы	Y	6	10101
22	Ф	F	Є / Э	10110
23	Ь	X	/	10111
24	А	A	–	11000
25	В	W	2	11001
26	Й	J	Ю	11010
27	Цифровий регістр			11011
28	У	U	7	11100
29	Я	O	1	11101
30	К	K	(	11110
31	Латинський регістр			11111
32	Національний регістр			00000

ний код № 2 (табл. 6.2). У цьому п'ятиелементному коді з 32 комбінацій 29 застосовуються для передачі літер, цифр, розділових і службових знаків у трьох регістрах (латинському, національному, цифровому), для яких призначено решту кодових комбінацій.

У зв'язку з тим що при спотворенні кодової комбінації, яка відповідає алфавітному регістру (латинському, національному), вся послідовність комбінацій, що передається після неї, декодується неправильно, для перемикання регістрів вибираються комбінації, найбільш захищені від дії завад (латинський регістр — 11111, національний регістр — 00000).

Коди для передачі даних та обміну інформацією. При передачі даних, крім літер, цифрових, арифметичних і службових знаків міжнародного телеграфного коду № 2, необхідно передавати також не тільки малі, а й великі літери, додаткові розділові, службові та керуючі знаки. Розроблений для цієї мети міжнародний семирозрядний стандартний код № 5 (табл. 6.3), який рекомендовано для передачі та оброблення інформації, побудовано так, щоб будь-який знак його кодової таблиці можна відобразити семиелементною послідовністю, яка містить три старші розряди, що відповідають стовпцю $a_7a_6a_5$, і чотири молодші розряди, які відповідають рядку $a_4a_3a_2a_1$ (наприклад, літері *В* відповідає кодова послідовність 1000010). При цьому в

Таблиця 6.3

a_7	0	0	0	0	1	1	1	1
a_6	0	0	1	1	0	0	1	1
a_5	0	1	0	1	0	1	0	1

№	a_4	a_3	a_2	a_1	0	1	2	3	4	5	6	7
0	0	0	0	0	NUL	(TC7)DLE	Space	0	@	P	\	p
1	0	0	0	1	(TC_1) SOH	DC_1	!	1	A	Q	a	q
2	0	0	1	0	(TC_2) STX	DC_2	"	2	B	R	b	r
3	0	0	1	1	(TC_3) ETX	DC_3	# £	3	C	S	c	s
4	0	1	0	0	(TC_4) EOT	DC_4	$	4	D	T	d	t
5	0	1	0	1	(TC_5) ENQ	(TC_8) NACK	%	5	E	U	e	u
6	0	1	1	0	(TC_6) ACK	(TC_9) SYNC	&	6	F	V	f	v
7	0	1	1	1	BEL	(TC_{10}) ETB	'	7	G	W	g	w
8	1	0	0	0	(FE_0) BS	CAN	(	8	H	X	h	x
9	1	0	0	1	(FE_1) HT	EM	)	9	I	Y	i	y
10	1	0	1	0	(FE_2) LF	SUB	*	:	J	Z	j	z
11	1	0	1	1	(FE_3) VT	ESC	+	;	K	[	k	
12	1	1	0	0	(FE_4) FF	(IS_4) FS	,	<	L		l	
13	1	1	0	1	(FE_5) CR	(IS_3) GS	—	=	M	]	m	
14	1	1	1	0	SO	(IS_2) RS	.	>	N	^	n	~
15	1	1	1	1	SI	(IS_1) US	/	?	O	–	o	DEL

ри необхідності простим виключенням старших розрядів можна дістати підмножини комбінацій меншої розрядності.

Цей код у США використовується під назвою коду ASCII (або USASCII) як засіб взаємодії з ЕОМ. Він дає змогу перетворювати машинні дані, записані в двійковому коді, на звичайні знаки (числа, літери), які можна роздрукувати та вивести на термінал. Оскільки восьмибітові коди застосовуються в ЕОМ значно частіше, ніж семибітові, в коді ASCII-8 восьмий (крайній лівий) біт використовується як біт парності (1 — якщо праві сім бітів складають непарне число одиниць і 0 — при парному їх числі для перевірки правильності передачі даних. При виявленні помилки користувачеві видається повідомлення Parity Error — помилка парності).

Міжнародний стандартний код № 5 логічно може бути поділений на чотири зони. До першої зони (колонки 0, 1) належать функціональні символи, за винятком символу DEL (витирання, перебій), для якого відведено останнє місце (комбінація 1111111). Основна частина функціональних символів цієї зони поділяється на чотири групи:

- перша — для керування передачею інформації по каналах зв'язку (TC_1... TC_{10});
- друга — для керування друком (FE_0... FE_5);
- третя — для керування кінцевими пристроями (DC_1... DC_4);
- четверта — для роздільників інформації (IS_1... IS_4).

Перша група містить символи зв'язку для керування передачею інформації по каналах зв'язку. Вони мають два призна-

чення: використовуються для обрамлення повідомлення у формат, який легко розпізнається, або в послідовність, яка може оброблятися споживачем; застосовуються для керування передачею даних у мережі.

Як відомо, текст — це інформаційний зміст повідомлення. Якщо воно досить тривале, то його, як правило, розбивають на кілька блоків, які передають один за одним по лінії зв'язку. Залежно від системи, що використовується для передачі, перед блоками може бути заголовок, який повинен мати адресну та керуючу інформацію, що супроводжує текст повідомлення. Заголовок може містити відомості про пріоритет повідомлення, дату та час його відправлення, ідентифікатор лінії зв'язку, по якій передається повідомлення, відомості про ступінь секретності тощо. Отже, в заголовку подаються відомості про те, як повідомлення має оброблятися на шляху від джерела до одержувача.

Рішення про використання заголовка повідомлення, а також його зміст приймається на основі характеристик конкретної системи передачі даних і програмного забезпечення. Якщо повідомлення містить, наприклад, чотири блоки тексту, то перший його блок передається із заголовком, у якому подаються необхідні характеристики повідомлення, а всі наступні блоки передаються без заголовка.

Розглянемо призначення окремих символів зв'язку більш докладно.

Перша група. Символ TC_1 (початок заголовка) використовується як перший символ заголовка інформаційного повідомлення і повідомляє одержувача, що інформація, яка подається після нього, має інтерпретуватися (тлумачитися) як заголовок повідомлення.

Аналогічно символ TC_2 (початок тексту) розміщується на початку тексту і застосовується для позначення кінця заголовка, а також показує, що інформація, яка подається після нього, є текстом повідомлення.

Символ TC_3 (кінець тексту) передається в кінці тексту й означає, що повідомлення було передано повністю.

Символ TC_4 (кінець передачі) використовується для зазначення кінця передачі одного або кількох текстів.

Символ TC_5 (хто там?) застосовується для запиту відповіді віддаленої станції, причому відповідь може містити ідентифікатор і статус станції.

Символ TC_6 (підтвердження) передається приймачем як підтвердження прийому запиту передавача.

Символ TC_7 (перший авторегістр) використовується для розширення функцій керування передачею даних; при цьому він змінює значення обмеженої кількості символів, які передаються безпосередньо за ним. До цих символів можуть належати тільки графічні символи та символи зв'язку.

Символ TC_8 (негативна квитанція) передається приймачем як негативна відповідь передавачу.

Символ TC_9 (синхронізація) застосовується в синхронних системах передачі за відсутності інформаційних символів для встановлення та підтримки синхронізації кінцевого устаткування.

Символ TC_{10} (кінець блока) використовується для зазначення кінця блока тоді, коли дані, що передаються, розбиваються на блоки.

Зазначимо, що в деяких системах при передачі багатоблокового повідомлення символ TC_3 (кінець тексту) має бути в кінці кожного блока для того, щоб одержувач міг скласти з блоків повне повідомлення.

Друга група. До цієї групи належать символи керування друком, які використовуються при розміщенні інформації на друкованому аркуші або на екрані пристрою візуального зображення з метою полегшення сприйняття даних. Перший символ FE_0 (повернення на крок) відповідає поверненню на крок і дає змогу повернути головку пристрою друку на один крок назад. Для пристроїв візуального зображення цей символ означає переведення покажчика на одну позицію ліворуч.

Символ FE_1 (горизонтальна табуляція) дає змогу перемістити головку друку в задане положення в горизонтальному напрямку, символ FE_2 (переведення рядка) — в те саме положення в наступному рядку; символ FE_3 (вертикальна табуляція) — в те саме положення через кілька рядків у межах однієї сторінки; символ FE_4 (переведення формату) — в те саме положення на умовленому рядку на іншій сторінці; символ FE_5 (повернення каретки) — в початкове положення рядка.

Третя група. До цієї групи належать символи керування пристроями, що встановлюють фізичні функції на терміналі. Так, символ DC_1 призначений для приєднання до терміналу касетного накопичувача; символ DC_2 — для вимикання цього накопичувача; символ DC_3 – для виведення інформації з пристрою візуального відображення на допоміжний пристрій друку, а символ DC_4 — для блокування пристрою візуального зображення, щоб оператор не зміг вивести з нього дані.

Четверта група. Ця група символів має чотири роздільники інформації, призначені для логічного розмежовування інформації з метою полегшення її оброблення на ЕОМ. Так, символ IS_1 використовується для розмежовування найменшого обсягу інформації і називається *роздільником одиниць*; символ IS_2 — для розмежовування підгруп інформації, які можуть містити кілька одиниць; символ IS_3 — для розмежовування підгруп ін-

формації, які можуть складатися з кількох підгруп, а символ IS_4 — для розмежовування файлів, які можуть містити кілька груп інформації.

Решту зон табл. 6.3 зайнято графічними знаками. Її другу зону (колонки 2, 3) виділено для спеціальних математичних знаків, а також знаків пунктуації та цифр. У двох останніх зонах цієї таблиці розміщуються великі й малі латинські літери відповідно до вимог лексико-графічної впорядкованості (положення цифр, знаків пунктуації та пробілу також вибрано з урахуванням цих вимог). Тут знаходяться й усі резервні позиції.

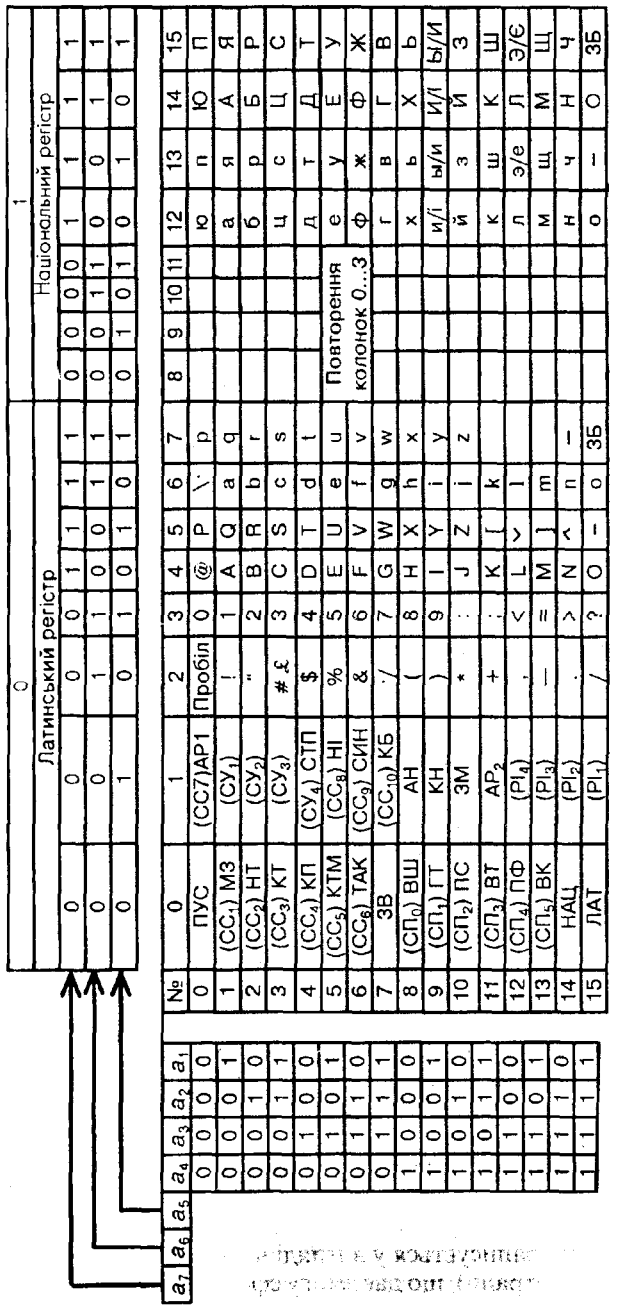
Знак «Пробіл» («Space») не друкується, але належить до числа графічних і застосовується для поділу слів і переміщення позиції друку на один крок уперед. Цифри кодуються звичайним двійковим кодом. Побудова чотирирозрядних комбінацій для них виконується відкиданням трьох старших розрядів.

Є два методи декодування знаків у міжнародному стандартному коді № 5. По-перше, можна застосувати їх двійкове подання (наприклад, послідовність 1010100 відповідає знаку Т). Інший метод полягає у використанні номерів рядків і стовпців для однозначного визначення конкретного знака (наприклад, запис 5/04 відповідатиме тому самому знаку Т , який знаходиться у п'ятому стовпці та четвертому рядку).

Відсутність у розглядуваному коді знаків, які відповідали б національному алфавіту (в тому разі, коли він відрізняється від латинського) не дає змоги широко використовувати його у країнах з нелатинськими алфавітами, в тому числі й у нашій країні. Тому на основі міжнародного стандартного коду № 5 було розроблено код [12, 43], в якому враховано особливості національного алфавіту (табл. 6.4). У цей код уведено національний регістр (український). Зручність використання такого семирозрядного коду в системах передачі даних полягає в простоті перетворення його на восьмирозрядний код з перевіркою на парність, що дає змогу підвищити завадостійкість передачі інформації по каналах із завадами.

Подальший розвиток кодів обчислювальних машин і апаратури передачі даних пов'язаний зі співробітництвом у 70-ті роки XX ст. країн Ради Економічної Взаємодопомоги у сфері розробки та вдосконалення обчислювальної техніки, завдяки чому було розроблено і впроваджено комплекс програмно-сумісних обчислювальних машин третього покоління — Єдиної системи ЕОМ.

Одними із засобів забезпечення їх сумісності є прийняті форма та формати подання даних. Структурною одиницею тут виступає восьмибітний байт. При цьому використовуються такі коди: КОІ-7 і КОІ-8 — для обміну інформацією; КПК-12 — для подання даних на перфокартах; ДКОІ — внутрішній.



Код КОІ-7 розроблено на основі семирозрядного стандартного коду [12,43], в якому враховано вимоги міжнародного стандарту ІСО МС 646. Семирозрядний код КОІ-7$Н_0$ містить 128 знаків міжнародного алфавітно-цифрового набору. Індекс «0» у позначенні $Н_0$ указує місце цього коду при перекодуванні його в код КОІ-8: він розміщується в лівій частині кодової таблиці КОІ-8. При цьому старшим буде розряд 0.

Семирозрядний код КОІ-7$Н_1$ також містить 128 знаків: у стовпцях 0... 3 розміщуються символи, що відповідають знакам кодової таблиці КОІ-7$Н_0$, а в стовпцях 4...7 — великі та малі літери національного алфавіту. Індекс «1» у позначенні $Н_1$ указує місце цього коду при перекодуванні його в код КОІ-8: він розміщується в правій частині кодової таблиці КОІ-8 під восьмим розрядом 1.

Семирозрядний код КОІ-7$С_1$ містить 32 кодові комбінації для передачі додаткових керуючих знаків. Індекс «1» у позначенні $С_1$ указує місце цього коду при перекодуванні його в код КОІ-8: він розміщується в правій частині кодової таблиці КОІ-8 під восьмим розрядом 1.

Якщо під час запису інформації в послідовності кодованих знаків необхідно виконати перехід від коду КОІ-7$Н_0$ до коду КОІ-7$Н_1$ і навпаки, то подаються спеціальні кодові комбінації: «ВИХ» — 0001110 та «ВХ» — 0001111. Відсутність на початку тексту керуючих знаків «ВИХ» і «ВХ» рівнозначно наявності там керуючого знака «ВХ».

Коди КОІ-7$Н_0$, КОІ-7$Н_1$ і КОІ-7$С_1$ дають змогу доповнювати їх восьмим розрядом для виявлення помилок в інформації. Це робиться перевіркою на парність семирозрядної кодової комбінації.

Сукупність названих кодів з додатковим восьмим розрядом у кодовій комбінації утворює восьмирозрядний код КОІ-8.

Двійковий код ДКОІ для оброблення інформації розроблено на основі універсального міжнародного коду EBCDIC (Exended Binary Disimal Interchage Code) із включенням літер національного алфавіту [43].

Код КПК-12 для перфокарт [43] призначений для занесення інформації на 12-позиційні перфокарти пристроїв підготовки введення-виведення даних. Він містить 256 кодових комбінацій, які будуються на основі матриці 16 × 16. У кодовій таблиці виключено літери національного алфавіту, що мають однакове написання з латинськими (наприклад С, К, Р та ін.).

6.3.2. ДВІЙКОВО-ДЕСЯТКОВІ КОДИ

У двійково-десятковому коді (ДДК) кожний розряд десяткового числа записується у вигляді чотирирозрядного двійкового числа (тетради), що дає змогу сформувати $2^4 = 16$ кодових

комбінацій. Через те, що в десятковій системі числення використовується 10 цифр, шість комбінацій є надмірними і, як наслідок, вибір 10 комбінацій ДДК, які застосовуються для його побудови, має $16!6! \approx 2 \cdot 9 \cdot 10^{10}$ варіантів [12].

Усі ДДК можна поділити на *вагові* та *невагові*, де вагою кожного розряду двійкового числа є її еквівалент у десятковій системі числення. Підсумовуючи вагу всіх чотирьох розрядів, дістають цифру десяткового числа. *Вагомим* називається такий ДДК, в якому вага кожного розряду для всіх 10 комбінацій залишається сталою. Будується він з урахуванням таких умов:

- вага найменшого розряду q_1 дорівнює 1;
- вага другого за мінімальним значенням розряду q_2 може дорівнювати 1 або 2;
- вага, що відповідає останнім двом розрядам q_3 та q_4 коду, підбирається так, щоб сума їх була більшою або дорівнювала шести (якщо $q_2 = 2$) чи семи (якщо $q_2 = 1$).

Згідно з цим можна дістати 17 варіантів ДДК з додатною вагою розрядів (табл. 6.5). Крім коду 8 4 2 1, решта 16 кодів не мають однозначності в зображенні десяткових чисел. Так, код 3 3 2 1 дає змогу записати цифру 5 у двійковій формі як 1010 або 0110.

В інших ДДК вага окремих розрядів може бути додатною або від'ємною (табл. 6.6) [30].

Залежно від коду, що використовується, кожна з десяткових цифр може бути подана однією з комбінацій чотирьох двійкових елементів. Так, цифра 7 у ДДК 6 4 –2 –1 з додатними та від'ємними числами записується як 1111, а в коді 4 4 3 2 — як 1010 або 0110.

Вибір того чи іншого ДДК залежить від конкретних умов його застосування та зручності реалізації. ДДК широко використовуються у вимірювальних пристроях, де вимірюваний параметр має відтворюватися на цифрових індикаторах.

Таблиця 6.5

№ пор.	Вага розрядів				№ пор.	Вага розрядів			
	q_4	q_3	q_2	q_1		q_4	q_3	q_2	q_1
1	8	4	2	1	10	3	3	2	1
2	7	4	2	1	11	6	2	2	1
3	6	4	2	1	12	5	2	2	1
4	5	4	2	1	13	4	2	2	1
5	4	4	2	1	14	6	3	1	1
6	7	3	2	1	15	5	3	1	1
7	6	3	2	1	16	4	3	1	1
8	5	3	2	1	17	5	2	1	1
9	4	3	2	1					

Таблиця 6.6

№ пор.	Вага розрядів				№ пор.	Вага розрядів				№ пор.	Вага розрядів			
	q_4	q_3	q_2	q_1		q_4	q_3	q_2	q_1		q_4	q_3	q_2	q_1
1	8	7	–4	–2	19	8	2	1	–4	37	7	3	1	–2
2	8	4	–3	–2	20	7	2	1	–4	38	6	3	1	–2
3	7	2	–4	–1	21	6	2	1	–4	39	5	3	1	–2
4	7	2	–3	–1	22	8	4	2	–3	40	6	2	1	–2
5	8	4	–2	–1	23	6	4	2	–3	41	8	4	2	–1
6	7	4	–2	–1	24	5	4	2	–3	42	7	4	2	–1
7	6	4	–2	–1	25	7	5	1	–3	43	6	4	2	–1
8	5	4	–2	–1	26	7	2	1	–3	44	5	4	2	–1
9	6	3	–2	–1	27	6	2	1	–3	45	4	4	2	–1
10	6	3	–1	–1	28	8	4	3	–2	46	7	3	2	–1
11	7	5	3	–6	29	6	4	3	–2	47	6	3	2	–1
12	6	4	3	–5	30	5	4	3	–2	48	5	3	2	–1
13	6	5	3	–4	31	4	4	3	–2	49	4	3	2	–1
14	6	5	2	–4	32	6	3	2	–2	50	6	2	2	–1
15	8	3	2	–4	33	8	4	1	–2	51	5	2	2	–1
16	6	3	2	–4	34	6	4	1	–2	52	6	3	1	–1
17	8	6	1	–4	35	5	4	1	–2	53	5	3	1	–1
18	7	5	1	–4	36	4	4	1	–2					

Усі ДДК мають деяку надмірність, що, як правило, застосовується для виявлення помилок. З цією метою, крім 10 робочих, мають фіксуватися й решта шість комбінацій, причому під час приймання останніх, які можуть бути тільки результатом помилок при передачі ДДК, має забезпечуватися заборона відтворення інформації, хоча при цьому виявляються далеко не всі, навіть найпростіші, однократні помилки.

6.3.3. ДВІЙКОВО-ДЕСЯТКОВІ КОДИ З САМОДОПОВНЕННЯМ

Необхідність заміни операції віднімання в ЕОМ операцією додавання, що виконується за допомогою спеціальних машинних кодів, привела до розробки ДДК, які мають властивість самодоповнення [12].

В оберненому коді з самодоповненням кожний розряд $[a_i]_2$ подається як доповнення до 2–1, тобто $[a_i]_2 = 2–1 – a_i$. Доповнення до 2–1=1 дорівнює 0, якщо $a_i = 1$, і 1, якщо $a_i = 0$, тобто є інверсією цифри a_i. Для десяткового коду потрібно знаходити доповнення до 9.

Зручність цих кодів полягає в тому, що ДДК цифри, яка є доповненням до 9, аналогічно двійковому коду знаходиться простою інверсією двійкових зображень десяткового числа в

Таблиця 6.7

Десяткове число	Код Айкена	Код із надмірністю 3	Десяткове число	Код Айкена	Код із надмірністю 3
0	0000	0011	5	1011	1000
1	0001	0100	6	1100	1001
2	0010	0101	7	1101	1010
3	0011	0110	8	1110	1011
4	0100	0111	9	1111	1100

коді, для якого відшукується доповнення. Таким чином, якщо розряд десяткового числа a_i подано тетрадою двійкових розрядів $q_4q_3q_2q_1$, то доповнення до 9 визначається як $[a_i]_{10} = \bar{q}_4\bar{q}_3\bar{q}_2\bar{q}_1$, де $\bar{q}_i$ — заперечення двійкової цифри q_i.

Найбільшого поширення з ДДК із самодоповненням дістали код Айкена (2 4 2 1) і код із надмірністю 3 (8 4 2 1), які наведено в табл. 6.7. Як випливає з таблиці, при заміні цифр усіх чотирьох розрядів коду з 0 на 1 (або навпаки) дістаємо доповнення до 9 для кодованої десяткової цифри.

6.3.4. ДВІЙКОВО-ШІСТНАДЦЯТКОВИЙ КОД

У двійково-шістнадцятковому коді для запису двійкового байта (вісім розрядів двійкового коду) використовується шістнадцяткова система числення. При цьому чотирирозрядні двійкові числа (тетради) записуються шістнадцятковим символом. Так, запис двійкових послідовностей 1100 0110, 0010 1101 має вигляд .C6 і .2D, де крапка перед символами вказує на відмінність дворозрядного двійково-шістнадцяткового числа від дворозрядних шістнадцяткових чисел. У цьому прикладі $(C6)_{16} = (198)_{10}$ і $(2D)_{16} = (45)_{10}$.

Двійково-шістнадцятковий код широко застосовується для скороченого запису кодових комбінацій байтової структури [23].

6.3.5. РЕФЛЕКСНІ КОДИ

Особливість побудови рефлексних кодів полягає в тому, що сусідні кодові комбінації на відміну від двійкових простих кодів різняться цифрою тільки в одному розряді, тобто кодова відстань між ними дорівнює одиниці.

Іншою особливістю цих кодів є те, що зміна елементів у кожному розряді при переході від комбінації до комбінації відбувається в два рази рідше, ніж у простому коді, завдяки чому

Таблиця 6.8

Десяткове число	Варіанти рефлексних кодів				
	перший	другий	третій	четвертий	п'ятий
0	000	000	000	000	000
1	010	100	100	001	001
2	011	101	110	011	101
3	001	001	010	010	100
4	101	011	011	110	110
5	111	111	111	111	111
6	110	110	101	101	011
7	100	010	001	100	010

значно спрощується кодер. Крім того, при додаванні двох сусідніх комбінацій рефлексного коду за модулем 2 кількість одиниць дорівнюватиме кількості розрядів мінус 3, тобто одиниці, що використовується для перевірки правильності прийнятої кодової комбінації.

Свою назву рефлексні коди дістали через наявність осей симетрії, відносно яких виразно проглядається ідентичність елементів у деяких розрядах. Вісь симетрії, що розміщується в n-значному рефлексному коді між комбінаціями, які відповідають рівням $(2^{n-1}-1)$ і 2^{n-1}, називається *головною*. Щодо неї є ідентичність елементів в $(n-1)$ розрядах симетричних кодових комбінацій.

Можна утворити велику кількість двійкових рефлексних кодів, у яких дві сусідні комбінації відрізнятимуться тільки одним символом (табл. 6.8).

Найбільшого поширення з рефлексних кодів дістав код Грея (табл. 6.9), який, на відміну від інших, простіший при перетворенні його на двійковий простий код. Обернене перетворення двійкового простого коду на код Грея виконується за алгоритмом

$$y_i = x_i \oplus x_{i+1},$$

Таблиця 6.9

Десяткове число	Двійковий простий код	Код Грея	Десяткове число	Двійковий простий код	Код Грея
0	0000	0000	8	1000	1100
1	0001	0001	9	1001	1101
2	0010	0011	10	1010	1111
3	0011	0010	11	1011	1110
4	0100	0110	12	1100	1010
5	0101	0111	13	1101	1011
6	0110	0101	14	1110	1001
7	0111	0100	15	1111	1000

де y_i — значення i-го розряду коду Грея; x_i, x_{i+1} — відповідні значення розрядів двійкового числа ($i = 1, 2, \ldots, n$, починаючи зліва).

Таким чином, для утворення комбінації коду Грея практично досить зсунути двійкову комбінацію простого коду на один розряд праворуч, порозрядно додати її за модулем 2 до початкової кодової комбінації без перенесення між розрядами і відкинути молодший розряд здобутої суми.

Декодування (обернене перетворення) коду Грея можна виконати двома способами:

• перший спосіб

$$\begin{cases} x_n = y_n; \\ x_i = x_{i+1} \oplus y_i, \end{cases}$$

де x_n і y_n — відповідно значення старшого розряду двійкового простого коду та коду Грея ($i = n-1, n-2, \ldots, 1$, починаючи зліва);

• другий спосіб

$$x_j = \sum_{j=1}^{n} y_j,$$

де y_j — значення розрядів коду Грея, а сума береться за всіма розрядами цього коду від i- до n-го (старшого, крайнього зліва).

Іншими словами, щоб перейти від коду Грея до двійкового простого коду, треба:

• залишити цифру старшого розряду без зміни;

• кожну наступну цифру інвертувати стільки разів, скільки одиниць є перед нею в коді Грея, або виконати послідовне порозрядне підсумовування за модулем 2 першого (старшого) та другого розрядів комбінації цього коду (1 ⊕ 2), після чого послідовно додати 1 ⊕ 2 ⊕ 3, 1 ⊕ 2 ⊕ 3 ⊕ 4 і т. д.

До характерних особливостей коду Грея належить те, що, по-перше, кожна наступна комбінація завжди відрізняється від попередньої тільки в одній позиції (одному розряді); по-друге, зміна значень елементів у кожному розряді при переході від комбінації до комбінації відбувається в два рази швидше, ніж у двійковому простому коді, тобто якщо в останньому зміна елемента першого (молодшого) розряду відбувається з чергуванням елементів 0-1-0-1-..., елемента другого розряду — з чергуванням елементів 00-11-00-11-..., елемента третього розряду — з чергуванням елементів 0000-1111-0000-1111-... і т. д., то в коді Грея відповідно маємо такі чергування елементів: для першого розряду 11-00-11-00-..., для другого 0000-1111-0000-1111-... і т. д., що дає змогу при тій самій швидкості кодера досягати вищої точності кодування порівняно з двійковим простим ко-

дом; по-третє, при додаванні двох сусідніх комбінацій за модулем 2 кількість одиниць дорівнюватиме кількості розрядів мінус 3, що використовується для перевірки наявності помилки в прийнятій кодовій комбінації; по-четверте, в цьому коді можна виділити кілька осей симетрії, відносно яких спостерігається ідентичність елементів у деяких розрядах. Так, має місце симетрія деяких розрядів відносно осей, проведених між числами 1 і 2, 3 та 4, 5 і 6, 7 та 8, 9 і 10, 11 та 12 (див. табл. 6.9).

Код Грея широко застосовується для аналого-цифрового перетворення різних неперервних повідомлень. Він дає змогу зменшити кількість помилок від завад, які виникають при передачі інформації по каналах зв'язку.

До недоліків цього коду належить «невагомість» кодової комбінації, коли вага одиниці в ній не визначається номером розряду, на місці якого вона знаходиться, а переведення кодової комбінації з двійкової системи числення в десяткову не визначатиме порядковий номер комбінації в коді Грея. Такі коди важко декодувати, тому перед декодуванням їх, як правило, перетворюють на двійковий простий код, після чого й обробляють останній.

6.4. НЕДВІЙКОВІ ПЕРВИННІ КОДИ

Основа (алфавіт) недвійкових кодів завжди більша від двох, тобто $q \geq 3$; тому для побудови їх використовують методи теорії комбінаторики: перестановки P_q з q елементів, розміщення A_q^m і сполучення C_q^m з q по m елементів.

Для кодів, які ґрунтуються на *перестановках* символів алфавіту, довжина кодової комбінації $n = q = \text{const}$. Загальна кількість перестановок визначається виразом

$$N = P_q = \prod_{i=1}^{n} i = 1 \cdot 2 \cdot 3 \ldots n = n! = q!.$$

Так, для коду з алфавітом $q = 3$ (a, b, c) всього буде шість перестановок: abc, acb, bac, bca, cab, cba, тобто $N = P_3 = \prod_{i=1}^{3} i = 1 \cdot 2 \cdot 3 = 6$. Збільшення алфавіту коду приводить до зростання кількості перестановок, а отже, й кількості кодових комбінацій N.

Відмітною особливістю такого коду є відсутність однакових символів у одній кодовій комбінації; тому їх можна віднести до кодів з виявленням однократних і деяких багатократних помилок. Дійсно, перехід будь-якого символу в комбінації призводить до появи в ній кількох однакових символів, що автоматично виявиться на приймальному боці і спричинить захисну відмову в прийманні комбінації.

Для кодів, які ґрунтуються на *розміщеннях* символів алфавіту, загальна кількість комбінацій визначається виразом

$$N = A_q^m = q(q-1)(q-2) \ldots (q-m+1) = q! / (q-m)!,$$

де m — кількість символів алфавіту q, які містить кодова комбінація. Завжди $q > m$, а довжина кодової комбінації $n = m$.

Так, для коду з алфавітом $q = 3$ (a, b, c) і $n = m = 2$ загальна кількість кодових комбінацій $N = A_3^2 = 3! / (3-2)! = 3! = 1 \cdot 2 \cdot 3 = 6$. Це комбінації *ab, ac, ba, ca, bc, cb.*

Ці коди на відміну від кодів на перестановки дають змогу виявляти тільки деякі однократні помилки, коли під дією завад виникає перетворення, що зумовлює подвоєння символів у комбінації. Решта перетворень символів у кодовій комбінації утворять іншу дозволену комбінацію, тому така помилка не буде виявлена на приймальному боці.

Для кодів, які ґрунтуються на *певному числі сполучень* символів алфавіту, загальна кількість комбінацій визначається виразом

$$N = C_q^m = q! / [(q-m)!\, m!].$$

При цьому $q > m$, а довжина кодової комбінації $n = m$. Від кодів на розміщення ці коди відрізняються відсутністю комбінацій, які різняться тільки порядком розташування символів, тобто до таких кодів належать кодові комбінації, що різняться тільки самими символами q.

Так, для коду з алфавітом $q = 3$ (a, b, c) та $n = m = 2$ загальна кількість кодових комбінацій $N = C_3^2 = 3! / [2!(3-2)!] = 1 \cdot 2 \cdot 3 / (1 \cdot 2) = 3$. Це комбінації *ab, ac, bc*. До комбінацій коду не можуть належати сполучення *ba, ca, cb,* оскільки в них використані ті самі сполучення символів, що й у дозволених комбінаціях.

У таких кодах, як і в кодах, що ґрунтуються на розміщеннях, в одній комбінації не може бути двох однакових символів. Ці комбінації легко виявляються на приймальному боці.

Коди, які ґрунтуються на *всіх сполученнях* символів алфавіту, в одній комбінації можуть містити будь-які, в тому числі й однакові, символи. При цьому загальна кількість комбінацій визначається виразом

$$N = q^n.$$

Для таких кодів можна відзначити збільшення кількості комбінацій порівняно з кодами на розміщення. Це пояснюється тим, що $q^n > A_q^m$.

Так, при $q = 3$ та $n = m = 2$ загальна кількість кодових комбінацій $N = q^n = 3^2 = 9$. Це комбінації *aa, ab, ac, bb, ba, bc, cc, ca,*

cb. Збільшення кількості комбінацій тут досягається завдяки використанню таких комбінацій, як *aa, bb, cc*.

Змінно-якісний код можна дістати з коду на всі сполучення символів, якщо накласти на нього деякі обмеження. Так, у комбінації змінно-якісного коду однакові символи не повинні знаходитися поруч. Загальна кількість комбінацій такого коду визначається виразом

$$N = q(q-1)^{n-1}.$$

Наприклад, при $q = 3$ та $n = 3$ можна утворити 12 кодових комбінацій: *aba, aca, abc, acb, bab, bcb, bac, bca, cac, cbc, cab, cba.*

До переваг змінно-якісного коду слід віднести можливість розрізнення кодових комбінацій і виявлення в них помилок, тому що в комбінаціях цього коду два однакових символи не повинні знаходитися поруч.

КОНТРОЛЬНІ ЗАДАЧІ

1. Закодувати ДДК 8 4 2 1 десяткове число 2987.

Розв'язання. З умови задачі відомо вагу кожного з розрядів ДДК: $q_1 = 1$, $q_2 = 2$, $q_3 = 4$, $q_4 = 8$. Відомо також, що кожне десяткове число в ДДК записується двійковою тетрадою. Таким чином, запис десяткового числа 2987 у ДДК матиме вигляд 0010 1001 1000 0111.

2. Закодувати ДДК 4 4 2 1 десяткове число 492.

3. Закодувати ДДК 7 3 2 1 десяткове число поточного року.

4. Закодувати ДДК 5 4 2 –3 з додатними та від'ємним числами десяткове число 364.

Розв'язання. З умови задачі відомо вагу кожного з розрядів ДДК: $q_1 = -3$, $q_2 = 2$, $q_3 = 4$, $q_4 = 5$. Відомо також, що кожне десяткове число в ДДК записується двійковою тетрадою. Таким чином, запис десяткового числа 364 в ДДК матиме вигляд 0111 1101 0100 або 0111 1101 1011.

5. Закодувати ДДК 8 7 –4 –2 з додатними та від'ємними числами десяткове число 237.

6. Закодувати ДДК 4 3 2 –1 з додатними та від'ємним числами десяткове число 13415.

7. Перетворити на код Грея двійковий простий код 000111001.

Розв'язання. Перетворення двійкового простого коду на код Грея виконується за алгоритмом $y_i = x_i \oplus x_{i+1}$, де y_i — значення i-го розряду коду Грея; x_i, x_{i+1} — відповідні значення розрядів двійкового коду ($i = 1, 2, \ldots, n$, починаючи зліва). Отже, щоб здійснити це перетворення, досить зсунути двійкову комбінацію простого коду на один розряд праворуч, порозрядно додати її за модулем 2 до початкової кодової комбінації без перенесення між розрядами і відкинути молодший розряд здобутої суми:

$$\begin{array}{r} \oplus\ 000111001 \\ \quad\ 00011100(1) \\ \hline 000100101, \end{array}$$

тобто шукана кодова комбінація має вигляд 000100101.

8. Перетворити на код Грея такі комбінації двійкового простого коду: 01100100, 110110, 101010001, 0001101101.

9. Розв'язати попередню задачу, якщо комбінаціями двійкового простого коду є 001100111, 11100001, 1011110, 010001.

10. Перетворити код Грея 1001011011 на двійковий простий код.

Розв'язання. Обернене перетворення коду Грея на двійковий простий код можна виконати двома способами. Для полегшення запису алгоритмів цього перетворення кожному розряду комбінації коду Грея присвоїмо значення $y_{10} \ldots y_1$, а двійкового простого коду — значення $x_{10} \ldots x_1$ залежно від номера розряду.

Перший спосіб: $x_{10} = y_{10} = 1$; $x_9 = x_{10} \oplus y_9 = 1 \oplus 0 = 1$; $x_8 = x_9 \oplus y_8 = 1 \oplus 0 = 1$; $x_7 = x_8 \oplus y_7 = 1 \oplus 1 = 0$; $x_6 = x_7 \oplus y_6 = 0 \oplus 0 = 0$; $x_5 = x_6 \oplus y_5 = 0 \oplus 1 = 1$; $x_4 = x_5 \oplus y_4 = 1 \oplus 1 = 0$; $x_3 = x_4 \oplus y_3 = 0 \oplus 0 = 0$; $x_2 + x_3 \oplus y_2 = 0 \oplus 1 = 1$; $x_1 = x_2 \oplus y_1 = 1 \oplus 1 = 0$, тобто маємо комбінацію 1110010010.

Другий спосіб: $x_{10} = \sum_{j=10}^{10} y_j = y_{10} = 1$; $x_9 = \sum_{j=9}^{10} y_j = y_{10} \oplus y_9 = 1 \oplus 0 = 1$;

$x_8 = \sum_{j=8}^{10} y_j = y_{10} \oplus y_9 \oplus y_8 = 1 \oplus 0 \oplus 0 = 1$; $x_7 = \sum_{j=7}^{10} y_j = y_{10} \oplus y_9 \oplus y_8 \oplus y_7 = 1 \oplus 0 \oplus 0 \oplus 1 = 0$; $x_6 = \sum_{j=6}^{10} y_j = y_{10} \oplus y_9 \oplus y_8 \oplus y_7 \oplus y_6 = 1 \oplus 0 \oplus 0 \oplus 1 \oplus 0 = 0$;

$x_1 = \sum_{j=1}^{10} y_j = y_{10} \oplus y_9 \oplus y_8 \oplus y_7 \oplus y_6 \oplus y_5 \oplus y_4 \oplus y_3 \oplus y_2 \oplus y_1 =$

$= 1 \oplus 0 \oplus 0 \oplus 1 \oplus 0 \oplus 1 \oplus 1 \oplus 0 \oplus 0 \oplus 1 \oplus 1 = 0$, тобто маємо комбінацію 1110010010.

Перевіряємо правильність виконаного перетворення, для чого утворений двійковий простий код 1110010010 перетворимо на код Грея:

$$\begin{array}{r} \oplus\ 1110010010 \\ \quad\ 111001001(0) \\ \hline 1001011011, \end{array}$$

Таким чином, здобута комбінація 1001011011 коду Грея збігається з даною за умовою задачі, що свідчить про правильність виконаних перетворень.

11. Перетворити комбінацію 01011100 коду Грея на двійковий простий код.

12. Розв'язати попередню задачу стосовно коду Грея з комбінацією 100001001.

13. Побудувати таблицю прямого й оберненого перетворень усіх комбінацій п'ятирозрядного двійкового простого коду на код Грея і навпаки.

14. Розв'язати попередню задачу стосовно шестирозрядного двійкового простого коду.

15. Виконати пряме й обернене перетворення на код Грея всіх комбінацій п'ятирозрядного двійкового простого коду з вагою $w = 2$.

16. Розв'язати попередню задачу стосовно семирозрядного двійкового простого коду з вагою $w = 6$.

17. Побудувати комбінації недвійкового коду на перестановки для алфавіту $q = 4$.

Р о з в ' я з а н н я. Загальна кількість перестановок недвійкового коду

$$N = q! = 4! = 24.$$

Довжина комбінацій цього коду $n = q$; тому за умови, що його алфавіт $q \in [0, 1, 2, 3]$, комбінаціями коду будуть 0123, 0231, 0321, 0132, 0213, 1023, 1032, 1203, 1230, 1302, 1320, 2013, 2031, 2103, 2130, 2301, 2310, 3012, 3021, 3102, 3120, 3201, 3210, 0312.

18. Розв'язати попередню задачу для $q = 5$.

19. Побудувати комбінації недвійкового коду на перестановки для алфавіту $q = 6$.

20. Побудувати комбінації недвійкового коду на розміщення для алфавіту $q = 4$ та $n = m = 2$.

Р о з в ' я з а н н я. Загальна кількість комбінацій недвійкового коду на розміщення

$$N = q! / (q - m)! = 4! / (4 - 2)! = 12.$$

Для цього коду за умови, що його алфавіт $q \in [0, 1, 2, 3]$, належатимуть такі комбінації: 01, 10, 02, 20, 03, 30, 12, 21, 13, 31, 23, 32.

21. Розв'язати попередню задачу для $q = 4$ та $n = m = 3$.

22. Побудувати комбінації недвійкового коду на розміщення для алфавіту $q = 5$ і $n = m = 2$.

23. Побудувати комбінації недвійкового коду на певне число сполучень для алфавіту $q = 4$ та $n = m = 3$.

Р о з в ' я з а н н я. Загальна кількість комбінацій недвійкового коду на певне число сполучень

$$N = q! / [(q - m)!\, m!] = 4! / [(4 - 3)!\, 3!] = 4.$$

До цього коду за умови, що його алфавіт $q \in [0, 1, 2, 3]$, належатимуть комбінації 012, 013, 023, 123.

24. Розв'язати попередню задачу для $q = 4$ та $n = m = 2$.

25. Побудувати комбінації недвійкового коду на певне число сполучень для алфавіту $q = 5$ і $n = m = 3$.

26. Побудувати комбінації недвійкового коду на всі сполучення для алфавіту $q = 3$ та $n = m = 3$.

Р о з в ' я з а н н я. Загальна кількість комбінацій недвійкового коду на всі сполучення

$$N = q^n = 3^3 = 27.$$

До цього коду за умови, що його алфавіт $q \in [0, 1, 2]$, належатимуть такі комбінації: 000, 001, 002, 010, 020, 012, 021, 011, 022, 100, 110, 111, 112, 101, 102, 120, 122, 121, 200, 201, 202, 210, 220, 221, 222, 212, 211.

27. Розв'язати попередню задачу для $q = 4$ та $n = m = 2$.

28. Побудувати комбінації недвійкового коду на всі сполучення для алфавіту $q = 5$ і $n = m = 2$.

29. Побудувати комбінації недвійкового змінно-якісного коду для алфавіту $q = 4$ та $n = 3$.

Р о з в ' я з а н н я. Загальна кількість комбінацій недвійкового змінно-якісного коду

$$N = q(q-1)^{n-1} = 4 \cdot 3^2 = 36.$$

До цього коду за умови, що його алфавіт $q \in [0, 1, 2, 3]$, належатимуть такі комбінації: 012, 010, 020, 030, 013, 021, 031, 023, 032, 101, 102, 103, 120, 130, 123, 121, 131, 132, 201, 202, 203, 212, 213, 231, 210, 230, 232, 301, 302, 303, 310, 320, 312, 313, 321, 323.

30. Розв'язати попередню задачу для $q = 5$ і $n = 3$.

31. Побудувати комбінації недвійкового змінно-якісного коду для алфавіту $q = 5$ і $n = 4$.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які коди належать до первинних?
2. Де використовується код Морзе?
3. Як будуються число-імпульсні коди?
4. Де застосовуються рівномірні двійкові первинні коди?
5. Які особливості побудови п'ятирозрядного міжнародного телеграфного коду № 2?
6. На які зони можна поділити міжнародний семирозрядний стандартний код № 5?
7. Які функції виконують чотири групи функціональних символів міжнародного семирозрядного стандартного коду № 5?
8. З якою метою розроблено коди КОІ-7 і КОІ-8 для обміну інформацією?
9. Як будуються ДДК з додатною вагою розрядів?
10. Як будуються ДДК з додатними та від'ємними вагами розрядів?
11. Яка особливість побудови ДДК із самодоповненням?
12. Де застосовуються ДДК?
13. Яка особливість побудови двійково-шістнадцяткового коду?
14. З якою метою застосовуються рефлексні коди?
15. Які характерні особливості коду Грея?
16. Чим відрізняються недвійкові первинні коди від двійкових?
17. Як будуються недвійкові коди на перестановки?
18. Які особливості побудови недвійкових кодів на розміщення?
19. Як будуються недвійкові коди на певне число сполучень?
20. Чим різняться недвійкові коди на певне число сполучень і на всі сполучення?
21. Які особливості побудови недвійкових змінно-якісних кодів?

РОЗДІЛ 7 КОДИ, ЩО ВИЯВЛЯЮТЬ ПОМИЛКИ

Особливість кодів, які виявляють помилки, полягає в тому, що кодові комбінації, які входять до складу цих кодів, різняться кодовою відстанню, не меншою ніж $d_{min}=2$.

Такі коди умовно можна поділити на дві групи: коди, в яких використовуються всі комбінації, але до кожної з них за обумовленим правилом додаються r перевірних елементів; коди, утворені зменшенням кількості дозволених комбінацій.

До першої групи кодів, що виявляють помилки, належать коди з перевіркою на парність і непарність; код із простим повторенням; інверсний та кореляційний коди; до другої — код зі сталою вагою. Код з кількістю одиниць у комбінації, кратною трьом, може належати до першої або другої групи кодів залежно від методики його побудови.

7.1. ДВІЙКОВІ КОДИ, ЩО ВИЯВЛЯЮТЬ ПОМИЛКИ

7.1.1. КОД ІЗ ПЕРЕВІРКОЮ НА ПАРНІСТЬ

Це найпоширеніший код, який застосовується для виявлення поодиноких помилок і всіх помилок непарної кратності. Код містить $(n-1)$ інформаційних й один перевірний елементи, належить до систематичних кодів і позначається як $(n, n-1)$-код.

Перевірний елемент коду визначається сумою за модулем 2 всіх інформаційних елементів:

$$b_1 = \sum_{i=1}^{n-1} \oplus a_i,$$

тобто він утворюється доповненням комбінації k-елементного первинного коду одним елементом таким чином, щоб кількість одиниць у новому n-розрядному $(n = k+1)$ коді була парною. Кодова відстань $d_{min}=2$.

Для виявлення помилки на приймальному боці перевіряють на парність усю прийняту кодову комбінацію, визначаючи кодовий синдром

$$s_1 = \sum_{i=1}^{n-1} \oplus a_i' \oplus b_1',$$

де a_i', b_1' — прийняті на приймальному боці відповідно інформаційні та перевірний елементи.

Вважається, що при $s_1 = 0$ помилки в комбінації немає, а при $s_1 = 1$ помилка є. Надмірність коду визначається виразом

$$R_{над} = 1 - k/(k+1) = 1/(k+1).$$

7.1.2. КОД ІЗ ПЕРЕВІРКОЮ НА НЕПАРНІСТЬ

Цей код відрізняється від попереднього тим, що кожна його комбінація має непарну кількість одиниць, тобто додатковий перевірний елемент формують, виходячи з кількості одиниць у початковій кодовій комбінації: при парній кількості перевірний елемент дорівнює одиниці, а при непарній — нулю.

При виявлення помилки в кодовій комбінації на приймальному боці її перевіряють на непарність. Код є подільним завдяки $n-1$ інформаційних й один перевірний елементи; він може так само виявляти помилки та має надмірність, як і код із перевіркою на парність.

7.1.3. КОД ІЗ ПРОСТИМ ПОВТОРЕННЯМ

Код із простим повторенням (без інверсії) є подільним лінійним кодом. Він містить k інформаційних і $r = k$ перевірних елементів. У цьому коді r перевірних елементів є простим повторенням k інформаційних елементів первинної кодової комбінації: $b_i = a_i$, де $i = 1 \ldots k$.

Неважко бачити, що код має відстань $d_{min} = 2$, він може використовуватися для виявлення поодиноких помилок. Ця процедура зводиться до порівняння однойменних інформаційних і перевірних елементів у прийнятій кодовій комбінації. Незбіг їх свідчить про наявність помилок у ній. Код дає змогу виявити не лише однократні помилки, а й деякі помилки більшої кратності, за винятком «дзеркальних», коли в інформаційній та перевірній послідовностях кодової комбінації внаслідок дії завад спотворюються елементи, що знаходяться на однакових за номером розрядах.

Надмірність коду визначається виразом

$$R_{над} = 1 - k/(2k) = 1/2.$$

7.1.4. ІНВЕРСНИЙ КОД

Інверсний код (із повторенням та інверсією) є подільним лінійним кодом, який має k інформаційних і стільки ж перевірних елементів. Його відмінність від попереднього коду полягає в тому, що значення перевірних елементів у ньому залежать від значення суми за модулем 2 всіх інформаційних елементів. За умови $\sum_{i=1}^{k} Oa_i = 0$, тобто при парній кількості одиниць у початковій кодовій комбінації, перевірні елементи просто повторюють інформаційні ($b_i = a_i$, де $i = 1 \ldots k$), а за умови $\sum_{i=1}^{k} Oa_i = 1$, тобто при непарній кількості зазначених одиниць, перевірні елементи повторюють інформаційні в інвертованому вигляді (в оберненому коді): $b_i = a_i \oplus 1$, де $i = 1 \ldots k$.

Для виявлення помилок на приймальному боці в послідовності, що складається з $2k$ елементів, спочатку підсумовують одиниці, які знаходяться в перших k елементах. Якщо їх кількість парна, то решту k елементів приймають у позитиві. Обидві зареєстровані частини комбінації поелементно порівнюють (перший елемент із першим, другий — з другим і т.д.). За наявності хоча б одного незбігу вся послідовність елементів бракується.

Якщо кількість одиниць серед перших k елементів непарна, то решту k елементів приймають у негативі (інвертують), після чого поелементно порівнюють їх. Наявність незбігу призводить до відбраковування всіх $2k$ елементів. Така побудова коду дає змогу виявляти майже всі випадки спотворення його елементів, крім двократних «дзеркальних» помилок.

Надмірність коду визначається виразом

$$R_{\text{над}} = 1 - k/(2k) = 1/2.$$

7.1.5. КОРЕЛЯЦІЙНИЙ КОД

У цьому коді кожний розряд двійкового початкового коду записується у вигляді двох елементів: 0 — як 01, а 1 — як 10. Так, початковій кодовій комбінації 010011 відповідатиме комбінація 011001011010 кореляційного коду. В технічній літературі такий двійковий запис дуже часто називається *Манчестер-кодом.*

Приймальний пристрій в кожному такті, що складається з двох сусідніх елементів кореляційного коду, має зафіксувати перехід $0 \rightarrow 1$ або $1 \rightarrow 0$. У разі прийняття двох нулів або одиниць приймальний пристрій фіксує наявність помилки.

Кореляційний код дає змогу виявляти помилки будь-якої кратності, але не здатний виявити двократні «дзеркальні» помилки, коли сусідні елементи одного такту під впливом завад змінюються на протилежні за значенням.

Надмірність коду визначається виразом

$$R_{\text{над}} = 1 - k/(2k) = 1/2.$$

До переваг кореляційного коду, крім відсутності постійної складової в напрузі кодового сигналу при передачі кодової комбінації по каналу зв'язку, можна віднести також можливість самосинхронізації генератора приймача, оскільки прийняття кожного біта супроводжується фронтом сигналу, що приймається, в центрі біта.

7.1.6. КОД ЗІ СТАЛОЮ ВАГОЮ

Код зі сталою вагою, тобто з незмінною кількістю одиниць і нулів у комбінаціях, часто називається *кодом на одне сполучення.* Повна кількість комбінацій цього коду визначається виразом

$$N = C_n^m = \frac{n!}{m!(n-m)!},$$

де m — кількість одиниць у кодовій комбінації завдовжки n.

Такий код утворюється з двійкового простого коду відбором комбінацій, що мають однакову кількість одиниць m. Приймальний пристрій, підраховуючи кількість одиниць у прийнятій кодовій комбінації, виявляє помилки, якщо кількість перших відрізнятиметься від m.

Код зі сталою вагою має мінімальну кодову відстань $d_{\min} = 2$. Він виявляє всі помилки непарної кратності, а також усі помилки парної кратності, що призводять до порушення умови $m = \text{const}$.

Надмірність коду визначається виразом

$$R_{\text{над}} = 1 - (\log_2 C_n^m)/n.$$

Порівняно з кодом із простим повторенням цей код при меншій його надмірності дає змогу виявляти помилки тієї самої кратності.

7.1.7. КОД ІЗ КІЛЬКІСТЮ ОДИНИЦЬ У КОМБІНАЦІЇ, КРАТНОЮ ТРЬОМ

Цей код можна утворити або додаванням до кожної комбінації початкового коду $r = 2$ перевірних елементів, або зменшенням кількості дозволених комбінацій початкового коду з

накладанням додаткової умови: кількість одиниць у кожній комбінації має бути кратною трьом.

У першому випадку до початкової кодової комбінації додаються два перевірних розряди, які мають такі значення, що сума одиниць у кодовій комбінації стає кратною трьом. Так, якщо початкова кодова комбінація має дві або п'ять одиниць, то для здобуття ваги $w = 3$ або 6 кодової комбінації треба доповнити її двома перевірними елементами 10. Якщо ж у початковій комбінації є одна або чотири одиниці, то вона доповнюється двома перевірними елементами 11. Так, комбінація 01010 початкового коду, закодована кодом із кількістю одиниць, кратною трьом, матиме вигляд 0101010, а $10000 \rightarrow 1000011$, $0110 \rightarrow 011010$, $101100 \rightarrow 10110000$, $110110 \rightarrow 11011011$, $0111011 \rightarrow 011101110$ тощо.

У другому випадку з усіх комбінацій початкового коду вибирають тільки ті, які мають вагу $w = 3$ та 6. Решту комбінацій використовувати не можна.

Код дає змогу виявити всі поодинокі помилки та деякі помилки більшої кратності, що призводять до порушення умови $w = 3$ або 6, де w — кількість одиниць у кодовій комбінації. Здатність коду виявляти помилкові комбінації майже така сама, як і коду зі сталою вагою.

Надмірність коду з доповненням до необхідної кількості одиниць визначається виразом

$$R_{\text{над}} = 1 - \frac{2}{k+2},$$

а коду, що утворюється відбором із загальної кількості комбінацій з відповідною кількістю одиниць (3 або 6), — виразом

$$R_{\text{над}} = 1 - \frac{\log_2\left(C_n^3 + C_n^6\right)}{n}.$$

7.2. НЕДВІЙКОВІ КОДИ, ЩО ВИЯВЛЯЮТЬ ПОМИЛКИ

Розрізняють два принципи побудови надмірних недвійкових (q-кодів, багатопозиційних) кодів, що виявляють помилки: введенням додаткових перевірних елементів, які утворюються після виконання лінійних операцій над елементами кодової комбінації; збільшенням надмірності завдяки зменшенню кількості дозволених і зростанню кількості недозволених кодових комбінацій. В обох випадках досягається збільшення кодової відстані до значення, що дає змогу виявити ту чи іншу кількість помилок у комбінації.

Як відомо, мінімальна кодова відстань для кодів, які виявляють помилки, визначається виразом $d_{\min} \geq \nu_в + 1$, де $\nu_в$ — кратність помилки, що виявляється.

Із недвійкових кодів з додатковими перевірними елементами, що виявляють помилки, найпростіше реалізуються коди з перевіркою на парність за модулем q та код із простим повторенням, а з недвійкових кодів, утворених збільшенням кількості дозволених кодових комбінацій, найбільшого поширення дістали незвідні змінно-позиційні коди (НЗ-коди) з елементами одиниці та різної ваг.

7.2.1. КОД ІЗ ПЕРЕВІРКОЮ ЗА МОДУЛЕМ q

Цей код будується аналогічно двійковому коду з перевіркою на парність за модулем 2. Відмінність у побудові полягає у доповненні комбінацій первинного q-коду перевірним розрядом до значення основи (алфавіту) q коду, тобто якщо кодова комбінація є множиною k елементів $\{a_1 a_2 \ldots a_k\}$, де $a_1, a_2, \ldots, a_k$ — інформаційні елементи комбінації, що набувають значень від 0 до $(q - 1)$, то перевірний розряд визначається сумою цих елементів за модулем q:

$$b_1 = (a_1 \oplus a_2 \oplus \ldots \oplus a_k) \bmod q.$$

Цей вираз і є алгоритмом побудови недвійкового коду з перевіркою за модулем q.

Якщо кожний розряд кодової комбінації має m позицій, тобто множиною позицій (знаків з алфавіту q), то перевірний розряд також повинен мати m позицій. Значення позицій перевірного розряду в цьому разі визначається сумою відповідних розрядів позицій всіх розрядів кодової комбінації за модулем q. Розглядуваний код має незначну надмірність $R_{\text{над}} = 1/(k + 1)$ і змогу виявляти наявність помилок у розрядах кодової комбінації при невідповідності значення перевірного розряду k інформаційних розрядів за модулем q.

7.2.2. КОД ІЗ ПОВТОРЕННЯМ

В основу побудови цього коду за аналогією з двійковим покладено повторення початкової кодової комбінації. Відмінність коду від аналогічного двійкового полягає в тому, що повторення кодової комбінації першого може виконуватися паралельно в часі введенням додаткової позиції надмірності. Так, при використанні багаточастотного коду подвоєння кількості частотних позицій забезпечує паралельну передачу комбінації цього коду.

При цьому з'являються додаткові переваги над двійковим кодом: передача інформаційної та перевірної частин комбінації багаточастотного коду виконується з рознесенням за частотою, що підвищує завадостійкість коду при селективних завмираннях, характерних для деяких типів безпроводових ліній зв'язку; при використанні кількох ознак сигналу передача інформаційної та перевірної частин кодової комбінації може виконуватися позиціями різних ознак (наприклад, інформаційна частина може передаватися частотними позиціями, а перевірна — фазовими, що застосовується для підвищення або вірогідності, або швидкості передачі інформації).

Надмірність коду з повторенням можна оцінити надмірністю позицій ознак сигналу. При цьому надмірність $R_{над} = 0{,}5$. Однак часова надмірність, яка згадувалася вище, може не підвищуватися.

Алгоритм побудови коду з повторенням має вигляд

$$a_i \Leftrightarrow b_i,\ i \in [1, k], \qquad (7.1)$$

де a_i, b_i — множини позицій, призначені для передачі i-х інформаційного та перевірного елементів кодової комбінації відповідно; k — кількість інформаційних елементів.

Розглядуваний код згідно з алгоритмом (7.1) дає змогу виявляти всі помилки, за винятком помилок у парних множинах позицій, призначених для передачі інформаційної та перевірної частин коду, що несуть одну й ту саму інформацію.

7.2.3. НЕЗВІДНІ ЗМІННО-ПОЗИЦІЙНІ КОДИ

Незвідні змінно-позиційні коди з елементами однієї ваги. Головна перевага цих кодів полягає в малій кількості елементів і відповідно часових позицій (інтервалів) у комбінації, що має виграш у швидкості передачі порівняно з двійковими кодами.

З класу змінно-позиційних кодів найбільшого поширення дістали НЗ-коди, при використанні яких відпадає необхідність у жорсткій синхронізації приймальної апаратури, що, цілком природно, підвищує надійність цих кодів. Розглянемо методику побудови їх.

Під *незвідним змінно-позиційним кодом* будемо розуміти код, який задовольняє такі умови:

- кожна кодова комбінація складається з однакової кількості елементів, які передаються послідовно;
- кожний елемент комбінації містить m позицій (знаків) з q;
- сусідні елементи кодової комбінації різняться хоча б однією позицією;

- останній елемент комбінації збігається з її першим елементом, тобто перші та основні елементи кодової комбінації мають різні багатопозиційні сполучення, що не збігаються.

Виконання останньої умови забезпечує незвідність коду, що дає змогу виконувати передачу елементів без пауз і в деяких випадках відмовитися від синхронізації, а також спрощує послідовність виконання операцій при декодуванні.

Послідовність побудови НЗ-коду загалом така:

- береться множина m з q позицій сигналу;
- визначається кількість сполучень позицій за заданою кількістю позицій в кожному сполученні;
- вся кількість сполучень позицій розбивається на n груп, де n — кількість елементів кодової комбінації;
- утворюються кодові комбінації з n елементів, для кожного з яких беруться сполучення позицій з закріпленої за елементом групи.

За методом побудови кодових комбінацій багатопозиційні НЗ-коди можна поділити на два класи: без поділу алфавіту коду на групи; з поділом алфавіту коду на групи. Останній клас, у свою чергу, поділяють на два підкласи, що містять: НЗ-код, кожний елемент якого має m позицій з різних груп; НЗ-код, кожний елемент якого містить m позицій з однієї групи.

Під *НЗ-кодом з елементами однієї ваги* будемо розуміти код, елементи якого складаються з однакової кількості позицій m.

НЗ-код без поділу алфавіту коду на групи. Кількість сполучень позицій алфавіту цього коду визначається виразом

$$C_q^m = \frac{q!}{m!(q-m)!},$$

кількість можливих кодових комбінацій в тому разі, якщо всі сполучення позицій розподілено між n групами порівну, — визначим

$$N = (C_q^m/n)^n. \qquad (7.2)$$

Якщо вся кількість сполучень розподіляється між групами не порівну через остачу деякої кількості сполучень, то

$$C_q^m/n = S + Q/n,$$

де S і Q — цілі додатні числа, причому $Q < n$, тобто Q — ціла остача від ділення C_q^m/n, а S — ціла частина частки.

Дістасмо S сполучень позицій в кожній групі й остачу Q сполучень. Раціонально, не відкидаючи остачу Q, використати її для збільшення кількості кодових комбінацій. Доцільність такого використання остачі Q підтверджує такий приклад: якщо

відкинути три сполучення позицій ($Q = 3$), що залишилося, з 19 для цілого поділу їх на чотири групи, то загальна кількість можливих кодових комбінацій при цьому зменшиться приблизно в два рази.

Таким чином, з урахуванням $n - Q$ груп без додаткових сполучень позицій з Q, що залишилися, дістаємо

$$N_{\max} = S^{n-Q}(S+1)^{Q}, \tag{7.3}$$

або

$$N_{\max} = \left[\frac{C_q^m - Q}{n}\right]^{n-Q}\left[\frac{C_q^m - Q + n}{n}\right]^{Q}. \tag{7.4}$$

Формули (7.3) та (7.4) тотожні й дійсні при кількості груп сполучень, що дорівнює кількості елементів n кодової комбінації.

При накладанні додаткових умов (наприклад, коли для обумовленого елемента потрібна менша кількість сполучень позицій, тому що він має передавати меншу кількість інформації, або коли необхідно передавати якийсь елемент з більшою завадостійкістю тощо) в групах буде неоднакова кількість сполучень позицій. Нехай i-та група містить S_i сполучень, тоді

$$N = \prod_{i=1}^{n} S_i = S_1 S_2 \ldots S_n, \tag{7.5}$$

причому

$$\sum_{i=1}^{n} S_i = C_q^m.$$

При цьому для утворення якомога більшої кількості кодових комбінацій сполучення позицій між групами треба поділити приблизно порівну, щоб різниця між кількістю їх у групах була мінімальною й не перевищувала одиниці, за винятком окремих груп.

НЗ-код із поділом алфавіту коду на групи. При поділі q позицій на $n_{\text{гр}}$ груп у тому разі, коли кодовий елемент містить m позицій з різних груп, а кількість позицій в кожній групі однакова ($q_1 = q_2 = \ldots = q_n = q_i$), за умови $m = n_{\text{гр}}$ кількість можливих сполучень позицій визначається виразом

$$n_{\text{сп}} = (q_i)^{n_{\text{гр}}}, \tag{7.6}$$

а кількість кодових комбінацій — виразом

$$N = \left[\frac{(q_i)^{n_{\text{гр}}}}{n}\right]^{n}. \tag{7.7}$$

Загалом при $m = n_{\text{гр}}$ і $m \neq n_{\text{гр}}$ вираз (7.6) набуває вигляду

$$C_{n_{\text{гр}}}^{m}\left(\frac{q_i}{n_{\text{гр}}}\right)^{m} = C_{n_{\text{гр}}}^{m}(q_i)^{m},$$

а вираз (7.7) — вигляду

$$N = \left[\frac{C_{n_{\text{гр}}}^{m}(q_i)^{m}}{n}\right]^{n}. \tag{7.8}$$

Якщо q позицій поділяються на $n_{\text{гр}}$ груп не порівну, тобто $q_1 \neq q_2 \neq \ldots \neq q_n$, то для здобуття максимальної кількості кодових комбінацій необхідно поділити позиції між групами по можливості рівномірно.

Загалом для визначення кількості сполучень позицій можна використатися формулою

$$n_{\text{сп}} = \sum_{i_1=1}^{n_{\text{гр}}}\sum_{i_2=i_1+1}^{n_{\text{гр}}}\sum_{i_3=i_2+1}^{n_{\text{гр}}}\ldots\sum_{i_m=i_{m-1}+1}^{n_{\text{гр}}}\left(q_{i_1}q_{i_2}\ldots q_{i_m}\right). \tag{7.9}$$

Через те, що ця формула громіздка, для приблизних розрахунків можна обійтися виразом

$$n_{\text{сп}} = C_{n_{\text{гр}}}^{m}\left(\prod_{i=1}^{n_{\text{гр}}} q_i\right)^{m/n_{\text{гр}}}. \tag{7.10}$$

Кількість можливих кодових комбінацій з урахуванням (7.10) визначається виразом

$$N = \left[\frac{C_{n_{\text{гр}}}^{m}\left(\prod_{i=1}^{n_{\text{гр}}} q_i\right)^{m/n_{\text{гр}}}}{n}\right]^{n}. \tag{7.11}$$

Для побудови комбінацій розглядуваного НЗ-коду можна застосувати інший підхід, який ґрунтується на поділі q позицій на групи та використанні сполучень позицій, здобутих за до-

помогою позицій, що беруться тільки з однієї групи для утво рення відповідного елемента кодової комбінації, тобто $n_{гр} = n$. Однак такий метод веде до повної надмірності коду. Кількість можливих кодових комбінацій при цьому визначається ви разом

$$N = C_{q_1}^{m} C_{q_2}^{m} \ldots C_{q_n}^{m}.$$

Максимальну кількість кодових комбінацій дістають при дотриманні рівності $q_1 = q_2 = \ldots = q_n = q/n_{гр}$. Тоді

$$N_{\max} = \left(C_{q/n_{гр}}^{m} \right)^{n_{гр}}. \qquad (7.12)$$

Незвідний змінно-позиційний код з елементами різної ваги Цей код на відміну від НЗ-кодів з елементами однієї ваги має властивість поелементної синхронізації завдяки тому, що його елементи містять різну кількість позицій. Це дає змогу утворювати значно більшу кількість кодових комбінацій.

У коді сусідні часові позиції складаються з m_1 і m_2 позицій, а кодова комбінація має n_1 елементів з m_1 позиціями та n_2 елементів з m_2 позиціями, причому $n_1 + n_2 = n$. Розглянемо кількість варіантів утворення сполучень позицій цього коду.

Для коду без поділу алфавіту на групи кількість сполучень позицій m_1 і m_2 визначається як $C_q^{m_1}$ і $C_q^{m_2}$. Щоб утворити кодові комбінації, виконують паралельний поділ сполучень позицій на групи окремо для m_1 і m_2, тобто при побудові кодової комбінації можуть зустрічатися випадки, коли елемент із меншою кількістю позицій складається з тих самих позицій, які використовуються в сусідньому елементі з більшою кількістю позицій.

Тоді за аналогією з (7.4) маємо

$$N_{\max} = \left\{ \left[\frac{C_q^{m_1} - Q_1}{n_1} \right]^{n_1 - Q_1} \left[\frac{C_q^{m_1} - Q_1 + n_1}{n_1} \right]^{Q_1} \right\} \times \left\{ \left[\frac{C_q^{m_2} - Q_2}{n_2} \right]^{n_2 - Q_2} \left[\frac{C_q^{m_2} - Q_2 + n_2}{n_2} \right]^{Q_2} \right\}, \qquad (7.13)$$

де Q_1, Q_2 — кількість сполучень позицій, що залишилися після їх поділу відповідно між групами елементів n_1 і n_2.

Для НЗ-коду, утвореного при поділі алфавіту на групи, максимальна кількість кодових комбінацій визначається виразом

$$N_{\max} = \left\{ \frac{C_{n_{гр}}^{m_1} \left[\left(\frac{q - Q_3}{n_{гр}} \right)^{n_{гр} - Q_3} \left(\frac{q - Q_3 - n_{гр}}{n_{гр}} \right)^{Q_3} \right]^{m_1/n_{гр}} - Q_1}{n_1} \right\}^{n_1 - Q_1} \times \left\{ \frac{C_{n_{гр}}^{m_1} \left[\left(\frac{q - Q_3}{n_{гр}} \right)^{n_{гр} - Q_3} \left(\frac{q - Q_3 - n_{гр}}{n_{гр}} \right)^{Q_3} \right]^{m_1/n_{гр}} - Q_1 + n_1}{n_1} \right\} \times \left\{ \frac{C_{n_{гр}}^{m_2} \left[\left(\frac{q - Q_4}{n_{гр}} \right)^{n_{гр} - Q_4} \left(\frac{q - Q_4 - n_{гр}}{n_{гр}} \right)^{Q_4} \right]^{m_2/n_{гр}} - Q_2}{n_2} \right\}^{n_2 - Q_2} \times \left\{ \frac{C_{n_{гр}}^{m_2} \left[\left(\frac{q - Q_4}{n_{гр}} \right)^{n_{гр} - Q_4} \left(\frac{q - Q_4 - n_{гр}}{n_{гр}} \right)^{Q_4} \right]^{m_2/n_{гр}} - Q_2 - n_2}{n_2} \right\}^{Q_2}, \qquad (7.14)$$

де Q_1, Q_2 — кількість сполучень позицій, що залишилися після їх поділу відповідно між групами елементів n_1 і n_2; Q_3, Q_3 — кількість позицій, що залишилися після їх поділу між елементами відповідно з m_1 і m_2 позиціями.

У цьому разі дістаємо значний виграш у кількості кодових комбінацій порівняно з НЗ-кодом з елементами однієї ваги та поділі алфавіту коду на групи.

У зв'язку з тим, що елементи кодових комбінацій всіх без винятку НЗ-кодів мають обумовлену вагу, такі коди можуть виявляти будь-які помилки, що призводять до зміни ваги елементів (зменшення або збільшення кількості позицій елемента) кодової комбінації, забезпечуючи її числовий захист.

Надмірність НЗ-кодів визначається виразом

$$R_{\text{над}} = 1 - \frac{\log_2 N_{\text{д}}}{\log_2 N_{\max}},$$

де $N_{\max}$ і $N_{\text{д}}$ — відповідно максимально можлива та дозволена до використання в коді кількість комбінацій.

7.3. ШТРИХОВІ КОДИ

Останнім часом найперспективнішим напрямком автоматизації процесу введення інформації в ЕОМ є застосування штрихових кодів (ШК), які використовуються для ідентифікації одиничних предметів і сконструйовані спеціально для побудови систем автоматизованого збирання первинної інформації для подальшого її комп'ютерного оброблення та оптимальної організації матеріальних потоків у найрізноманітніших галузях господарства (виробництві, торгівлі, медицині тощо) [1].

Штрихові коди можуть мати двійковий або недвійковий еквівалент запису, тому й розглянемо їх окремо.

Перші ШК з'явилися в 1949 р., а вже на початку 60-х років XX ст. почали широко застосовуватися для ідентифікації залізничних вагонів у США. З 1973 р. ШК у США використовуються в торгівлі для кодування товарів, а в Європі з цією самою метою вони застосовуються, починаючи з 1977 р.

У загальному вигляді ШК — це послідовність штрихів і пробілів, розташованих у напрямку уявної прямої. У ШК інформацію можуть нести як штрихи та пробіли різної ширини, так і штрихи різної висоти. Існують ШК, виконані у вигляді концентричних кіл різної ширини, що забезпечує можливість декодування їх при проходженні траєкторії зчитування через центр зображення незалежно від напрямку руху оптичного пристрою зчитування. Проте найпоширенішим способом запису ШК є запис у вигляді штрихів і пробілів різної ширини [1].

До основних термінів, які характеризують ШК, належать:

- *штрих* — темна зона зображення на однотонному фоні, обмежена прямими паралельними лініями;
- *пробіл* — простір між штрихами;

- *висота* та *ширина штриха* — розміри зображення, подані у відносних одиницях (модулях) або в лінійних одиницях (міліметрах, дюймах тощо);
- *модуль* — основний розмір, якому кратні всі величини, що визначають параметри елементів зображення ШК;
- *знак* — сукупність штрихів і пробілів, яка несе закодовану інформацію про символ (повідомлення), що кодується;
- *код неперервний* — код, у якому знаки не розділені між собою роздільовими знаками;
- *код дискретний* — код, у якому знаки розділені між собою пробілами або знаками, що обмежують знак зліва та справа;
- *код двобічний* — код, побудований на основі обмежених зліва і справа знаків, у якому парні позиції кодового слова кодуються штрихами різної ширини, а непарні — пробілами різної ширини, тобто інвертованим зображенням штрихів. Необхідною умовою для таких кодів є парна кількість знаків у кодовому слові;
- *код двоспрямований* — код, який дає змогу провадити зчитування в двох напрямках — зліва направо і навпаки;
- *код контрольований* — код, у зображенні знаків і кодових слів якого закладено надмірну інформацію (один або кілька контрольних символів), що забезпечує виявлення помилок.

Взагалі існує багато різновидів ШК [1]. Так, за кольоровою гамою ШК поділяються на *багато-* та *двокольорові*. Реалізація перших не дуже проста (нанесення зображення коду, складність зчитувальних пристроїв тощо), тому на практиці використовують двокольорові (чорно-білі) ШК. За призначенням ШК поділяються на *цифрові* (використовуються для подання десяткових чисел) та *алфавітно-цифрові*; за побудовою — на *дискретні*, в яких знаки розділені між собою пробілами, та *неперервні*, в яких сусідні знаки не розділяються роздільником.

Деякі ШК мають вузькоспеціальне призначення. Так, коди 12, CIP, «Кодабар» використовуються в медицині; код 128 — у техніці зв'язку; коди 39 і 93 — у видавничій справі та промисловості; коди «2 з 5» та ITF — у промисловості й торгівлі [1].

Найпоширенішими ШК є коди EAN (Europian Article Number) та UPC (Uniform Product Code), рекомендовані створеною в 1988 р. Європейською асоціацією EAN для застосування в стандартизованій мережі зв'язку між членами цієї асоціації з використанням правил електронного обміну інформацією для управління, торгівлі та транспорту — UN/EDIFACT (United Nation Rules for Electronic Data Interchange for Administration, Commerce and Transport), що дістала назву EANCOM [1].

Розглянемо побудову деяких ШК.

Штрихові коди EAN призначені для кодування 10 цифр (0...9) і п'яти додаткових символів (СТАРТ, СТОП і розділові знаки).

Код EAN двосприямований і може мати кодове слово завдовжки 4...8, 10, 12...14 знаків. Однак існують два основних різновиди цього коду: EAN-13 і EAN-8, де цифрою позначено довжину коду (кількість знаків у кодовому слові). Так, код EAN-13 має структуру, зображену на рис. 7.1.

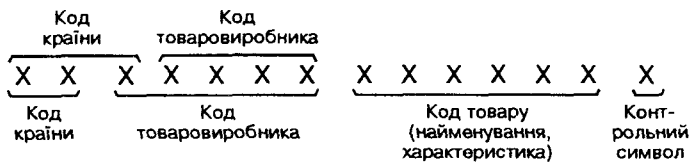


Рис. 7.1

До Європейської асоціації EAN зараз входять близько 80 країн. Коди деяких з них наведено нижче. За Україною EAN закріпила тризначковий код 482.

Країна	Код країни	Країна	Код країни
США, Канада	00...09	Португалія	560
Резерв EAN	20...29	Ісландія	569
Франція	30...37	Данія	57
Болгарія	380	Польща	590
Словенія	383	Румунія	594
Хорватія	385	Угорщина	599
Боснія-Герцеговина	387	Південно-Африканська Республіка	600, 601
Німеччина	40...43	Маврікій	609
Японія	45, 49	Марокко	611
Росія	460...469	Алжир	613
Тайвань	471	Туніс	619
Естонія	474	Фінляндія	64
Латвія	475	Китай	690, 691
Литва	477	Норвегія	70
Шрі-Ланка	479	Ізраїль	729
Філіппіни	480	Швеція	73
Україна	482	Гватемала, Сальвадор, Гондурас, Нікарагуа, Коста-Ріка, Панама	740...745
Молдова	484	Домініканська Республіка	746
Гонконг	489	Мексика	750
Велика Британія та Північна Ірландія	50	Венесуела	759
Греція	520	Швейцарія	76
Кіпр	529	Колумбія	770
Македонія	531	Уругвай	773
Мальта	535	Перу	775
Ірландія	539	Болівія	777
Бельгія та Люксембург	54	Аргентина	779
		Чилі	780
		Парагвай	784
		Еквадор	786
		Бразилія	789
		Італія	80...83
		Іспанія	84
		Куба	850
		Словаччина	858
		Чехія	859
		Югославія	860
		Туреччина	869
		Нідерланди	87
		Південна Корея	880
		Таїланд	885
		Сінгапур	888
		Індонезія	889
		Індія	890
		В'єтнам	893
		Австрія	90, 91
		Австралія	93
		Нова Зеландія	94
		Малайзія	955
		Папуа-Нова Гвінея	959

Зазначимо, що код країни-виробника може мати не два, а три знаки. В цьому разі код товаровиробника має не п'ять, а чотири знаки.

Код товаровиробника ідентифікує підприємство (присвоюється регіональною організацією EAN), яке випустило товар; решта цифр присвоюються товару безпосередньо самим підприємством. Вони характеризують найменування товару, його споживчі якості, розміри, масу, колір. Остання, 13-та, цифра є контрольною, вона формується за спеціальним алгоритмом:

- крок 1 — знаходять суму цифр, розташованих на непарних позиціях кодового слова (перегляд виконують справа наліво), і множать здобутий результат на 3;
- крок 2 — знаходять суму цифр, розташованих на парних позиціях кодового слова;
- крок 3 — додають результати, здобуті в попередніх кроках;
- крок 4 — контрольна цифра дорівнює найменшому числу, що не перевищує 9, яке, якщо його додати до результату, здобутому на попередньому кроці, дає число, кратне 10.

У кодах EAN використовуються чотири набори знаків A, B, C, D (табл. 7.1) для кодування десяткових цифр, а також знаків СТАРТ, СТОП (Н1, Н2, Н3) і розділових знаків (Н4 та Н5). Кожний знак складається з двох штрихів і двох пробілів. Довжина кожного знака для кодування цифр дорівнює семи модулям, а допоміжні знаки мають довжину три, п'ять і шість модулів. Як знак СТАРТ використовуються знаки Н1 і Н2, а як знак СТОП — ці самі знаки залежно від символів початку та кінця кодового слова (табл. 7.2, де 0 — пробіл, 1 — штрих).

Штрихові коди EAN, що мають довжину 4...7 знаків, кодуються набором А (див. табл. 7.1) й обмежуються знаками СТАРТ (Н1) і СТОП (Н2), а кодові слова завдовжки 8, 10, 12 і

Таблиця 7.1

Символ	Набір A		Набір B		Набір C		Набір D	
	Штриховий код	Двійковий еквівалент	Штриховий код	Двійковий еквівалент	Штриховий код	Двійковий еквівалент	Штриховий код	Двійковий еквівалент
0		0001101		0100111		1110010		1011000
1		0011001		0110011		1100110		1001100
2		0010011		0011011		1101100		1100100
3		0111101		0100001		1000010		1011110
4		0100011		0011101		1011100		1100010
5		0110001		0111001		1001110		1000110
6		0101111		0000101		1010000		1111010
7		0111011		0010001		1000100		1101110
8		0110111		0001001		1001000		1110110
9		0001011		0010111		1110100		1101000
H1		101	Обмежувальні знаки СТАРТ і СТОП				Роздільові знаки	
H2		010101						
H3		101010						
H4		01010						
H5		10101						

Таблиця 7.2

Знак СТАРТ	Кодове слово	Знак СТОП	Знак СТАРТ	Кодове слово	Знак СТОП
H1	0...0	H1	H1	0...1	H2
H3	1...1	H2	H3	1...0	H1

H4 знаків складаються з двох частин з однаковою кількістю знаків у кожній з них, розділених знаком H4. Для зображення знаків лівої частини кодового слова використовуються набори A та B, а правої — набори C та D. Такі кодові слова мають обмежувальні знаки СТАРТ і СТОП типу H1 [1].

Нехай деякий вид товару має номер 869047210056, тобто цей товар виготовлено в Туреччині (869). 12 цифр, які вже є, необхідно справа доповнити контрольною цифрою K, яку визначають згідно з алгоритмом, наведеним вище.

Процес кодування товару кодом EAN-13 показано на рис. 7.2, де K визначається так:

$$(6 + 0 + 1 + 7 + 0 + 6) \cdot 3 = 20 \cdot 3 = 60;$$
$$(5 + 0 + 2 + 4 + 9 + 8) = 28;$$
$$60 + 28 = 88;$$
$$88 + 2 = 90 \rightarrow K = 2.$$

У коді EAN-13 штрихове зображення складається з двох частин по шість знаків у кожній, які розділено знаком H4, і має зліва та справа обмежувальні знаки H1 (СТАРТ і СТОП). Перша цифра (U_{12}) товарного номера у вигляді штрихів і пробілів не кодується, а тільки пишеться зліва вище і визначає неявне кодування цифр, розташованих у лівій частині кодового слова між знаками H1 і H4 (табл. 7.3). Літерами A та B в табл. 7.3 позначено набори з табл. 7.1, якими кодуються відповідні знаки лівої частини кодового слова.

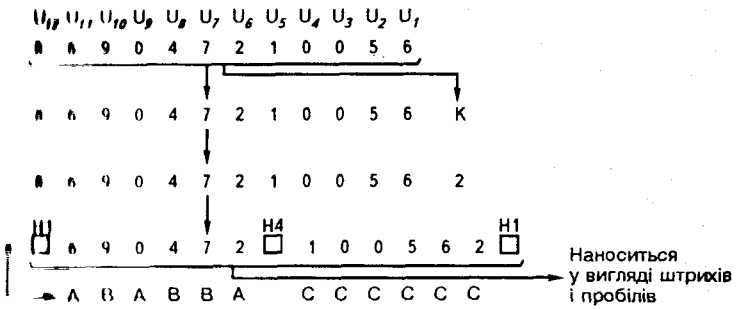


Рис. 7.2

Таблиця 7.3

U_{12}	U_{11}	U_{10}	U_9	U_8	U_7	U_6	U_{12}	U_{11}	U_{10}	U_9	U_8	U_7	U_6
0	A	A	A	A	A	A	5	A	B	B	A	A	B
1	A	A	B	A	B	B	6	A	B	B	B	A	A
2	A	A	B	B	A	B	7	A	B	A	B	A	A
3	A	A	B	B	B	A	8	A	B	A	B	B	A
4	A	B	A	A	B	B	9	A	B	B	A	B	A

Щодо розглядуваного прикладу рядок 8 ($U_{12} = 8$) табл. 7.3 має вигляд ABABBA й визначає процес кодування цифр $U_{11}... U_6$, тобто цифри $U_{11} = 6$, $U_9 = 0$ і $U_6 = 2$ кодуються знаками набору A, а цифри $U_{10} = 9$, $U_8 = 4$ та $U_7 = 7$ — знаками набору B (див. табл. 7.1). Цифри $U_5... U_1$, а також контрольна цифра K = 2 кодуються відповідними знаками набору C. Знаки набору D, а також знаки H2, H3 та H5 в коді EAN-13 не використовуються.

Таким чином, товарному номеру 869047210056 відповідає код EAN-13, показаний на рис. 7.3, де обмежувальні (H1) і розділовий (H4) знаки зображено подовженими по висоті штрихами.

Деякі товари можуть мати короткий номер, який складається з семи цифр. Після доповнення його контрольною цифрою за наведеним вище алгоритмом для коду EAN-13 дістають вісім цифр, які кодують кодом EAN-8. Кодове слово останнього складається зі знака СТАРТ (H1), чотирьох знаків набору A (див. табл. 7.1), розділового знака H4, трьох знаків набору C, знака контрольної цифри в наборі C, а також знака СТОП (H1). У коді EAN-8 перша цифра U_7 неявне кодування не визначає, а кодується, як і наступні цифри $U_6... U_4$, набором A (див. табл. 7.1).

Штрихові коди UPC, що використовуються в США та Канаді для ідентифікації товарів, також призначені для кодування 10 цифр (0...9) і п'яти додаткових знаків. Ці коди сумісні з кодами EAN, оскільки застосовується одна й та сама таблиця наборів знаків (табл. 7.1).

Існує кілька різновидів кодів UPC, з яких найбільшого поширення дістали коди UPC-A та UPC-E. Кодове слово перших складається з 12 цифр (остання цифра — контрольна), тобто на одну цифру менше, ніж у коді EAN-13. Це пояснюється тим, що код країн (США та Канади) має тільки дві цифри. Відмітними особливостями коду UPC-A відносно коду EAN-13 є те, що:

Рис. 7.3

• старша цифра (U_{11}) товарного номера в штриховому зображенні кодується явно;

• штрихове зображення кодового слова в коді UPC-A складається зі знака

СТАРТ (H1), шести знаків набору A, розділового знака (H4), п'яти знаків набору C, знака контрольної цифри набору C та знака СТОП (H1) (див. табл. 7.1);

Рис. 7.4

• у штриховому зображенні цифри U_{11} і K мають висоту, що дорівнює висоті знаків H1 і H4, причому значення цифр під цими знаками не наносяться;

• зліва від штрихового зображення друкується цифра 0, що ідентифікує код UPC-A.

Контрольна цифра кодів UPC визначається за тим самим алгоритмом, що й у коді EAN-13.

Нехай номер виду товару буде 00135792468, його контрольна цифра дорівнює 1. Тоді штрихове зображення номера 001357924681 кодом UPC-A матиме вигляд, показаний на рис. 7.4.

Товарний номер у коді UPC-E складається з шести цифр; при цьому він також поділяється на дві частини по три цифри в кожній, перша з яких кодується набором A, а друга (в тому числі й контрольний знак) — набором C (див. табл. 7.1).

При декодуванні кодів EAN і UPC насамперед визначається контрольний знак, який має збігатися з переданим. Крім того, сума всіх цифр прийнятого на приймальному боці кодового слова має бути кратною 10. За цих умов помилки немає. Інакше виявляється помилка, причому вона буде й при неправильному прийманні знаків, які не відповідають наборам, установленим першим кодуванням за старшим знаком у коді EAN-13 цифр лівої половини кодового слова, тому що набори A та B (див. табл. 7.1) не збігаються.

КОНТРОЛЬНІ ЗАДАЧІ

1. Закодувати комбінацію 0110110 двійкового простого коду ($k = 7$) двійковими кодами, що виявляють помилки з перевіркою на парність і простим повторенням. Виявити однократну помилку та порівняти надмірності цих кодів.

Розв'язання. Комбінація коду з перевіркою на парність матиме вигляд $A_1 = 01101100$, а коду з простим повторенням — вигляд $A_2 = 01101100110110$.

Нехай в комбінації коду з перевіркою на парність виникла однократна помилка, вектор якої $E_1 = 00100000$. Тоді сума $A_1 \oplus E_1 = 01001100$. У цьому разі сума за модулем 2 елементів утвореної кодової комбінації дорівнює 1, тобто непарна, що вказує на наявність у ній помилки. Надмірність коду $R_{код} = 1 - 7/8 = 0{,}125$.

Нехай в комбінації коду з простим повторенням вектор однократної помилки буде $E_2 = 00010000000000$. Тоді сума $A_2 \oplus E_2 = 01111100110110$.

Перевіряючи першу та другу частини кодової комбінації (додаючи їх за модулем 2), дістаємо остачу, яка не дорівнюватиме нулю:

$$\begin{array}{r} \oplus\ 0111110 \\ \underline{0110110} \\ 0001000, \end{array}$$

що вказує на наявність помилки в прийнятій кодовій комбінації. Надмірність коду $R_{над2} = 0{,}5$.

Таким чином, $R_{над2} > R_{над1}$.

2. Закодувати комбінацію 110010 двійкового простого коду ($k = 6$) двійковим кодом, що виявляє помилки з кількістю одиниць в комбінації кратною трьом, та інверсним кодом. Виявити однократну помилку та порівняти надмірності цих кодів.

Розв'язання. Комбінація коду з кількістю одиниць, кратною трьом, матиме вигляд A_1 = 11001000, а інверсного коду — вигляд A_2 = 10010001101.

Нехай в комбінації коду з кількістю одиниць, кратною трьом, виникла однократна помилка, вектор якої E_1=00010000. Тоді сума $A_1 \oplus E_1$=11011000. У цьому разі вага утвореної кодової комбінації $w = 4$, тобто відрізняється від $w = 3$, що свідчить про наявність у ній помилки. Надмірність коду $R_{над1} = 1 - 6/8 = 0{,}25$.

Нехай в комбінації інверсного коду виникла однократна помилка, вектор якої E_2 =000010000000. Тоді сума $A_2 \oplus E_2$ = 110000001101. При прийманні перевіряємо кількість одиниць у першій половині кодової комбінації, яка дорівнює 2. Це означає, що друга половина комбінації має прийматися в позитиві. Порівнюючи першу та другу (неінвертовану) частини прийнятої кодової комбінації, дістаємо незбіг у п'яти розрядах, що вказує на наявність у ній помилки. Надмірність коду $R_{над2} = 0{,}5$.

Таким чином, $R_{над2} > R_{над1}$.

3. Закодувати комбінацію 01100 двійкового простого коду ($k = 5$) двійковим кодом, що виявляє помилки з перевіркою на непарність, і кореляційним кодом. Виявити однократну помилку та порівняти надмірності цих кодів.

Розв'язання. Комбінація коду з перевіркою на непарність матиме вигляд A_1 = 011001, а кореляційного — вигляд A_2 =0110100101.

Нехай в комбінації коду з перевіркою на непарність виникла однократна помилка, вектор якої E_1 =001000. Тоді сума $A_1 \oplus E_1$ = 010001. При прийманні перевіряємо за модулем 2 суму елементів утвореної кодової комбінації. Вона дорівнює нулю, тобто парна, що свідчить про наявність помилки в комбінації. Надмірність коду $R_{над1}$ =1 – 5/6 = 0,166(6).

Нехай в комбінації кореляційного коду виникла однократна помилка, вектор якої E_2 = 0000100000. Тоді сума $A_2 \oplus E_2$ = 0110000101. Як відомо, декодування комбінації при її прийманні провадиться тактами по два елементи в кожному. При цьому елементи одного такту не повинні мати однакове значення, тобто не повинно бути сполучень 00 і 11. У випадку, що розглядається в задачі, в третьому такті (парі елементів) є сполучення 00, що вказує на наявність помилки в комбінації. Надмірність коду $R_{над2} = 0{,}5$.

Таким чином, $R_{над2} > R_{над1}$.

4. Закодувати комбінацію 011010011011 двійкового простого коду двійковим кодом, що виявляє помилки з перевіркою на парність, та інверсним кодом. Виявити однократну помилку та порівняти надмірності цих кодів.

5. Закодувати комбінацію 1110010 двійкового простого коду двійковими кодами, що виявляють помилки з перевіркою на непарність і простим повторенням. Виявити однократну помилку та порівняти надмірності цих кодів.

6. Закодувати комбінацію 10001111 двійкового простого коду двійковими кодами, що виявляють помилки: інверсним і кореляційним. Виявити однократну помилку та порівняти надмірності цих кодів.

7. Закодувати комбінацію 01010101 двійкового простого коду двійковими кодами, що виявляють помилки з кількістю одиниць, кратною трьом, і простим повторенням. Виявити однократну помилку та порівняти надмірності цих кодів.

8. З комбінацій двійкового простого коду при $k = 6$ побудувати двійкові коди, що виявляють помилки зі сталими вагами $w = 2$ та 4. Виявити однократні помилки та порівняти надмірності цих кодів.

9. З комбінацій двійкового простого коду при $k = 8$ побудувати двійкові коди, що виявляють помилки з постійними вагами $w = 3$ та 5. Виявити однократні помилки та порівняти надмірності цих кодів.

10. Закодувати всіма двійковими кодами, що виявляють помилки, двійкове подання числа поточного дня тижня. Виявити однократну помилку та порівняти надмірності цих кодів.

11. Закодувати всіма двійковими кодами, що виявляють помилки, двійкове подання порядкового номера поточного місяця року. Виявити однократну помилку та порівняти надмірності цих кодів.

12. Закодувати комбінацію 021 трійкового коду на всі сполучення недвійковими кодами, що виявляють помилки з перевіркою за модулем q (при q = 3) та з повторенням. Виявити однократну помилку та порівняти надмірності цих кодів.

Розв'язання. Комбінація трійкового коду з перевіркою за модулем 3 матиме вигляд A_1 = 0210, а трійкового коду з повторенням — вигляд A_2 = 021021.

Нехай в комбінації трійкового коду з перевіркою за модулем 3 виникла однократна помилка, вектор якої E_1 =2000. Тоді сума за модулем 3 $A_1 \oplus E_1$ = 2210. При прийманні перевіряємо за модулем 3 суму елементів утвореної кодової комбінації. Вона дорівнює 2, тобто відрізняється від нуля, що свідчить про наявність помилки в комбінації. Надмірність коду $R_{над1} = 1/(k + 1) = 1/4 = 0{,}25$.

Нехай в комбінації трійкового коду з повторенням виникла однократна помилка, вектор якої E_2 = 002000. Тоді сума за модулем 3 $A_2 \oplus E_2$ = = 020021. Порівнюючи першу (розряди 1...3) та другу (розряди 4...6) частини кодової комбінації, бачимо, що вони різняться в третьому й шостому розрядах. Це вказує на наявність помилки в прийнятій кодовій комбінації. Надмірність коду $R_{над2} = 0{,}5$.

Таким чином, $R_{над2} > R_{над1}$.

13. Закодувати комбінацію 615 вісімкового коду на всі сполучення недвійковими кодами, що виявляють помилки з перевіркою за модулем 8 і з повторенням. Виявити однократну помилку та порівняти надмірності цих кодів.

14. Закодувати комбінацію 8921 шістнадцяткового коду на всі сполучення недвійковим кодом із перевіркою за модулем 16. Виявити однократну помилку та визначити надмірність коду.

15. Закодувати комбінацію 413 п'ятіркового коду на всі сполучення недвійковим кодом із повторенням. Виявити однократну помилку та визначити надмірність коду.

16. Знайти максимальну кількість комбінацій незвідного змінно-позиційного коду зі сталою вагою та поділом алфавіту коду на групи, яку можна дістати для алфавіту $q = 8$ коду при кількості елементів у кодовій комбінації $m = 2$ та кількості позицій в одному елементі, що дорівнює кількості груп ($m = n_{гр} = 2$).

Розв'язання. Розрізняють два види цього НЗ-коду: код із поділом q позицій на $n_{гр}$ груп з m позиціями в одному елементі з різних груп й аналогічний код, у якого m позицій беруться тільки з однієї групи.

Отже, треба визначити максимальну кількість комбінацій для двох варіантів побудови НЗ-коду.

Для НЗ-коду, побудованого за першим варіантом, максимальна кількість комбінацій визначається виразом (7.8), тобто

$$N_1 = \left[\frac{C^m_{n_{гр}} (q_i)^m}{n} \right]^n,$$

де q_i — кількість позицій в кожній групі.

У даному разі $q_i = q / n_{гр} = 8/2 = 4$. Тоді

$$N_1 = \left[\frac{C_2^2 \cdot 4^2}{2} \right]^2 = 64.$$

Для НЗ-коду, побудованого за другим варіантом, максимальна кількість комбінацій визначається виразом (7.12), тобто

$$N_2 = \left(C^m_{q/n_{гр}} \right)^{n_{гр}} = \left(C^2_{8/2} \right)^2 = 36$$

Отже, максимальна кількість комбінацій НЗ-коду зі сталою вагою та поділом алфавіту коду на групи може бути утворена при використанні для формування m-позиційного елемента позицій алфавіту різних груп.

Таким чином, $N_1 > N_2$.

17. Знайти максимальну кількість комбінацій незвідного змінно-позиційного коду зі сталою вагою без поділу алфавіту коду на групи.

18. Знайти максимальну кількість комбінацій незвідного змінно-позиційного коду з елементами різної ваги.

19. Необхідно забезпечити передачу 24 команд незвідним змінно-позиційним кодом зі сталою вагою без поділу алфавіту коду на групи. Визначити потрібне значення алфавіту коду, якщо кількість елементів у кодовій комбінації $n = 2$, а кількість позицій, які передаються в одному елементі, $m = 2$.

20. Необхідно забезпечити передачу 80 команд незвідним змінно-позиційним кодом з елементами різної ваги без поділу алфавіту коду на групи. Визначити потрібну кількість позицій в елементах m_1 і m_2 коду, якщо його алфавіт $q = 6$, а кількість елементів у кодовій комбінації $n = 2$.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Які коди належать до кодів, що виявляють помилки?
2. Де використовуються коди, що виявляють помилки?
3. Як впливає мінімальна кодова відстань коду на його здатність виявляти помилки?
4. Які особливості побудови двійкового коду з перевіркою на парність?
5. Чим відрізняється двійковий код із перевіркою на непарність від аналогічного коду з перевіркою на парність?
6. Які особливості побудови двійкового інверсного коду?
7. Чим відрізняється двійковий код із простим повторенням від інверсного?
8. Яка надмірність двійкових кодів із перевіркою на парність, інверсного та з простим повторенням?
9. Які особливості побудови двійкового кореляційного коду?
10. Які особливості побудови двійкових кодів зі сталою вагою?
11. Чим відрізняється двійковий код із кількістю одиниць в комбінації, кратною трьом, від коду зі сталою вагою?
12. Які особливості побудови недвійкових кодів, що виявляють помилки?
13. Чим відрізняється недвійковий код із перевіркою за модулем q від аналогічного двійкового коду?
14. Чим відрізняється недвійковий код із повторенням від аналогічного двійкового коду?
15. Які коди належать до незвідних змінно-позиційних кодів?
16. Які особливості побудови незвідних змінно-позиційних кодів з елементами однієї ваги?
17. Які особливості побудови незвідних змінно-позиційних кодів з елементами різної ваги?
18. Чим різняться недвійкові незвідні коди з елементами однієї та різної ваг?
19. Де застосовується ШК?
20. У чому полягають відмінності між кодами EAN-13 та UPC-A?

РОЗДІЛ 8

КОДИ, ЩО ВИПРАВЛЯЮТЬ ПОМИЛКИ

Коди, що виправляють помилки (або коректувальні коди) повинні мати мінімальну кодову відстань $d_{min} \geq 3$. Її зростання досягається збільшенням кількості n розрядів коду або зменшенням кількості N_0 дозволених кодових комбінацій, які використовуються для передачі повідомлень, тобто підвищенням надмірності коду.

Найбільшого поширення серед двійкових коректувальних кодів [9, 12, 19, 20, 23, 39] *дістали систематичні та несистематичні блокові коди (вони, у свою чергу, поділяються на лінійні та нелінійні), а також рекурентні коди. З недвійкових кодів для захисту інформації від помилок* [6, 8] *застосовуються узагальнений код Хеммінга, ланцюговий та ітеративний коди.*

8.1. ДВІЙКОВІ ГРУПОВІ КОДИ

8.1.1. ЛІНІЙНИЙ СИСТЕМАТИЧНИЙ ГРУПОВИЙ (БЛОКОВИЙ) КОД

Застосувавши теорію лінійних просторів, можна дати таке визначення [4, 40]: *лінійним систематичним груповим двійковим* (n, k)*-кодом* називається код, у якого перевірні елементи b_j (де $j = 1 \ldots r$) знаходяться як суми за модулем 2 обумовлених інформаційних елементів a_i (де $i = 1 \ldots k$). У нелінійних двійкових (n, k)-кодах перевірні елементи знаходяться інакше.

Таким чином, у лінійному коді перевірні елементи визначаються як

$$\begin{aligned} b_1 &= \sum_{i=1}^{k} \mathrm{O}\alpha_{ji} a_i; \\ &\cdots\cdots\cdots \\ b_r &= \sum_{i=1}^{k} \mathrm{O}\alpha_{ri} a_i . \end{aligned} \qquad (8.1)$$

де $\sum \mathrm{O}$ — знак суми за модулем 2; α_{ji} — коефіцієнти, значення яких дорівнюють 0 або 1.

Як випливає з (8.1), закон побудови лінійного коду визначається вибором kr коефіцієнтів α_{ij}.

Одна з властивостей лінійних кодів полягає в тому, що сума за модулем 2 будь-яких двох дозволених кодових комбінацій також є дозволеною комбінацією цього коду.

Нехай є дві дозволені кодові комбінації лінійного коду

$$V^{(1)} = \left(a_1^{(1)}, \ldots, a_k^{(1)}, b_1^{(1)}, \ldots, b_r^{(1)}\right);$$

$$V^{(2)} = \left(a_1^{(2)}, \ldots, a_k^{(2)}, b_1^{(2)}, \ldots, b_r^{(2)}\right),$$

перевірні символи яких визначаються відповідно виразами

$$b_j^{(1)} = \sum_{i=1}^{k} \mathrm{O}\alpha_{ij} a_i^{(1)}; \quad b_j^{(2)} = \sum_{i=1}^{k} \mathrm{O}\alpha_{ji} a_i^{(2)}.$$

При додаванні за модулем 2 комбінацій $V^{(1)}$ і $V^{(2)}$ дістанемо комбінацію $V^{(3)}$, яка також буде комбінацією цього коду:

$$V^{(3)} = V^{(1)} \oplus V^{(2)} = a_1^{(1)} + a_1^{(2)}, \ldots,$$

$$a_k^{(1)} \oplus a_k^{(2)}, b_1^{(1)} \oplus b_1^{(2)}, \ldots, b_r^{(1)} \oplus b_r^{(2)},$$

де

$$b_j^{(1)} \oplus b_j^{(2)} = \sum_{i=1}^{k} \mathrm{O}\alpha_{ji} a_i^{(1)} \oplus \sum_{i=1}^{k} \mathrm{O}\alpha_{ji} a_i^{(2)} = \sum_{i=1}^{k} \mathrm{O}\alpha_{ji} \left(a_i^{(1)} \oplus a_i^{(2)}\right),$$

тому що при додаванні за модулем 2 діють асоціативний, комутативний та дистрибутивний закони [32].

Таким чином,

$$V^{(3)} = (a_1^{(3)}, \ldots, a_k^{(3)}, b_1^{(3)}, \ldots b_r^{(3)}),$$

де

$$a_i^{(3)} \quad a_i^{(1)} \oplus a_i^{(2)}, \; b_j^{(3)} = b_j^{(1)} \oplus b_j^{(2)} \text{ або } b_j^{(3)} = \sum_{i=1}^{k} \mathrm{O}\alpha_{ji} a_i^{(3)}.$$

Така властивість дає можливість побудувати всі дозволені комбінації лінійного коду, маючи тільки обмежену кількість їх. При цьому побудова лінійного коду виконується на основі твірної (породжувальної) матриці. Ця матриця будується так, щоб:

1) кількість початкових кодових комбінацій дорівнювала k, тобто кількості інформаційних елементів первинного коду;

2) всі початкові кодові комбінації були різними;

3) нульова комбінація не входила до складу початкових;

4) всі початкові комбінації були лінійно незалежними;

5) кількість одиниць в кожній початковій комбінації була не меншою ніж d_{min};

6) кодова відстань між будь-якими парами початкових комбінацій також була не меншою ніж $d_{\min}$.

Підібрані за цим правилом початкові комбінації записуються у вигляді *твірної матриці* $G_{(n,\,k)}$, яка містить k рядків і n стовпців:

$$G_{(n,k)} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} & b_{11} & b_{12} & \dots & b_{1r} \\ a_{21} & a_{22} & \dots & a_{2k} & b_{21} & b_{22} & \dots & b_{2r} \\ . & . & . & . & . & . & . & . \\ a_{k1} & a_{k2} & \dots & a_{kk} & b_{k1} & b_{k2} & \dots & b_{kr} \end{bmatrix}. \qquad (8.2)$$

Матрицю (8.2) можна подати також у вигляді двох підматриць: *інформаційної* E_k та перевірної $C_{(r,\,k)}$. Першу зручно записати в канонічній формі як одиничну підматрицю, що має k стовпців і стільки ж рядків:

$$E_k = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$

Перевірна підматриця $C_{(r,\,k)}$ будується підбором r-розрядних комбінацій з кількістю одиниць в рядку не меншою від $(d_{\min} - 1)$. При цьому необхідно враховувати, що сума за модулем 2 будь-яких розрядів повинна мати більш як $(d_{\min} - 2)$ одиниць.

Ураховуючи викладене, в межах теорії матриць можна дати таке визначення [40]: *лінійним систематичним груповим двійковим* (n, k)*-кодом* називається код, у якого всі його $(2^k - 1)$ ненульові комбінації можуть бути утворені як суми за модулем 2 рядків деякої матриці $G_{(n,\,k)}$ розміром $k \times n$, яка називається *твірною матрицею* коду.

Розглянемо приклад побудови матриці систематичного групового коду, здатного виправляти однократну помилку ($\nu_{вп} = 1$) при передачі 16 повідомлень ($N_д = 16$).

Відомо, що $N_д = 2^k$, звідки $k = 4$ ($N_д = 16 = 2^4$). Отже, кількість рядків твірної матриці $G_{(n,\,k)}$ дорівнює 4, а кількість стовпців — довжині n коду ($n = k + r$, де $r = 3$ для $\nu_{вп} = 1$ і $d_{\min} = 2\nu_{вп} + 1 = 3$), тобто кількість стовпців підматриці $G_{(r,\,k)}$ становить 3. Кількість перевірних розрядів можна визначити, користуючись табл. 8.1.

Згідно з правилом побудови підматриці $C_{(r,\,k)}$ кількість одиниць в ній має бути не меншою ніж $d_{\min} - 1 = 3 - 1 = 2$. З триелементних комбінацій для підматриці $C_{(r,\,k)}$ вибираємо тільки ті, в яких кількість одиниць більша двох: 110, 101, 011, 111.

Таблиця 8.1

$d_{\min}$	k									
	1	2	3	4	5	6	7	8	9	10
3	1	1	1	1	1	1	1	1	1	1
4	2	3	3	3	4	4	4	4	4	4
5	3	5	5	6	6	7	7	7	7	8
6	4	7	8	8	9	9	10	10	10	11
7	5	9	10	11	11	12	12	13	13	14

Через те, що як інформаційна E_k використовується одинична підматриця, твірна матриця лінійного коду для розглядуваного прикладу, тобто коду, який виправляє одну помилку, остаточно має вигляд

$$G_{(7,4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}. \qquad (8.3)$$

Рядки в перевірній підматриці $C_{(r,\,k)}$ можна міняти місцями. При цьому можна записати кілька варіантів твірних матриць для прикладу, що розглядається.

Побудуємо за допомогою поданої раніше твірної матриці всі комбінації розглядуваного (7, 4)-коду. Оскільки перша комбінація — нульова, друга — п'ята комбінації є рядками твірної матриці, решту 11 знайдемо підсумовуванням за модулем 2 всіх можливих комбінацій рядків твірної матриці. Всі 16 комбінацій (7, 4)-коду з цією матрицею матимуть такий вигляд:

1. 0000000;
2. 1000011;
3. 0100101;
4. 0010110;
5. 0001111;
6. 1100110 (2 ⊕ 3);
7. 1010101 (2 ⊕ 4);
8. 1001100 (2 ⊕ 5);
9. 0110011 (3 ⊕ 4);
10. 0101010 (3 ⊕ 5);
11. 0011001 (4 ⊕ 5);
12. 1110000 (2 ⊕ 3 ⊕ 4);
13. 0111100 (3 ⊕ 4 ⊕ 5);
14. 1011010 (2 ⊕ 4 ⊕ 5);
15. 1101001 (2 ⊕ 3 ⊕ 5);
16. 1111111 (2 ⊕ 3 ⊕ 4 ⊕ 5).

При побудові комбінацій лінійного коду за твірною матрицею необхідно вміти знаходити перевірні розряди кодових комбінацій за інформаційними. Алгоритм утворення перевірних розрядів за допомогою твірної матриці $G_{(n,\,k)}$ за відомим інформаційним має вигляд

$$b_1 \quad b_{11}a_1 \oplus b_{21}a_2 \oplus \dots \oplus b_{k1}a_k;\ b_2 = b_{12}a_1 \oplus b_{22}a_2 \oplus \dots \oplus b_{k2}a_k,$$
$$\dots,\ b_r = b_{1r}a_1 \oplus b_{2r}a_2 \oplus \dots \oplus b_{kr}a_k.$$

Система перевірок для кожної конкретної матриці $G_{(n,\,k)}$ формулюється так: у першу перевірку разом із перевірним розря-

дом b_1 входять інформаційні розряди, що відповідають одиницям першого стовпця підматриці $C_{(r,k)}$; у другу — перевірний розряд b_2 та інформаційні розряди, які відповідають одиницям другого стовпця підматриці $C_{(r,k)}$ і т. д. Кількість перевірок дорівнює кількості перевірних розрядів коду, тобто кількості стовпців підматриці $C_{(r,k)}$.

Система перевірок для наведеної вище твірної матриці $G_{(7,4)}$ має вигляд $b_1 = a_2 \oplus a_3 \oplus a_4$, $b_2 = a_1 \oplus a_3 \oplus a_4$, $b_3 = a_1 \oplus a_2 \oplus a_4$.

Опираючись на твірну матрицю, можна побудувати перевірну матрицю H, яка містить r рядків, n стовпців і складається з двох підматриць: $D_{(k,r)}$, що має k стовпців і r рядків, кожний рядок якої відповідає транспонованому стовпцю перевірних розрядів підматриці $C_{(r,k)}$ твірної матриці $G_{(n,k)}$; одиничної підматриці E_r:

$$H_{(n,r)} = [D_{(k,r)}; E_r] = \begin{bmatrix} b_{11} & b_{21} & \dots & b_{k1} & 1\,0\,0\,\dots\,0 \\ b_{12} & b_{22} & \dots & b_{k2} & 0\,1\,0\,\dots\,0 \\ \dots & \dots & \dots & \dots & \dots\dots\dots \\ b_{1r} & b_{2r} & \dots & b_{kr} & 0\,0\,0\,\dots\,1 \end{bmatrix}. \quad (8.4)$$

Ця перевірна матриця дає змогу спростити операції кодування та декодування. Позиції, зайняті одиницями в i-му рядку підматриці $D_{(k,r)}$, визначають ті інформаційні розряди, які мають брати участь у формуванні i-го перевірного розряду.

Кількість розрядів кодового синдрому і , як наслідок, кількість перевірних елементів при виправленні однократних помилок визначається нерівністю $2^r - 1 \geq n$ або $2^r - 1 \geq C_n^1$. Для виправлення не тільки однократних, а й помилок більшої кратності необхідно виконати умову

$$2^r - 1 \geq C_n^1 + C_n^2 + \dots + C_n^{\nu_{вп}},$$

де $\nu_{вп}$ — кратність виправленої помилки.

Як приклад перетворимо твірну матрицю (8. 3) на перевірну

$$H_{(7,3)} = \begin{bmatrix} 0\,1\,1\,1 & 1\,0\,0 \\ 1\,0\,1\,1 & 0\,1\,0 \\ 1\,1\,0\,1 & 0\,0\,1 \end{bmatrix}. \quad (8.5)$$

Перевірними елементами, що визначаються за допомогою матриці, будуть $b_1 = a_2 \oplus a_3 \oplus a_4$, $b_2 = a_1 \oplus a_3 \oplus a_4$, $b_3 = a_1 \oplus a_2 \oplus a_4$. Повідомлення, наприклад 0110, закодоване таким кодом, матиме вигляд 0110011, оскільки $b_1 = 1 \oplus 1 \oplus 0 = 0$, $b_2 = 0 \oplus 1 \oplus 0 = 1$, $b_3 = 0 \oplus 1 \oplus 0 = 1$.

Методи виправлення помилок у систематичному груповому коді. Розрізняють два основних таких методи: за допомогою кодового синдрому та за допомогою кодів-супутників.

Використання кодового синдрому. Перевірка кодової комбінації, що приймається, при цьому виконується зіставленням її перевірних розрядів з перевірними розрядами, які обчислюються на основі прийнятих інформаційних.

Кодові комбінації , що передаються та приймаються, запишемо відповідно у вигляді

$$V = (a_1, \dots, a_k, b_1, \dots, b_r);\ V' = (a_1', \dots, a_k', b_1', \dots, b_r').$$

З метою виявлення помилки в кодовій комбінації, що приймається, зробимо перевірки

$$b_j'' = \sum_{i=1}^{k} O\alpha_{ji} a_i' \quad (8.6)$$

та порівняємо ці значення зі значеннями прийнятих перевірних елементів b_j'. У результаті порівняння дістанемо r-елементний набір $(S_1, \dots, S_r)$, що визначається як сума за модулем 2 двох r-елементних наборів:

$$(b_1'', \dots, b_r'') \ E \ (b_1', \dots, b_r') = (S_1, \dots, S_r). \quad (8.7)$$

Після підстановки (8.6) у (8.7) знайдемо значення кожного елемента S_j в наборі (8.7) :

$$S_j = \sum_{i=1}^{k} O\alpha_{ji} a_i' + b_j'. \quad (8.8)$$

Набір елементів $(S_1, \dots, S_r)$ називається *синдромом* (пізнавачем) помилок.

За відсутності помилки комбінація синдрому складається з одних нулів. Присутність хоча б одного ненульового елемента в комбінації синдрому вказує на спотворення елемента у прийнятій кодовій комбінації.

Між комбінацією синдрому та помилковою комбінацією, що спричинила створення синдрому, немає взаємної однозначної відповідності, тому за кожним синдромом закріплюється така комбінація помилок, що виправляється, виникнення якої в каналі є найімовірнішим.

Значення синдрому збігаються з комбінацією результатів перевірки на парність, які визначаються перевірною матрицею H. Так, синдром для систематичного (7,4)-коду з перевірною матрицею (8.5) визначається виразами $S_1 = b_1' \oplus b_1''$, $S_2 = b_2' \oplus b_2''$, $S_3 = b_3' \oplus b_3''$, де b_i' – i-й перевірний елемент прийнятої кодової комбінації, b_i'' – i-й перевірний елемент, обчислений за

інформаційними елементами на приймальному боці. Для (7,4)-коду маємо

$$b_1'' = a_2' \oplus a_3' \oplus a_4',\ b_2'' = a_1' \oplus a_3' \oplus a_4',\ b_3'' = \\ = a_1' \oplus a_2' \oplus a_4'.$$

Якщо значення b_i'' підставити у вираз S_i, то дістанемо

$$a_2' \oplus a_3' \oplus a_4' \oplus b_1' = S_1;\\ a_1' \oplus a_3' \oplus a_4' \oplus b_2' = S_2;\\ a_1' \oplus a_2' \oplus a_4' \oplus b_3' = S_3.$$

Установимо відповідність виду синдрому з видом помилки, що виправляється, для перевірної матриці H (табл. 8.2).

Якщо при передачі по каналах у кодовій комбінації виникне одноразова помилка в одному з розрядів, то обчислений на приймальному боці синдром укаже номер спотвореного розряду.

Побудова синдромів для виправлення подвійних, потрійних і помилок більшої кратності складна. Тому метод виправлення помилок за допомогою кодового синдрому використовується головним чином для виправлення поодиноких помилок, імовірність виникнення яких значно більша, ніж помилок більшої кратності.

Розглянутий вище лінійний систематичний груповий код належить до досконалих кодів [8, 12]. *Досконалим лінійним кодом* називається такий оптимальний за $d_{\min}$ лінійний код, у якого кількість ненульових комбінацій синдрому $(2^r - 1)$ дорівнює кількості всіх можливих комбінацій помилок із вагою $v_{вп}$ й менше, а кожна з $(2^r - 1)$ груп із 2^k комбінаціями помилок, які утворюють ненульові комбінації синдрому, має тільки одну із зазначених комбінацій помилок. Так, лінійний (7,4)-код, який задається перевірною матрицею (8,5), є досконалим лінійним кодом, тому що при $d_{\min} = 3$ та здатності виправлення поодиноких помилок кількість ненульових комбінацій синдрому $(2^3 - 1 = 7)$ дорівнює кількості всіх можливих комбінацій однократних помилок (див. табл. 8.2).

Таблиця 8.2

Комбінація (вектор) помилки	Номер спотвореного розряду	Синдром	Комбінація (вектор) помилки	Номер спотвореного розряду	Синдром
1000000	1	011	0000100	5	100
0100000	2	101	0000010	6	010
0010000	3	110	0000001	7	001
0001000	4	111			

Використання кодів-супутників. Цей метод передбачає побудову кодової таблиці з кодами-супутниками за таким правилом (табл. 8.3): в першому її рядку розташовують усі кодові комбінації V_i, для яких необхідно знайти коди-супутники; в другому рядку записують вектори, утворені підсумовуванням за модулем 2 кодових комбінацій V_i з вектором помилки e_1, вага якого $w = 1$, а одиниця знаходиться в першому розряді; третій рядок є результатом підсумовування за модулем 2 кодових комбінацій V_i з вектором e_2, вага якого $w = 1$, а одиниця розміщується в другому розряді, і т. д. Так діють доти, доки не будуть підсумовані з кодовими комбінаціями V_i всі вектори e_i вагою $w = 1$ з одиницями в кожному з n розрядів, потім підсумовують за модулем 2 вектори e_i вагою $w = 2$ з послідовним перекриванням усіх можливих розрядів (вага векторів e_i визначає кількість помилок, що виправляються, а розрядність цих векторів відповідає розрядності кодової комбінації).

Таким чином, для кожної кодової комбінації V_i лінійного систематичного коду дістають свою групу кодів-супутників, розташованих у відповідному стовпці (див. табл. 8.3), які зберігаються в пам'яті ЕОМ. У разі приймання комбінації, що збігається з одним із кодів-супутників, спотворена комбінація розшифровується як початкова робоча комбінація, до якої належить цей код-супутник.

Недоліками розглянутого методу є велика ємність пам'яті ЕОМ і значні затрати часу на перебір комбінацій кодів-супутників.

Укорочені лінійні систематичні коди. Розрізняють повні та укорочені лінійні систематичні коди. До перших належать лінійні (n, k)-коди, що містять 2^k комбінацій, а до других — лінійні (n, k, i) коди (де $i = 1 \dots k - 1$), які мають 2^{k-i} комбінацій. Укорочений лінійний систематичний код утворюють з повного коду, при цьому дістають кодову відстань $d_{\min}$, не меншу ніж у почат-

Таблиця 8.3

e_i	V_i				
	V_1	V_2	V_3	…	$V_{(2^k-1)}$
e_1	$e_1 \oplus V_1$	$e_1 \oplus V_2$	$e_1 \oplus V_3$	…	$e_1 \oplus V_{(2^k-1)}$
e_2	$e_2 \oplus V_1$	$e_2 \oplus V_2$	$e_2 \oplus V_2$	…	$e_2 \oplus V_{(2^k-1)}$
		—	—	—	—
		—	—	—	—
$e_{(2^n-1)}$	$e_{(2^n-1)} \oplus V_1$	$e_{(2^n-1)} \oplus V_2$	$e_{(2^n-1)} \oplus V_3$	…	$e_{(2^n-1)} \oplus V_{(2^k-1)}$

кового лінійного систематичного коду зі збереженням такої самої кількості перевірних елементів, тобто $r = (n - i) - (k - i) = n - k$.

Здобуття вкороченого лінійного систематичного коду ґрунтується на тому, що оскільки з загальної кількості 2^k комбінацій первинного коду, які потрібно закодувати лінійним систематичним кодом, 2^{k-1} комбінацій починаються з 0, а 2^{k-1} — з 1, після кодування лінійного (n, k)-коду 2^{k-1} його комбінацій також починатимуться з 0.

Ці 2^{k-1} комбінацій розглядатимемо як новий лінійний систематичний код. Тоді, через те що вони належать початковому (n, k)-коду, кодова відстань $d_{\min}$ нового коду буде не меншою, ніж початкового. Нульовий символ на початку кодової комбінації зберігається й після кодування її лінійним (n, k)-кодом, причому в утворенні перевірних елементів лінійного систематичного коду участі він не бере. З урахуванням цього нульовий символ можна відкинути. При цьому дістанемо лінійний $(n-1, k-1)$-код, тобто код, що містить $n-1$ елементів, з яких $k-1$ є інформаційними, а r — перевірними.

Укорочений лінійний систематичний код легко дістати з повного лінійного (n, k)-коду, що подається у вигляді твірної матриці $G_{(n, k)}$ розміром $k \times n$ виду (8.2) виключенням з матриці перших рядка та стовпця. В результаті дістанемо твірну матрицю $G_{(n-1, k-1)}$ нового вкороченого лінійного систематичного коду розміром $(k-1) \times (n-1)$, що утворює 2^{k-1} його комбінацій. Для здобуття перевірної матриці $H_{(n-1, r)}$ цього коду досить виключити перший стовпець із матриці $H_{(n, r)}$ відповідного повного коду.

Так, твірну й перевірну матриці вкороченого лінійного систематичного (6, 3)-коду можна дістати з відповідних матриць $G_{(7, 4)}$ (8.3) та $H_{(7, 3)}$ (8.5) повного коду:

$$G_{(6,3)} = \begin{bmatrix} 1\,0\,0 & 1\,0\,1 \\ 0\,1\,0 & 1\,1\,0 \\ 0\,0\,1 & 1\,1\,1 \end{bmatrix}; \qquad H_{(6,3)} = \begin{bmatrix} 1\,1\,1 & 1\,0\,0 \\ 0\,1\,1 & 0\,1\,0 \\ 1\,0\,1 & 0\,0\,1 \end{bmatrix}. \qquad (8.9)$$

Після вкорочення на один символ лінійного систематичного $(n-1, k-1)$-коду здобудемо вкорочений $(n-2, k-1)$- код і т. д.

8.1.2. КОДИ ХЕММІНГА

Це одні з найпоширеніших систематичних кодів, які виправляють помилки. До кодів Хеммінга належать коди з мінімальною кодовою відстанню $d_{\min} = 3$, що виправляють всі поодинокі помилки.

Формування r перевірних елементів у комбінаціях цих кодів виконують за k інформаційними елементами. Таким чином, довжина кодової комбінації $n = k + r$. Перевірними елементами є лінійні комбінації інформаційних елементів, тобто зважені суми інформаційних елементів з ваговими коефіцієнтами 1 та 0.

Послідовність одиниць і нулів у кодовій комбінації називають ще *кодовим вектором*. Кодам Хеммінга притаманні властивості лінійних кодів: сума (різниця) векторів лінійного коду дає вектор, який належить цьому коду; лінійні коди утворюють алгебричну групу відносно операції додавання за модулем 2; мінімальна кодова відстань між векторами групового коду дорівнює мінімальній вазі ненульових кодових векторів.

При передачі кодового вектора може бути спотворений будь-який елемент, кількість таких ситуацій $C_n^1 = n$. До цього слід додати ще одну ситуацію, коли помилка не виникає. Таким чином, загальна кількість 2^r комбінацій перевірних елементів має перевищувати кількість можливих помилкових ситуацій в коді з урахуванням відсутності помилок для правильного розпізнання їх і визначення місць помилки:

$$2^r \geq n + 1. \qquad (8.10)$$

Оскільки $2^n = 2^{k+r} = 2^k \cdot 2^r$, можна записати

$$2^n \geq (n + 1) \cdot 2^k, \qquad (8.11)$$

де 2^n — повна кількість комбінацій коду.

Мінімальне співвідношення коректувальних та інформаційних розрядів, нижче якого код не може зберігати задані коректувальні властивості, визначається виразом

$$2^r - 1 = n.$$

Для розрахунку основних параметрів кодів Хеммінга можна задати кількість перевірних елементів r; тоді з останнього виразу визначається n, а кількість інформаційних елементів $k = n - r$. Співвідношення між r, n і k для кодів Хеммінга наведено в табл. 8.4.

Характерна особливість перевірної матриці коду з $d_{\min} = 3$ полягає у тому, що її стовпці є різними ненульовими комбіна-

Таблиця 8.4

k	1	1	2	3	4	4	5	6	7	8	9	10	11	11
r	2	3	3	3	3	4	4	4	4	4	4	4	4	5
n	3	4	5	6	7	8	9	10	11	12	13	14	15	16

ціями завдовжки r. Наприклад, при $r = 4$, $n = 15$ перевірна матриця (15, 11)-коду може мати такий вигляд:

$$H_{(15,4)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 & u_8 & u_9 & u_{10} & u_{11} & u_{12} & u_{13} & u_{14} & u_{15} \end{bmatrix}. \quad (8.12)$$

Таким чином, якщо взяти комбінації чотириелементного двійкового простого коду й відкинути ненульову комбінацію, можна досить легко дістати перевірну матрицю, записавши всі кодові комбінації послідовно в стовпці матриці H.

Після переставлення стовпців, які мають одну одиницю, матриця (8.12) набуває вигляду

$$H_{(15,4)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 & a_8 & a_9 & a_{10} & a_{11} & b_1 & b_2 & b_3 & b_4 \end{bmatrix}. \quad (8.13)$$

Згідно з матрицею (8.13) дістаємо систему перевірних рівнянь, за допомогою яких знаходимо перевірні розряди:

$$\left.\begin{aligned} b_1 &= a_5 \oplus a_6 \oplus a_7 \oplus a_8 \oplus a_9 \oplus a_{10} \oplus a_{11}; \\ b_2 &= a_2 \oplus a_3 \oplus a_4 \oplus a_8 \oplus a_9 \oplus a_{10} \oplus a_{11}; \\ b_3 &= a_1 \oplus a_3 \oplus a_4 \oplus a_6 \oplus a_7 \oplus a_{10} \oplus a_{11}; \\ b_4 &= a_1 \oplus a_2 \oplus a_4 \oplus a_5 \oplus a_7 \oplus a_9 \oplus a_{11}. \end{aligned}\right\} \quad (8.14)$$

Поява помилки в кодовій комбінації веде до невиконання тих перевірних співвідношень (8.14), в які входить значення помилкового розряду.

Так, якщо помилка виникла в сьомому інформаційному розряді (спотвореним є елемент a_7), то не виконуються перше, третє та четверте співвідношення (8.14), тобто синдром дорівнює 1011 [збігається з сьомим стовпцем матриці H (8.13)]. Таким чином, місцезнаходження стовпця матриці H, що збігається зі знайденим синдромом, визначає місце помилки.

Обчислене значення синдрому обов'язково збігається з одним із стовпців матриці H, тому що як стовпці вибираються всі можливі r-розрядні двійкові комбінації. Р. Хеммінг запропонував розташувати стовпці перевірної матриці так, щоб номер i-го

овпця матриці H і номер розряду кодової комбінації відповідали двійковому поданню числа i. Тоді синдром, знайдений з перевірних рівнянь, буде двійковим поданням номера розряду кодової комбінації, в якій виникла помилка. Для цього перевірні розряди мають знаходитися не в кінці кодової комбінації, а на номерах позицій, які подаються степенем двійки (2^1, 2^2, 2^3, ... ,2^{r-1}), як у матриці (8.12), тому що кожний з них входить тільки до одного з перевірних рівнянь. В останньому випадку перевірні розряди розміщуються між інформаційними.

Синдром відповідно до перевірної матриці (8.12) визначається з системи рівнянь

$$\left.\begin{aligned} S_1 &= u_1 \oplus u_3 \oplus u_5 \oplus u_7 \oplus u_9 \oplus u_{11} \oplus u_{13} \oplus u_{15}; \\ S_2 &= u_2 \oplus u_3 \oplus u_6 \oplus u_7 \oplus u_{10} \oplus u_{11} \oplus u_{14} \oplus u_{15}; \\ S_3 &= u_4 \oplus u_5 \oplus u_6 \oplus u_7 \oplus u_{12} \oplus u_{13} \oplus u_{14} \oplus u_{15}; \\ S_4 &= u_8 \oplus u_9 \oplus u_{10} \oplus u_{11} \oplus u_{12} \oplus u_{13} \oplus u_{14} \oplus u_{15}. \end{aligned}\right\} \quad (8.15)$$

Як перевірні вибираються розряди u_1, u_2, u_4 та u_8, що зустрічаються в системі рівнянь (8.15) по одному разу.

Наприклад, якщо необхідно закодувати повідомлення 11001010110 двійкового простого коду ($k = 11$) у коді Хеммінга, то потрібно визначити перевірні розряди в комбінації $u_1u_21u_4100u_81010110$.

З перевірної матриці (8.12) маємо

$$\left.\begin{aligned} u_1 &= u_3 \oplus u_5 \oplus u_7 \oplus u_9 \oplus u_{11} \oplus u_{13} \oplus u_{15}; \\ u_2 &= u_3 \oplus u_6 \oplus u_7 \oplus u_{10} \oplus u_{11} \oplus u_{14} \oplus u_{15}; \\ u_4 &= u_5 \oplus u_6 \oplus u_7 \oplus u_{12} \oplus u_{13} \oplus u_{14} \oplus u_{15}; \\ u_8 &= u_9 \oplus u_{10} \oplus u_{11} \oplus u_{12} \oplus u_{13} \oplus u_{14} \oplus u_{15}. \end{aligned}\right\} \quad (8.16)$$

Обчислюючи u_1, u_2, u_4 та u_8 згідно з (8.16), дістаємо

$$\begin{aligned} u_1 &= 1 \oplus 1 \oplus 0 \oplus 1 \oplus 1 \oplus 1 \oplus 0 = 1; \\ u_2 &= 1 \oplus 0 \oplus 0 \oplus 0 \oplus 1 \oplus 1 \oplus 0 = 1; \\ u_4 &= 1 \oplus 0 \oplus 0 \oplus 0 \oplus 1 \oplus 1 \oplus 0 = 1; \\ u_8 &= 1 \oplus 0 \oplus 1 \oplus 0 \oplus 1 \oplus 1 \oplus 0 = 0. \end{aligned}$$

комбінація коду Хеммінга для розглядуваного повідомлення вигляд 111110001010110.

припустимо, що шостий елемент цієї комбінації приймається тобто одержано повідомлення 11111 1 001010110. Обчислюючи за допомогою системи рівнянь (8.15), знаходимо $S_1 = 0$; $S_2 = 1$; $S_4 = 0$, тобто синдром має вигляд 0110. Якщо це двійкове перевести в десяткове, то дістанемо $6(0 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 0 \cdot 2^0 =$ Таким чином, необхідно виправити шостий розряд повідомлення, одержано.

Розширений код Хеммінга. Код Хеммінга з кодовою відстанню $d_{min} = 4$ називається *розширеним*. Він забезпечує виправлення всіх однократних і виявлення всіх дво- та трикратних помилок. Для цього вводиться додатковий перевірний розряд b_0, який дописується до перевірної матриці Хеммінга з кодовою відстанню $d_{min} = 3$, завдяки чому остання збільшується до 4.

Додатковий перевірний розряд займає останній стовпець одиничної підматриці перевірної матриці. Крім того, збільшення кількості перевірних розрядів у комбінації коду Хеммінга веде до зростання кількості рядків перевірної матриці. Додатковий рядок утворюється доповненням стовпців перевірної матриці до непарності, як це показано на прикладі перетворення перевірної матриці (8.13) коду Хеммінга з $d_{min} = 3$ на перевірну матрицю коду з $d_{min} = 4$:

$$H_{(16,5)} = \left[\begin{array}{cccccccccccccccc} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{array}\right] \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{array}.$$

$$a_1\ a_2\ a_3\ a_4\ a_5\ a_6\ a_7\ a_8\ a_9\ a_{10}\ a_{11}\ b_1\ b_2\ b_3\ b_4\ b_0$$

На практиці доцільніше застосовувати інший метод утворення розширеного коду Хеммінга з коду Хеммінга, в якого кодова відстань $d_{min} = 3$. Для цього кодову комбінацію останнього просто доповнюють додатковим перевірним елементом b_0, який знаходять за допомогою перевірки кодової комбінації на парність. При цьому перевірний елемент має дорівнювати одиниці, якщо кількість одиниць в закодованій комбінації непарна, й нулю, якщо ця кількість парна.

Розширений код Хеммінга декодують у зворотній послідовності: спочатку виконують загальну перевірку прийнятої кодової комбінації, а потім — її перевірку без елемента b_0. При цьому можуть виникнути такі варіанти:

1) помилок немає (дві перевірки — загальна й без елемента b_0 дають нульові кодові синдроми);

2) є однократна помилка (загальна перевірка свідчить про наявність помилки — кодовий синдром не дорівнює нулю, а перевірка без елемента b_0 дає синдром, який вказує номер спотвореного елемента);

3) є двократна помилка (загальна перевірка свідчить про відсутність помилки — кодовий синдром дорівнює нулю, а перевірка без елемента b_0 дає синдром, який вказує номер позиції, де нібито помилка виникла, проте її виправляти не слід — треба тільки констатувати наявність двох помилок);

4) є трикратна помилка (загальна перевірка свідчить про наявність помилки — кодовий синдром не дорівнює нулю, а перевірка без елемента b_0 дає синдром, що може набувати будь-якого значення, в тому числі й нульового).

З урахуванням викладеного коди Хеммінга з $d_{min} = 4$ використовуються, як правило, для виявлення одно-, дво- та трикратних помилок.

Для утворення вкороченого коду Хеммінга з мінімальною кодовою відстанню $d_{min} = 3$ або 4 керуються правилами, розглянутими при формуванні аналогічних лінійних систематичних кодів.

8.1.3. ЦИКЛІЧНІ КОДИ

Ці коди також широко застосовуються для захисту інформації від помилок. Подання комбінацій в циклічних кодах використовують у вигляді поліномів формальної змінної x, що дає змогу [illegible] дії над кодовими комбінаціями до дій над поліномами (с. п. 5.3).

Лінійні систематичні (n, k)-коди, в яких циклічний зсув a_{n-2}, $a_{n-3}, a_{n-4}, \ldots, a_2, a_1, a_0, a_{n-1}$ дозволеної комбінації $a_{n-1}, a_{n-2}, a_{n-3}, \ldots$, a_1, a_0 також є дозволеною комбінацією, що належить цьому коду, називаються *циклічними*. Така циклічна перестановка елементів виникає після множення полінома на x. Якщо $V(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \ldots + a_1x + a_0$, то $xV(x) = a_{n-1}x^n + a_{n-2}x^{n-1} + \ldots \ldots + a_1x^2 + a_0x$. Щоб степінь полінома не перевищував $n-1$, член $a_{n-1}x^n$ замінюється одиницею. Тому

$$xV(x) = F(x) = a_{n-2}x^{n-1} + \ldots + a_1x^2 + a_0x + a_{n-1}.$$

Особливу роль у теорії циклічних кодів відіграють *твірні поліноми*. Помічено, що комбінації лінійного коду мають властивість циклічності, коли як твірні використовуються поліноми, що є дільниками двочлена $x^n + 1$. Кожний такий двочлен може бути розкладений на кілька незвідних поліномів, тобто таких, які можуть бути подані у вигляді добутку поліномів нижчих степенів (вони діляться самі на себе або на одиницю).

Як твірні поліноми різних циклічних кодів можуть бути застосовані всі незвідні поліноми та добутки їх, тому що вони також є дільниками двочлена $x^n + 1$.

З урахуванням викладеного можна дати ще одне визначення лінійного циклічного коду [40]: *циклічним* називається лінійний систематичний (n, k)-код, всі 2^k комбінацій якого подано поліномами степеня $x - 1$ і менше, що діляться на деякий поліном $P(x)$ степеня $r = n - k$, який є дільником двочлена $x^n + 1$. Деякі твірні поліноми наведено в табл. 8.5.

Таблиця 8.5

r	Твірний поліном $P(x)$	Двійковий запис полінома
1	$x + 1$	11
2	$x^2 + x + 1$	111
3	$x^3 + x + 1$	1011
4	$x^3 + x^2 + 1$	1101
4	$x^4 + x + 1$	10011
4	$x^4 + x^3 + 1$	11001
4	$x^4 + x^3 + x^2 + x + 1$	11111
5	$x^5 + x^2 + 1$	100101
5	$x^5 + x^3 + 1$	101001
5	$x^5 + x^3 + x^2 + x + 1$	101111
5	$x^5 + x^4 + x^2 + x + 1$	110111
6	$x^6 + x^5 + x^4 + 1$	1110001
8	$x^8 + x^7 + x^6 + x^5 + x^2 + x + 1$	111100111
9	$x^9 + x^5 + x^3 + 1$	1000101001
15	$x^{15} + x^{14} + x^{13} + x^{12} + x^4 + x^3 + x^2 + x + 1$	1111000000011111
16	$x^{16} + x^{12} + x^5 + 1$	10001000000100001

Розрізняють *алгебричні* та *матричні* методи побудови циклічного коду. Побудова дозволеної кодової комбінації (алгоритм кодування) перших зводиться ось до чого:

- подати інформаційну частину з k елементів у вигляді полінома $Q(x)$ степеня $k - 1$;
- помножити $Q(x)$ на x^r (що еквівалентно зсуву k-розрядної кодової комбінації на r розрядів);
- поділити поліном $x^rQ(x)$ на вибраний твірний поліном $P(x)$, степінь якого дорівнює r, і визначити остачу від ділення $R(x)$, тобто

$$\frac{x^rQ(x)}{P(x)} = C(x) \oplus \frac{R(x)}{P(x)}, \tag{8.17}$$

де $C(x)$ — частка від ділення, яка має той самий степінь, що й поліном $Q(x)$; $R(x)$ — остача від ділення, яка має степінь, не більший від $r - 1$ [менший, ніж степінь дільника $P(x)$].

Зазначимо, що r розрядів остачі є r перевірними елементами кодової комбінації, причому

$$x^rQ(x) = C(x)P(x) \oplus R(x), \tag{8.18}$$

або

$$F(x) = C(x)P(x) = x^rQ(x) \oplus R(x), \tag{8.19}$$

де $F(x)$ — комбінація циклічного коду.

З виразу (8.19) випливають два рівноцінних алгебричних методи побудови комбінацій циклічного коду:

$$F_1(x) = x^rQ(x) \oplus R(x); \tag{8.20}$$

$$F_2(x) = C(x)P(x). \tag{8.21}$$

Ще один метод можна дістати, замінивши в (8.21) частку від ділення $C(x)$ на поліном $Q(x)$ кодової комбінації r-елементного двійкового простого коду, що подається для кодування циклічним кодом. Ця заміна цілком слушна, оскільки поліноми $C(x)$ і $Q(x)$ мають однаковий найбільший степінь (однакову кількість розрядів). Отже,

$$F_3(x) = Q(x)P(x). \tag{8.22}$$

У комбінаціях циклічних кодів, побудованих за першим і другим методами [див. вирази (8.20) і (8.21)], розташування інформаційних і перевірних елементів підпорядковується такому правилу: k старших розрядів комбінації є інформаційними, решта $n - k = r$ розрядів — перевірними.

При використанні третього методу побудови циклічних кодів [див. вираз (8.22)] дістають комбінації неподільного циклічного коду, в яких інформаційні та перевірні елементи не відокремлені один від одного, що ускладнює процес декодування. Тому на практиці найпоширенішими є перші два алгебричні методи побудови циклічного коду.

Як видно, що комбінація $F(x)$ має ділитися на поліном $P(x)$ без остачі. На цьому й ґрунтується перевірка комбінації на наявність помилок при її прийманні. Якщо прийнята комбінація $F(x)$ ділиться на поліном $P(x)$ без остачі, то вона визнається безпомилковою.

Розглядають два матричних методи побудови циклічного коду на основі твірної матриці $G_ц$. Перший з них ґрунтується на відомому виразі (8.17). Відповідний рядок твірної матриці записується у вигляді $x^rQ(x) + R_i(x)$. При цьому вся матриця розбивається на дві підматриці, як і в лінійному систематичному груповому коді (див. п.8.11): $G_ц = [E_k{}^T, C_{r,k}]$, де $E_k{}^T$ — транспонована одинична інформаційна підматриця; $C_{r,k}$ — перевірна підматриця з кількістю стовпців r і рядків k, що утворюється остачами від ділення $R_i(x)$. Такий метод дає можливість дістати твірну матрицю в канонічному вигляді.

Для побудови твірної матриці (формування її рядків) беруть не довільні комбінації $Q(x)$ двійкового простого коду, а тільки ті з них, які містять одиницю в одному розряді $Q_i(x)$, де $i = 1, 2, \ldots, k$. Ці комбінації множать на x^r і знаходять остачу від ділення $Q_i(x)x^r/P(x)$, що дорівнює $R_i(x)$.

Так, для циклічного (7, 4)-коду з $n = 7$, $k = 4$ і твірним поліномом $P(x) = x^3 + x^2 + 1$, щоб побудувати твірну матрицю, вибирають чотириелементні одиничні комбінації Q_i двійкового простого коду: $Q_1(x) = 1$ (0001); $Q_2(x) = x$ (0010); $Q_3(x) = x^2$ (0100); $Q_4(x) = x^3$ (1000). Вибрані комбінації $Q_i(x)$ множать на x^3 ($r = n - k = 7 - 4 = 3$), ділять на $P(x) = x^3 + x^2 + 1$ і знаходять остачі:

$$\frac{1x^3}{x^3+x^2+1} \to R_1(x) = x^2 + 1 \to 101;$$

$$\frac{xx^3}{x^3+x^2+1} \to R_2(x) = x^2 + x + 1 \to 111;$$

$$\frac{x^2x^3}{x^3+x^2+1} \to R_3(x) = x + 1 \to 011;$$

$$\frac{x^3x^3}{x^3+x^2+1} \to R_4(x) = x^2 + x \to 110;$$

Таким чином, твірна матриця для розглядуваного прикладу матиме вигляд

$$G_{\text{ц}(7,4)} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{E_k^T} \quad \underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}}_{C_{r,k}} \tag{8.23}$$

Твірна матриця має змогу дістати k комбінацій коду. Решту $2^k - k - 1$ комбінацій, крім нульової, знаходять додаванням за модулем 2 рядків твірної матриці в усіх можливих сполученнях. Остання комбінація коду — нульова.

За аналогією з лінійним систематичним груповим (n, k)-кодом твірну матрицю $G_{\text{ц}(7,4)}$ (8.23) можна перетворити на перевірну матрицю $H_{\text{ц}}$ циклічного коду, яка має такий вигляд:

$$H_{\text{ц}(7,3)} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ a_1 & a_2 & a_3 & a_4 & b_1 & b_2 & b_3 \end{bmatrix}. \tag{8.24}$$

Перевірними елементами, що визначає ця матриця, будуть

$$b_1 = a_1 \oplus a_3 \oplus a_4;\ b_2 = a_1 \oplus a_2 \oplus a_3;\ b_3 = a_2 \oplus a_3 \oplus a_4.$$

Інший матричний метод побудови циклічного коду ґрунтується на твірному поліномі $P(x)$. За цим методом при побудові твірної матриці $G_{\text{ц}}$ як її рядки беруть k лінійно незалежних комбінацій, що відповідають поліномам $x^0P(x), x^1P(x), \ldots, x^{k-1}P(x)$ [40]:

$$G_{\text{ц}} = \begin{bmatrix} x^0P(x) \\ x^1P(x) \\ \ldots\ldots\ldots \\ x^{k-1}P(x) \end{bmatrix}. \tag{8.25}$$

Додаванням по два, три і т. д. до k рядків за модулем 2 дістають усі $2^k - k - 1$ ненульові комбінації циклічного коду, що залишилися.

Так, якщо задатися твірним поліномом $P(x) = x^3 + x^2 + 1$ (див. табл. 8.5), можна побудувати циклічний (7,4)-код, який виправляє однократні помилки.

Як $k = 4$ рядки матриці $G_{\text{ц}}$ використовуються поліноми

$$x^0P(x) = x^3 + x^2 + 1;\ x^1P(x) = x^4 + x^3 + x;\ x^2P(x) = x^5 + x^4 + x^2;$$
$$x^3P(x) = x^6 + x^5 + x^3,$$

тобто

$$G_{\text{ц}(7,4)} = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{8.26}$$

$2^4 - 2^4 = 16$ комбінацій циклічного (7,4)-коду, що описується ...ірною матрицею, перша комбінація — нульова, друга — ... с рядками твірної матриці, а решта 11 — сумами за ...м 2 всіх можливих сполучень (комбінацій) рядків твір... риці.

...ну матрицю (8.26) можна звести до вигляду твірної мат... $_{(7,4)}$ (8.3) лінійного систематичного групового коду, для ...конаємо деякі нескладні операції: переставимо перший ...ртий, другий та третій рядки, потім додамо до утворено... ...ого рядка другий і третій, до другого рядка — третій та ...ий, до третього рядка — четвертий, а четвертий рядок ...шемо без змін. У результаті матимемо

$$G_{\text{ц}(7,4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}. \tag{8.27}$$

Звідси за аналогією з лінійним систематичним груповим (n, k)-кодом цю матрицю легко перетворити на перевірну матрицю (8.24) циклічного коду та знайти перевідні елементи:

$$b_1 = a_1 \oplus a_3 \oplus a_4;\ b_2 = a_1 \oplus a_2 \oplus a_3;\ b_3 = a_2 \oplus a_3 \oplus a_4.$$

Циклічний код можна побудувати також за допомогою перевірного полінома $H(x)$ степеня k, який дістають діленням двочлена $x^n + 1$ на твірний поліном $P(x)$, тобто

$$H(x) = (x^n + 1)/P(x).$$

Так, для циклічного (7,4)-коду, твірний поліном якого $P(x) = x^3 + x + 1$, перевірний поліном має вигляд

$$H(x) = (x^7 + 1)/(x^3 + x + 1) = x^4 + x^2 + x + 1,$$

або у двійковому запису $H(0,1) = 10111$.

Перевірна матриця при цьому будується за аналогією з твірною матрицею (8.25):

$$H_{\text{ц}} = \begin{bmatrix} x_1^0 H(x) \\ x_1^1 H(x) \\ \cdots\cdots\cdots \\ x^{r-1} H(x) \end{bmatrix}. \qquad (8.28)$$

Для розглядуваного прикладу [$P(x) = x^3 + x + 1$] перевірна матриця має такий вигляд:

$$H_{\text{ц}(7,3)} = \begin{bmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{bmatrix}. \qquad (8.29)$$

Для запису перевірних співвідношень матриця (8.29) має бути доведена до канонічного вигляду (з одиничною перевірною підматрицею), що виконується аналогічно утворенню подібних твірних матриць циклічного коду:

$$H_{\text{ц}(7,3)} = \begin{bmatrix} 1 & 0 & 1 & 1 & & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & & 0 & 0 & 1 \end{bmatrix}.$$
$$a_1\ a_2\ a_3\ a_4 \qquad b_1\ b_2\ b_3$$

Користуючись останньою матрицею, легко дістати перевірні співвідношення для кодування. Для розглядуваного прикладу

мони мають вигляд

$$b_1 = a_1 \oplus a_3 \oplus a_4;\ b_2 = a_1 \oplus a_2 \oplus a_3;\ b_3 = a_2 \oplus a_3 \oplus a_4.$$

Звідси маємо такі рівняння декодування:

$$a_1 \oplus a_3 \oplus a_4 \oplus b_1 = 0;\ a_1 \oplus a_2 \oplus a_3 \oplus b_2 = 0;\ a_2 \oplus a_3 \oplus a_4 \oplus b_3 = 0.$$

Декодування циклічних кодів. Перевірну матрицю $H_{\text{ц}}$ можна застосувати для виявлення та виправлення помилок у циклічному коді, визначивши кодовий синдром (за аналогією з лінійним систематичним груповим кодом). Однак із цією метою частіше використовуються методи, що ґрунтуються на застосуванні твірного полінома $P(x)$.

При цьому виявлення помилок при декодуванні зводиться до ділення прийнятої кодової комбінації на той самий твірний поліном, що й при кодуванні. Остача, яка називається *синдром*, коли вона має r нульових коефіцієнтів, свідчить про відсутність помилки. Якщо ж остача не нульова, то помилка є.

Оскільки поліном остачі має степінь, менший від r, а кількість степенів у ній дорівнює r, кількість різних ненульових остач може досягати $2^r - 1$. Для коду, що виправляє одну помилку ($\nu_{\text{вп}} = 1$), кількість таких остач дорівнює довжині n кодової комбінації ($2^r - 1 \geq n$) або перевищує її. Номер розряду комбінації, в якій виникла помилка, однозначно пов'язаний з вигляду остачі.

Для виправлення помилки виконують умову, за якої кількість різних нульових остач дорівнюватиме кількості елементів при кратності виправлення $\nu_{\text{вп}} = 1$) або кількості комбінацій по $\nu_{\text{вп}}$, де $\nu_{\text{вп}}$ — кількість помилок, яка виправляється коду. Це значить, що, наприклад, при $n = 15$ і $\nu_{\text{вп}} = 2$ треба $C_{15}^2 = 105$ ненульових остач для однозначного виправлення всіх будь-яких помилок у коді завдовжки n.

Для цього необхідно вибрати поліном $P(x)$ степеня $r = 7$ і побудувати код завдовжки $n = 15$ при $r = 7$ перевірних елементах (при цьому $k = 8$).

Розглянемо випадок виправлення однократної помилки (багатократні помилки виправляються аналогічно, але обсяг роботи в кілька разів більший). Визначення місця помилки в циклічному коді ґрунтується на порівнянні результатів деякої послідовності виконаної операції з синдромом помилки. Ця операція базується на тій властивості, що при діленні прийнятої з помилкою комбінації $F'(x)$ на поліномі $P(x)$ та частина $F'(x)$, яка дорівнює переданій комбінації $F(x)$, ділиться без остачі, а власне остача визначається вектором помилки $E(x)$:

$$F'(x) = F(x) \oplus E(x).$$

Оскільки $F(x) = C(x)P(x)$ згідно з (8.19), маємо

$$\frac{F'(x)}{P(x)} = C(x) \oplus \frac{E(x)}{P(x)}.$$

Знаючи заздалегідь або виконуючи послідовно операції ділення векторів помилки [чи циклічно зсунутих комбінацій $F'(x)$, що дозволяється з урахуванням кількості зсувів] на поліном $P(x)$, можна однозначно вказати місцезнаходження помилки за збігом остач.

Для виправлення помилок у прийнятій на приймальному боці комбінації $F'(x)$ циклічного коду можна скористатися різними методами. Однак найпростіше реалізується алгоритм визначення помилки, який ґрунтується на методі гіпотез, суть якого полягає ось у чому.

Як зазначено вище, за наявності, наприклад, однократної помилки в кодовій комбінації $F'(x)$ остача від ділення цієї комбінації на твірний поліном $P(x)$ не дорівнює нулю. Для знаходження місця помилки виконуємо такі операції:

К р о к 1. Будуємо гіпотезу про помилку в молодшому розряді комбінації $F'(x)$, тобто припускаємо, що вектор помилки $E_1(x) = 1$ ($E_1 = 00\ldots001$). Підсумовуємо $F'(x) \oplus E_1(x)$ і ділимо цю суму на поліном $P(x)$ із метою підтвердження (в разі нульової остачі) або спростування (в разі ненульової остачі) гіпотези. Якщо остача $R(x)$ не дорівнює нулю, то гіпотеза відкидається.

К р о к 2. Будуємо гіпотезу про помилку в другому розряді комбінації $F'(x)$, тобто припускаємо, що вектор помилки $E_2(x) = x$ ($E_2 = 00\ldots010$). Підсумовуємо $F'(x) + E_2(x)$ і ділимо цю суму на поліном $P(x)$ із метою підтвердження або спростування цієї гіпотези. При $R(x) \neq 0$ гіпотеза відкидається.

К р о к 3. Будуємо послідовно гіпотези про наявність помилки в третьому, четвертому і т. д. розрядах комбінації $F'(x)$, для чого відповідні вектори помилок $E_i(x)$ ($i = 3,4,\ldots, n$) додаємо до комбінації $F'(x)$ і результат ділимо на поліном $P(x)$ до здобуття остачі $R(x) = 0$, тобто до підтвердження гіпотези.

Остача $R(x) = 0$ свідчить про те, що помилку виправлено. Сума прийнятої комбінації $F'(x)$ з вектором помилки $E_i(x)$, яка дає остачу $R(x) = 0$, відповідає початковій комбінації $F(x)$ циклічного коду, яка передається в канал, тобто $F(x) = F'(x) \oplus E_i(x)$.

Інший поширений метод виявлення та виправлення помилок у прийнятій на приймальному боці кодовій комбінації $F'(x)$ циклічного коду після здобуття остачі $R(x) \neq 0$ від ділення прийнятої комбінації на твірний поліном, зводиться до виконання таких процедур:

• підраховується вага w остачі, тобто кількість одиниць в ній. Якщо $w \leq \nu_{вп}$, де $\nu_{вп}$ — кількість помилок, яку дає змогу випра-

ти код, то прийнята комбінація $F'(x)$ додається за модулем 2 до остачі. Сума й дає виправлену комбінацію;

• якщо $w > \nu_{вп}$, то виконується циклічний зсув прийнятої комбінації $F'(x)$ на один розряд ліворуч і здобута комбінація ділиться на твірний поліном $P(x)$. Якщо при цьому вага цієї остачі $w \leq \nu_{вп}$, то циклічно зсунута комбінація додається за модулем 2 до остачі, а потім циклічно зсувається на один розряд праворуч (повертається в попередній стан). Утворена таким чином комбінація вже не містить помилок;

якщо ж після першого циклічного зсуву й наступного ділення одержана остача, як і раніше, $w > \nu_{вп}$, то виконуються додаткові циклічні зсуви ліворуч. При цьому після кожного зсуву утворена комбінація ділиться на поліном $P(x)$ і перевіряється вага остачі. Робиться доти, доки не буде здобуто вагу остачі $w \leq \nu_{вп}$. Ця комбінація, утворена внаслідок останнього циклічного зсуву, додається за модулем 2 до остачі від ділення цієї комбінації на поліном $P(x)$, після чого виконується циклічний зсув праворуч на стільки розрядів, на скільки було зсунуто її відносно прийнятої комбінації $F'(x)$. Як результат буде здобуто виправлену комбінацію $F(x)$ циклічного коду.

Циклічні коди з кодовою відстанню $d_{min} = 4$. Вони можуть виправляти однократні та виявляти одно-, дво- і трикратні помилки.

Для збільшення кодової відстані до $d_{min} = 4$ кількість перевірних елементів у комбінації такого коду має бути на одиницю більшою, ніж у коді з $d_{min} = 3$. Твірний поліном цього коду з $d_{min} = 4$ дорівнює поліному $P(x)d_{min} = 3$, що забезпечує кодову відстань $d_{min} = 3$, помноженому на двочлен $(x + 1)$:

$$P(x)\, d_{min} = 4 = P(x)\, d_{min} = 3(x + 1). \qquad (8.30)$$

Це пояснюється тим, що двочлен $(x + 1)$ дає змогу виявляти одно- та трикратні помилки, а поліном $P(x)\, d_{min} = 3$ — дво-кратні. Так, із циклічного (7,4)-коду при $d_{min} = 3$ і твірному поліномі $P(x) = x^3 + x^2 + 1$, що розглядався вище, можна утворити циклічний код із $d_{min} = 4$, який виявлятиме трикратні помилки, якщо вибрати твірний поліном

$$P(x)\, d_{min} = 4 = (x^3 + x^2 + 1)(x + 1) = x^4 + x^2 + x + 1.$$

Загалом степінь цього полінома дорівнює кількості перевірних елементів у кодовій комбінації.

Подальша процедура кодування та декодування залишається такою самою, як і для циклічного коду з $d_{min} = 3$.

Укорочені циклічні коди. Циклічні (n, k)-коди, що містять 2^k комбінацій, називаються *повними*. З кожного повного можна

утворити вкорочений (неповний) циклічний $(n-i, k-i)$-код (де $i = 1 \ldots k-1$). Такий код має 2^{k-i} комбінацій і кодову відстань $d_{\min}$ не меншу, ніж у повного циклічного коду.

Процес укорочення повного циклічного коду полягає у виборі з 2^k його поліномів тільки частини із членами не вище x^{n-i-1}. Ці поліноми мають вигляд

$$V_y(x) = a_{n-i-1}x^{n-i-1} + \ldots + a_1 x + a_0.$$

Усі комбінації вкороченого циклічного коду діляться на твірний поліном $P(x)$ повного циклічного коду, з якого він був утворений. Однак циклічний зсув комбінацій вкороченого коду не завжди приводить до створення комбінації, що належить цьому коду, тобто даний код не має властивості циклічності. Отже, вкорочені коди не є циклічними і тому часто називаються *псевдоциклічними*.

Твірну матрицю G_y вкороченого циклічного $(n-i, k-i)$-коду дістають з матриці $G_ц$ повного циклічного коду виключенням з неї перших i рядків і стовпців, а перевірну матрицю H_y — з матриці $H_ц$ повного коду виключенням перших стовпців.

Так, твірна та перевірна матриці вкороченого циклічного (6,3)-коду можуть бути здобуті з відповідних матриць (8.23) і (8.24) циклічного (7,4)-коду з твірним поліномом $P(x) = x^3 + x^2 + 1$:

$$G_{(6,3)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}; \quad H_{(6,3)} = \begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Ґрунтуючись на твірній матриці $G_{(6,3)}$, дістаємо такі комбінації вкороченого циклічного коду: 000000, 100011, 010111, 001101, 110100, 101110, 011010, 111001. Цей код є псевдоциклічним, тому що циклічний зсув ліворуч на один елемент, наприклад комбінації 100011, дає комбінацію 111000, яка не належить до даного коду.

8.1.4. КОДИ БОУЗА — ЧОУДХУРІ — ХОКВІНГЕМА

Ці коди є різновидом циклічних кодів з кодовою відстанню $d_{\min} \geq 5$. Вони дають змогу виявляти та виправляти будь-яку кількість помилок. При кодуванні задаються кількістю помилок, яку слід виправити, або мінімальною кодовою відстанню та загальною кількістю n елементів у кодовій комбінації. Кількість інформаційних k і перевірних r елементів визначають при побудові коду Боуза — Чоудхурі — Хоквінгема (БЧХ). Розглянемо деякі правила цієї побудови.

Довжину n комбінації кодів БЧХ можна визначити так:

$$n = 2^h - 1 \text{ або } n = (2^h - 1)/g, \tag{8.31}$$

де $h > 0$ — ціле число; g — непарне додатне число, при діленні на яке n стає цілим непарним числом. Таким чином, довжина n може мати тільки непарну кількість елементів.

Керуючись (8.31), установлюємо, що n може дорівнювати 3, 7, 15, 31, 63, 127, 255, 511, 1023 розрядам і т.д.

Так, розкладаючи 2^{h-1} на співмножники, знаходимо такі значення n і g:

$2^3 - 1 = 7$; $63 = 2^6 - 1 = 7 \cdot 3 \cdot 3 = 21 \cdot 3$; $511 = 2^9 - 1 = 73 \cdot 7$;
$15 = 2^4 - 1 = 5 \cdot 3$; $127 = 2^7 - 1 = 127$; $1023 = 2^{10} - 1 = 31 \cdot 11 \cdot 3$;
$31 = 2^5 - 1 = 31$; $255 = 2^8 - 1 = 17 \cdot 5 \cdot 3$; $2047 = 2^{11} - 1 = 89 \cdot 23$.

Звідси випливає, що при $h = 6$ довжина n кодової комбінації може дорівнювати не тільки 63, а й 21 (при $g = 3$).

Кількість перевірних елементів коду визначається виразом

$$r \leq \frac{h(d-1)}{2} = [\log_2(n+1)]\frac{d-1}{2}, \tag{8.32}$$

а кількість інформаційних елементів — виразом

$$k \geq \left(2^h - 1\right) - \frac{h(d-1)}{2} \text{ або } k = n - r. \tag{8.33}$$

Твірний поліном коду БЧХ є найменшим спільним кратним (НСК) мінімальних поліномів $M_i(x)$, де $i = 1, 3, 5, \ldots, d_{\min} - 2$ — порядок полінома $P(x) = \text{НСК}\,[M_1(x)M_3(x)\ldots M_{d-2}(x)]$. Отже, кількість L мінімальних поліномів визначається кількістю помилок $\nu_{вп}$, які виправляються кодом: $L = \nu_{вп}$.

Найбільше значення степеня x мінімального полінома є найменшим цілим числом, при якому $2^l - 1$ ділиться на n або ng без остачі, тобто $n = 2^l - 1$ або $ng = 2^l - 1$. Звідси випливає, що $l = h$.

У табл. 8.6 наведено деякі мінімальні поліноми кодів БЧХ.

Степінь b твірного полінома залежить від НСК і не перевищує добуток $l\nu_{вп}$ або lL тому, що $L = \nu_{вп}$. Так, для коду БЧХ довжини $n = 15$, що виправляє $\nu_{вп} = 2$ помилки, кількість мінімальних поліномів $L = 2$, а найбільший степінь мінімального полінома l залежить від довжини n коду ($n = 2^l - 1$), тобто $l = 4$. Цьому твірний поліном $P^b(x)$, де $b = lL = 4 \cdot 2 = 8$, визначається виразом

$$P^8(x) = \text{НСК}\,[M_1{}^4(x)M_3{}^4(x)],$$

Таблиця 8.6

Номер мінімального полінома	Мінімальні поліноми різного степеня l			
	2	3	4	5
$M_1(x)$	$x^2 + x + 1$	$x^3 + x + 1$	$x^4 + x + 1$	$x^5 + x^2 + 1$
$M_3(x)$		$x^3 + x^2 + 1$	$x^4 + x^3 + x^2 + x + 1$	$x^5 + x^4 + x^3 + x^2 + 1$
$M_5(x)$			$x^2 + x + 1$	$x^5 + x^4 + x^2 + x + 1$
$M_7(x)$			$x^4 + x^3 + 1$	$x^5 + x^3 + x^2 + x + 1$
$M_9(x)$				$x^5 + x^4 + x^2 + x + 1$
$M_{11}(x)$				$x^5 + x^4 + x^3 + x + 1$
$M_{13}(x)$				

Закінчення табл. 8.6

Номер мінімального полінома	Мінімальні поліноми різного степеня l			
	6	7	8	9
$M_1(x)$	$x^6 + x + 1$	$x^7 + x^3 + 1$	$x^8 + x^4 + x^3 + x^2 + 1$	$x^9 + x^4 + 1$
$M_3(x)$	$x^4 + x^4 + x^2 + x + 1$	$x^7 + x^3 + x^2 + x + 1$	$x^8 + x^6 + x^5 + x^4 + x^2 + x + 1$	$x^9 + x^6 + x^4 + x^3 + 1$
$M_5(x)$	$x^6 + x^5 + x^2 + x + 1$	$x^7 + x^4 + x^3 + x^2 + 1$	$x^8 + x^7 + x^6 + x^5 + x^4 + x + 1$	$x^9 + x^8 + x^5 + x^4 + 1$
$M_7(x)$	$x^6 + x^3 + 1$	$x^7 + x^6 + x^5 + x^4 + x^2 + x + 1$	$x^8 + x^6 + x^5 + x^3 + 1$	$x^9 + x^7 + x^4 + x^3 + 1$
$M_9(x)$	$x^3 + x^2 + 1$	$x^7 + x^5 + x^4 + x^3 + x^2 + x + 1$	$x^8 + x^7 + x^5 + x^4 + x^3 + x^2 + 1$	$x^9 + x^8 + x^4 + x + 1$
$M_{11}(x)$	$x^6 + x^5 + x^3 + x^2 + 1$	$x^7 + x^6 + x^4 + x^2 + 1$	$x^8 + x^7 + x^6 + x^5 + x^2 + x + 1$	$x^9 + x^5 + x^3 + x^2 + 1$
$M_{13}(x)$		$x^7 + x + 1$	$x^8 + x^5 + x^3 + x + 1$	$x^9 + x^6 + x^5 + x^4 + x^2 + x + 1$

який після підстановки значень $M(x)$ набуває вигляду

$$P^8(x) = (x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1) = \\ = x^8 + x^7 + x^6 + x^4 + 1 \rightarrow 111010001.$$

Найбільший степінь твірного полінома $P(x)$ визначає кількість перевірних елементів у комбінації ($r = 8$), а кількість її інформаційних елементів $k = n - r = 15 - 8 = 7$. Маємо (15, 7)-код БЧХ з $\nu_{вп} = 2$.

За необхідності твірну матрицю коду БЧХ можна побудувати за правилами побудови такої матриці для циклічного коду. Так, для прикладу, що розглядається, за аналогією з (8.25) і (8.26) твірна матриця коду БЧХ матиме вигляд

$$G_{БЧХ(15,7)} = \begin{bmatrix} 1&0&0&0&0&0&0 & 1&1&1&0&1&0&0&0 \\ 0&1&0&0&0&0&0 & 0&1&1&1&0&1&0&0 \\ 0&0&1&0&0&0&0 & 0&0&1&1&1&0&1&0 \\ 0&0&0&1&0&0&0 & 0&0&0&1&1&1&0&1 \\ 0&0&0&0&1&0&0 & 1&1&1&0&0&1&1&0 \\ 0&0&0&0&0&1&0 & 0&1&1&1&0&0&1&1 \\ 0&0&0&0&0&0&1 & 1&1&0&1&0&0&0&1 \end{bmatrix}.$$

Усі подальші процедури кодування виконуються аналогічно методикам, викладеним вище для циклічних кодів.

У табл. 8.7 наведено параметри деяких кодів БЧХ завдовжки до $n = 255$. Подані в таблиці параметри визначено згідно з викладеною вище методикою. Для зручності запису твірні поліноми $P(x)$ записано у вісімковій системі числення.

Щоб дістати твірний поліном у звичайному вигляді, тобто в тій формі, яка використовується для побудови кодів БЧХ, кожну цифру треба перевести у двійковий трибіт. Наприклад, поліном $P(x) = 45$ двійковими числами запишеться так: $4 \rightarrow 100$ і $5 \rightarrow 101$. Отже, маємо двійкове число 100101, яке можна подати поліномом $P(x) = x^5 + x^2 + 1$.

Як показано вище, коди БЧХ мають непарне значення мінімальної кодової відстані d_{min}. Для того щоб збільшити d_{min} на одиницю, досить домножити твірний поліном коду БЧХ на двочлен $(x + 1)$.

Коди БЧХ характеризують деякі закономірності. По-перше, співвідношення між максимальною кодовою відстанню та числом h може бути подане як

$$d_{max} = 2^{h-1} - 1.$$

Дійсно, для прикладу, що розглядався вище, при $n = 15$ ($h = 4$) $d_{max} = 2^{4-1} - 1 = 7$ (це підтверджується табл. 8.7), а кількість інформаційних розрядів, яка може бути використана при обумовлених значеннях h і d_{max}, дорівнює $(h + 1)$.

По-друге, кількість кодів, що різняться своєю коректувальною здатністю і мають однакову довжину кодової комбінації ($n = 2^h - 1$), на дві одиниці менша від кількості всіх незвідних поліномів, на які розкладається двочлен $x^{2^h-1} + 1$.

Таблиця 8.7

n	k	r	d_{min}	Твірний поліном $P(x)$
7	4	3	3	13
15	11	4	3	23
	7	8	5	721
	5	10	7	2467
31	26	5	3	45
	21	10	5	3551
	16	15	7	107657
	11	20	11	5423325
	6	25	15	313365047
63	57	6	3	103
	51	12	5	12471
	45	18	7	1701317
	39	24	9	166623567
	36	27	11	1033500423
	30	33	13	1574641656547
	24	39	15	17323260404441
	18	45	21	1363026512351725
127	120	7	3	211
	113	14	5	41567
	106	21	7	11554743
	99	28	9	3447023271
	92	35	11	624730022327
	85	42	13	130704476322273
	78	49	15	26230002166130115
	71	56	19	6255010713253127753
	64	63	21	1206534025570773100045
255	247	8	3	435
	239	16	5	267543
	231	24	7	156720665
	223	32	9	75626641375
	215	40	11	23157564726421
	207	48	13	16176560567636227
	199	56	15	7633031270420722341
	191	64	17	2663470176115333714567
	187	68	19	5275531354000132223635 1
	179	76	21	22624710717340432416300455

Так, для $n = 15$ дістанемо $h = 4$ та двочлен $(x^{15} + 1)$. Цей двочлен не є найпростішим, тому $h = 4$ буде старшим степенем незвідного полінома, на який розкладається двочлен $x^{15} + 1$. Таким чином, двочлен $x^{15} + 1$ буде розкладатися на незвідні поліноми четвертого степеня та на незвідні поліноми тих степенів,

показники яких є дільниками числа 4, тобто 1 і 2:

$$x^{15} + 1 = (x + 1)(x^2 + x + 1)(x^4 + x + 1)(x^4 + x^3 + 1)(x^4 + x^3 + x^2 + x + 1).$$

Як випливає з цього розкладання, кількість незвідних поліномів дорівнює п'яти, а кількість циклічних кодів для $n = 15$ становить $5 - 2 = 3$, що підтверджує табл. 8.7.

Декодування кодів БЧХ (виявлення та виправлення помилок) завдовжки $n = 15$ може виконуватися з використанням методики, викладеної для циклічних кодів з $d_{min} < 5$. При декодуванні кодів БЧХ з довжиною комбінацій $n > 15$ можуть виникнути деякі труднощі, пов'язані з великим обсягом обчислень для виявлення та виправлення помилок. У таких випадках при $k > n/2$, де k — кратність зсуву, рекомендується комбінацію, утворену після k-кратного зсуву і підсумовування з осередком, зсувати не праворуч, а ліворуч на n циклічних кроків.

Кількість помилок, які можуть виправляти коди БЧХ, не обмежена, але зі збільшенням кратності помилки значно зростає складність пристроїв декодування, що призводить до зменшення швидкості передачі інформації.

8.1.5. КОД ФАЙРА

Коди БЧХ розраховані на виправлення кількох помилок, які обов'язково знаходяться поруч; тому вони потребують значної кількості перевірних елементів. Двійковий код Файра призначений для виправлення поодиноких пачок помилок, для чого потребує значно меншої кількості перевірних елементів порівняно з кодами БЧХ.

Під *пачкою* (пакетом) помилок розуміють не тільки групу помилок, розташованих поруч, а й групу або кілька спотворених, неспотворених елементів, які знаходяться між двома спотвореними елементами. В останньому випадку до пачки помилок разом двох крайніх спотворених елементів пачки, належать спотворені та неспотворені елементи, розташовані між ними, тобто в середині пачки.

Твірний поліном коду Файра [12, 25] визначається виразом

$$P_{\Phi}(x) = P(x)(x^c + 1), \qquad (8.34)$$

де $P(x)$ — незвідний поліном степеня l, що належить h; c — просте число, яке не повинно ділитися на h без остачі.

Поліном $P(x)$ має деякий степінь h, якщо h — найменше додатне число таке, що двочлен $x^h + 1$ ділиться на $P(x)$ без остачі.

Для будь-якого l існує принаймні один **незвідний поліном** $P(x)$ степеня l, який належить числу

$$h = 2^l - 1. \qquad (8.35)$$

Незвідний поліном $P(x)$ вибирається з табл. 8.5 так, щоб виконувалась умова (8.35), причому $l \geq b$, де $b = \nu_{вп}$ — довжина пачки помилок. Так, якщо $P(x) = x^3 + x^2 + 1$ (l=3), то $h = 2^l - 1 = 7$, а число c може мати значення, які не діляться на 7, тобто 15, 16, 17, 18, 19, 20, 22 тощо.

Довжина коду Файра визначається виразом

$$n = \text{НСК}(c, h), \qquad (8.36)$$

тобто є НСК чисел c та h, тому що тільки в цьому разі двочлен $x^n + 1$ буде ділитися на поліном $P_{\Phi}(x)$ без остачі.

Кількість перевірних елементів цього коду визначається так:

$$r = c + l, \qquad (8.37)$$

а інформаційних — так:

$$k = n - c - l. \qquad (8.38)$$

Якщо скористатися методом утворення вкорочених циклічних кодів, викладеним вище, то можна дістати код Файра меншої довжини з тією самою кількістю перевірних елементів.

Код Файра виправляє будь-яку поодиноку пачку помилок завдовжки b або менше й одночасно виявляє будь-яку пачку помилок завдовжки $B \geq b$ або менше, якщо $c \geq b + B - 1$ і $l \geq b$.

Якщо користуватися цим кодом тільки для виявлення помилок, то можна виявити будь-яку комбінацію з двох пачок помилок, довжина найменшої з яких не перевищує l, а сума довжин обох пачок менша, ніж $c + 1$. Можна також виявити будь-яку поодиноку пачку помилок з довжиною, не більшою від $r = c + l$, де r — кількість перевірних елементів.

Так, згідно з [12], якщо твірний поліном коду Файра $P_{\Phi}(x) = (x^4 + x + 1)(x^7 + 1) = x^{11} + x^8 + x^7 + x^4 + x + 1$, то степінь незвідного полінома $P(x)$ становить $l = 4$, а $c = 7$. Значення h, n, r і k знаходимо за (8.35) – (8.38): $h = 2^4 - 1 = 15$; n = НСК (7, 15) = $= 7 \cdot 15 = 105$; $r = 4 + 7 = 11$; $k = 105 - 7 - 4 = 94$. Цей код може бути використаний, наприклад, для виправлення пачки помилок завдовжки до $b = 4$ та виявлення будь-якої пачки помилок завдовжки $b > 4$ або для виправлення пачки помилок завдовжки $b \leq 2$ та виявлення пачок помилок завдовжки $b \leq 6$. Якщо застосувати його тільки для виявлення помилок, то ним можна виявити будь-яку поодиноку пачку помилок завдовжки $b \leq$

(4 + 7) і будь-яку комбінацію з двох пачок помилок, довжина найменшої з яких $l \leq 4$, а загальна сума довжин їх не перевищує 4 + 4 = 8.

Порівняння коду Файра з аналогічним кодом БЧХ за здатністю виправляти помилки буде не на користь останнього. Дійсно, кількість перевірних елементів коду БЧХ, що виправляє чотири поодинокі помилки і має довжину n, близьку до довжини коду Файра, становитиме $r = 28$ (див. табл. 8.7) при загальній довжині $n = 127$ і кількості його інформаційних елементів $k = 99$ порівняно з $r = 11$ та $k = 94$ коду Файра. Отже, надмірність коду БЧХ ($R_{над} = 28/127 = 0{,}22$) буде значно вищою, ніж надмірність коду Файра ($R_{над} = 11/105 = 0{,}1$). З цього випливає, що виправити чотири помилки, які знаходяться в одному місці, набагато простіше, ніж ті самі чотири помилки, випадково розташовані по всій довжині комбінації.

Останнє зумовлює використання коду Файра при передачі інформації по каналах з великою ймовірністю виникнення пачок помилок.

8.1.6. КОД ІЗ БАГАТОКРАТНИМ ПОВТОРЕННЯМ

Код із багатократним повторенням (без інверсії) є подільним лінійним кодом. Він містить k інформаційних і $n_R k$ перевірних елементів, де $n_R \geq 2$ — кількість повторень початкової кодової комбінації. В цьому коді кожні k перевірних елементів є просто повторами інформаційними елементами

$$b_j = b_{j+2k} = b_{j+3k} = \ldots = b_{j+(n_R-1)k} = a_j, j = 1 \ldots k.$$

Через те, що код має кодову відстань $d_{\min} = n_R + 1$, він може використовуватися для виявлення та виправлення помилок. Процедура виявлення помилок у прийнятій кодовій комбінації полягає в порівнянні однойменних інформаційних і перевірних елементів. Незбіг їх свідчить про наявність помилок у прийнятій комбінації.

При виправленні помилок у кодовій комбінації застосовується мажоритарний принцип виправлення для кожного інформаційного елемента, тобто «голосування за більшістю», коли за істинне значення приймається те, яке найчастіше зустрічається в цьому інформаційному та відповідних перевірних елементах. Код дає змогу виправити помилки кратністю від 1 до $(d_{\min} - 1)/2$ та деякі помилки більш високої кратності залежно від кількості повторень їх.

Надмірність коду визначається виразом

$$R_{над} = n_R/(n_R + 1).$$

8.1.7. ІТЕРАТИВНІ КОДИ

Уперше ітеративні коди були запропоновані П. Еліасом [48]. Вони характеризуються двома або більшою кількістю перевірок усередині кодової комбінації, а властивості цих кодів повністю визначаються параметрами їх.

Так, довжина n кодової комбінації, кількість інформаційних параметрів k та мінімальна кодова відстань $d_{\min}$ визначаються виразами.

$$n=\prod_{i=1}^{S} n_i; \quad k=\prod_{i=1}^{S} k_i; \quad d_{\min}=\prod_{i=1}^{S} d_{\min i},$$

де $n_i, k_i, d_{\min i}$ — параметри ітерованих кодів; S — кратність ітерування; $\prod$ — знак множення.

На практиці широко застосовуються двовимірні лінійні ітеративні коди з кодуванням за рядками та стовпцями з однією перевіркою на парність. Дозволяється використовувати коди з кількістю перевірних елементів 8, 9 і 16. Для коду з $r = 8$ застосовується блок інформаційних елементів розміром 3×4 (з $k_1 = 3$ рядками та $k_2 = 4$ стовпцями). При цьому кількість інформаційних елементів $k = k_1k_2 = 3 \cdot 4 = 12$, а перевірних $r = 8$; $n = 20$. Для коду з $r = 9$ беруть $k = k_1k_2 = 4 \cdot 4 = 16$, $n = 25$; для коду з $r = 16$ або $k = k_1k_2 = 8 \cdot 7 = 56$, $n = 72$, або $k = k_1k_2 = 7 \cdot 8 = 56$, $n = 72$.

Ці коди мають мінімальну кодову відстань $d_{\min} = 2 \cdot 2 = 4$ і дають змогу виявити помилки будь-якої кратності, за винятком деяких чотири-, шести- та восьмикратних помилок, якщо вони розміщуються на вершинах прямокутників або попарно в певному порядку. В режимі виправлення та виявлення помилок код виправляє будь-які поодинокі помилки і виявляє всі подвійні та деякі помилки більшої кратності.

При виявленні помилок на приймальному боці виконується перевірка на парність кожних рядка та стовпця. Невиконання умови парності в якомусь стовпці свідчить про наявність спотворених елементів у прийнятій кодовій комбінації.

При виправленні та виявленні помилок на приймальному боці визначаються рядки і стовпці, для яких не виконується умова парності. Спотворений інформаційний елемент знаходиться на місці перетину рядка та стовпця, для яких не виконується перевірка на парність.

Надмірність двовимірних ітеративних кодів становить:

$$R_{\text{над}} = 1 - k/n = r/n = 8/20 = 2/5 \text{ при } r = 8;$$
$$R_{\text{над}} = 9/25 \text{ при } r = 9;$$
$$R_{\text{над}} = 16/72 = 2/9 \text{ при } r = 16.$$

Нехай, наприклад, двовимірним ітеративним кодом треба закодувати комбінацію 110111000110 двійкового простого коду з $k = 12$ інформаційними елементами.

Розбиваємо цю комбінацію на три частини й записуємо її в три рядки:

1	1	0	1
1	1	0	0
0	1	1	0

Перевіряємо на парність елементи кожних рядка та стовпця й дописуємо перевірні елементи:

1	1	0	1	1	(b_1)
1	1	0	0	0	(b_2)
0	1	1	0	0	(b_3)
0	1	1	1	1	
b_4	b_5	b_6	b_7	b_8	

Таким чином, маємо двовимірний ітеративний код із перевіркою на парність. У лінію (канал) передається така послідовність двійкових елементів: 1101111000011001111.

Припустимо, що при передачі внаслідок спотворень виникла помилка і на приймач прийшла комбінація 1101111000111001111. При декодуванні на приймальному боці прийняту двійкову послідовність знову записуємо у вигляді матриці

1	1	0	1	1
1	1	0	0	0
1	1	1	0	0
0	1	1	1	1

і перевіряємо на парність кожні її рядок та стовпець:

1	1	0	1	1	0
1	1	0	0	0	0
1	1	1	0	0	1
0	1	1	1	1	0
1	0	0	0	0	

При відсутності спотворень усі перевірні елементи, сформовані декодером, мають дорівнювати нулю. Однак у розглядуваному прикладі виникла помилка в першому стовпці та третьому рядку, причому в двох випадках брав участь елемент a_9. Якщо значення цього елемента змінити на протилежне, то всі перевірні елементи, сформовані декодером, дорівнюватимуть нулю. Таким чином, помилку буде виявлено й виправлено.

При побудові ітеративних кодів для кодування елементів по рядках і стовпцях можна використовувати не тільки код із перевіркою на парність (або непарність), а й інші коди (наприклад, коди Геммінга). При цьому мінімальна кодова відстань $d_{\min}$ збільшується, а значить, зростає й здатність коду виправляти помилки.

Суттєвим недоліком ітеративних кодів є порівняно висока надмірність їх, яка значно перевищує надмірність циклічних

кодів, здатних виявляти та виправляти ту саму кількість помилок за інших однакових умов. Однак їх використання в системах передачі даних зумовлює більш просте порівняно з циклічними кодами кодування та декодування за допомогою ЕОМ.

Ітеративні коди знайшли широке застосування для виявлення та виправлення помилок, які виникають при запису, зберіганні та зчитуванні цифрової інформації на магнітних носіях.

8.1.8. КАСКАДНІ КОДИ

Збільшення мінімальної кодової відстані $d_{\min}$ і, як наслідок, здатності коду виправляти помилки можна досягти, якщо застосувати кілька ступенів кодування (каскадний принцип кодування). Такі коди дістали назву *каскадних*.

На практиці поширеними є каскадні коди, що складаються з двох кодів (два ступені кодування), які називаються *внутрішнім* і *зовнішнім*. При цьому зовнішній код використовується для кодування повідомлень, що надходять від джерела у вигляді первинного коду, а внутрішній — для кодування комбінацій зовнішнього коду перед передачею їх у канал зв'язку. На рис. 8.1 показано спрощену схему системи передачі з каскадним принципом кодування повідомлень.

Як зовнішній код звичайно використовуються коди Ріда - Соломона або коди БЧХ [27, 37]. Вибір внутрішнього коду залежить від характеристик каналу зв'язку та інтенсивності виникнення помилок. Це може бути код БЧХ, код Хеммінга чи інший код. Взагалі задачею внутрішнього коду є забезпечення прийнятої ймовірності помилки, а зовнішнього — зниження результуючої ймовірності неправильного декодування до заданого значення.

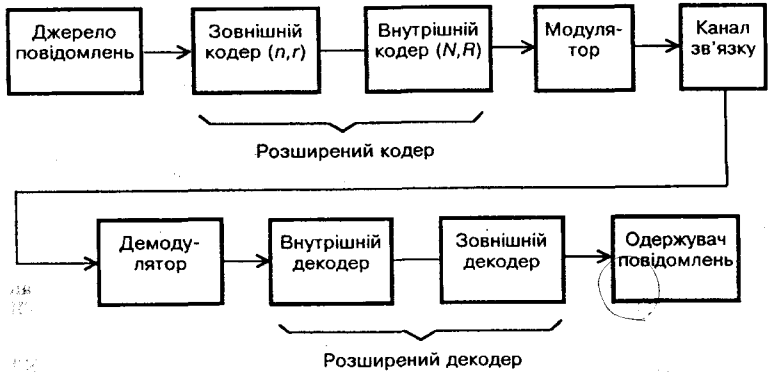


Рис. 8.1

При кодуванні зовнішнім кодом як інформаційні елементи комбінацій первинного коду приймаються всі елементи комбінації первинного коду, що надходять від джерела повідомлень, а при кодуванні внутрішнім кодом — усі елементи (інформаційні та перевірні) комбінації зовнішнього коду. Декодування виконується в зворотному порядку — спочатку декодується комбінація внутрішнього коду, а потім — зовнішнього.

Загалом, тобто при S ступенях (каскадах) кодування, мінімальна кодова відстань каскадного коду визначається кодовими відстанями кодів, які використовуються для його побудови:

$$d_{\min} = \prod_{i=1}^{S} d_{\min i}.$$

Швидкість і надмірність каскадного коду також залежать від кодів, що використовуються при його побудові.

Каскадні коди знайшли широке застосування для передачі повідомлень по радіоканалах із великим рівнем завад, зокрема в супутникових лініях зв'язку.

8.2. РЕКУРЕНТНІ КОДИ

Рекурентними (неперервними) називаються коди, що подаються неперервною послідовністю кодових елементів без поділу на окремі комбінації.

Рекурентні коди дають суттєвий ефект при захисті інформації, що передається по каналах, де можливе виникнення помилок кратності та початок помилок. Найпростіше ці коди будуються при надмірності $R_{\text{над}} = 1 - a/l = 0{,}5$, де a — кількість інформаційних елементів; l — довжина ділянки послідовності елементів, що передаються.

На відміну від блокових рекурентні коди відрізняються тим, що дають змогу кодувати інформаційну послідовність неперервно, не розбиваючи її на блоки фіксованої довжини n з k інформаційними та r перевірними елементами. Такі коди ще називаються *згортковими*. В них при передачі кожний перевірний елемент формується додаванням за модулем 2 двох інформаційних елементів, відстань між якими дорівнює кроку додавання $t_{\text{кр}} = k - i$ (рис. 8.2):

$$a_i \oplus a_k = b_{i,k};\ a_{i+1} \oplus a_{k+1} = b_{i+1,\,k+1};$$

$$a_k \oplus a_k + t_{\text{кр}};\ a_{k+1} \oplus a_{k+1+t_{\text{кр}}} = b_{k+1,k+1+t_{\text{кр}}};\ldots.$$

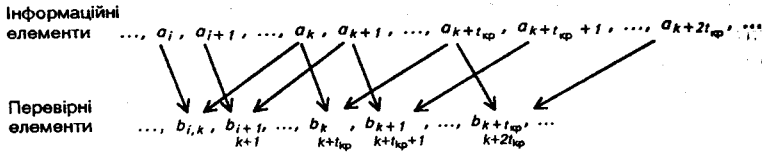


Рис. 8.2

Кількість перевірних елементів, сформованих за час T, дорівнює кількості інформаційних елементів, які надійшли за той самий час. Ці елементи передаються через один (a, b, a, b, a, b, ...). На приймальному боці вони розділяються й реєструються незалежно.

Із прийнятої послідовності інформаційних елементів формуються контрольні елементи b_i'' за тим самим алгоритмом, що й елементи b_i при кодуванні. При цьому кожний контрольний елемент b_i'' порівнюється із прийнятим перевірним елементом b_i'. Якщо спотворень не було, то $b_i' = b_i''$ (перевірний елемент збігається із відповідним контрольним). Наявність двох незбігів контрольних і перевірних елементів, зсунутих один відносно одного на $t_{кр}$ елементів, свідчить про спотворення інформаційного елемента, спільного для обох перевірних елементів, і його значення необхідно змінити на протилежне.

При спотворенні тільки перевірного елемента й правильному прийманні інформаційних елементів a_i' та a_k' буде тільки один незбіг контрольних і перевірних елементів, що вказує на помилкове приймання перевірного елемента, і ніяких виправлень роботи не потрібно. З принципу виправлення помилок у ланцюговому коді випливає, що правильне виправлення помилок можливе тільки в тому разі, коли два елементи з трьох, охоплені перевіркою, прийняті правильно.

Коректувальна здатність ланцюгового коду залежить від кроку додавання $t_{кр}$. Якщо кожний перевірний елемент перед передачею в канал затримати на час $T_з$ і пачки помилок, розташовані поруч, розділити захисним інтервалом A, який не містить спотворених елементів ($A = 6t_{кр} + 1$; $T_з = 3(t_{кр} + 1)\tau_е$, де $\tau_е$ — тривалість одного елемента), то ланцюговий код може виправити пачку помилок завдовжки $2t_{кр}$. Зміною довжини кроку $t_{кр}$ коректувальну здатність коду можна узгоджувати з характеристиками каналу зв'язку, зменшуючи чи збільшуючи допустиму частість помилок [12, 45].

8.3. НЕДВІЙКОВІ КОДИ

Ці коди поділяються на *блокові* та *неперервні*. Як відомо [2, 8, 12], двійкові блокові коди призначені в основному для виправлення незалежних помилок. Відповідні їм q-коди також виправляють помилки аналогічного походження. Проте слід урахувати, що один елемент q-коду несе $\log_2 q$ (при $m = 1$) або $\log_2 C_q^m$ (при $m \geq 2$) бітів інформації залежно від методу побудови конкретного коду.

Зазначена особливість q-коду дає підставу стверджувати, що навіть недвійковий блоковий код дає змогу виправляти умовний пакет помилок із $\log_2 q$ або $\log_2 C_q^m$ бітів інформації, який, якщо б він виник у аналогічному двійковому коді, не міг бути ним виправлений. Це є однією з переваг використання недвійкових кодів, які виправляють помилки.

Розглянемо алгоритм побудови деяких недвійкових кодів, що виправляють помилки: з багатократним повторенням, узагальненого коду Хеммінга, кодів БЧХ, кодів Ріда — Соломона, ітеративних і ланцюгового.

8.3.1. КОД ІЗ БАГАТОКРАТНИМ ПОВТОРЕННЯМ

Відмінність цього коду від q-коду з повторенням полягає в багатократному ($n_R \geq 3$) повторенні початкової кодової комбінації. Даний метод кодування застосовується при передачі інформації по каналах з високим рівнем завад для істотного підвищення вірогідності, коли немає можливості для цієї мети використати зворотний канал.

Якщо застосування двійкового коду з багатократним повторенням потребує відповідного аналогічного q-коду дає змогу або зберегти колишню швидкість передачі, або зменшити її несуттєво. Це пояснюється тим, що при використанні q-коду можна ввести аналог n_R-кратного повторення, збільшивши кількість позицій та ознак сигналу (алфавіт коду). В разі передачі інформації по радіоканалах, де, крім селективних завмирань, діють також завмирання сигналу в часі, аналог багатократного повторення можна ввести за кілька часових позицій (інтервалів).

Код із багатократним повторенням містить k інформаційних елементів, а кількість перевірних елементів залежить від кількості повторень n_R, причому кожний перевірний елемент збігається з відповідним йому інформаційним. Таким чином, довжина кодової комбінації $n = k + kn_R = k(1 + n_R)$, де $r = kn_R$.

При n_R-кратному повторенні надмірність коду визначається виразом

$$R_{над} = n_R/(1 + n_R).$$

Алгоритм побудови q-коду з багатократним повторенням має вигляд

$$a_i \Leftrightarrow b_j^i;\ i = \overline{1,k},\ j = \overline{1,n_R},$$

де a_i, b_i — множини позицій, призначені для передачі i-х інформаційного та перевірного елементів кодової комбінації відповідно; j — порядковий номер повторення.

Цей код є характерним представником класу низькошвидкісних кодів, тому що його швидкість $R = k/n = 1/(n_R + 1)$.

8.3.2. УЗАГАЛЬНЕНИЙ КОД ХЕММІНГА

Серед q-кодів найпростішими кодами, які мають алгебричну структуру й забезпечують нескладні процедури кодування та декодування, є лінійні блокові коди, що виправляють одну помилку. В класі двійкових кодів існує аналог їх — код Хеммінга. Хоча між цими кодами є суттєві відмінності, q-код часто називають *узагальненим кодом Хеммінга* [8], маючи на увазі узагальнення коду на недвійковий алфавіт $q > 2$.

На відміну від двійкового символами q-коду є елементи q-поля: його перевірна матриця H не є множиною всіх послідовних номерів позицій елементів коду в блоці завдовжки n. В q-коді, крім визначення місця помилки в блоці, необхідно ще знати її значення e [8].

Розглянемо матричне подання узагальненого коду Хеммінга. Його перевірна матриця

$$H = [h_{ij}],\ i = \overline{|1,r|},\quad j = \overline{|1,r|} \tag{8.39}$$

має розмір $r \times n$. Тут i — номер рядка; j — номер стовпця; r — кількість перевірних елементів у блоці; n — довжина блока.

На відміну від двійкового коду матриця H є підматрицею матриці

$$A = [a_{ir}],\ i = \overline{|1,r|},\ z = \overline{|1,q_r|}$$

розміром $r \times q^r$, у стовпцях якої послідовно записуються всі r-компонентні q-вектори ($H = A$ при $q = 2$). Стовпці H мають бути ненульовими, різними та лінійно незалежними. Для цього в матрицю H необхідно ввести всі вектори-стовпці матриці A, що мають однакову першу ненульову компоненту δ.

Оскільки ця компонента відповідає умові $1 \le \delta \le q - 1$, кількість її можливих значень дорівнює $q - 1$. Тому, виключивши з A нульовий вектор-стовпець, дістанемо кількість векторів-стовпців у матриці H (отже, й довжину кодового блока) $n = (q^r - 1)/(q - 1)$, що відповідає [32].

Із побудови матриці H випливає, що r її перших векторів-стовпців, кожний з яких містить єдину ненульову компоненту δ, утворюють діагональну підматрицю розміром $r \times r$. Ця обставина вказує на зручні позиції для розміщення r перевірних елементів у блоці.

Загалом кодовий вектор має вигляд

$$\bar{X} = b_1 \ldots b_i \ldots b_r a_1 \ldots a_j \ldots a_k, \tag{8.40}$$

де a_j – q-інформаційні символи джерела; $j = \overline{|1,k|}$; k — кількість інформаційних елементів вектора $X (k = n - r)$; b_i — перевірні елементи кодового блока ($i = \overline{|1,r|}$), причому

$$b_i = \delta^{-1} \sum_{j=1}^{\bar{k}} a_j h_{i,j+r}. \tag{8.41}$$

Вираз (8.41) можна дістати з матричного рівняння $H\bar{X}^T = 0$ відносно перевірних елементів, де $\bar{X}^T$ — транспонований вектор $\bar{X}$.

Указане у (8.40) розміщення перевірних елементів не є єдино можливим, але воно забезпечує мінімальний обсяг обчислень при кодуванні та декодуванні. При цьому вираз (8.41) установлює оптимальний алгоритм кодування.

Кодовий блок (8.40) покомпонентно передається в канал зв'язку, де він може бути спотворений завадою. Не вдаючись до суті цього явища, процес спотворень зручно подати як

$$\bar{Y} = \bar{X} + \bar{E},$$

де $\bar{Y}$ — спотворений n-вимірний кодовий вектор на виході каналу передачі; $\bar{E}$ — n-вимірний вектор помилки з єдиною ненульовою компонентою e (в припущенні одиничної помилки).

Першим кроком при декодуванні за аналогією з двійковим кодом є обчислення r-компонентного перевірного синдрому

$$\bar{S} = H\bar{Y}^T = \bar{L}e.$$

Вектор $\bar{S}$ — це значення помилки $e (1 \le e \le q - 1)$, помножене на стовпець $\bar{L}$ перевірної матриці H, який відповідає позиції спотвореного елемента в $\bar{Y}$. Його називають *локатором помилки*. Оскільки всі вектори-стовпці в H мають першу ненульову компоненту δ, в S перша ненульова компонента визначається добутком

$$s_1 = \delta e.$$

Це зумовлює другий крок процедури декодування — визначення помилки

$$e = s_1 \delta^{-1} = s_1/\delta.$$

Із визначення $\bar{S} = H\bar{Y}^T = \bar{L}e$ випливає третій крок процедури декодування — знаходження локатора помилки

$$\bar{L} = \bar{S}e^{-1} = \bar{S}/e.$$

На четвертому кроці процедури декодування впорядкованим перебором стовпців перевірної матриці H і порівнянням їх із локатором $\bar{L}$ за збігом визначається позиція спотвореного елемента в кодовому блоці.

П'ятим й останнім кроком декодування є виправлення помилки, яке загалом виконується відніманням значення помилки e від спотвореного елемента, знайденого в $\bar{Y}$ за його локатором $\bar{L}$. Далі спотворений елемент у блоці замінюється результатом віднімання і після виключення перевірних елементів одержувач дістає інформаційну частину кодового блока.

Усе описане вище стосується коду з довільною основою q. Зазначимо, що q-коди прийнято поділяти на дві великі групи: коди з простою основою $q = p$, де $p \in \{3,5,7,11,13,\ldots\}$; коди з основою q, що розкладається. Найбільший практичний інтерес викликає тільки окремий випадок цих кодів при $q = 2^l$, коли символи мають інформаційну ємність l бітів і можуть бути зіставлені з усіма l-розрядними двійковими числами.

Вибір q впливає на визначення операцій додавання, віднімання, множення і ділення під час виконання процедур кодування та декодування. Якщо основа q — просте число, то зручно використати апарат обчислень за модулем цього числа. Якщо ж $q = 2^l$, то необхідно звернутися до алгебричного апарата обчислень за модулем незвідного полінома (див. п.5.3). Символи коду при цьому ставляться у відповідність елементам скінченного поля порядку q.

8.3.3. КОДИ БОУЗА — ЧОУДХУРІ — ХОКВІНГЕМА

Недвійкові коди БЧХ є різновидом циклічних кодів. Як і двійкові, недвійкові коди БЧХ будуються за допомогою твірних поліномів $P(x)$, які визначаються за заданою мінімальною кодовою відстанню $d_{\min}$ і довжиною n кодової комбінації.

Одним із найпоширеніших підкласів недвійкових кодів БЧХ є коди, для побудови яких застосовуються поле елементів $GF(q)$ і розширене поле локаторів $GF(q^h)$, де $h > 0$ — ціле число.

Так, якщо $h > 0$ і n ділиться без остачі на $q^h - 1$, то в розширеному полі $GF(q^h)$ завжди знайдеться елемент $\beta \in GF(q^h)$ порядку n (за умови, що $\beta^n = 1$). У цьому разі всі елементи β^i, де $i = \overline{1,n}$, будуть різними і лінійно незалежними. Така властивість дає змогу з q^h елементів відібрати тільки n елементів для задання локаторів кодового блока завдовжки n. Усі ці локатори є елементами розширеного поля $GF(q^h)$ і становлять основу для побудови перевірної матриці H коду (першого її рядка), а також твірного полінома.

Твірний поліном коду БЧХ може задаватися з урахуванням того, що вибір перших r степенів елемента β визначається значенням β^{2^j}, де $j = 0,1,2,3,\ldots$ (тобто $\beta^1, \beta^2, \beta^4, \beta^8, \ldots$), яке використовується як спектр твірних коренів. При цьому твірний поліном має вигляд

$$P(x) = \prod_{j=0}^{j=r-1} (x - \beta^{2^j})$$

й обов'язково містить коефіцієнти поля локаторів $GF(q^h)$.

Твірний поліном можна задавати також виразом

$$P(\bar{x}) = \prod_{i=1} (\bar{x} - \beta^i),$$

де β^i набуває значень $\beta^1, \beta^2, \beta^3, \beta^4, \ldots$.

Той чи інший спосіб задання твірного полінома залежить від вибраного алгоритму декодування.

Послідовність локаторів $\beta^i (i = \overline{1,n})$ коду БЧХ за деяких умов відзначається особливою властивістю. Так, якщо розташувати локатори β^i у вигляді кільця, то ця структура буде кільцем класів остач з теорії чисел і полінома. Тоді будь-якому елементу даної структури обов'язково відповідатиме один протилежний елемент (їх добуток дорівнює 1), а сума цих елементів є елементом розширеного поля $GF(q^h)$, тобто належить полю $GF(q)$.

Вибравши першу половину твірних коренів довільно, а другі як елементи поля локаторів, що відповідають вибраним елементам, дістанемо твірний поліном $P(x)$ у вигляді добутку лінійних множників $(x - \beta^v)$, де β^v — всі вибрані твірні корені коду.

Характерна властивість цього методу побудови твірного полінома полягає в тому, що такий поліном завжди виходить «самодвоїстим» і його коефіцієнти є числами поля $GF(q^h)$, тобто елементами поля $GF(q)$. Остання обставина дає змогу при незначних основах q-коду будувати довгі (за n) коди БЧХ для виявлення помилок. При цьому як кодування, так і обчислення синдрому виконується тільки в числовому полі $GF(q)$, що значно простіше, ніж у розширеному полі $GF(q^h)$.

При побудові кодів БЧХ, ґрунтуючись на цьому методі, завжди забезпечується мінімально досяжна для вказаного класу кодів надмірність ($r = 2v_{вп}$), тоді як за методом [32] надмірність коду визначається виразом $r = 2v_{вп} + 1$.

Таким чином, викладений метод побудови твірного полінома $P(x)$ дає змогу в блоці тієї самої довжини n при тій самій кількості $\nu_{вп}$ дістати на один надмірний елемент менше. При цьому швидкість коду $R = 1 - r/n = 1 - 2\nu_{вп}/n$, а його надмірність — $R_{над} = 2\nu_{вп}/n$. Код дає змогу виявити $2\nu_в$ і виправити $\nu_{вп}$ помилок, тому що мінімальна кодова відстань $d_{min} = 2\nu_{вп} + 1$.

8.3.4. КОДИ РІДА — СОЛОМОНА

Ці коди використовуються для передачі інформації по каналах з високою інтенсивністю завад, коли виникають помилки кратності два й більше, пачки помилок, а також сполучення пачок і однократних помилок.

Коди Ріда — Соломона (РС) можна розглядати як окрему гілку циклічних кодів БЧХ, головна відмінність якої полягає в тому, що поле локаторів коду РС збігається з полем його елементів. Отже, якщо поле $GF(q)$ коду БЧХ має q окремих елементів (потужність поля дорівнює q), а поле $GF(q^h)$ його локаторів — q^h елементів і є h-розширенням поля $GF(q)$, то в коді РС і елементи, і локатори їх знаходяться в одному полі, тобто належать полю $GF(q)$ [2, 3, 24, 25, 32]. Іншими словами, код РС — це вироджена форма коду БЧХ, у якого $h = 1$.

Довжина коду РС визначається виразом

$$n = q - 1,$$

де q — основа (алфавіт) коду.

На підставі цього виразу коди РС називаються *відносно короткими*.

Код РС, як і код БЧХ, може задаватися твірною чи перевірною матрицею або твірним чи перевірним поліномом. Найпоширенішим є метод побудови коду РС на основі твірного полінома $P(x)$. Перевірну матрицю H часто використовують для вивчення деяких властивостей коду РС і його зв'язку з систематичними кодами.

Твірний поліном коду РС із виправленням $\nu_{вп}$ помилок є добутком кількості перевірних елементів $r = 2\nu_{вп}$ і мінімальних поліномів для спектра елементів поля $GF(q)$ [2, 24]:

$$P(x) = \left(x - \beta^{j_0}\right)\left(x - \beta^{j_0+1}\right)\ldots\left(x - \beta^{j_0+r-1}\right), \qquad (8.42)$$

де β^{j_0}, β^{j_0+1}, ... β^{j_0+r-1} — спектр твірних коренів полінома.

Степінь полінома (8.42) дорівнює кількості $r = 2\nu_{вп} = n - k$ перевірних елементів.

Для спрощення побудови коду РС вибирають $j_0 = 1$ і дістають звичайну форму коду РС:

$$P(x) = (x - \beta^1)(x - \beta^2)\ldots(x - \beta^r) = \prod_{j=1}^{r}(x - \beta^j). \qquad (8.43)$$

Перевірний поліном $H(x)$ коду РС знаходять як частку від ділення $(x^{q-1} + 1)$ на $P(x)$.

Мінімальна кодова відстань коду РС становить

$$d_{min} = r + 1 = 2\nu_{вп} + 1,$$

тобто код РС — це код з максимально досяжною кодовою відстанню.

Надмірність коду РС визначається виразом

$$R_{над} = \frac{r}{n} = \frac{2\nu_{вп}}{q-1},$$

а його швидкість — виразом

$$R = \frac{\overline{k}}{n} = \frac{n-r}{n} = \frac{q - 2\nu_{вп} - 1}{q-1}.$$

Перевірна матриця коду РС має вигляд [2,24,32].

$$H = \begin{bmatrix} \beta^{j_0(n-1)}\beta^{j_0(n-2)}\ldots\beta^{j_0}1 \\ \beta^{(j_0+1)(n-1)}\beta^{(j_0+1)(n-2)}\ldots\beta^{j_0+1}1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \beta^{(j_0+r-1)(n-1)}\beta^{(j_0+r-1)(n-2)}\ldots\beta^{j_0+r-1}1 \end{bmatrix}, \qquad (8.44)$$

де β — примітивний елемент поля $GF(q)$; j_0 — примітивний корінь твірного полінома; n — довжина коду; r — кількість його перевірних елементів.

Слід звернути увагу на те, що другий праворуч стовпець у матриці (8.44) відповідає спектру коренів твірного полінома (8.42). У матриці (8.44) перелік стовпців по i йде справа наліво, причому $0 \le i \le n - 1$. Перелік елементів у кодовому блоці має бути таким самим. Перелік рядків у матриці (8.44) йде зверху вниз по $j_0 \le j \le j_0 + r - 1$. Можливий й інший порядок переліку, треба лише додержуватися збігу послідовності твірних коренів $P(x)$ із другим праворуч стовпцем матриці (8.44).

Найпростіше коди РС реалізуються для алфавіту $q = 2^l$, де $l = 2,4,8,\ldots$.

Якщо поле має основу $2(q = 2^l)$, то операції віднімання збігаються з операціями додавання; тому скрізь знак « — » можна замінити на знак « + », зокрема у виразах (8.42) та (8.43).

Декодування кодів РС виконується відповідно до загальних прикладів декодування [2, 3, 24, 25, 32]. При виправленні помилок дістають значення локаторів $L^j = \beta^j$, що відповідають спотвореним елементам, і значення помилок для кожного спотвореного елемента. Помилки виправляють, віднімаючи значення помилки від значення відповідного елемента.

Розглянемо кілька прикладів побудови коду РС.

Нехай $q = 2^4 = 16$. Тоді код РС може мати довжину $n = q - 1 = 16 - 1 = 15$. Примітивний елемент $\beta \in GF(q)$ такий, що $\beta^{15} = 1$. Всі ненульові елементи β^j $(0 \le j \le 14)$ утворюють множину локаторів коду, яких буде 15. Вони збігаються з першим рядком матриці (8.44) при $j_0 = 1$ або з другим рядком при $j_0 = 0$. Всі вони разом із нулем є елементами поля $GF(q)$ локаторів й елементів коду.

Для виконання розрахунків з елементами поля $GF(q)$ треба користуватися адитивною та мультиплікативною формами їх подання (див. табл. 5.3 для $q = 16$). Нехай $\nu_{вп} = 2$. Виберемо $j_0 = 1$, тоді згідно з (8.43) матимемо

$$P(x) = (x - \beta)(x - \beta^2)(x - \beta^3)(x - \beta^4) = (x + \beta)(x + \beta^2)(x + \beta^3)(x + \beta^4) = \\ = x^4 + \beta^{13}x^3 + \beta^6x^2 + \beta^3x + \beta^{10} = x^4 + 13x^3 + 12x^2 + 8x + 7.$$

При кодуванні таким поліномом $r = 2\nu_{вп} = 4$. Знаючи, що $n = 15$, дістаємо $k = n - r = 15 - 4 = 11$, тобто це — систематичний (15, 11)-код із кодовою відстанню $d_{min} = r + 1 = 4 + 1 = 5$. Інформаційний блок цього коду має 11 шістнадцяткових елементів (44 біт).

Якщо ж при тій самій основі коду РС $(q = 16 = 2^4)$ та довжині $n = 15$ побудувати код, що виправляє три помилки $(\nu_{вп} = 3)$, то $r = 2\nu_{вп} = 2 \cdot 3 = 6$ і при $j_0 = 0$ твірний поліном коду матиме вигляд

$$P(x) = (x + \beta^0)(x + \beta^1)(x + \beta^2)(x + \beta^3)(x + \beta^4)(x + \beta^5) = x^6 + \beta^9x^5 + \\ + \beta^{12}x^4 + \beta^1x^3 + \beta^2x^2 + \beta^4x + \beta^5 = x^6 + 10x^5 + 15x^4 + 2x^3 + 4x^2 + 3x + 6.$$

У цьому коді $k = n - r = 15 - 6 = 9$, $d_{min} = r + 1 = 6 + 1 = 7$, тобто його інформаційний блок має дев'ять елементів по чотири біти кожний, а всього 36 біт.

Кодування кодом РС виконують одним із відомих способів кодування циклічного коду. Можливе також несистематичне кодування множенням інформаційного полінома $Q(x)$ степеня 8 на твірний поліном $P(x)$ степеня 6. При цьому дістають поліном степеня 8 + 6 = 14 завдовжки $n = 15$ елементів. Однак більш поширеним є кодування діленням добутку $x^rQ(x)$ інформаційного полінома $Q(x)$ на твірний поліном $P(x)$, як це характерно для циклічних кодів з подальшим додаванням остачі $R(x)$ від ділення:

$$F(x) = x^rQ(x) \oplus R(x).$$

8.3.5. БАГАТОВИМІРНІ ІТЕРАТИВНІ КОДИ

Для передачі інформації широко використовуються ітеративні коди, які відзначаються високою здатністю виявляти помилки. Так, відомий двійковий ітеративний код [12], що реалізує

перевірку на парність за рядками та стовпцями, виявляє всі одно-, дво- й трикратні помилки, деякі чотирикратні помилки і помилки більшої кратності, розташовані в кутах прямокутника, а також деякі шести-, восьмикратні та інші помилки.

Для усунення вказаного недоліку в даний двовимірний ітеративний код вводиться додаткова перевірка по діагоналі. Утворений таким чином тривимірний ітеративний код дає змогу виявляти більшість помилок кратністю до п'яти та деякі помилки більшої кратності.

Підвищення швидкості передачі інформації тривимірним ітеративним кодом можна досягнути збільшенням базового основа його інформаційних елементів.

Подальше підвищення здатностей кодів цього класу виявляти та виправляти помилки можливе при збільшенні кількості перевірок, тобто вимірності ітеративного коду. Так, якщо базовий блок інформаційних елементів сформувати у вигляді куба або прямокутного паралелепіпеда, то можна збільшити здатність ітеративного коду виявляти та виправляти помилки, ввівши додаткові перевірки по діагоналі. Через великі обсяги інформації, що потребують зберігання й оброблення, функції кодування та декодування таких кодів бажано доручити спеціалізованій ЕОМ.

Підвищення швидкості ітеративного коду можна досягти збільшенням основи q коду. При $q > 2$ зростає обсяг інформації, що міститься, оскільки кількість інформації, яка міститься в одному елементі кодової комбінації, визначається основою q коду.

При визначенні перевірних елементів ітеративних q-кодів використовують додавання елементів комбінацій базового коду за модулем q. Помилки виявляються за аналогією з двійковим кодом порівнянням перевірних елементів кожного рядка (стовпця), прийнятих з каналу і здобутих обчисленнями.

Спотворений елемент виправляється так. Якщо не виконується перевірка i-го рядка та j-го стовпця, то елемент, що знаходиться на їх перетині, заміняється елементом, який був обчислений додаванням цього прийнятого елемента (помилкового) до значення, знайденого відніманням значення, прийнятого з каналу, й обчисленням перевірних елементів i-го рядка (або j-го стовпця).

При виникненні кількох помилок в одному рядку (стовпці) вони виправляються послідовно для тих стовпців (рядків), де помилки є поодинокими.

Слід відзначити, що кратність помилок, які виправляються q-кодом, вища порівняно з двійковим кодом, оскільки виправлення одного елемента q-коду відповідає виправленню $\log_2 q$ двійкових одиниць.

8.3.6. НЕДВІЙКОВИЙ ЛАНЦЮГОВИЙ КОД

Рекурентні (неперервні) коди призначені в основному для виправлення пачок (пакетів) помилок. Для таких кодів характерним є те, що операції кодування та декодування виконуються над неперервною послідовністю елементів. Це є перевагою рекурентних кодів, оскільки дає більші можливості для використання надмірності, що вводиться.

Найпростіше рекурентні коди реалізуються при надмірності, що дорівнює 0,5, коли кількість перевірних множин (елементів кодових комбінацій) дорівнює кількості інформаційних. Такий код називається *ланцюговим*. З усіх рекурентних кодів цей код дістав найбільшого поширення.

Розглянемо принципи побудови недвійкових ланцюгових кодів. У ланцюговому q-коді на передачі кожна перевірна множина формується за двома інформаційними, розташованими одна від одної на відстані $t_{кр} \geq 1$.

Через те що кожна інформаційна множина бере участь у формуванні двох перевірних множин, а кожна перевірна множина формується за двома інформаційними, кількість перевірних множин, сформованих за час T, дорівнюватиме кількості інформаційних множин, що надходять за той самий час. У канал інформаційні та перевірні множини передаються послідовно через одну: a, b, a, b, a і т. д.

На приймальному боці множини розділяються і фіксуються незалежно. З прийнятої послідовності інформаційних множин a' перевірні множини b'' формуються таким самим чином, як і при передачі. Після цього вони порівнюються, а помилки виправляються за аналогією з відомими двійковими ланцюговими кодами.

Часова надмірність ланцюгового коду визначається структурою множин, які передаються за один часовий інтервал. У разі суміщення інформаційних і перевірних множин вони розміщуються на одному часовому інтервалі й часова надмірність відсутня. Однак при цьому зростає позиційна надмірність, яка збільшується в два рази. При розміщенні ж інформаційних і перевірних множин на окремих часових інтервалах у два рази зростає часова надмірність, а позиційна може бути відсутньою.

При передачі повідомлень по каналах, у яких помилки (спотворення елементів кодових комбінацій) групуються в пачки по $t_{кр}$ помилок у кожній, перевірні множини формуються так, щоб вони зв'язували інформаційні множини, що відстоять одна від одної на відстань $t_{кр}$, яка визначається станом каналу зв'язку.

Зв'язок між інформаційними та перевірними елементами ланцюгового коду при $t_{кр} = 1$ можна подати такою системою рівнянь:

$$\left.\begin{array}{l} a_1 \otimes a_2 = b_{1,2}; \\ a_2 \otimes a_3 = b_{2,3}; \\ \ldots\ldots\ldots\ldots\ldots \\ a_i \otimes a_{i+1} = b_{i,i+1}; \\ a_{k-1} \otimes a_k = b_{k-1,k} \end{array}\right\} \qquad (8.45)$$

а алгоритм кодування при $t_{кр} = 1$ записати у вигляді

$$a_i \otimes a_{i+1} = b_{i,i+1}, \quad i = \overline{|1,k|}.$$

Загалом (при $t_{кр} \geq 1$) цей алгоритм має вигляд

$$a_i \otimes a_{i+t_{кр}} = b_{i,i+t_{кр}} \quad i = \overline{|1,k|}. \qquad (8.46)$$

Ланцюговий код виявляє та виправляє стирання будь-яких **множин**, зв'язаних між собою перевірною множиною, що випливає з алгоритму кодування (8.46).

Помилка виявляється, але не може бути виправлена, якщо **стирається** три або більше множин виду : $b_{i,i-t_{кр}}$, a_i, $b_{i,i+t_{кр}}$, куди **входять** дві перевірні множини. Помилка не виявляється при **помилковому** переході трьох зв'язаних між собою множин, дві з **яких** перевірні, у відповідні дозволені множини.

Слід звернути увагу на те, що для можливості виявлення та **виправлення** спотворених крайніх (перших й останніх) інформаційних елементів у неперервну послідовність елементів на **початку** та в кінці треба вводити $t_{кр}$ додаткових елементів, які **заздалегідь** відомі на приймальному боці.

До **недоліків** ланцюгового q-коду треба віднести те, що він **не виправляє** послідовні помилки кратністю, більшою від $2\,t_{кр}$; **надмірність** цього коду дорівнює 0,5.

КОНТРОЛЬНІ ЗАДАЧІ

1. Побудувати твірну матрицю та визначити всі комбінації лінійного **систематичного** двійкового групового коду, здатного виправляти поодинокі помилки ($\nu_{вп} = 1$), при передачі восьми повідомлень.

Розв'язання. Через те що $N_д = 2^k = 2^3 = 8$, маємо $k = 3$. Отже, кількість рядків твірної матриці $G_{n,k}$ дорівнює 3, а кількість її стовпців визначається довжиною коду $n = k + r$. Кількість r перевірних розрядів можна знайти з табл. 8.1, урахувавши, що $d_{min} = 2\nu_{вп} + 1 = 3$. Тоді $r = 3$, тобто кількість стовпців підматриці $C_{r,k}$ дорівнює 3, а твірної матриці $G_{n,k}$ – 6.

Згідно з правилом побудови підматриці $C_{r,k}$ кількість одиниць у кожному її рядку має бути не меншою ніж $d_{min} - 1 = 3 - 1 = 2$, а кодова від-

стань між рядками — не меншою ніж $d_{\min} - 2 = 3 - 2 = 1$. Тому з триелементних комбінацій для підматриці $C_{r,k}$ вибираємо тільки ті, які задовольняють ці умови, тобто 101, 011, 110.

Оскільки як інформаційна підматриця E_k твірної матриці $G_{n,k}$ вибирається одинична підматриця, дописавши до неї перевірну підматрицю, дістанемо твірну матрицю лінійного систематичного групового коду, здатного виправляти однократні помилки:

$$G_{(6,3)} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix}.$$

За допомогою цієї матриці визначаємо всі вісім комбінацій даного коду:

1. 000000;
2. 100101;
3. 010011;
4. 001110;
5. 110110 (2 ⊕ 3);
6. 101011 (2 ⊕ 4);
7. 011101 (3 ⊕ 4);
8. 111000 (2 ⊕ 3 ⊕ 4).

2. Побудувати перевірну матрицю лінійного систематичного двійкового групового коду, здатного виправляти однократні помилки ($\nu_{вп} = 1$) при передачі восьми повідомлень. Закодувати за допомогою перевірної матриці цього коду повідомлення 101 і 001.

Розв'язання. Для побудови перевірної матриці лінійного систематичного двійкового групового коду, здатного виправляти однократні помилки, скористаємось твірною матрицею, побудованою в попередній задачі.

Перевірна матриця H повинна мати $r = 3$ рядки та $n = 6$ стовпців. Вона складається з двох підматриць: $D_{(3,3)}$, що містить по три стовпці та рядки, кожний рядок якої відповідає стовпцю перевірної підматриці $C_{(3,3)}$ твірної матриці $G_{(6,3)}$; одиничної підматриці $E_{(3)}$. Отже,

$$H_{(6,3)} = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Перевірні елементи коду згідно з цією матрицею визначаються як

$$b_1 = a_1 \oplus a_3;\ b_2 = a_2 \oplus a_3;\ b_3 = a_1 \oplus a_2.$$

Користуючись перевірною матрицею $H_{(6,3)}$, виконуємо кодування повідомлення 101 і 001, для чого визначаємо перевірні елементи для них. Для повідомлення 101: $b_1 = 1 \oplus 1 = 0$; $b_2 = 0 \oplus 1 = 1$; $b_3 = 1 \oplus 0 = 1$; для повідомлення 001: $b_1 = 0 \oplus 1 = 1$; $b_2 = 0 \oplus 1 = 1$; $b_3 = 0 \oplus 0 = 0$.

Таким чином, кодові комбінації для заданих повідомлень матимуть вигляд 101011 і 001110.

3. Користуючись твірною матрицею лінійного систематичного двійкового групового (n, k)-коду при $k = 4$ та $n = 7$, що виправляє однократні помилки:

$$G_{(7,4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix},$$

побудувати перевірну матрицю $H_{(7,3)}$ й закодувати за її допомогою комбінації 0010, 1010, 1110 двійкового простого коду. Визначити синдром для виправлення однократних помилок у комбінаціях заданого коду. Навести приклад виправлення однократної помилки.

Розв'язання. Згідно з правилом побудови перевірної матриці остання матиме $r = n - k = 7 - 4 = 3$ рядки та $n = 7$ стовпців і складатиметься з двох підматриць: $D_{(4,3)}$ з $k = 4$ стовпцями та $r = 3$ рядками, кожний рядок якої відповідає транспонованому стовпцю перевірної підматриці твірної матриці $G_{(7,4)}$; одиничної матриці $E_{(3)}$. Отже,

$$H_{(7,3)} = \begin{bmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Перевірні елементи коду згідно з цією матрицею визначаються як

$$b_1 = a_1 \oplus a_3 \oplus a_4;\ b_2 = a_1 \oplus a_2 \oplus a_4;\ b_3 = a_2 \oplus a_3 \oplus a_4.$$

Користуючись перевірною матрицею $H_{(7,3)}$, кодуємо задані комбінації простого коду. Для цього визначаємо перевірні елементи:

для комбінації 0010 $b_1 = 0 \oplus 1 \oplus 0 + 1$; $b_2 = 0 \oplus 0 \oplus 0 = 0$; $b_3 = 0 \oplus 1 \oplus 0 = 1$. Комбінацією систематичного коду буде 0010101;

для комбінації 1010 $b_1 = 1 \oplus 1 \oplus 0 = 0$; $b_2 = 1 \oplus 0 \oplus 0 = 1$; $b_3 = 0 \oplus 1 \oplus 0 = 1$. Комбінація систематичного коду має вигляд 1010011;

для комбінації 1110 $b_1 = 1 \oplus 1 \oplus 0 = 0$; $b_2 = 1 \oplus 1 \oplus 0 = 0$; $b_3 = 1 \oplus 1 + 0 = 0$. Комбінацією систематичного коду буде 1110000.

На приймальному боці для виявлення та виправлення однократної помилки в прийнятій кодовій комбінації систематичного коду виконується перевірка — визначається синдром помилки. Для матриці $H_{(7,3)}$ маємо

$$S_1 = a_1 \oplus a_3 \oplus a_4 \oplus b_1;\ S_2 = a_1 \oplus a_2 \oplus a_4 \oplus b_2;$$
$$S_3 = a_2 \oplus a_3 \oplus a_4 \oplus b_3.$$

Знайдемо та виправимо однократну помилку, наприклад, у комбінації 1010101 систематичного коду. Для цього визначимо кодовий синдром помилки:

$$S_1 = 1 \oplus 1 \oplus 0 \oplus 1 = 1;\ S_2 = 1 \oplus 0 \oplus 0 \oplus 0 = 1;\ S_3 = 0 \oplus 1 \oplus 0 \oplus 1 = 0,$$

тобто синдром має вигляд 011, що відповідає першому стовпцю перевірної матриці $H_{(7,3)}$. Синдром показує, що помилка знаходиться в першому розряді прийнятої кодової комбінації. Для її виправлення інвертуємо значення цього розряду, тобто замість 1 записуємо 0. Виправлена комбінація систематичного коду матиме вигляд 0010101.

4. Побудувати коди-супутники для комбінацій A та B лінійного систематичного двійкового групового коду, що виправляє одно- та двократні помилки, якщо $A = 00011$, $B = 11100$.

Розв'язання. Відомо, що цей код виправляє двократні помилки ($\nu_{вп} = 2$). Тому будуємо коди-супутники, які відрізняються від заданих кодових комбінацій на e_j з вагами $w = 1$ і 2. Для цього користуємося табл. 8.3, а результати заносимо в табл. 8.8.

Таблиця 8.8

e_j	A = 00011	B = 11100	e_j	A = 00011	B = 11100
00001	00010	11101	00110	00101	11010
00010	00001	11110	01001	01010	10101
00100	00111	11000	01010	01001	10110
01000	01011	10100	01100	01111	10000
10000	10011	01100	10001	10010	01101
			10010	10001	01110
00011	00000	11111	10100	10111	01000
00101	00110	11001	11000	11011	00100

Прийняті кодові комбінації A та B з помилками в одному або двох розрядах відповідатимуть кодам-супутникам, які належать цим комбінаціям, і тому будуть розшифровані як задані. Наприклад, прийнята спотворена комбінація 11011 згідно з табл. 8.8 розшифровується як повідомлення A, а комбінація 10100 — як повідомлення B.

5. Побудувати твірну матрицю лінійного систематичного двійкового групового коду, здатного виправляти однократні помилки ($v_{вп}$ = 1) при передачі 64 повідомлень ($N_д$ = 64).

6. Побудувати перевірну матрицю коду, заданого в попередній задачі, й навести приклад кодування за допомогою цієї матриці.

7. Закодувати лінійним систематичним двійковим груповим кодом, здатним виправляти однократні помилки, комбінацію ДДК, що відповідає порядковому номеру поточного дня тижня. Виправити за допомогою перевірної матриці будь-яку поодиноку помилку в утвореній комбінації систематичного коду.

8. Побудувати коди-супутники для комбінацій 10010101 і 10111010 лінійного систематичного двійкового групового коду, здатного виправляти одно- та двократні помилки.

9. Визначити, які з комбінацій 1101001, 0110011, 0111100, 0011110 лінійного систематичного двійкового групового блокового (7, 4)-коду містять помилку, коли відомо, що код побудовано за твірною матрицею

$$G_{(7,4)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

10. Закодувати двійковим кодом Хеммінга комбінацію A = 10111 двійкового простого коду та показати на прикладі виправлення будь-якої однократної помилки. Визначити надмірність коду Хеммінга.

Р о з в ' я з а н н я. Виконуємо кодування заданої комбінації A. Згідно з табл. 8.4 при k = 5 мінімальна кількість перевірних елементів r = 4 становить $n = k + r = 5 + 4 = 9$. Перевірні елементи знаходяться на позиціях 1,2,4 та 8 [див. перевірну матрицю (8.12)].

Записавши кодовий вектор коду Хеммінга у вигляді $u_1u_2 1u_4 011u_8 1$, з урахуванням матриці (8.12) і рівняння (8.16) визначаємо значення u_1, u_2, u_4, u_8:

$$u_1 = u_3 \oplus u_5 \oplus u_7 \oplus u_9 = 1 \oplus 0 \oplus 1 \oplus 1 = 1; \; u_2 = u_3 \oplus u_6 \oplus u_7 = 1 \oplus 1 \oplus 1 = 1;$$
$$u_4 = u_5 \oplus u_6 \oplus u_7 = 0 \oplus 1 \oplus 1 = 0; \; u_8 = u_9 = 1.$$

Тоді комбінація коду Хеммінга матиме вигляд 111001111.

Щоб виконати декодування цієї комбінації з виправленням однократної помилки, припустимо, що при передачі сталося спотворення і замість 111001111 була прийнята кодова комбінація 111001011.

Для виявлення та виправлення помилки зробимо ті самі перевірки на парність, що й при кодуванні, але з урахуванням перевірних елементів, тобто знайдемо синдром помилки згідно з перевірною матрицею (8.12) і системою рівнянь (8.15):

$$S_1 = u_1 \oplus u_3 \oplus u_5 \oplus u_7 \oplus u_9 = 1 \oplus 1 \oplus 0 \oplus 0 \oplus 1 = 1;$$
$$S_2 = u_2 \oplus u_3 \oplus u_6 \oplus u_7 = 1 \oplus 1 \oplus 1 \oplus 0 = 1;$$
$$S_3 = u_4 \oplus u_5 \oplus u_6 \oplus u_7 = 0 \oplus 0 \oplus 1 \oplus 0 = 1;$$
$$S_4 = u_8 \oplus u_9 = 1 \oplus 1 = 0.$$

Маємо синдром 0111. Отже, спотворено елемент за номером $0111_2 = 7_{10}$, тобто елемент u_7. Виправляємо його за допомогою інверсії: замість помилкового елемента $u_7 = 0$ записуємо значення $u_7 = 1$ і дістаємо правильну кодову комбінацію 111001111.

Надмірність коду Хеммінга $R_{над} = r/n = 4/9$.

11. Побудувати перевірну матрицю двійкового коду Хеммінга для k = 7 і її допомогою закодувати комбінації 0110010 та 1000111 двійкового простого коду. Показати на прикладі виправлення будь-якої однократної помилки в утворених комбінаціях коду Хеммінга та визначити надмірність цього коду.

12. Закодувати двійковим кодом Хеммінга комбінацію двійкового простого коду та виправити будь-яку однократну помилку, якщо комбінацією простого коду є запис поточного року в двійковій системі числення.

13. Закодувати двійковим кодом Хеммінга восьмиелементну кодову комбінацію та виправити будь-яку однократну помилку, якщо ця комбінація є записом останніх двох цифр поточного року в ДДК 7 2 1 1.

14. Побудувати двійковий код Хеммінга для виправлення однократної помилки при k = 9 і 17. Побудувати такий самий код для передачі прозорядної (k = 10) інформаційної комбінації 1101011101. Показати процес виявлення та виправлення помилки, що виникла в дев'ятому розряді комбінації коду Хеммінга.

15. Закодувати двійковим циклічним кодом, що виправляє однократні помилки, комбінацію 0111 двійкового простого коду і виправити будь-яку однократну помилку в утвореній комбінації циклічного коду. Визначити надмірність цього коду.

Р о з в ' я з а н н я. Щоб закодувати задану комбінацію циклічним кодом, необхідно вибрати твірний поліном $P(x)$. Його степінь визначається числом r перевірних елементів у комбінації циклічного коду, причому значення r при d_{min} = 3 можна знайти з виразу $2^r - 1 \ge n$ або $2^r \ge k + r + 1$. Так, при k = 4 маємо r = 3. Користуючись табл. 8.5, вибираємо такий твірний поліном: $P(x) = x^3 + x + 1$.

Виконаємо кодування початкової комбінації двійкового простого коду:

$$Q(x) = x^2 + x + 1 \rightarrow 0111.$$

З цією метою:

• помножимо $Q(x)$ на x^r. Оскільки r = 3, дістаємо $Q(x)x^3 = (x^2 + x + 1)x^3 = x^5 + x^4 + x^3$;

• поділимо $Q(x)x^3$ на $P(x)$ для визначення остачі $R(x)$, коефіцієнти при степенях x якої є перевірними елементами комбінації циклічного коду. Отже,

$$\begin{array}{l} \oplus\ x^5 + x^4 + x^3 \quad \big|\, x^3 + x + 1 \\ \underline{x^5 + x^3 + x^2} \quad\quad\ \big|\, x^2 + x \\ \oplus\ x^4 + x^2 \\ \underline{x^4 + x^2 + x} \\ x \end{array}$$

Дістаємо остачу $R(x) = x$, якій відповідає трирозрядний ($r = 3$) вектор 010;

• додамо остачу $R(x)$ до $Q(x)x^3$ й утворимо комбінацію $F(x) = Q(x)x^3 + R(x) = x^5 + x^4 + x^3 + x \to F = 0111010$ двійкового циклічного коду.

Припустимо, що при передачі інформації виникла однократна помилка, вектор якої $E(x) = x^3 \to 0001000$. Тоді поліном прийнятої комбінації циклічного коду $F'(x) = F(x) + E(x) = x^5 + x^4 + x \to 0110010$.

На приймальному боці декодер виконує перевірне ділення комбінації $F'(x)$ на поліном $P(x)$, використаний при кодуванні інформації:

$$\begin{array}{l} \oplus\ 0110010 \quad \big|\, 1011 \\ \underline{1011} \quad\quad\ \big|\, 0111 \\ \oplus\ 1111 \\ \underline{1011} \\ \oplus\ 1000 \\ \underline{1011} \\ 011 \to x + 1 \end{array}$$

Отже, остача $R(x) = x + 1$, або $R = 011$.

Оскільки остача від ділення не дорівнює нулю, робимо висновок про наявність помилки в прийнятій кодовій комбінації.

Для визначення місця помилки користуємося методом гіпотез:

• будуємо гіпотезу про помилку в молодшому розряді комбінації $F'(x)$, тобто вважаємо, що вектор помилки $E_1(x) = 1 \to E_1 = 0000001$. Виконавши додавання $F' \oplus E_1$ і поділивши результат на поліном $P(x)$ з метою підтвердження (в разі нульової остачі) або спростування (в разі ненульової остачі) гіпотези, дістанемо

$$\begin{array}{lll} \oplus\ \begin{array}{l}0110010\\ \underline{0000001}\\ 0110011\end{array} & \oplus\ \begin{array}{l}0110011\\ \underline{1011}\\ \oplus\ 1111\\ \underline{1011}\\ \oplus\ 1001\\ \underline{1011}\\ 010 \to R(x) = x,\end{array} & \begin{array}{|l}1011\\ \hline 0111\end{array} \end{array}$$

тобто остача ненульова й гіпотеза відкидається;

• будуємо гіпотезу про помилку в другому розряді комбінації F', тобто вважаємо, що вектор помилки $E_2(x) = x \to E_2 = 0000010$. Виконавши додавання $F' \oplus E_2$ та поділивши результат на поліном $P(x)$ з метою підтвердження або спростування гіпотези, матимемо

$$\begin{array}{lll} \oplus\ \begin{array}{l}0110010\\ \underline{0000010}\\ 0110000\end{array} & \oplus\ \begin{array}{l}0110000\\ \underline{1011}\\ \oplus\ 1110\\ \underline{1011}\\ \oplus\ 1010\\ \underline{1011}\\ 001 \to R(x) = 1,\end{array} & \begin{array}{|l}1011\\ \hline 0111\end{array} \end{array}$$

тобто остача ненульова й гіпотеза відкидається;

• будуємо гіпотезу про помилку в третьому розряді комбінації F', тобто вважаємо, що вектор помилки $E_3(x) = x^2 \to E_3 = 0000100$. Виконавши додавання $F' \oplus E_3$ та поділивши результат на поліном $P(x)$ з метою підтвердження або спростування гіпотези, знайдемо

$$\begin{array}{lll} \oplus\ \begin{array}{l}0110010\\ \underline{0000100}\\ 0110110\end{array} & \oplus\ \begin{array}{l}0110110\\ \underline{1011}\\ \oplus\ 1101\\ \underline{1011}\\ \oplus\ 1100\\ \underline{1011}\\ 111 \to R(x) = x^2 + x + 1,\end{array} & \begin{array}{|l}1011\\ \hline 0111\end{array} \end{array}$$

тобто остача ненульова й гіпотеза відкидається;

• будуємо гіпотезу про помилку в четвертому розряді комбінації F, тобто вважаємо, що вектор помилки $E_4(x) = x^3 \to E_4 = 0001000$. Виконавши додавання $F' + E_4$ та поділивши результат на поліном $P(x)$ з метою підтвердження або спростування гіпотези, дістанемо

$$\begin{array}{lll} \oplus\ \begin{array}{l}0110010\\ \underline{0001000}\\ 0111010\end{array} & \oplus\ \begin{array}{l}0111010\\ \underline{1011}\\ \oplus\ 1101\\ \underline{1011}\\ 0 \to R(x) = 0,\end{array} & \begin{array}{|l}1011\\ \hline 011\end{array} \end{array}$$

тобто помилка дійсно є в четвертому розряді, а початкова комбінація циклічного коду має вигляд $F = 0111010 \to F(x) = x^5 + x^4 + x^3 + x$.

Надмірність цього коду $R_{над} = 3/7$.

16. Закодувати двійковим циклічним кодом, що виправляє однократні помилки, комбінацію $Q(x) = x^4 + x$ простого двійкового коду та виправити будь-яку однократну помилку в утвореній комбінації циклічного коду. Визначити надмірність цього коду.

Розв'язання. Щоб закодувати комбінацію $Q(x) = x^4 + x$ циклічним кодом, необхідно вибрати твірний поліном $P(x)$. Його степінь визначається кількістю r перевірних елементів, яку знаходимо з виразу $2^r - 1 \geq n$ (для $d_{min} = 3$). При $k = 5$ дістаємо $r = 4$ і вибираємо поліном четвертого степеня $P(x) = x^4 + x + 1$ (див. табл. 8.5).

Виконаємо кодування початкової комбінації $Q(x)$. З цією метою:

• помножимо $Q(x)$ на x^r. Оскільки $r = 4$, дістаємо $Q(x)x^4 = (x^4 + x)x^4 = x^8 + x^5$;

• поділимо $Q(x)x^4$ на $P(x)$ для визначення частки від ділення $C(x)$.

Отже,

$$\begin{array}{l} \oplus\dfrac{x^8 + x^5}{x^8 + x^5 + x^4} \quad \Big|\dfrac{x^4 + x + 1}{x^4 + 1 \to C(x)} \\ \oplus\dfrac{x^4}{x^4 + x + 1} \\ \hline x + 1 \end{array}$$

• помножимо $C(x)$ на $p(x)$ і знайдемо кодову комбінацію $F(x) = C(x)P(x) = (x^4 + 1)(x^4 + x + 1) = x^8 + x^4 + x^5 + x + x^4 + 1 = x^8 + x^5 + x + 1$.

Припустимо, що при передачі інформації виникла однократна помилка, вектор якої $E(x) = x^2$. Тоді поліном прийнятої комбінації циклічного коду $F'(x) = x^8 + x^5 + x^2 + x + 1$.

На приймальному боці декодер виконує перевірне ділення комбінації $F'(x)$ на поліном $P(x)$:

$$\begin{array}{l} \oplus\dfrac{x^8 + x^5 + x^2 + x + 1}{x^8 + x^5 + x^4} \quad \Big|\dfrac{x^4 + x + 1}{x^4 + 1} \\ \oplus\dfrac{x^4 + x^2 + x + 1}{x^4 + x + 1} \\ \hline x^2 \to 0100,\ w = 1 \end{array}$$

Отже, остача $R(x)$ не дорівнює нулю; тому робимо висновок про наявність помилки в прийнятій кодовій комбінації $F'(x)$. Оскільки вага остачі ($w = 1$) не перевищує кількості помилок $\nu_{вп}$, яку може виправляти даний код, для виправлення помилки та утворення початкової комбінації $F(x)$ циклічного коду прийняту кодову комбінацію $F'(x)$ треба додати до остачі $R(x)$:

$$F(x) = F'(x) \oplus R(x) = x^8 + x^5 + x^2 + x + 1 + x^2 = x^8 + x^5 + x + 1.$$

Надмірність цього коду $R_{над} = 4/9$.

17. Закодувати двійковим циклічним кодом, що виправляє однократні помилки, комбінацію $Q(x) = x^7 + x^4 + 1$ двійкового коду і виправити будь-яку однократну помилку. Визначити надмірність цього коду.

18. Закодувати двійковим циклічним кодом і виправити будь-яку однократну помилку для повідомлення, що відповідає порядковому номеру поточного дня тижня в ДДК 6 3 2 1. Визначити надмірність двійкового циклічного коду.

19. Закодувати двійковим циклічним кодом, що виявляє трикратні помилки ($d_{min} = 4$), комбінацію $Q(x) = x^5 + x^2 + x$ двійкового простого коду та виявити будь-яку трикратну помилку. Визначити надмірність двійкового циклічного коду.

20. Побудувати твірну матрицю двійкового циклічного коду, здатного виправляти однократну помилку, твірний поліном якого $P(x) = x^4 + x^3 + 1$.

21. Визначити кількість перевірних елементів двійкового циклічного коду, що виправляє однократну помилку, якщо кількість його інформаційних елементів $k = 9$.

22. Визначити, яка з чотирьох прийнятих комбінацій 0111011, 1011010, 1111110, 0111010 двійкового коду помилкова, якщо твірний поліном коду $P(x) = x^3 + x + 1$. Виправити виявлену помилку.

23. Знайти твірний поліном $P(x)$ двійкового коду БЧХ, здатного виправляти трикратні помилки, для передачі 25 повідомлень у двійковому коді.

Розв'язання. Мінімальна кодова відстань коду БЧХ, здатного виправляти три помилки, $d_{min} = 2\nu_{вп} + 1 = 2 \cdot 3 + 1 = 7$. Для передачі 25 повідомлень досить мати $k = 5$ ($2^5 > 25$) інформаційних елементів.

Щоб визначити твірний поліном $P(x)$ коду БЧХ з $d = 7$ і $k = 5$, скористуємося табл. 8.7. Мінімальна довжина коду з заданими параметрами становить $n = 15$ ($k = 5$, $r = 10$ і $d = 7$), твірний поліном якого $P(x) = 2467$ у вісімковій системі числення [або $P(0,1) = 010100110111$ у двійковій системі числення] визначається виразом $P(x) = x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1$.

24. Побудувати двійковий код БЧХ завдовжки $n = 15$, що виправляє три помилки або виявляє до шести помилок, і виявити шість помилок. Визначити надмірність цього коду.

Розв'язання. Твірний поліном $P(x)$ коду БЧХ визначається як НСК добутку $L = \nu_{вп} = 3$ мінімальних поліномів. Найбільше значення степеня l мінімального полінома можна знайти з виразу $n = 2^l - 1$, тобто $l = 4$ ($15 = 2^4 - 1$). Взявши з табл. 8.6 три мінімальних поліноми $[M_1(x), M_3(x), M_5(x)]$, дістанемо

$$\begin{array}{l} P(x) = \text{НСК}\,[M_1(x)M_3(x)M_5(x)] = \\ = (x^4 + x + 1)(x^4 + x^3 + x^2 + x + 1)(x^2 + x + 1) = \\ = x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1 \to 10100110111. \end{array}$$

Найбільший степінь цього полінома визначає кількість перевірних елементів коду ($r = 10$). Згідно з виразом $k = n - r$ кількість інформаційних елементів становить $k = 15 - 10 = 5$.

Закодуємо комбінацію двійкового простого коду з $k = 5$ [$Q(x) = x^4 + x^2 + 1 \to 10101$] кодом БЧХ, для чого:

• помноживши $Q(x)$ на x^{10}, дістанемо $(x^4 + x^2 + 1)x^{10} = x^{14} + x^{12} + x^{10}$;

• поділивши $Q(x)x^{10}$ на $P(x)$, знайдемо остачу $R(x)$, тобто

$$\begin{array}{l} \oplus\dfrac{x^{14} + x^{12} + x^{10}}{x^{14} + x^{12} + x^9 + x^8 + x^6 + x^5 + x^4} \quad \Big|\dfrac{x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1}{x^4 + 1} \\ \oplus\dfrac{x^{10} + x^9 + x^8 + x^6 + x^5 + x^4}{x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1} \\ \hline x^9 + x^6 + x^2 + x + 1 \to R(x); \end{array}$$

• додавши $Q(x)x^{10}$ до остачі $R(x)$, матимемо комбінацію коду БЧХ

$$F(x) = x^{14} + x^{12} + x^{10} + x^9 + x^6 + x^2 + x + 1 \to \underbrace{10101}_{k}\ \underbrace{1001000111}_{r}.$$

При виявленні шести помилок [вектор шестикратної помилки $E(x) = x^{13} + x^{12} + x^{10} + x^7 + x^3 + 1$] прийнята комбінація коду БЧХ має вигляд $x^{14} + x^{13} + x^9 + x^7 + x^6 + x^3 + x^2 + x$.

На приймальному боці декодер виконує перевірне ділення комбінації на поліном $P(x)$, використаний при кодуванні інформації:

$$\begin{array}{l} x^{13} + x^9 + x^7 + x^6 + x^3 + x^2 + x \quad \Big|\dfrac{x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1}{x^4 + x^3 + x^2 + x + 1} \\ x^{12} + x^9 + x^8 + x^6 + x^5 + x^4 \\ \hline x^{13} + x^{12} + x^8 + x^7 + x^5 + x^4 + x^3 + x^2 + x \\ x^{13} + x^{11} + x^8 + x^7 + x^5 + x^4 + x^3 \\ \hline \oplus\dfrac{x^{12} + x^{11} + x^2 + x}{x^{12} + x^{10} + x^7 + x^6 + x^4 + x^3 + x^2} \\ \oplus\dfrac{x^{11} + x^{10} + x^7 + x^6 + x^4 + x^3 + x}{x^{11} + x^9 + x^6 + x^5 + x^3 + x^2 + x} \\ \oplus\dfrac{x^{10} + x^9 + x^7 + x^5 + x^4 + x^2}{x^{10} + x^8 + x^5 + x^4 + x^2 + x + 1} \\ \hline x^9 + x^8 + x^7 + x + 1 \to R(x). \end{array}$$

Отже, остача $R(x)$ не дорівнює нулю, що свідчить про наявність помилок у прийнятій кодовій комбінації $F(x)$. Здатність коду не дає змоги виправити шість помилок.

Надмірність цього коду $R_{над} = 10/15 = 2/3$.

25. Побудувати двійковий циклічний код БЧХ завдовжки $n = 31$, здатний виправляти п'ять помилок.

26. Побудувати двійковий циклічний код БЧХ завдовжки $n = 63$, здатний виправляти семикратну помилку. Показати процес виправлення семикратної помилки.

27. Побудувати двійковий код Файра, здатний виправляти пачки помилок завдовжки до $b = 5$. Визначити надмірність цього коду.

28. Закодувати комбінацію 10011 двійкового простого коду двійковим кодом з багатократним повторенням, здатним виправляти двократні помилки. Виправити будь-яку двократну помилку та визначити надмірність цього коду.

Розв'язання. Кількість повторень n_R при виправленні двох помилок ($v_{вп} = 2$) визначається з виразу $d = n_R + 1$. Мінімальна кодова відстань при $v_{вп} = 2$ становить $d_{min} = 2v_{вп} + 1 = 2 \cdot 2 + 1 = 5$; тому $n_R = d - 1 = 5 - 1 = 4$.

Комбінація двійкового коду з багатократним повторенням матиме вигляд 10011 10011 10011 10011.

Припустимо, що при передачі комбінації двійкового коду з багатократним повторенням виникла двократна помилка, вектор якої має вигляд 00100000000100000000. Тоді прийнятою кодовою комбінацією буде 10111100111101110011.

На приймальному боці декодер розбиває цю комбінацію на чотири частини по 5 біт у кожній і виконує порозрядне порівняння їх:

10111
10011
11011
10011,

завдяки чому виявляються помилки в третьому та четвертому розрядах. Застосувавши «голосування за більшістю», можна виправити ці помилки. Виправлена комбінація початкового двійкового коду матиме вигляд 10011.

Надмірність цього коду $R_{над} = 15/20 = 3/4$.

29. Закодувати комбінацію 0110011 двійкового простого коду двійковим кодом із багатократним повторенням, здатним виправляти трикратні помилки. Виправити будь-яку трикратну помилку та визначити надмірність цього коду.

30. Закодувати комбінацію 1110001100110011 двійкового простого коду двовимірним двійковим ітеративним кодом, здатним виправляти однократні помилки. Виправити будь-яку однократну помилку та визначити надмірність цього коду.

31. Закодувати комбінацію 101110 двійкового простого коду двоступінчастим каскадним двійковим кодом, здатним виправляти двократні помилки, та визначити надмірність цього коду.

32. Закодувати двійковим ланцюговим кодом із кроком $t_{кр} = 4$ інформаційну послідовність 1111111000000111000. Виправити будь-яку чотирикратну помилку та визначити надмірність цього коду.

Розв'язання. Для побудови ланцюгового коду з кроком $t_{кр} = 4$ та виправленням чотирьох помилок треба виконати такі дії:

- записати інформаційну послідовність двійкових елементів;
- визначити перевірні елементи додаванням за модулем 2 двох інформаційних елементів через крок $t_{кр} = 4$ (рис. 8.3);
- утворену послідовність інформаційних і перевірних елементів передати в канал зв'язку.

Припустимо, що при передачі виникла чотирикратна помилка в прийнятій послідовності інформаційних елементів і на декодер подається помилкова послідовність з чотирма помилками в інформаційних розрядах. На приймальному боці декодером виконуються такі операції:

- інформаційні та перевірні розряди прийнятої послідовності розділяються;
- з прийнятої інформаційної послідовності формуються контрольні елементи за правилом формування перевірних елементів на передачі;
- здобуті контрольні та прийняті з каналу перевірні елементи порівнюються між собою порозрядно, незбіг їх указує на наявність помилок.

Виправлення помилок виконується так. Якщо не збігаються дві пари перевірних і контрольних елементів, відстань між якими дорівнює кроку додавання $t_{кр} = 4$, то інвертується той інформаційний елемент, який утворює ці елементи (рис. 8.4).

Надмірність коду $R_{над} = 1/2$.

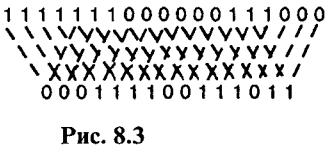
33. Закодувати двійковим ланцюговим кодом із кроком додавання $t_{кр} = 4$ інформаційну послідовність 01111001010101110010. Виправити будь-яку двократну помилку та визначити надмірність цього коду.

34. Закодувати двійковим ланцюговим кодом із кроком додавання $t_{кр} = 4$ інформаційну послідовність 10011110000011111000101010111. Виправити будь-яку пачку з п'яти помилок.

35. Закодувати двійковим ланцюговим кодом, виправити будь-яку однократну помилку та визначити надмірність коду при передачі всіх цифр поточного року в ДДК 5 3 2 1 із кроком додавання $t_{кр} = 4$.

36. Закодувати комбінацію A = 5426713 недвійкового коду з алфавітом $q = 8$ узагальненим кодом Хеммінга та виправити будь-яку однократну помилку.

Розв'язання. Для побудови узагальненого коду Хеммінга необхідно задати проверівну матрицю H. Перший кількість її перевірних



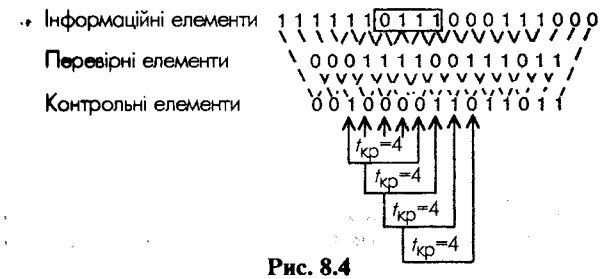


Рис. 8.4

елементів $r = 2$ [інформаційна ємність одного елемента комбінації цього коду становить $l = 3$ біт ($q = 2^l = 8 = 2^3$), тому двох перевірних елементів досить, щоб виправити однократну помилку].

Виходячи з виразу $n = (q^r - 1)/(q - 1)$ при $q = 8$ маємо $n = 5$. Отже, матриця H згідно з (8.39) має розмір $r \otimes n$, тобто $2 \otimes 9$. Довільно вибираємо $\delta = 4$ й будуємо цю матрицю

$$H = \begin{bmatrix} 4 & 0 & 4 & 4 & 4 & 4 & 4 & 4 & 4 \\ 0 & 4 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{bmatrix}.$$

Кодовий вектор при передачі кодової комбінації $\bar{A} = 5426713$ має вигляд $\bar{X} = (b_1\ b_2\ 5426713)$. Обчислюємо значення перевірних елементів за (8.41).

Виконавши множення та додавання згідно з табл. 5.3 та 5.2, дістанемо $b_1 = 0$, $b_2 = 3$. Таким чином, на передачу подається кодовий вектор $\bar{X} = (035426713)$.

Виправимо однократну помилку. Для цього припустимо, що вектор прийнятої комбінації $\bar{Y} = \bar{X} + \bar{E} = (035426113)$, де $\bar{E}$ — вектор помилки [$\bar{E} = (000000600)$].

Декодування починаємо з обчислення перевірного синдрому $\bar{S} = L\bar{e}$, де L — локатор помилки; e — її значення. Отже,

$$\bar{S} = H\bar{Y}^T = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}.$$

Відповідно до співвідношення $e = S_1/\delta$ маємо $e = 5/4 = 6$, а локатор помилки $\bar{L} = \bar{S}/e$ набуває вигляду

$$\bar{L} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} \Big/ 6 = \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Виконуючи впорядкований перебір стовпців матриці H і порівнюючи їх з локатором $\bar{L}$, виявляємо за збігом, що спотвореним є сьомий елемент у комбінації $\bar{Y}$, значення якого $y_7' = 1$. Для виправлення помилки до значення y_7' прийнятої кодової комбінації треба додати значення помилки $e = 6$. Тоді істинне значення сьомого елемента становитиме

$$y_7 = y_7' + e = 1 + 6 = 7.$$

Після заміни спотвореного значення y_7' елемента на істинне видаємо одержувачеві повідомлення інформаційну частину $\bar{A} = 5426713$ виправленої кодової комбінації.

37. Закодувати комбінацію $\bar{A} = 130267$ недвійкового коду з алфавітом $q = 8$ узагальненим кодом Хеммінга та виправити будь-яку однократну помилку.

38. Закодувати комбінацію $\bar{A} = 5671204322$ недвійкового коду з алфавітом $q = 8$ узагальненим кодом Хеммінга та виправити будь-яку однократну помилку.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що таке коректувальна здатність коду?
2. Як класифікуються двійкові коди, що виправляють помилки?
3. Яка особливість побудови лінійного систематичного двійкового групового коду?

4. Для чого в лінійному систематичному груповому коді використовуються твірна та перевірна матриці?
5. Що таке кодовий синдром і як він визначається?
6. Як виявляються та виправляються помилки в лінійному систематичному двійковому груповому коді?
7. Чим відрізняються вкорочені лінійні систематичні двійкові групові коди від повних?
8. Для чого застосовуються коди-супутники?
9. Як визначається склад перевірних елементів у двійковому коді Хеммінга?
10. Де розміщуються перевірні елементи в двійковому коді Хеммінга?
11. Чим різняться двійкові коди Хеммінга з кодовими відстанями $d_{min} = 3$ та 4?
12. Які коди належать до циклічних?
13. Як вибирається твірний поліном у двійкових циклічних кодах?
14. Які є методи побудови двійкових циклічних кодів?
15. Як виявляються та виправляються помилки в двійкових циклічних кодах?
16. Які основні різновиди двійкових циклічних кодів і які властивості їх?
17. Як вибирається твірний поліном $P(x)$ у двійкових циклічних кодах з $d_{min} = 4$?
18. Чим відрізняються вкорочені двійкові циклічні коди від повних та які особливості побудови їх?
19. Як визначається довжина комбінації в двійкових кодах БЧХ?
20. Як вибирається твірний поліном у двійкових кодах БЧХ?
21. Що таке мінімальний поліном?
22. Як побудувати твірну матрицю двійкового коду БЧХ?
23. Яка процедура декодування двійкових кодів БЧХ?
24. Де застосовуються двійкові коди Файра?
25. Яка послідовність побудови двійкового коду Файра?
26. Що таке пачка (пакет) помилок?
27. У чому полягають переваги двійкових кодів Файра над кодами БЧХ?
28. Як будуються двійкові коди з багатократним повторенням?
29. Чим визначається вимірність ітеративних і каскадних кодів?
30. Яка процедура виявлення та виправлення помилок в ітеративних кодах?
31. Як виправляються помилки в каскадному коді?
32. Які коди належать до рекурентних?
33. Яке основне призначення рекурентних кодів і який принцип побудови їх?
34. Чим визначається здатність виправлення ланцюгового коду?
35. Як класифікуються недвійкові коректувальні коди?
36. Чим відрізняється недвійковий код з багатократним повторенням від аналогічного двійкового?
37. Які особливості побудови узагальненого коду Хеммінга?
38. Що визначає локатор помилки в узагальненому коді Хеммінга?
39. Де застосовуються та як будуються коди РС?
40. Як будуються багатовимірні ітеративні коди?
41. Чим відрізняється недвійковий ланцюговий код від аналогічного двійкового?

РОЗДІЛ 9 ЕФЕКТИВНІСТЬ КОДУВАННЯ ТА ПЕРЕДАЧІ ІНФОРМАЦІЇ

При проектуванні та розробленні систем збирання, передавання й оброблення інформації значну увагу приділяють таким чинникам, як вибір коду для захисту інформації під час передачі по каналах зв'язку, можливість її стиснення для збільшення швидкості передачі повідомлень, необхідність додаткового захисту інформації при передачі по каналах із значним рівнем завад та ін. Про все це йтиметься в цьому розділі.

9.1. ВІРОГІДНІСТЬ ПЕРЕДАЧІ КОДОВАНИХ ПОВІДОМЛЕНЬ

Вірогідністю передачі кодованих повідомлень оцінюється відповідність між кодовими комбінаціями, прийнятими на приймальному боці, та комбінаціями, переданими в канал зв'язку.

У свою чергу, оцінка вірогідності обміну інформацією визначається ймовірністю спотворень повідомлень, тобто прийняттям $P_{п}$ помилкового повідомлення при відомій імовірності $P_{е}$ спотворення елемента повідомлення, що передається. Таким чином, вірогідність передачі є похідною характеристикою, яка залежить як від коректувальної здатності коду, так і від типу каналу (лінії) зв'язку та умов передачі по ньому елементів кодової комбінації.

Для двійкових кодів дія завад на елемент кодової комбінації може призвести до різних наслідків у несиметричних і симетричних каналах. Так, у несиметричних каналах імовірності переходу під впливом завад 0 в 1 (P_{01}) та 1 в 0 (P_{10}) не будуть однаковими, а в симетричних вони збігаються. Проте для спрощення розрахунків при оцінюванні вірогідності повідомлень, що передаються, вважатимемо двійковий канал симетричним, тобто $P_{01} = P_{10} = P_{е}$.

Крім того, розрізняють канали з *незалежними помилками*, коли виникнення однієї помилки не залежить від появи іншої

(канал без пам'яті), та канали з *пакетним розподілом помилок*, в яких є залежність імовірності спотворення наступного елемента, що передається, від спотворення попереднього (канал має пам'ять).

Розглянемо оцінку вірогідності при передачі кодованих повідомлень по цих двійкових симетричних каналах.

У разі передачі повідомлень по каналах із незалежними помилками ймовірність відсутності спотворень двійкового елемента повідомлення позначимо як $(1 - P_{е})$. Тоді для двійкової послідовності завдовжки n елементів [12] імовірність правильно прийнятої послідовності

$$P_{пр} = (1 - P_{е})^{n}, \tag{9.1}$$

а імовірність помилки в прийнятій послідовності

$$P_{п} = 1 - (1 - P_{е})^{n}. \tag{9.2}$$

Останній вираз можна подати так:

$$P_{п} = C_{n}^{1}P_{е} + C_{n}^{2}P_{е}^{2} + C_{n}^{3}P_{е}^{3} + \ldots + C_{n}^{\nu}P_{е}^{\nu}, \tag{9.3}$$

де $C_{n}^{\nu} = \dfrac{n!}{\nu!(n-\nu)!}$, а $\nu = 1, 2, 3, \ldots$ — кратність помилки.

Через те, що $P_{е} << 1$, ймовірність помилки в n-елементній кодовій послідовності можна записати як

$$P_{п} \approx nP_{е}.$$

Як відомо, коректувальні коди дають змогу виявляти або виправляти (залежно від кодової відстані) ту чи іншу кількість помилок. Тому для оцінювання ефективності кодів необхідно знати ймовірність виникнення в кодовій комбінації помилок тієї чи якої кратності.

При незалежних помилках імовірність появи ν-кратних помилок визначається за формулою Бернуллі

$$P_{п}(\nu) = C_{n}^{\nu}P_{е}^{\nu}(1 - P_{е})^{n-\nu}. \tag{9.4}$$

Для кодів з $d_{min} > 1$, що використовуються з метою виявлення всіх помилок кратністю $\nu_{в} = d_{min} - 1$ і меншою, ймовірність помилкової (неправильної) комбінації на виході декодера легко знайти, врахувавши, що при передачі комбінації можуть бути лише чотири ситуації:

- комбінація, яка приходить з каналу, прийнята без помилок (правильно) з імовірністю $P_{пр}$;
- комбінація містить не більш як $\nu_{в}$ спотворених елементів, що виявляються кодом (імовірність такої події $P_{в.п}$);

• комбінація містить $\nu > \nu_в$ помилок, але вони розташовані так, що виявляються кодом (імовірність такої події $P'_{в.п}$);

• комбінація містить $\nu > \nu_в$ помилок, які кодом не виявляються з імовірністю $P_{нв.п}$.

Оскільки сума ймовірностей всіх перелічених подій дорівнює 1, а ймовірність $P'_{в.п} \approx 0$,

$$P_п = P_{в.п} + P_{нв.п}. \tag{9.5}$$

З урахуванням того, що спотворені комбінації, які виявляються декодером, не видаються споживачеві інформації, ймовірність здобуття ним помилкових комбінацій оцінюватиметься тільки ймовірністю невиявленої помилки $P_{нв.п}$, яку можна записати у вигляді [12]

$$P_{нв.п} = \sum_{\nu=d_{\min}}^{n} W(w) P_e^{\nu} (1-P_e)^{n-\nu}, \tag{9.6}$$

де $W(w)$ — вагова характеристика коду (кількість його комбінацій вагою w, тобто кількість варіантів помилок, які не виявляються цим кодом); $d_{\min}$ — мінімальна кодова відстань.

Значення $W(w)$ визначається за спеціальними методиками для різних кодів [14] . Так, доведено, що коли двійковий (n, k)-код має кодову відстань d, то

$$W(w) = \begin{cases} = 0, \nu < d; \\ \le \dfrac{C_n^{\nu-t}}{C_n^t}, \nu > d, \end{cases} \tag{9.7}$$

де $t = (d-1)/2$.

З урахуванням (9.7) маємо

$$P_{нв.п} \le \frac{1}{(1-P_e)^t C_{n+t}^t} \left[1 - \sum_{\nu=0}^{2t} C_{n+t}^t P_e^t (1-P_e)^{n+t-\nu} - \right. \\ \left. - \sum_{\nu=n+1}^{n+t} C_{n+t}^{\nu} P_e^{\nu} (1-P_e)^{n+t-\nu} \right]. \tag{9.8}$$

Права частина (9.8) при $0 < P_e < 0{,}5$ обмежена зверху значенням $P_e = 0{,}5$; тому в будь-якому каналі з незалежними помилками

$$P_{нв.п} \le 2^t / C^t_{n+t}.$$

Цю оцінку можна застосувати для коротких кодів з невеликою надмірністю. Для кодів з великою надмірністю бажано користуватися виразом [14]

$$P_{нв.п} \le 2^k (d/n)^d (1 - d/n)^{n-d},$$

k — кількість інформаційних елементів; d — мінімальна кодова відстань; n — довжина коду.

На практиці поширенішою є формула для приблизного оцінювання ймовірності виникнення невиявленої помилки [12]

$$P_{нв.п} \approx \frac{1}{2^r} \sum_{\nu=d}^{n} C_n^{\nu} P_e^{\nu} (1-P_e)^{n-\nu}, \tag{9.9}$$

де r — кількість перевірних елементів коду; ν — кратність помилки; d — мінімальна кодова відстань.

Для кодів з $d > 2$, що використовуються для виправлення всіх помилок кратністю $\nu_{вп} = [(d-1)/2]$, де $[\cdot]$ означає цілу частину й менше, ймовірність $P_п$ виникнення помилкової комбінації на виході декодера можна знайти, врахувавши, що після проходження кодової комбінації по каналу можливими є такі чотири ситуації:

• комбінація прийнята без помилок (правильно) з імовірністю $P_{пр}$;

• комбінація містить не більш як $\nu_{вп}$ спотворених елементів, які виправляються кодом з імовірністю $P_{вп.п}$;

• комбінація містить $\nu > \nu_{вп}$ помилок, розташованих так, що вони виправляються кодом з імовірністю $P'_{вп.п}$;

• комбінація містить $\nu > \nu_{вп}$ помилок, які не виправляються кодом (імовірність такої події $P_{н.вп.п}$).

Оскільки ймовірність $P'_{вп.п} \approx 0$, маємо

$$P_п = P_{вп.п} + P_{н.вп.п}, \tag{9.10}$$

звідки

$$P_{н.вп.п} = P_п - P_{вп.п}.$$

При незалежних помилках імовірність виправлення помилок кратністю до ν кодами, що виправляють помилки, визначається виразом

$$P_{вп.п} = \sum_{i=1}^{\nu} C_n^i P_e^i (1-P_e)^{n-i}. \tag{9.11}$$

З урахуванням виразів (9.2) та (9.11) дістаємо

$$P_{н.вп.п} = 1 - (1-P_e)^n - \sum_{i=1}^{\nu} C_n^i P_e^i (1-P_e)^{n-i}. \tag{9.12}$$

На основі здобутих значень $P_п$ можна визначити коректувальний код, оптимальний для даного каналу.

Двійковий коректувальний код завдовжки n з N комбінаціями називається *оптимальним* для двійкового симетричного каналу, якщо ймовірність $P_п$ виникнення помилкової комбінації на

виході декодера не перевищує такої самої імовірності для будь-якого іншого двійкового коду тієї самої довжини n із тією самою кількістю комбінацій N.

У разі передачі повідомлень по каналах із пакетним розподілом помилок визначення ймовірності їх за формулами (9.3) — (9.12) дають значення, які набагато відрізняються від реальних. Це пояснюється тим, що в цих каналах на проходження сигналів (а це в основному радіоканали) сильно впливають сезонні та добові зміни метеорологічних умов, промислові завади, інтенсивність яких змінюється протягом доби та тижня, тощо [5, 12]. Все це призводить до виникнення в каналах пакетів (пачок) помилок. Визначити ймовірність їх за таких умов досить складно, оскільки необхідно провести дослідження реальних характеристик каналів.

Формула, що дає приблизне значення ймовірності помилок при пакетному розділі їх і передачі двійкової послідовності n елементів [21], має вигляд

$$P_{\text{п}} \approx \frac{P_{\text{e}}}{\bar{l}} \sum_{b=1}^{b_{\max}} \left(1+\frac{n-1}{b}\right) \frac{bP_b}{\sum\limits_{b=1}^{b_{\max}} bP_b}, \tag{9.13}$$

де P_{e} — імовірність спотворення двійкового елемента; $\bar{l}$ — щільність помилок у пакеті, яка визначається відношенням кількості помилок у ньому до довжини пакета b; P_b — умовна ймовірність виникнення пакета помилок завдовжки b.

Для кодів, що виявляють пакети помилок, імовірності помилок можна знайти за формулами [21]:

$$P_{\text{в.п}} \approx \frac{P_{\text{e}}}{\bar{l}} \left\{ \sum_{b=1}^{b_{\max}} \left(1+\frac{n-1}{b}\right) \frac{bP_b}{\sum\limits_{b=1}^{b_{\max}} bP_b} - \frac{1}{2^r} \sum_{b=l_k+1}^{b_{\max}} \left[1+ \frac{n-(2l_k+1)}{b}\right] \frac{bP_b}{\sum\limits_{b=1}^{b_{\max}} bP_b} \right\}; \tag{9.14}$$

$$P_{\text{нв.п}} \approx \frac{1}{2^r} \frac{P_{\text{e}}}{\bar{l}} \sum_{b=l_k+1}^{b_{\max}} \left[1+\frac{n-(2l_k+1)}{b}\right] \frac{bP_b}{\sum\limits_{b=1}^{b_{\max}} bP_b}, \tag{9.15}$$

де l_k довжина пакета помилок, яка виявляється.

Значення усіх величин, які входять у ці формули, дістають експериментально, визначаючи характер розподілу помилок, або беруть із літератури для каналів аналогічного типу.

У [46] реальні канали описано за допомогою двох параметрів: імовірності P_{e} спотворення двійкового елемента та показника α групування помилок. При цьому наближені формули для визначення ймовірності невиявлених помилок мають такий вигляд:

- для кодів, що виявляють помилки,

$$P_{\text{нв.п}} \approx \frac{P_{\text{e}}}{2^r} \left(\frac{n}{d}\right)^{1-\alpha};$$

- для кодів, що виправляють помилки,

$$P_{\text{нв.п}} \approx \left(\frac{n}{\nu+1}\right)^{1-\alpha} P_{\text{e}};$$

- для кодів, що виявляють і виправляють помилки,

$$P_{\text{нв.п}} \approx \frac{\sum\limits_{i=0}^{\nu} C_n^i}{2^r} \left(\frac{n}{d-\nu}\right)^{1-\alpha} P_{\text{e}},$$

де d — мінімальна кодова відстань; ν — кратність помилки, що виправляється; r — кількість перевірних елементів.

Ці формули дають непогані результати при $\nu < 0{,}3n$, де n — кількість елементів кодової комбінації.

9.2. СТИСНЕННЯ ІНФОРМАЦІЇ*

Стиснення інформації застосовується для прискорення та зниження витрат на її оброблення, зберігання й пошук, а також для зменшення ємності пам'яті, зайнятої в ЕОМ.

Під *стисненням інформації* розумітимемо операцію, внаслідок якої певному коду чи повідомленню ставиться у відповідність код або повідомлення меншої довжини.

Способи стиснення інформації поділяють за призначенням, характером і ступенем стиснення, швидкістю та ступенем відновлення початкового стану інформації (втратами).

За *призначенням* розрізняють дві великі групи способів стиснення: для передачі даних і для їх архівації. Різниця між ними полягає в тому, що перші оперують з незначними інформацій-

* Написано спільно з канд. техн. наук *Б. Ю. Жураковським.*

ними масивами (до кількох десятків, сотень байтів), а другі — зі значно більшим обсягом інформації (мегабайти).

За *характером* стиснення інформації розрізняють лінійні, матричні, комбіновані та каскадні способи. До *лінійних* належать способи, за якими стиснення елементів інформаційного масиву виконується в одному з напрямків (горизонтальному або вертикальному). Залежно від цього лінійними способами можуть виконуватися поздовжнє (горизонтальне) та поперечне (вертикальне) стиснення інформації.

До *матричних* належать способи стиснення інформації, за якими елементи інформаційного масиву стискаються з використанням матричного принципу заміни повторюваних елементів.

Комбіновані способи поєднують одночасне використання для стиснення інформаційного масиву двох чи більше лінійних або/та матричних способів.

До *каскадних* належать способи стиснення інформації, за якими воно виконується послідовно різними способами.

За ступенем стиснення інформації розрізняють низькоефективні (з коефіцієнтом стиснення до 1,5), середньоефективні (1,51...3) та високоефективні (понад 3) способи; за швидкістю стиснення/розпаковування — низько-, середньо- та високошвидкісні, при яких швидкість стиснення/розпаковування змінюється від кількох кілобайтів за секунду (низькошвидкісні) до кількох мегабайтів за секунду (високошвидкісні).

За *ступенем відновлення початкового стану інформації* (втратами) способи стиснення поділяють на без відновлення початкового стану інформації, з частковою її втратою та без втрати інформації (з повним її відновленням). Що стосується останнього поділу способів стиснення інформації, то до першої групи належать найпримітивніші, а до другої та третьої груп — складніші й ефективніші способи.

Так, до відомих способів стиснення інформації без відновлення її початкового стану можна віднести стиснення за допомогою поділу кодової комбінації на кілька частин [19, 22, 42] і з порозрядним зсувом [42]. Ці способи застосовуються дуже рідко, оскільки не гарантують повного відновлення стисненої інформації з точки зору неоднозначності утвореної при стисненні послідовності символів.

У той же час способи з частковою втратою інформації мають специфічне застосування [22] , коли часткова її втрата майже не позначається на якості відновлюваної інформації. Тому в підручнику розглядаються тільки способи стиснення, що гарантують повне відновлення стисненої інформації й які широко використовуються в системах збирання, передавання та оброблення інформації.

9.2.1. СПОСОБИ СТИСНЕННЯ ДАНИХ ПРИ ПЕРЕДАЧІ

Розглянемо кілька найпоширеніших способів стиснення даних, які застосовуються при передачі їх. До них належать лінійні, матричні, комбіновані та каскадні способи, що гарантують повне відновлення початкового стану стисненої та переданої інформації.

Лінійні способи. ***Стиснення даних із використанням замість повторень додаткових символів.*** Ці способи ґрунтуються на заміні повторюваних елементів деякими умовними символами. Вони є ефективними в тому разі, коли масиви інформації, які подаються у вигляді рядків або стовпців, розташованих у зростаючому порядку, мають однакові значення елементів в одних і тих самих розрядах, що характерно для техніко-економічної інформації. Таке стиснення даних дає змогу скоротити масив у кілька разів [42].

Так, якщо елементи повторюються на початку рядків (стовпців) відносно попередніх, то замість виключених розрядів у масив уводиться знак поділу ρ , який дає можливість відокремити елементи в згорнутому масиві. При розгортанні замість знака ρ поновлюють всі пропущені розряди, які були до елемента, що знаходився безпосередньо за ρ в стисненому тексті. Запис знаків, які знаходяться після ρ, виконується з кінця рядка (стовпця).

Як приклад розглянемо масив, у якому інформацію записано у вигляді рядків, що складаються з восьми десяткових знаків:

```
3 9 1 4 5 6 8 0   5 6 7 1 8 3 2 9
3 9 1 4 5 6 8 6   5 6 7 1 8 3 4 3
3 9 1 6 7 5 9 6   5 6 7 2 9 4 6 2
3 9 1 4 5 7 2 1   5 6 7 1 8 3 4 8
3 9 1 4 5 6 3 8   5 6 7 1 7 6 3 1
```

Згорнутий масив інформації матиме такий вигляд:

```
3 9 1 4 5 6 8 0   2 9 ρ 4 3 ρ 2 9
ρ 6 ρ 6 7 5 9 6   7 6 2 ρ 1 8 3 4
ρ 4 5 7 2 1 ρ 6   8 ρ 7 6 3 1
3 8 5 6 7 1 8 3
```

Розгортку виконуємо з кінця масиву. При цьому на наступний рядок переходимо або після заповнення рядка, або при знаку поділу ρ:

```
3 9 1 4 5 6 8 0   5 6 7 1 8 3 2 9
. . . . . . . 6   . . . . . . 4 3
. . . 6 7 5 9 6   . . . 2 9 7 6 2
. . . 4 5 7 2 1   . . . 1 8 3 4 8
. . . . . 6 3 8   . . . . 7 6 3 1
```

Після заповнення пропущених цифр за аналогічними розрядами попереднього рядка дістаємо масив інформації, який був до стиснення.

Недоліком цього способу стиснення інформації є неможливість його застосування до впорядкованих масивів, у яких повторювані розряди зустрічаються не на початку рядків (стовпців).

Для згортання масивів, у яких в одному рядку (стовпці) є тільки одна повторювана ділянка, можна використати вищезгаданий спосіб з уведенням додаткового символу К (кінець рядка, стовпця). При цьому розгортання масиву ведеться від К до К. В разі фіксованої довжини рядка (стовпця) всі розряди, які знаходяться між К, разом із пропущеними розрядами мають утворювати повний рядок (стовпець). Так, початковий і згорнений по рядках масиви матимуть вигляд

4 7 5 3 6 8 9 1	4 7 5 3 6 8 9 1
4 7 5 3 6 4 3 2	К ρ 4 3 2 К 3 5
3 5 5 3 6 6 9 0	ρ 6 9 0 К 6 9 ρ
6 9 5 3 6 2 4 1	2 4 1 К 7 ρ 1 0
7 9 5 3 6 2 1 0	К 8 ρ 8 0 2 К 1
8 9 5 3 6 8 0 2	6 ρ К 6 3 ρ К 4
1 6 5 3 6 8 0 2	6 9 ρ 5 К ρ 3 4
6 3 5 3 6 8 0 2	8 К
4 6 9 3 6 8 0 5	
4 6 9 3 6 3 4 8	

Поновлення масиву може виконуватися з початку або з кінця. За наявності в рядку (стовпці) кількох повторюваних ділянок замість ρ вводять спеціальні символи, що вказують необхідну кількість пропусків. При цьому необхідність у символі К, який визначає кінець рядка (стовпця), відпадає. Так, якщо для позначення заданої кількості пропусків у рядку ввести відповідно символи $X-2$, $Y-3$, $Z-5$, то початковий і згорнений масиви десяткових знаків матимуть такий вигляд:

2 3 4 5 6 7 8 5	2 3 4 5 6 7 8 5
6 3 4 5 6 7 9 8	6 *Z* 9 8 *X* 7 *X X*
6 3 7 5 6 7 9 2	2 2 4 *Y Y* 8 9 3
2 4 7 5 6 7 9 2	*X* 2 3 1 1 *Z X Z*
8 9 3 5 6 2 3 1	4 8 9 3 8 *Y* 5 *X*
1 9 3 5 6 2 3 1	2 *Y* 7 *Y* 1 7 2 *X*
1 9 3 5 6 4 8 9	*X* 0
3 8 3 5 6 5 8 9	
2 8 3 5 7 5 8 9	
1 7 2 5 7 5 8 0	

Розгортання масивів при цьому способі стиснення виконується з їх початку або з кінця, а заповнення відповідної кількості пропусків замість додаткових символів X, Y, Z — перенесенням відповідної кількості символів, які знаходяться на однойменних розрядах попереднього рядка:

2 3 4 5 6 7 8 5
6 . . . . . 9 8
. . 7 2
2 4 . . . . .
8 9 3 . . 2 3 1
1 . . . . .
. . . . . 4 8 9
3 8 . . . 5 . .
2 . . . 7 . . .
1 7 2 0

Стиснення інформації виключенням даних, які повторюються в різних файлах [42]. За цим способом виконується така індексація (адресація) даних, яка дає змогу досягти логічної та фізичної незалежності їх. При цьому повторювані дані можуть відновлюватися багатократним зверненням до одного й того самого поля запису.

Стиснення інформації виключенням повторюваних символів [42]. Виконується воно введенням двох додаткових символів, один з яких указує повторення, а інший — кількість їх (кількість літер, що повторюються в десятковій системі числення).

Наприклад, перелік виробів приладобудівного заводу:

1. Генератор сигналів низькочастотний ГЗ-109.
2. Генератор сигналів низькочастотний ГЗ-112/1.
3. Генератор сигналів спеціальної форми Г6-26.
4. Осцилограф С1-81.
5. Осцилограф С1-83.
6. Осцилограф С1-65

у стисненому вигляді можна подати так:

1. Генератор сигналів низькочастотний ГЗ-109.
2. * 36 — 12/1.
3. * 17 спеціальної форми Г6-26.
4. Осцилограф С1-81.
5. * 14 — 3.
6. * 13 — 65,

де * — знак повторення, а цифри після нього — кількість повторюваних знаків.

При відновленні даних розгортання виконується зверху вниз і справа наліво.

Зонне стиснення інформації [42]. У цьому разі враховується те, що запис даних у пам'ять ЕОМ здійснюється за допомогою восьми восьмибітових слів, яким ставляться у відповідність літери та цифри повідомлення. Використання восьмибітового

слова (байта) дає змогу закодувати $2^8 = 256$ знаків, тоді як реальні алфавіти з урахуванням цифр і деяких допоміжних символів містять до 50...60 знаків, тобто для кодування їх потрібні п'яти-шестибітові комбінації та аналогічні структури комірок пам'яті. Два-три біти, що залишаються, не вирішують проблему стиснення інформації, оскільки за їх допомогою можна записати тільки 4...8 знаків.

Однак, якщо використати півбайта для запису $2^4 = 16$ знаків деякого абстрактного алфавіту, а потім закодувати повідомлення в цьому $q = 16$-знаковому алфавіті по $m = 2$ знаки в кодовому слові, то одним байтом можна передавати ті самі $N = q^m = 16^2 = 256$ знаків.

Цей 16-знаковий алфавіт можна побудувати так, щоб 13 якісних ознак використовувалися як основні символи, а три — як допоміжні. Такий алфавіт матиме вигляд

0.	0000	4.	0100	8.	1000	C.	1100
1.	0001	5.	0101	9.	1001	D.	1101
2.	0010	6.	0110	A.	1010	E.	1110
3.	0011	7.	0111	B.	1011	F.	1111

Перші 12 символів умовно називатимемо ЦИФРА, а решту 4 — ЗОНА (співвідношення «цифр» і «зон» можна змінювати від 8 : 8 до 15 : 1).

Умовимось, що в кодовому слові вторинного алфавіту перші чотири розряди завжди становитимуть зону, а чотири інші — цифри. Кількість можливих комбінацій вторинного алфавіту в даному разі буде $N = 4 \cdot 12 = 48$.

Для виконання зонного стиснення інформації потрібно знаки вторинного алфавіту розбити на зони. При цьому, якщо в тексті зустрічаються поруч знаки, які належать до однієї зони, номер її вказується тільки перед першим знаком, а запис наступних знаків обмежується записом цифрової частини їх.

Для того щоб знаки, які мають однакові зони, утворювали більш довгі послідовності при створенні кодових слів, у вторинному алфавіті необхідно враховувати статистичні характеристики алфавіту, з якого складаються тексти, які слід обробляти. Бажано також врахувати ймовірність різних сполучень деяких знаків.

Ефективність розбиття на зони встановлюють за допомогою потенціального коефіцієнта стиснення, який визначається виразом

$$K_{ст} = N_1/N_2,$$

де N_1, N_2 — відповідно кількість байтів у первинному та стисненому текстах.

Стиснення інформації відбудеться тільки тоді, коли наступний знак у тексті належатиме до тієї самої зони. Тому ймовірність цієї події

$$P_1 = a^2 + b^2 + c^2 + d^2,$$

де a, b, c, d — ймовірності того, що наступний знак належить відповідно першій, другій, третій та четвертій зонам.

Імовірність відсутності стиснення інформації

$$P_2 = 1 - P_1.$$

При стисненні на кожний знак відводиться півбайта, а при його відсутності — повний байт. Таким чином,

$$N_2 = N_1P_1/2 + N_1(1 - P_1);$$

$$K_{ст} = \frac{N_1}{N_1P_1/2 + N_1(1-P_1)} =$$

$$= \frac{1}{\left(a^2+b^2+c^2+d^2\right)/2 + 1 - a^2 - b^2 - c^2 - d^2} =$$

$$= \frac{2}{2 + a^2 + b^2 + c^2 + d^2 - 2a^2 - 2b^2 - 2c^2 - 2d^2} =$$

$$= \frac{2}{2 - \left(a^2 + b^2 + c^2 + d^2\right)}.$$

Розподіл літер українського алфавіту та знаків на чотири зони наведено в табл. 9.1, де враховано частість появи їх у тексті.

Таблиця 9.1

Зона				Код знака в зоні
C	D	E	F	
Літера	Літера	Літера (знак)	Цифра (знак)	
Пробіл	М	Ф	—	Ø
О	П	Щ	1	1
Е	Є	Ш	2	2
А	З	Ц	3	3
Р	К	Ч	4	4
Л	Д	Ж	5	5
Т	Я	Х	6	6
Н	У	Ю	7	7
В	Ь	,	8	8
І	Б	.	9	9
И	Й	;	!	A
С	Г	:	?	B

Розглянемо стиснення інформації на конкретному прикладі.

Закодуємо такий текст: «Постійний запам'ятовуючий пристрій є невід'ємною частиною електронної обчислювальної машини». У закодованій за допомогою табл. 9.1 формі він матиме вигляд

D1C1B69DAC7ADACOD3C3D1C3D0E8D6C618D7E74CA
DAC0D1C4AB649DAC0D2C07289D5E8D20C71E7C0E4C3
B6A71E7C0252D4C641771901D9E4CAB5E7C835D8C7190D0
C3E2CA7AE9

Коефіцієнт стиснення тексту $K_{ст} = 1, 3$.

Стиснення інформації зменшенням розрядності кодованих слів [42]. За цим способом ефект стиснення інформації досягається завдяки поділу послідовності наперед упорядкованих чисел на кілька однакових відрізків, всередині яких відлік ведеться не за їх абсолютним значенням, а від межі попереднього відрізка. При цьому утворена розрядність чисел буде завжди менша від розрядності відповідних реальних чисел, для чого в пам'яті ЕОМ необхідно зберігати ці стиснені числа з розрядністю меншою, ніж розрядність реальних чисел.

Ємність пам'яті (двійкових розрядів) ЕОМ для розміщення N кодів визначається виразом

$$Q = N\log_2 M, \qquad (9.16)$$

де M — кількість кодів, що розміщуються в пам'яті ЕОМ.

Як випливає з (9.16), зі збільшенням M зростає довжина кодової комбінації ($\log_2 M$). Тому для економії ємності пам'яті число $2^{(\log_2 M)}$ (де $[\log_2 M]$ округляється до найближчого цілого числа) розбивається на L однакових частин. Максимальне число в утвореному інтервалі чисел буде не більше $\log_2 \frac{M}{L}$, а ємність пам'яті для їх зберігання $Q_L = N\log_2 \frac{M}{L}$.

Якщо в пам'яті ЕОМ зберігати також адреси меж відрізків і порядкові номери чисел, що зберігаються, відлік яких ведеться від чергової межі, то розрядність чисел для запису номеру межі можна визначити як $\log_2(N-1)$, а ємність пам'яті для зберігання номерів меж — як $(L-1)\log_2(N-1)$, де $(L-1)$ — кількість меж між відрізками, причому $(L-1) \le (N-1)$.

Загальна ємність пам'яті ЕОМ з урахуванням викладеного визначається виразом

$$Q' = N\log_2 \frac{M}{L} + (L-1)\log_2(N-1). \qquad (9.17)$$

Якщо продиференціювати (9.17) по L і прирівняти похідну до нуля, то дістанемо значення Q_{min}, яке буде при $L_{опт} = \frac{N}{\ln N}$.

Ємність пам'яті при оптимальній кількості зон, на які розбиваються числа, що зберігаються в пам'яті ЕОМ, знайдемо після підставлення в (9.17) значення $L_{опт}$:

$$Q' = N\log_2 \frac{eM\ln N}{N} - \log_2(N-1). \qquad (9.18)$$

При $N > 100$ можна користуватися приблизною формулою

$$Q' = M\log_2 \frac{eM\ln N}{N}.$$

Щоб знайти інформацію, записану в пам'ять ЕОМ, спочатку визначають $L_{опт}$, а потім — значення інтервалів між двома межами

$$C = \frac{2^{[\log_2 M]}}{L},$$

де $[\log_2 M]$ — округлене до найближчого цілого числа значення $\log_2 M$. Після цього встановлюють інтервал, у якому знаходиться шукане число X:

$$K = X/C.$$

Адреса шуканого числа визначається як різниця між його абсолютним значенням і числом, що є граничним для цього інтервалу.

Виграш у ємності пам'яті визначається як

$$\Delta Q = Q - Q' = N\log_2 M - N\log_2 \frac{eM\ln N}{N} - \log_2(N-1).$$

Розглянемо конкретний приклад.

Якщо $M = 800$, $X = 300$ і $N = 25$, то пошук числа 500 ведемо за такою послідовністю:

$$L = \frac{25}{\ln 25} = \frac{25}{3{,}22} \approx 8; \qquad C = \frac{2^{[\log_2 800]}}{L} = \frac{1024}{8} = 128;$$

$$K = \frac{X}{C} = \frac{300}{128}; \qquad 3 > K > 2.$$

Отже, шукане число лежить у третьому інтервалі, де знаходяться числа від $128 \cdot 2 = 256$ до $383 = 256 + 127$ (додаємо 127, а не 128, оскільки в інтервалі знаходиться всього 128 чисел, з яких одне — нуль). Порядковий номер числа визначаємо як різницю між шуканим числом і числом, що є граничним для шуканого інтервалу:

$$300 - 256 = 44.$$

Знаходимо виграш у ємності пам'яті ΔQ:

$$Q = N \log_2 M = 25\log_2 800 \approx 25 \cdot 10 = 250 \text{ двійкових розрядів;}$$

$$Q'' = N\log_2 \frac{eM\ln N}{N} = 25\log_2 \frac{2{,}3 \cdot 800 \ln 25}{25} \approx$$

$$\approx 25\log_2 237 \approx 200 \quad \text{двійкових розрядів;}$$

$$\Delta Q = Q - Q'' = 250 - 200 = 50 \text{ двійкових розрядів.}$$

Стиснення інформації заміною деяких комбінацій літер одиничними символами [42]. Цей спосіб ґрунтується на заміні деяких сполучень літер, які найчастіше зустрічаються в тексті, одиничними символами у вигляді двійкових кодових комбінацій, що не використовуються для подання знаків і символів при кодуванні як в ЕОМ, так і при обміні даними за допомогою ліній зв'язку.

Так, для подання алфавітно-цифрової та службової інформації в кодах КОІ-8 і ДКОІ з 256 можливих не застосовуються 167 символів, які не мають графічних еквівалентів. За цим способом стиснення інформації пропонується використати цей резерв для кодування біграм, що найчастіше зустрічаються в тексті. Ефективність стиснення науково-технічних текстів при цьому досягає 40 % і більше. Даний спосіб можна поширити також на заміну в тексті сполучень з трьох і більше літер.

Стиснення інформації використанням адаптивного кодування [8]. Воно з успіхом застосовується при передачі техніко-економічної, статистичної та інших видів інформації, де, як правило, дуже висока вірогідність передачі потрібна не завжди.

Вимоги до вірогідності інформації можуть коливатися в широких межах залежно від характеру повідомлень. Так, при передачі статистичної інформації про кількість виробленої продукції помилка в молодших розрядах повідомлення (десятки та одиниці) менше впливатиме на правильність переданої інформації, ніж спотворення старших розрядів (мільйони та тисячі). При передачі текстових повідомлень вплив помилок на передачу ще менший.

Таким чином, доцільно передавати повідомлення з заданою вірогідністю у відповідних межах, використовуючи адаптивне кодування відносно джерела повідомлень, що дає змогу забезпечити вірогідність інформації, яка передається, залежно від вимог джерела повідомлень.

Для цього повідомлення необхідно поділити за категоріями з заданою вірогідністю передачі. На початку кожної категорії в інформацію вводяться додаткові службові комбінації, які при надходженні в кодер (під час передавання) та декодер (під час приймання) системи передачі перебудовують їх. Перебудова

кодера може виконуватися також за сигналами від аналізатора інформації, якщо він установлений на вході системи передачі. Аналізатор обробляє інформацію, що надходить від джерела повідомлень, і поділяє її на категорії.

Для передачі доцільно використовувати три категорії інформації з різною вірогідністю (з різною ймовірністю помилок): $10^{-3}...10^{-4}$ — для передачі текстів; $10^{-5}...10^{-6}$ — для передачі цифрових повідомлень від одиниць до тисячі; $< 10^{-6}$ — для передачі цифрових повідомлень більше тисячі та будь-яких дуже важливих цифрових і текстових повідомлень. Принципи поділу інформації на категорії можуть змінюватися залежно від конкретних умов і вимог.

Передачу інформації з різною вірогідністю можна забезпечити зміною способу кодування. Так, для передачі інформації по проводових каналах з вірогідністю $10^{-3}...10^{-4}$ досить застосувати двійковий первинний код без уведення будь-якого захисту інформації; з вірогідністю $10^{-5}...10^{-6}$ — двійковий код, що виявляє помилки, а з вірогідністю $< 10^{-6}$ — код, який виправляє відповідну кількість помилок. Тоді для передачі, наприклад, 32 кодових комбінацій їх довжина в першому випадку становитиме $n_1 = 5$, у другому — $n_2 = 6...7$, у третьому — $n_3 = 8...12$ елементів.

Середня довжина кодової комбінації при використанні адаптивного кодування залежатиме від процентного співвідношення повідомлень різної категорії і загалом може бути визначена як

$$n_{\text{сер}} = \frac{n_1 N_1 + n_2 N_2 + n_3 N_3}{N},$$

де n_1, n_2, n_3 — довжини кодових комбінацій при передачі повідомлень першої, другої та третьої категорій; N_1, N_2, N_3 — відповідна кількість цих комбінацій; N — загальна кількість їх.

Коефіцієнт стиснення інформації можна визначити як відношення максимальної довжини n кодової комбінації до її середньої довжини $n_{\text{сер}}$:

$$K_{\text{ст}} = n/n_{\text{сер}}.$$

Наприклад, при передачі $N = 1000$ повідомлень, з яких $N_1 = 700$, $N_2 = 200$ і $N_3 = 100$, якщо $n_1 = 5$, $n_2 = 6$, $n_3 = 8$, матимемо

$$n_{\text{сер}} = \frac{5 \cdot 700 + 6 \cdot 200 + 8 \cdot 100}{1000} = 5{,}5 \text{ елемента}$$

і коефіцієнт стиснення

$$K_{\text{ст}} = \frac{8}{5{,}5} \approx 1{,}45.$$

Стиснення інформації збільшенням основи коду. Ґрунтується воно на перекодуванні кодованих послідовностей символів системи числення з меншою основою в систему з більшою основою. Так, якщо інформаційний масив, який складається з двійкових елементів, подати у вісімковій або шістнадцятковій системі числення (тобто перекодувати його), то це приведе до значного зменшення кількості елементів масиву (в кілька разів):

$$\left.\begin{array}{cccccccccccccccc} 0&0&0&0&0&1&0&1&0&0&1&1&1&0&0&1\\ 0&0&1&0&1&0&0&1&1&1&0&0&1&0&1&1\\ 0&1&0&0&1&1&1&0&0&1&0&1&1&1&0&1\\ 0&1&1&1&0&0&1&0&1&1&1&0&1&1&1&0\\ 1&0&0&1&0&1&1&1&0&1&1&1&0&0&0&0\\ 1&0&1&1&1&0&1&1&1&0&0&0&0&0&1&0\\ 1&1&0&1&1&1&0&0&0&0&0&1&0&1&0&0\\ 1&1&1&0&0&0&0&0&1&0&1&0&0&1&1&1 \end{array}\right\} N_1 = 192;$$

$$\left.\begin{array}{cccccccc} 0&1&2&3&4&5&6&7\\ 1&2&3&4&5&6&7&0\\ 2&3&4&5&6&7&0&1\\ 3&4&5&6&7&0&1&2\\ 4&5&6&7&0&1&2&3\\ 5&6&7&0&1&2&3&4\\ 6&7&0&1&2&3&4&5\\ 7&0&1&2&3&4&5&6 \end{array}\right\} N_2 = 64; \qquad \left.\begin{array}{cccccc} 0&5&3&9&7&7\\ 2&9&C&B&B&8\\ 4&E&5&D&C&1\\ 7&2&E&E&0&A\\ 9&7&7&0&5&3\\ B&B&8&2&9&C\\ D&C&1&4&E&5\\ E&0&A&7&2&E \end{array}\right\} N_3 = 48;$$

При цьому коефіцієнти стиснення первинного масиву двійкових елементів у вісімкову і шістнадцяткову системи числення становитимуть

$$K_{ст8} = \frac{N_1}{N_2} = \frac{192}{64} = 3; \qquad K_{ст16} = \frac{N_1}{N_3} = \frac{192}{48} = 4,$$

тобто залежать від кількості бітів, що несе один елемент інформаційного масиву (у вісімковій системі $8 = 2^3 \rightarrow 3$ біт/елемент, у шістнадцятковій — $16 = 2^4 \rightarrow 4$ біт/елемент).

Матричні способи. До ефективних матричних способів стиснення інформації при передачі даних належить ***спосіб зі зберіганням атрибутів у вигляді бітової матриці*** [42]. При цьому способі скінченна кількість атрибутів (табл. 9.2) виноситься в першу частину («шапку») матриці, тілом якої є набір двійкових елементів, з яких 1 означає наявність, а 0 — відсутність атрибута (табл. 9.3). «Шапка» бітової матриці та її тіло можуть зберігатися на різних ділянках пам'яті ЕОМ.

Значний ефект стиснення дає інший ***матричний спосіб із заміною елементів,*** що повторюються на деякій площі інформаційного масиву, одиничними елементами, які несуть ознаки цих ділянок масиву [8]. Так, якщо ввести символи, що відбивають

Таблиця 9.2

Номер виробничої дільниці	Тип комплектувального виробу					
0001	0301	1017	2216	1100		
0012	1008	0011	2001	2216	1017	
0015	1100	2001	2000	1008	0301	
0117	1100	0011	0301			
1206	2001	2216	0011	0301	1100	1017

Таблиця 9.3

Номер виробничої дільниці	Тип комплектувального виробу							
	0011	0301	1008	1017	1100	2002	2001	2216
0001	0	1	0	1	1	0	0	1
0012	1	0	1	1	0	0	1	1
0015	0	1	1	0	1	1	1	0
0117	1	1	0	0	1	0	0	0
1206	1	1	0	1	1	0	1	1

деякі обмежені площі інформаційного масиву, в яких елементи повторюються порівняно з однойменними розрядами попереднього рядка (наприклад, R – 2 × 2, X – 2 × 5, Y – 3 × 2, Z – 5 × 3), то інформаційний масив

1	2	3	4	5	6	7	8	9	0
2	2	3	4	8	6	7	5	9	0
3	2	3	4	7	6	7	3	9	0
3	2	3	4	7	6	7	3	9	0
3	2	3	4	7	6	7	3	9	0
3	2	3	4	7	8	9	1	9	0

після згортання матиме вигляд

1	2	3	4	5	6	7	8	9	0
2	Y	8	R	5	X	3	7	3	Z
Y	8	9	1						

Розгортка виконується з початку масиву:

1	2	3	4	5	6	7	8	9	0
2	.	.	.	8	.	.	5	.	.
3	.	.	.	7	.	.	3	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	8	9	1	.	.

Коефіцієнт стиснення $K_{ст} = 60/24 = 2,5$.

Стиснення інформації з типовими матрицями застосовується головним чином для відносно великих масивів. При цьому в пам'яті ЕОМ зберігаються матриці деяких найпоширеніших наборів символів, які зустрічаються в інформаційному масиві, що стискається. Для цього заздалегідь аналізують інформаційний масив, утворюють типові матриці і при стисненні замість наборів символів, які відповідають цій типовій матриці, в послідовність символів інформаційного масиву вводять адреси цих матриць.

Комбіновані та каскадні способи. З комбінованих способів як приклад можна навести ***лінійно-матричне стиснення інформації,*** при якому первинний інформаційний масив стискається за допомогою додаткових елементів, якими позначають повторення елементів у рядках порівняно з попереднім (наприклад, символами K – 2, L – 3, M – 5 та в обмежених ділянках — матрицях символами X – 2 × 2, Y – 2 × 5, Z – 3 × 5). Тоді, якщо інформаційний масив має вигляд

1 2 3 4 5 6 7 8 9 0
2 2 3 4 5 6 7 5 9 0
2 2 3 4 5 6 4 6 9 0
2 2 3 4 5 6 4 4 9 0
3 2 3 4 5 6 4 2 9 0
3 2 3 4 5 7 4 2 9 0
3 2 3 4 5 7 4 2 9 0,

то згорненим він матиме такий вигляд:

1 2 3 4 5 6 7 8 9 0
2 2 Z K 5 Y X 6 4 6
X 4 3 2 2 X 7 K L M.

Розгортання виконується з початку масиву:

1 2 3 4 5 6 7 8 9 0
2 2 . . . . . 5 . .
. . . . . 6 4 6 . .
. . . . 4 . .
3 2 . . . . . 2 . .
. . . . . 7 . . . .
. . . . . .

Коефіцієнт стиснення $K_{ст} = 70/30 = 2,33(3)$.

Каскадні способи стиснення інформації застосовуються для збільшення коефіцієнта стиснення. При цьому, як правило, використовуються комбінації лінійних і матричних способів. Можна утворити такі каскадні способи стиснення, як кодо-зонний, кодо- та зонно-матричні тощо.

9.2.2. СПОСОБИ СТИСНЕННЯ ДАНИХ ПРИ АРХІВАЦІЇ

При архівації даних в ЕОМ, на дисках, як правило, оперують зі значно більшим обсягом інформації, ніж при передачі даних. При цьому для стиснення застосовують архіватори [16, 26]. Бажано, щоб останні працювали в реальному часі. Крім того, вони мають задовольняти дві вимоги: забезпечувати максимально високі ступені стиснення даних і швидкості довільного доступу до них.

Для виконання першої вимоги, з одного боку, логічно було б стискати всі дані на диску в цілому, тобто як один неперервний потік інформації. Проте з іншого боку, в цьому разі для забезпечення доступу необхідно перепаковувати (розпаковувати) весь диск. Тому на практиці весь логічний простір диска розбивають на окремі області, дані в яких стискаються незалежно від інших, а файли подають як ланцюги дискового простору (кластери).

Щоб записувати дані на диск і зберігати координати положення їх для подальших операцій зчитування та модифікації, виникає необхідність уведення додаткових структур для планування областей стиснених даних диска. До цих даних належать: початкове положення (номер стартового сектора) та розмір сегмента (кількість секторів). Ці дані для кожного кластера оформляються у вигляді додаткової таблиці його дескрипторів.

Таким чином, для читання деякого кластера досить за допомогою таблиці дескрипторів визначити, де знаходиться фрагмент з його стисненими даними, прочитати та розпакувати його.

Однак при записуванні кластера в область даних користуватися інформацією про дисковий простір у тому вигляді, як її подано в таблиці дескрипторів, дуже складно. Тому, як правило, її подають у вигляді списку вільного простору області стиснених даних. Як альтернатива може бути утворена також бітова карта зайнятих секторів області стиснених даних [26] .

При обробленні запитів драйвером стисненого диска можна виділити три групи алгоритмів, які мають самостійне значення: планування дискового простору, хешування (організація структур даних, що забезпечує ефективний пошук та доповнення), стиснення/розпаковування.

При архівації стискають, як правило, блоки даних ємністю 1...8 Кбайт, але при цьому самі алгоритми стиснення мають використовувати мінімальну ємність пам'яті ЕОМ.

Із *лінійних* способів стиснення інформації, які широко застосовуються при архівації, можна виділити ***кодування потоків символів*** у деякому алфавіті з різною частотою появи символів

у потоці. Метою кодування є перетворення цього потоку на потік бітів мінімальної довжини. Це досягається завдяки зменшенню надмірності вхідного потоку врахуванням частоти появи символів на вході та довжини коду, яка має бути пропорційною інформації, що міститься у вхідному потоці даних. Такі способи ґрунтуються на використанні оптимального кодування.

Якщо розподіл імовірностей появи символів у вхідному потоці даних наперед невідомий, то можна скористатися одним із двох підходів. Згідно з першим підходом слід переглянути вхідний потік даних і побудувати оптимальний код (наприклад, код Хаффмена), ґрунтуючись на наявній статистиці. При цьому дістають вихідний потік даних, закодований оптимальним кодом. Як приклад використання такого підходу можна назвати ***статистичне кодування Хаффмена*** [16]. У цьому разі вхідним символам, поданим послідовностями бітів однакової довжини, зіставляються послідовності бітів змінної довжини. Довжина коду для символу пропорційна (з округленням до цілого) двійковому логарифму частоти його появи, що береться з оберненим знаком. Це кодування є префіксним, що дає змогу декодувати його однопрохідним алгоритмом. Префіксний код зручно подавати у вигляді двійкового кодового дерева, в якого шлях від кореня до вершини визначає код символу (див. п.5.7).

Нехай, наприклад, вхідний алфавіт складається з чотирьох символів *a*, *b*, *c* та *d*, ймовірності появи яких відповідно дорівнюють 1/2, 1/4, 1/8, 1/8. Тоді кодом Хаффмена для цього алфавіту замість рівномірного коду 00, 01, 10, 11 може бути код *a* – 0, *b* – 10, *c* – 110, *d* – 111, і вхідна комбінація *abaabbaaacd*, замість 00010000010100000001011 матиме вигляд 01000101000011011 1, тобто замість 22 біт на вході на виході буде 18 біт.

Згідно з другим підходом використовується адаптивний кодер, який змінює схему кодування оптимальним кодом, залежно від первинних даних. Декодер, що декодує кодований потік, синхронно з кодером змінює схему декодування, починаючи з деякої заданої наперед.

Адаптивне кодування інформації дає більший ступінь її стиснення, оскільки враховуються локальні зміни ймовірностей у вхідному потоці даних. Прикладом такого кодування є ***адаптивне (динамічне) кодування Хаффмена*** [16]. При такому кодуванні потрібне постійне коректування кодового дерева відповідно до зміни статистики вхідного потоку. При реалізації цього способу виникає складність перебудови кодового дерева згідно з новими ймовірностями символів на кожному кроці стиснення.

Кодування Хаффмена відзначається мінімальною надмірністю за умови, що кожний символ кодується окремою послідовністю в алфавіті {0,1} і забезпечує досить високі швидкість та якість стиснення інформації. До недоліків цього способу слід

віднести залежність ступеня стиснення від близькості ймовірностей символів до від'ємного степеня двійки, що пов'язано з необхідністю округлення до цілого числа в бік збільшення, оскільки кожний символ кодується цілим числом бітів. Це призводить до того, що, наприклад при двосимвольному алфавіті, стиснення інформації взагалі неможливе, якщо кожний з символів округляється до 1/2.

Більший ступінь стиснення дає ***арифметичне кодування*** [16, 32]. У цьому разі текст, стиснений арифметичним кодером, розглядається як деякий двійковий дріб з інтервалу [0, 1]. Результат стиснення можна подати як послідовність двійкових символів цього дробу. Первинним текстом є запис дробу, в якому кожний вхідний символ має вагу, пропорційну ймовірності його появи.

Нехай, наприклад, алфавіт складається з двох символів *a* та *b* з імовірностями появи їх 3/4 і 1/4 відповідно. Розглянемо відкритий праворуч інтервал [0, 1). Поділимо цей інтервал на кілька частин, довжина яких пропорційна ймовірностям появи символів, тобто [0, 3/4) та [3/4, 1). Кожному слову у вхідному алфавіті відповідає деякий підінтервал з [0,1), а порожньому слову — весь інтервал [0, 1). Після приймання кожного наступного символу арифметичний кодер зменшує інтервал, вибираючи ту його частину, яка відповідає поданому в даний момент на вхід символу. Кодом послідовності є інтервал, що виділяється після оброблення всіх її символів, тобто двійковий запис координати будь-якої точки з цього інтервалу. Таким чином, довжина утвореного інтервалу пропорційна ймовірності появи кодованої послідовності символів.

При реалізації цього алгоритму необхідна дійсна арифметика необмеженої точності. Крім того, результат кодування стає відомим тільки після закінчення приймання вхідного потоку символів.

Найбільшого поширення зі способів стиснення при архівації даних дістали алгоритми Лемпеля — Зіва та їхні різновиди (Лемпеля— Зіва — Велча LZW, Міллера — Ведмана MW, AP, Y тощо) [16, 26].

Розглянемо як приклад *алгоритм* LZ77, запропонований А. Лемпелем та Я. Зівом у 1977 р. Роботу LZ77-кодера за цим алгоритмом пояснює рис. 9.1 [26]. Основною операцією кодера є пошук рядка у вхідному буфері, який якомога краще збігався б з поточною позицією вхідного потоку. Для прискорення процедури цього пошуку застосовується хешування. При цьому за першими двома або трьома символами вхідного потоку даних визначається деяка величина, яка, незважаючи на те, що є обмеженою зверху (наприклад, числом 2000), досить добре характеризує рядок, з якого вона була утворена. Значення цієї величини використовується далі як індекс до деякої таблиці, в якій розміщуються адреси рядків з певними значеннями хеш-індекса в порядку появи їх у вхідному буфері.

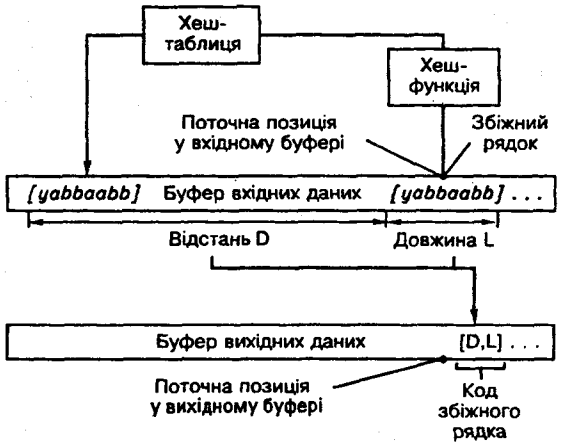


Рис. 9.1

Якщо тепер при обновленні змісту хеш-таблиці з'ясується, що в ній вже зберігалася адреса деякого рядка, то при цьому залишиться тільки перевірити, наскільки цей рядок збігається з поточною послідовністю символів.

Для збільшення ймовірності знаходження потрібного рядка застосовується також перелік дозволених колізій — ланцюга адрес рядків, які мають однакові хеш-індекси, тобто якщо найближчий рядок, занесений до хеш-таблиці, в дійсності не збігається з поточною послідовністю символів, то процес пошуку рядка, що найбільше підходить, продовжується вже в переліку колізій. Після закінчення цього аналізу в хеш-таблицю заноситься адреса поточного рядка, а перелік колізій доповнюється рядком, який витискується з хеш-таблиці.

Якщо такий пошук є вдалим, тобто довжина збігу одного з підрядків з поточною послідовністю символів у вхідному буфері перевищує деякий поріг (як правило, два або три символи), то у вихідний потік пересилається її код, який складається зі значень відстані D та довжини L (див. рис. 9.1).

У тому разі, коли жоден з рядків, які розглядаються, не підходить, у вихідний буфер пересилається поточний символ вхідного потоку й увесь цей процес повторюється для подальших символів.

Для поділу кодів рядків, які збігаються, та окремих символів у вихідний потік подаються також біти ознак рядок/символ.

Ефективність стиснення інформації певним способом залежить від її виду. Так, ступінь стиснення може змінюватися від 1,2...1,5 раза для текстової інформації до 6...8 разів для чорно-білих зображень, а швидкість стиснення — відповідно від 0,07...0,2 Мбайт/с до 1,5...1,9 Мбайт/с [26] .

9.3. ЗБІЛЬШЕННЯ ОСНОВИ КОДУ

Підвищення ефективності кодування, що поліпшує якість передачі інформації, можна розуміти по-різному, але прийняті в теорії кодування показники якості зводяться до трьох наступних тверджень, які є критеріями ефективності завадостійких кодів:

1) серед кодів з однаковими довжиною n та мінімальною кодовою відстанню d_{min} найкращі коди мають найбільшу довжину інформаційної частини блока;

2) серед кодів з однаковими довжиною n і кількістю k інформаційних елементів найкращі коди мають найбільшу мінімальну кодову відстань d_{min};

3) серед кодів з однаковими k та d_{min} найкращі коди мають найменшу довжину n блока.

Відповідно до першого твердження у кращому коді ефективніше використовується коректувальний потенціал* d_{min}, який захищає більшу кількість k інформаційних елементів у блоці, ніж в іншому коді.

Відповідно до другого твердження кращий код має більший коректувальний потенціал d_{min}, що дає змогу йому виправляти більшу кількість $\nu_{вп}$ спотворених елементів у кодовому блоці.

Відповідно до третього твердження кращий код має меншу кількість r надмірних елементів у блоці ($r = n - k$), а тому й меншу надмірність r/n.

Оскільки наведені твердження стосуються трьох параметрів коду (n, k та d_{min}), переформулюємо їх у термінах відносних величин, що дасть змогу, замаскувавши один із цих параметрів, проводити порівняння кодів у площинній формі двох вимірювань.

Для цього, поділивши k та d_{min} на довжину n блока, дістанемо швидкість $R = k/n$ коду і відносну кодову відстань d_{min}/n. Додамо до визначення останньої коефіцієнт 1/2, щоб проводити аналіз у єдиних з відомими джерелами одиницях [25, 32]. Останнє робиться з урахуванням функціональної залежності $d_{min} = 2\nu_{вп} + 1$, звідки випливає, що $\nu_{вп} = (d_{min} - 1)/2$.

У відносному поданні ν/n для досить великих значень n маємо

$$\frac{\nu_{вп}}{n} = \frac{d_{min} - 1}{2n} \approx \frac{d_{min}}{2n}. \qquad (9.19)$$

---

* Під *коректувальним потенціалом* розумітимемо коректувальну здатність коду, яка визначається показником d_{min}, а втілюється в кількість $\nu_{вп}$ спотворених елементів у блоці, які можуть бути виправлені цим кодом, причому $d_{min} \geq 2\nu_{вп} + 1$ [12, 25].

Таким чином, наступний розгляд зводиться до двох параметрів: $R = k/n$ і $\nu_{вп}/n \approx d_{min}/(2n)$. Для спрощення відкинемо індекс min біля d, маючи на увазі, що d і є мінімальною кодовою відстанню.

Тоді наведені вище твердження можуть бути викладені так:

1) серед кодів з однаковою відносною кодовою відстанню $d/(2n) = \text{const}$ кращим є код із більшою швидкістю R;

2) серед кодів з однаковою швидкістю $R = k/n = \text{const}$ кращим є код із більшою кодовою відстанню $d/(2n)$;

3) серед кодів з однаковими значеннями k та d кращим є код із більшими значеннями $R = k/n$ і $d/(2n)$.

Розглянемо з позицій цих критеріїв вплив на ефективність кодування основи q коду. Загальновизнаним методом оцінювання ефективності кодування взагалі є відшукання граничних залежностей типу $R = f[d/(2n)]$ для найкращих кодів, якими є не якісь конкретні коди, а коди завдовжки $n \to \infty$ [25]. Остання умова дає змогу таким кодам відповідно до (9.19) мати досить велику коректувальну здатність $\nu_{вп}$. Ці граничні залежності $R[d/(2n)]$ називаються межами (верхніми або нижніми) для мінімальної кодової відстані d.

У [25, 32] межу Плоткіна записано у вигляді

$$k \le n - \frac{qd-1}{q-1} + 1 - \log_q d. \qquad (9.20)$$

Поділивши цю нерівність на n, дістанемо

$$R = \frac{k}{n} \le 1 - \frac{q}{q-1}\frac{d}{n} + \frac{1}{n(q-1)} + \frac{1}{n} + \frac{\log_q d}{n}.$$

Якщо збільшувати n до $n \to \infty$, то матимемо нерівність

$$R \le 1 - \frac{2q}{q-1}\frac{d}{2n}, \qquad (9.21)$$

в якій відкинуто доданки вищих ступенів.

Лінії межі Плоткіна $П_q$, що відповідають (9.21), показано на рис. 9.2. Для $q = 2$ це лінія $П_2$, а для $q = \infty$ — лінія $П_\infty$. З рисунка випливає, що зі збільшенням основи q коду в повній відповідності до трьох наведених вище критеріїв збільшується ефективність кодування. Так, при фіксованому значенні $d/(2n)$ збільшення основи q коду приводить до зростання $R = k/n$ і навпаки. Це одна з верхніх меж для d, зміст якої полягає в тому, що будь-який код відображається на рис. 9.2 у вигляді точки зі своїми координатами R та $d/(2n)$ і немає жодного коду, координа-

ти якого перевищували б відповідні координати будь-якої точки верхньої межі, тобто не існують коди, кращі від тих, що відображаються верхньою межею.

Є ще дві верхні та одна нижня межі для d, поширені завдяки своїй невеликій складності також тому, що вони адекватно відображають багато відомих кодів [25, 32], уточнюючи та доповнюючи одна одну. Розглянемо їх докладніше. Це межі Хеммінга, Еліаса та нижня межа Варшамова — Гільберта. Всі вони пов'язані з нерівністю Чернова [32], де використовується функція виду

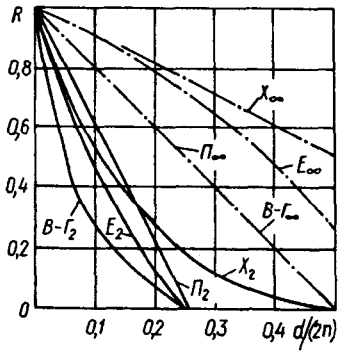


Рис. 9.2

$$\varphi(x) = x\log_q(q-1) - x\log_q x - (1-x)\log_q(1-x), \qquad (9.22)$$

яка тісно пов'язана з ентропією $H(x)$ джерела і при $q = 2$ збігається з нею.

Межа Хеммінга має такий вигляд [25]:

$$R \le 1 - \varphi(\nu_{вп}/n). \qquad (9.23)$$

Якщо врахувати (9.19) і $n \to \infty$, то дістанемо

$$R \le 1 - \varphi\left(\frac{d}{2n}\right) = 1 - \varphi(x), \qquad (9.24)$$

де $x = d/(2n)$.

Включивши до (9.24) вираз (9.22), матимемо рівняння межі

$$R = 1 - x\log_q(q-1) + x\log_q x + (1-x)\log_q(1-x). \qquad (9.25)$$

Проаналізуємо вираз (9.25) детальніше. Для малих значень $x \to 0$ при $q \to \infty$ деякі складові його прямують до нуля й тоді $R \le 1 - x$. В іншому випадку, коли $x \to 1$ та $q \to \infty$, з тих самих причин $R \le 1 - x$. Середня область для x визначається виразом (9.25), з якого випливає межа Хеммінга X_q; для двійкового коду це лінія X_2 на рис. 9.2. Якщо збільшувати основу коду до $q \to \infty$, то лінія межі X_q піднімається вище над лінією X_2 — при $q = \infty$ займає положення X_∞, даючи той самий асимптотичний вираз межі Хеммінга при великих значеннях основи q, тобто

$$R \le 1 - x. \qquad (9.26)$$

Верхня межа Еліаса визначається виразом [25]

$$d/n < \delta(R)\left[2 - \frac{\delta(R)q}{q-1}\right] + \varepsilon, \qquad (9.27)$$

де $\delta(R)$ – розв'язок рівняння $\varphi(x) = 1 - R$, утвореного з (9.23), а $\varphi(x)$ — функція (9.22); $\varepsilon > 0$ — як завгодно мала складова.

Позначимо $x = \delta(R)$, тоді $\varphi[\delta(R)] = 1 - R$. Таким чином, за допомогою межі Хеммінга (9.24) або (9.25) дістаємо парні значення $x = \delta(R)$ та R. Після цього для кожного R за відомим $x = \delta(R)$ розраховуємо значення $d/(2n)$ межі Еліаса, користуючись виразом (9.27), перетвореним до вигляду

$$\frac{d}{2n} < \frac{1}{2}\delta(R)\left[2 - \frac{q}{q-1}\delta(R)\right] + \varepsilon. \qquad (9.28)$$

Лінії E_2 (для $q = 2$) та E_∞ (для $q = \infty$) верхньої межі Еліаса відповідно до (9.28) зображено на рис. 9.2. Всі проміжні лінії E_q для $2 < q < \infty$ розташовуються між цими двома лініями.

Нарешті нижня межа Варшамова — Гільберта визначається виразом [25]

$$q^{n-k} \le \sum_{i=0}^{d-2} C_n^i (q-1)^i, \qquad (9.29)$$

де C_n^i – біномні коефіцієнти. Щоб виключити їх із розгляду, скористаємося асимптотичною оцінкою цих коефіцієнтів у вигляді нерівності Чернова

$$\sum_{i=0}^{\nu} C_n^i (q-1)^i \le q^{n\varphi(\nu/n)}, \qquad (9.30)$$

де $\varphi(x)$ — функція (9.22).

З урахуванням (9.30) вираз (9.29) набуває вигляду

$$q^{n-k} \le q^{n\varphi\left(\frac{d-2}{n}\right)}. \qquad (9.31)$$

Після логарифмування обох частин (9.31) матимемо

$$n - k \le n\varphi\left(\frac{d-2}{n}\right),$$

звідки

$$R = k/n \ge 1 - \varphi\left(\frac{d}{n} - \frac{2}{n}\right).$$

Збільшуючи n до $n \to \infty$, дістаємо асимптотичну форму нижньої межі Варшамова — Гільберта

$$R \ge 1 - \varphi(d/n). \qquad (9.32)$$

Якщо (9.32) порівняти з межею Хеммінга (9.24), то можна побачити, що

$$R \ge 1 - \varphi\left(2\frac{d}{2n}\right) = 1 - \varphi(2x), \qquad (9.33)$$

де $x = d/(2n)$ — спільний аргумент усіх розглянутих асимптотичних меж, а значення R у формі (9.33) може бути розраховане за (9.25), якщо аргумент x межі Хеммінга замінити аргументом $y = 2x$ межі Варшамова — Гільберта.

Остаточно після перетворення (9.25) знаходимо

$$R \ge 1 + \log_q\left[\left(\frac{2x}{q-1}\right)^{2x}(1-2x)^{(1-2x)}\right], \qquad (9.34)$$

де $x = d/(2n)$.

Лінії $B — \Gamma_2$ (для $q = 2$) та $B — \Gamma_\infty$ (для $q = \infty$) нижньої межі Варшамова — Гільберта, що відповідають (9.34), показано на рис. 9.2. Всі проміжні лінії $B — \Gamma_q$ для $2 < q < \infty$ розташовуються між цими двома лініями.

Різниця між наведеними асимптотичними межами для мінімальної кодової відстані пояснюється різницею між існуючими і ще не відшуканими в явному вигляді кращими кодами. Проте більш важливими є загальні риси цих меж. Роль усіх проміжних меж $2 < q < \infty$ полягає в тому, що для кожного значення q немає кодів, точки відображення яких зі своїми координатами R і $d/(2n)$ на рис. 9.2 лежали б вище відповідної лінії верхньої або нижче відповідної лінії нижньої меж. Це дає змогу оцінити із загальних позицій граничні можливості кодів взагалі і вказати ті параметри, коди з якими потрібно відшукувати і будувати як найефективніші. Разом із тим ці асимптотичні (для $n \to \infty$) межі при різних значеннях q дають змогу виявити великий позитивний вплив основи коду на ефективність кодування інформації відповідно до зазначених вище критеріїв.

9.4. ВИКОРИСТАННЯ ЗВОРОТНОГО ЗВ'ЯЗКУ ДЛЯ ПІДВИЩЕННЯ ЕФЕКТИВНОСТІ ПЕРЕДАЧІ ІНФОРМАЦІЇ

При передачі повідомлень первинним кодом вірогідність приймання їх в основному визначається лінією (каналом) зв'язку, а також рівнем і видом завад, які діють в ній. Здобута при

цьому вірогідність дуже часто не задовольняє вимоги споживачів інформації; тому для підвищення вірогідності передачі повідомлень, як правило, вживаються спеціальні заходи.

Одним із них є використання зворотного зв'язку. При цьому поряд із підвищенням вірогідності повідомлень досягається також збільшення ефективності передачі інформації.

Дійсно, якщо в системах передачі інформації без зворотного зв'язку для підвищення вірогідності передачі повідомлень потрібно застосовувати більш складні коректувальні коди (такі як, наприклад, циклічні коди БЧХ, Файра та інші, що, у свою чергу, призводить до збільшення довжини кодових комбінацій і, як наслідок, до зменшення швидкості передачі інформації), то в системах зі зворотним зв'язком дуже складні (й дуже довгі) коректувальні коди використовувати не треба. В таких системах для передачі повідомлень досить користуватися кодами, що виявляють помилки та які мають значно меншу довжину порівняно з коректувальними кодами, що виправляють помилки.

У разі виявлення помилки в прийнятій по прямому каналу кодовій комбінації в системі передачі зі зворотним зв'язком по зворотному каналу посилається сигнал запиту, за яким передавальний пристрій системи повторює передачу інформації по прямому каналу. У зв'язку з цим передавальний пристрій системи має зберігати інформацію про передані кодові комбінації на період часу, достатній для аналізу їх приймальним пристроєм й одержання можливого запиту при виявленні помилок.

Кількість повторень залежить від стану каналу зв'язку і, як правило, обмежується з урахуванням вимог замовника. Зі збільшенням повторень швидкість передачі інформації зменшується, але це помітно тільки при поганому стані каналу зв'язку.

Таким чином, ефективність передачі повідомлень у системах зі зворотним зв'язком залежатиме як від вибору коду, що виявляє помилки, так і від інтенсивності та виду помилок у каналі зв'язку.

Критерієм ефективності методу підвищення вірогідності передачі повідомлень може бути вираз

$$k_{\text{еф}} = \log_2 \frac{a}{g}, \tag{9.35}$$

де $a = P_{\text{п}k}/P_{\text{п}n}$ — виграш у захисті від помилок (тут $P_{\text{п}k}$, $P_{\text{п}n}$ — відповідно ймовірності помилки в повідомленні без надмірності та з надмірністю); $g = g_{\text{ін}} + g_{\text{сх}} = \frac{V_{\text{с}}}{V_{\text{k}}} + \mu \frac{C_{\text{пд}}}{C_0}$ — надмірність [тут $g_{\text{ін}}$, $g_{\text{сх}}$ — відповідно інформаційна та схемна надмірності; $V_{\text{с}}$, V_{k} – відповідно сумарна та корисна (без надмірності) швидкості передачі інформації; μ — ваговий коефіцієнт, що зводить інформаційну та схемну надмірності до еквівалентних техніко-економічних показників; $C_{\text{пд}}$, $C_{\text{о}}$ – відповідно об'єми передавальної апаратури з пристроями підвищення вірогідності передачі повідомлень і без цих пристроїв в еквівалентних одиницях].

Для кодів, що виявляють помилки з наступним запитом, інформаційна надмірність [12] визначається виразом

$$g = \left(1 + \frac{r}{k}\right)\left(\frac{1}{1 - P'_{\text{п}n}}\right). \tag{9.36}$$

Другий множник у (9.36) показує збільшення надмірності через повторення.

Оцінимо ефективність використання двійкового (7, 4)-коду Хеммінга для виправлення однократних помилок у системі передачі повідомлень без зворотного зв'язку та цього самого коду для виявлення двократних помилок із запитом у системах зі зворотним зв'язком. Умовимося, що ймовірність спотворення одиничного елемента $P_{\text{е}} = 5 \cdot 10^{-3}$.

При біномному характері розподілу помилок маємо

$$P_{\text{п}n} = 1 - (1 - P_{\text{е}})^n = 1 - (1 - 5 \cdot 10^{-3})^7 \approx 3{,}5 \cdot 10^{-2}.$$

Інформаційна надмірність:

• при виправленні однократної помилки

$$g_{\text{ін1}} = (1 + r/k) = 1 + 3/4 = 1{,}75;$$

• у разі виявлення помилок

$$g_{\text{ін2}} = (1 + r/k)\left(\frac{1}{1 - P'_{\text{п}n}}\right) = 1{,}75 \cdot \frac{1}{1 - 3{,}5 \cdot 10^{-2}} = 1{,}81.$$

Якщо умовно взяти схемну надмірність [12] для систем передачі повідомлень при виправленні помилок $q_{\text{сх1}} = 1{,}5$ і в разі їх виявлення $q_{\text{сх2}} = 2{,}5$, то

$$g_1 = g_{\text{ін1}} + g_{\text{сх1}} = 1{,}75 + 1{,}5 = 3{,}25;$$
$$g_2 = g_{\text{ін2}} + g_{\text{сх2}} = 1{,}81 + 2{,}5 = 4{,}31.$$

Імовірність появи помилки в повідомленні без надмірності

$$P_{\text{п}k} = 1 - (1 - P_{\text{е}})^k = 1 - (1 - 5 \cdot 10^{-3})^4 \approx 0{,}02,$$

а в повідомленні з використанням двійкового (7, 4)-коду Хеммінга, що виправляє одну помилку,

$$P_{\text{п}n}(1) = 1 - (1 - P_{\text{е}})^n - nP_{\text{е}}(1 - P_{\text{е}})^{n-1} =$$
$$= 1 - (1 - 5 \cdot 10^{-3})^7 - 7 \cdot 5 \cdot 10^{-3}(1 - 5 \cdot 10^{-3})^6 \approx 0{,}002.$$

Імовірність виникнення помилки при застосуванні цього самого коду для виявлення двократних помилок

$$P_{пп}(2) = 7P_e^3(1-P_e)^4 + 7P_e^4(1-P_e)^3 + P_e^7 = \\ = 7(5\cdot 10^{-3})^3(1-5\cdot 10^{-3})^4 + \\ + 7(5\cdot 10^{-3})^4(1-5\cdot 10^{-3})^3 + (5\cdot 10^{-3})^7 \approx 0{,}86\cdot 10^{-6}.$$

Виграш у захисті від помилок

$$a_1 = \frac{P_{пк}}{P_{пп}(1)} = \frac{0{,}02}{0{,}002} = 10;$$

$$a_2 = \frac{P_{пк}}{P_{пп}(2)} = \frac{0{,}02}{0{,}86\cdot 10^{-6}} \approx 23000.$$

Критерії ефективності

$$k_{1еф} = \log_2 \frac{a_1}{g_1} = \log_2 \frac{10}{3{,}25} \approx 1{,}6;$$

$$k_{2еф} = \log_2 \frac{a_2}{g_2} = \log_2 \frac{23000}{4{,}31} \approx 12{,}3,$$

тобто використання двійкового (7, 4)-коду Хеммінга для виявлення помилок більш ефективне, ніж для їх виправлення.

КОНТРОЛЬНІ ЗАДАЧІ

1. Визначити ймовірності виникнення виявлених $P_{в.п}$ і невиявлених $P_{нв.п}$ помилок у двійковому коді, що виявляє помилки, якщо він складається з таких трьох комбінацій: 001, 010, 100.

2. Визначити ймовірності виникнення виявлених $P_{в.п}$ і невиявлених $P_{нв.п}$ помилок у n-елементному двійковому коді що виявляє помилки, якщо він складається з таких кодових комбінацій: 0001, 0010, 0100, 1000. Мінімальна кодова відстань $d_{min} = 2$. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 10^{-4}$).

3. Визначити ймовірності виникнення виявлених $P_{в.п}$ і невиявлених $P_{нв.п}$ помилок у двійковому коді завдовжки $n = 6$ з перевіркою на парність, якщо канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 5\cdot 10^{-4}$).

Розв'язання. Заданий код виявляє всі помилки непарної кратності й не виявляє помилки парної кратності. Тому ймовірність виявлення помилок

$$P_{в.п} = P(1) + P(3) + P(5) = \\ = C_6^1 P_e(1-P_e)^5 + C_6^3 P_e^3(1-P_e)^3 + C_6^5 P_e^5(1-P_e).$$

Очевидно, найбільша ймовірність виявлення помилки припадає тут на перший член, тобто

$$P_{в.п} \approx P(1) = C_6^1 P_e(1-P_e)^5 \approx 6\cdot 5\cdot 10^{-4} = 3\cdot 10^{-3}.$$

Імовірність виникнення невиявлених помилок

$$P_{нв.п} = \frac{1}{2^r}[P(2)+P(4)+P(6)] = \\ = \frac{1}{2}\left[C_6^2 P_e^2(1-P_e)^4 + C_6^4 P_e^4(1-P_e)^2 + C_6^6 P_e^6(1-P_e)^0\right],$$

але через те, що $P(2) > P(4) > P(6)$, маємо

$$P_{нв.п} \approx \frac{1}{2}P(2) = \frac{1}{2}C_6^2 P_e^2(1-P_e)^4 \approx \frac{15}{2}\left(5\cdot 10^{-4}\right)^2 \approx 1{,}88\cdot 10^{-6}.$$

4. Визначити ймовірності виникнення виявлених і невиявлених помилок у двійковому коді завдовжки $n = 5$ з перевіркою на парність. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 2\cdot 10^{-4}$).

5. Визначити ймовірності виникнення виявлених і невиявлених помилок у двійковому коді зі сталою вагою при $w = 3$ та $n = 5$. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 0{,}5\cdot 10^{-4}$).

6. Визначити ймовірності виникнення виявлених і невиявлених помилок у двійковому інверсному коді при $n = 4$. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 0{,}25\cdot 10^{-3}$).

7. Визначити ймовірності виникнення виявлених і невиявлених помилок у двійковому кореляційному коді при $n = 6$. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 0{,}8\cdot 10^{-3}$).

8. Визначити ймовірності виникнення виправлених $P_{вп.п}$, виявлених $P_{в.п}$ і невиявлених $P_{нв.п}$ помилок у двійковому (7, 4)-коді Хеммінга ($n = 7$, $k = 4$). Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 10^{-3}$), мінімальна кодова відстань $d_{min} = 3$.

Розв'язання. Заданий код виправляє однократні помилки, а виявляє двократні. Тому імовірність виправлення помилок

$$P_{вп.п} = P(1) = C_7^1 P_e(1-P_e)^6 \approx 7\cdot 10^{-3},$$

імовірність виявлення їх

$$P_{в.п} = P(2) = C_7^2 P_e^2(1-P_e)^5 \approx 2{,}1\cdot 10^{-5},$$

а ймовірність виникнення невиявлених помилок

$$P_{нв.п} = P(3) = C_7^3 P_e^3(1-P_e)^4 \approx 1{,}05\cdot 10^{-7}.$$

9. Розв'язати попередню задачу стосовно двійкового (15, 11)-коду Хеммінга ($n = 15$, $k = 11$).

10. Визначити ймовірності виникнення виправлених і невиправлених помилок у двійковому циклічному (7, 4)-коді ($n = 7$, $k = 4$). Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 2\cdot 10^{-3}$), мінімальна кодова відстань $d_{min} = 3$.

11. Визначити ймовірності виникнення виправлених і невиправлених помилок у двійковому циклічному (11, 7)-коді. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 5\cdot 10^{-4}$), мінімальна кодова відстань $d_{min} = 3$.

12. Визначити ймовірності виникнення виправлених і невиправлених помилок у двійковому (15, 7)-коді БЧХ. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 4\cdot 10^{-3}$), мінімальна кодова відстань $d_{min} = 5$.

13. Визначити ймовірності виникнення виправлених і невиправлених помилок у двійковому (15, 5)-коді БЧХ. Канал зв'язку симетричний ($P_{01} = P_{10} = P_e = 0{,}5\cdot 10^{-2}$), мінімальна кодова відстань $d_{min} = 7$.

14. Стиснути лінійним способом, використовуючи знак поділу ρ, масив інформації, який має вигляд

```
4 1 5 2 3 3 8 6 7 1
4 1 5 2 3 3 8 6 7 2
4 1 5 2 3 3 8 6 7 4
4 1 5 2 3 3 8 6 7 0
4 1 5 2 3 3 8 6 7 2
4 1 5 2 3 3 8 6 7 5
4 1 5 2 3 3 8 6 7 1
```

15. Розв'язати попередню задачу та визначити коефіцієнт стиснення стосовно такого масиву інформації:

```
2 8 1 4 5 7 5 2
2 8 1 4 5 7 3 6
2 8 1 4 5 7 4 1
2 8 1 4 5 7 6 9
2 8 1 4 5 7 2 2
2 8 1 4 5 7 3 8
```

16. Розгорнути стиснений масив інформації

```
3 5 6 2 3 3 7 8
ρ 2 1 ρ 3 9 ρ 4
4 2 5 ρ 4 ρ 6 ρ
ρ ρ 3 8
```

17. Відновити первинний масив інформації, якщо на проміжному етапі розгортання стиснений масив мав такий вигляд:

```
6 4 4 5 3 9 8 8
. . . . . . . 7
. . . . . . . 3
. . . . . 4 5 6
. . . . . . 2 1
. . . . 3 1 0 8
4 5 6 1 1 8 4 2
. . . . . 3 5 4
. . . . . . . 6
. . . . . . 2 2
```

18. Використовуючи знак поділу ρ та знак K кінця рядка, стиснути масив інформації

```
3 1 8 8 4 4 1 5 2 0
4 0 8 8 4 4 1 5 2 4
5 6 8 8 4 4 1 5 2 5
7 5 8 8 4 4 1 5 2 1
2 8 8 8 4 4 1 5 2 3
9 7 8 8 4 4 1 5 2 8
7 3 8 8 4 4 1 5 2 2
6 2 8 8 4 4 1 5 2 9
```


19. Розгорнути стиснений масив інформації

```
3 1 8 8 4 4 1 5 2 0
K 4 0 ρ 4 K 5 6 ρ 5
K 7 5 ρ 1 K 2 8 ρ 3
K 9 7 ρ 8 K 7 3 ρ 2
K 6 2 ρ 9 K
```

20. Розв'язати попередню задачу стосовно таких двох масивів інформації:

```
4 7 3 9 3 5 5 6 6 1
K 3 ρ 2 K 9 ρ 8 K ρ
K ρ K 7 ρ 1 K 2 ρ 0
1 3 5 0 2 K 6 ρ 8 K
```

```
1 5 4 6 8 7 7 6
K 4 ρ 8 K 6 ρ 4
7 K 8 ρ 1 4 4 K
5 6 ρ 0 K
```

21. Стиснути масив інформації, заданий в задачі 18, використовуючи символи X = 2 та Z = 5.

22. Використовуючи символи X = 2, Y = 3 та Z = 5, стиснути масив інформації

```
0 2 3 9 4 5 6 7 1 8 6 1 5
2 2 3 9 4 5 6 7 1 8 6 2 6
3 2 3 9 4 5 6 7 1 8 6 3 3
3 2 3 9 7 5 6 7 0 7 6 3 3
6 2 3 9 7 5 5 1 0 7 6 3 4
5 1 3 9 7 4 5 1 0 7 6 3 4
```

23. Розгорнути такі два стиснені масиви інформації:

```
0 2 3 9 4 5 6 7 1 8 6 1 5
2 Z Z 2 6 3 Z Z 3 3 3 Y 7
Y 0 7 Y 6 Z 5 1 X X 4 5 1
Y 4 Z X
```

```
1 2 2 1 1 5 6 7 1 2
3 Z X X 4 6 Y Y 5 3
X 2 Y 6 Y 3 Z X X 2
X 7 Z 2 1 7 Y 4 X X
```

24. Використовуючи символи X, Y і Z, стиснути масив інформації

```
0 9 8 7 6 5 4 3 2 1 0
0 9 8 7 6 8 4 3 2 1 3
1 9 8 7 4 8 4 3 2 0 6
4 3 8 7 3 8 4 3 2 0 1
4 3 8 7 5 8 4 3 2 0 1
5 3 8 7 5 3 4 3 6 0 1
5 3 7 7 5 3 4 3 6 0 1
7 1 7 7 4 3 4 3 2 0 1
```

25. Стиснути масив інформації, заданий у попередній задачі, використовуючи знак поділу ρ та знак K кінця рядка.

26. Первинний масив інформації має вигляд

8	7	0	1	2	3	4	5	6	8	8	2	3	5	6
8	7	0	1	2	3	4	5	6	8	8	1	3	4	2
4	7	0	1	2	3	4	5	6	8	8	1	4	3	1
4	7	0	1	2	3	4	5	6	8	8	1	4	1	7
2	6	0	1	2	3	4	5	6	8	8	1	4	1	5
2	6	0	1	2	3	4	5	6	8	8	1	5	4	7

Стиснути цей масив, використовуючи: а) знак поділу ρ; б) знак поділу ρ та знак К кінця рядка; в) символи X, Y, Z. Порівняти утворені масиви між собою.

27. Яким способом найкраще стиснути такий масив інформації:

0	9	8	7	6	5	4	3
0	9	8	7	6	5	4	3
0	9	8	7	6	5	4	3
0	9	8	7	6	5	4	3
0	9	8	7	6	5	4	3?

Показати процес згортання та розгортання цього масиву.

28. Способом зонного стиснення інформації з використанням табл. 9.3 закодувати такий текст: «Стиснення інформації використовується для прискорення процесів оброблення, зберігання та пошуку інформації». Визначити коефіцієнт стиснення.

29. Розв'язати попередню задачу стосовно такого тексту: «Майже всі сучасні ЕОМ побудовано на дискретних елементах».

30. Розгорнути масив інформації

D1C4A08AD9C1490D4C1D5C980D5C5D6C0D1C242D5C3
E4C9097E0C14D0C3E3C990D4C24D7E7C6D8CBD6C08A
D0C1DBC3D0CA0D5C108941DBC9D5C71B69097E0C14D
0C3E3C99E8C0E1C10D1C242D5C3D2C6D8CBD6E8C0
630E2C8AD54C1B690990D1C242D5C3E4C9E9,

який був згорнений способом зонного стиснення з використанням табл. 9.1, та визначити коефіцієнт стиснення.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Чим оцінюється вірогідність кодованих повідомлень?
2. Як визначається ймовірність виникнення *i*-кратних помилок у разі їх незалежності?
3. Як оцінюються ймовірності виникнення виявлених і невиявлених помилок у симетричному каналі для кодів, що виявляють помилки?
4. Як оцінюються ймовірності виникнення виправлених і невиправлених помилок для кодів, що виправляють помилки?
5. Як визначається ймовірність виникнення помилок у разі передачі повідомлень по каналах з пакетним розподілом помилок?
6. Що розуміють під стисненням інформації?
7. Як класифікуються способи стиснення інформації?
8. Який основний недолік способів стиснення інформації без відновлення її початкового стану?
9. Які способи стиснення належать до лінійних?
10. Що характеризує матричні, комбіновані та каскадні способи стиснення інформації?
11. Як виконується стиснення інформації з використанням замість повторень додаткових символів?
12. Які способи стиснення використовуються для архівації даних на дисках?
13. На чому грунтується стиснення інформації виключенням даних, які повторюються в різних файлах?
14. Які переваги має спосіб зонного стиснення інформації?
15. У чому полягає спосіб стиснення інформації зменшенням розрядності кодованих слів?
16. Які переваги має спосіб стиснення інформації заміною деяких комбінацій літер одиничними символами?
17. На чому грунтується спосіб стиснення інформації використанням адаптивного кодування?
18. У чому полягає спосіб стиснення інформації зберіганням атрибутів у вигляді бітової матриці?
19. Які основні критерії ефективності завадостійких кодів?
20. Що показують верхні та нижні межі для кодової відстані?
21. Як впливає збільшення основи коду на його ефективність?
22. Які коди рекомендується використовувати в системах передачі інформації зі зворотним зв'язком?
23. Як впливає на вірогідність передачі повідомлень уведення зворотного зв'язку в системах передачі інформації?
24. Як оцінюється ефективність методів підвищення вірогідності передачі інформації?

Додаток 1. **Фрагмент таблиці значень двійкових логарифмів цілих чисел**

x	$\log_2 x$	x	$\log_2 x$	x	$\log_2 x$	x	$\log_2 x$
1	0,000	36	5,170	71	6,150	106	6,728
2	1,000	37	5,209	72	6,170	107	6,741
3	1,585	38	5,248	73	6,190	108	6,755
4	2,000	39	5,285	74	6,209	109	6,768
5	2,322	40	5,322	75	6,229	110	6,781
6	2,585	41	5,358	76	6,248	111	6,794
7	2,807	42	5,392	77	6,267	112	6,807
8	3,000	43	5,426	78	6,285	113	6,820
9	3,170	44	5,459	79	6,304	114	6,833
10	3,332	45	5,492	80	6,322	115	6,845
11	3,459	46	5,524	81	6,340	116	6,858
12	3,585	47	5,555	82	6,358	117	6,870
13	3,700	48	5,585	83	6,375	118	6,883
14	3,807	49	5,615	84	6,392	119	6,895
15	3,907	50	5,644	85	6,409	120	6,907
16	4,000	51	5,672	86	6,426	121	6,919
17	4,087	52	5,700	87	6,443	122	6,931
18	4,170	53	5,728	88	6,459	123	6,943
19	4,248	54	5,755	89	6,476	124	6,954
20	4,322	55	5,781	90	6,492	125	6,966
21	4,392	56	5,807	91	6,508	126	6,977
22	4,459	57	5,833	92	6,524	127	6,989
23	4,524	58	5,858	93	6,539	128	7,000
24	4,585	59	5,883	94	6,555	200	7,644
25	4,644	60	5,907	95	6,570	256	8,000
26	4,700	61	5,931	96	6,585	300	8,229
27	4,755	62	5,951	97	6,600	400	8,644
28	4,807	63	5,977	98	6,615	500	8,966
29	4,858	64	6,000	99	6,629	512	9,000
30	4,907	65	6,022	100	6,614	600	9,229
31	4,954	66	6,044	101	6,658	700	9,451
32	5,000	67	6,066	102	6,672	800	9,644
33	5,044	68	6,087	103	6,687	900	9,814
34	5,087	69	6,109	104	6,700	1000	9,965
35	5,129	70	6,129	105	6,714	10 000	13,288

Додаток 2. **Фрагмент таблиці значень функції $-p\log_2 p$**

p	$-p\log_2 p$	p	$-p\log_2 p$	p	$-p\log_2 p$	p	$-p\log_2 p$
0,001	0,0099	0,027	0,1407	0,150	0,4105	0,600	0,4432
0,002	0,0179	0,028	0,1444	0,160	0,4230	0,610	0,4350
0,003	0,0251	0,029	0,1481	0,175	0,4400	0,625	0,4238
0,004	0,0319	0,030	0,1518	0,180	0,4453	0,650	0.4040
0,005	0,0382	0,032	0,1589	0,190	0,4552	0,675	0,3828
0,006	0,0443	0,035	0,1693	0,200	0,4644	0,700	0,3602
0,007	0,0501	0,037	0,1760	0,210	0,4728	0,710	0,3508
0,008	0,0557	0,040	0,1858	0,225	0,4842	0,725	0,3364
0,009	0,0612	0,042	0,1941	0,250	0,5000	0,750	0,3113
0,010	0,0664	0,045	0,2013	0,275	0,5122	0,775	0,2850
0,011	0,0716	0,047	0,2073	0,300	0,5211	0,800	0,2575
0,012	0,0766	0,050	0,2161	0,310	0,5238	0,810	0,2462
0,013	0,0814	0,055	0,2301	0,325	0,5270	0,825	0,2290
0,014	0,0862	0,060	0,2435	0,350	0,5301	0,850	0,1993
0,015	0,0909	0,065	0,2563	0,375	0,5306	0,875	0,1810
0,016	0,0954	0,070	0,2686	0,400	0,5288	0,900	0,1368
0,017	0,0999	0,075	0,2803	0,410	0,5274	0,910	0,1238
0,018	0,1043	0,080	0,2915	0,425	0,5246	0,925	0,1040
0,019	0,1086	0,085	0,3023	0,450	0,5184	0,950	0,0703
0,020	0,1129	0,090	0,3127	0,475	0,5102	0,975	0,0356
0,021	0,1170	0,095	0,3226	0,500	0,5000	0,980	0,0286
0,022	0,1211	0,100	0,3322	0,510	0,4954	0,990	0,0143
0,023	0,1252	0,110	0,3503	0,525	0,4880	0,095	0,0072
0,024	0,1291	0,125	0,3750	0,550	0,4744	0,097	0,0043
0,025	0,1330	0,130	0,3826	0,575	0,4591	0,999	0,0014
0,026	0,1369	0,140	0,3971				

СПИСОК ВИКОРИСТАНОЇ ТА РЕКОМЕНДОВАНОЇ ЛІТЕРАТУРИ

1. *Арманд В. А., Железнов В. В.* Штриховые коды в системах обработки информации. — М.: Радио и связь, 1989. — 92 с.
2. *Блейхут Р.* Теория и практика кодов, контролирующих ошибки. — М.: Мир, 1986. — 576 с.
3. *Галлагер Р.* Теория информации и надежная связь. — М.: Сов. радио, 1974. — 720 с.
4. *Горяинов О. А., Хохлов Г. И.* Элементы теории информации и кодирования / МИРЭА. — М., 1985. — 117 с.
5. *Емельянов Г. А., Шварцман В. О.* Передача дискретной информации. — М.: Радио и связь, 1982. — 240 с.
6. *Жураковский Ю. П., Волошин В. И.* Многочастотные системы передачи дискретных сигналов. — К.: Техніка, 1981. — 120 с.
7. *Жураковский Ю. П., Назаров В. Д.* Каналы связи. — К.: Вища шк. Головное изд-во, 1985. — 236 с.
8. *Жураковский Ю. П.* Передача информации в ГАП. — К.: Вища шк., 1991. — 216 с.
9. *Злотник Б. М.* Помехоустойчивые коды в системах связи. — М.: Радио и связь, 1989. — 232 с. — (Стат. теория связи; Вып. 31).
10. *Зюко А. Г., Коробов Ю. Ф.* Теория передачи сигналов. — М.: Связь, 1972. — 282 с.
11. *Игнатов В. А.* Теория информации и передачи сигналов. — М.: Радио и связь, 1991. — 280 с.
12. *Кодирование* информации (двоичные коды) / Н. Т. Березюк, А. Г. Андрущенко, С. С. Мощицкий и др. — Харьков: Вища шк. Изд-во при Харьк. ун-те, 1978. — 252 с.
13. *Колесник В. Д., Полтырев Г. Ш.* Курс теории информации. — М.: Наука, 1982. — 416 с.
14. *Коржик В. И., Финк Л. М.* Помехоустойчивое кодирование дискретных сообщений в каналах со случайной структурой. — М.: Связь,1975. — 272 с.
15. *Котельников В. А.* Теория потенциальной помехоустойчивости. — М.: Госэнергоиздат, 1956. — 152 с.
16. *Кохманюк Д.* Сжатие данных: Как это делается. Ч. 2 // Index PRO. — 1993. — № 2. — С. 30—49.
17. *Кричевский Р. Е.* Сжатие и поиск информации. — М.: Радио и связь, 1989. — 168 с.
18. *Кузьмин И. В., Кедрус В. А.* Основы теории информации и кодирования. — К.: Вища шк. Головное изд-во, 1986. — 238 с.
19. *Курбаков К. И.* Кодирование и поиск информации в автоматическом словаре. — М.: Сов. радио, 1968.
20. *Логинов В. М., Цепков Г. В., Чинаев П. И.* Экономичное кодирование. — К.: Техніка, 1976. — 176 с.
21. *Мельников Ю. Н.* Достоверность информации в сложных системах. — М.: Сов. радио, 1973. — 192 с.
22. МикроЭВМ в информационно-измерительных системах / С. М. Переверткин, Н. И. Гаранин, Ю. Н. Костин, И. И. Миронов. — М.: Машиностроение, 1987. — 248 с.
23. *Митюшин К. Г.* Телеконтроль и телеуправление в энергосистемах. — М.: Энергоатомиздат, 1990. — 288 с.
24. *Муттер В. М.* Основы помехоустойчивой телепередачи информации. — Л.: Энергоатомиздат. Ленингр. отд-ние, 1990. —288 с.
25. *Питерсон У., Уэлдон Э.* Коды, исправляющие ошибки. — М.: Мир, 1976. — 590с.
26. *Резник Ю.* Алгоритмы в системах сжатия реального времени // Index PRO. — 1994. — № 1. — С.22—33.
27. *Самойленко С. И.* Помехоустойчивое кодирование. — М.: Наука, 1966. — 240 с.
28. *Советов Б. Я.* Теория информации. — Л.: Изд-во Ленингр. ун-та, 1977. — 184 с.
29. *Стратонович Р. Л.* Теория информации. — М.: Сов. радио, 1975. — 420 с.
30. *Субье-Ками А.* Двоичная техника и обработка информации: Пер. с фр. / Под ред. Д. Ю.Панова. — М.: Мир, 1964. — 500с.
31. *Темников Ф. Е., Афонин В. А., Дмитриев В. И.* Теоретические основы информационной техники. — М.: Энергия, 1979. — 512 с.
32. *Теория* кодирования: Пер. с яп. / Т. Касами, Н. Токура, Е. Ивадари, Я. Инагаки. — М.: Мир, 1978. — 576 с.
33. *Теория* передачи сигналов / А. Г.Зюко, Д. Д. Кловский, М. В. Назаров, Л. М. Финк. — М.: Радио и связь,1986. — 304 с.
34. *Тутевич В. Н.* Телемеханика. — М.: Высш. шк., 1985. — 423 с.
35. *Файнстейн А.* Основы теории информации. — М.: Изд-во иностр. лит., 1960. — 140 с.
36. *Фано Р.* Передача информации. Статистическая теория связи. — М.: Мир, 1965. — 483 с.
37. *Форни Д.* Каскадные коды / Пер с англ. под ред. С. И. Самойленко. — М.: Мир, 1970. — 208с.
38. *Харкевич А. А.* Очерки общей теории связи // Избр.тр. — М.: Наука, 1973 — Т.3. — 194 с.
39. *Хайслей Т.* Передача данных и системы телеобработки: Пер. с англ. — М.: Радио и связь, 1982. — 200 с.
40. *Хохлов Г. И.* Элементы теории корректирующих кодов / МИРЭА. — М., 1980. — 136 с.
41. *Цымбал В. П.* Задачник по теории информации и кодированию. — К: Вища шк. Головное изд-во, 1976. — 276 с.
42. *Цымбал В. П.* Теория информации и кодирование. — К.: Вища шк., 1992. — 263 с.
43. *Четвериков В. Н.* Подготовка и телеобработка данных в АСУ. — М.: Высш. шк., 1981. — 320 с.
44. *Шеннон К.*-Э. Работы по теории информации и кибернетики. — М.: Изд-во иностр. лит., 1963. — 830 с.
45. *Шляпоберский В. И.* Элементы дискретных систем связи. — М.: Воениздат, 1965. — 304 с.
46. *Элементы* теории передачи дискретной информации / Л. П. Пуртов, А. С. Замрий, А. И. Захаров, В. М. Охорзин; Под ред. Л. П. Пуртова. — М.: Связь, 1972. — 232 с.
47. *Элементы* технической кибернетики. Терминология. — М.: Наука, 1968. — 52 с.
48. *Элиас П.* Безошибочное кодирование: Коды с обнаружением и исправлением ошибок / Под ред. А.М.Петровского. — М., 1956. — С.59—71.
49. *Яглом А. М., Яглом И. М.* Вероятность и информация. — М.: Наука, 1973. - 511 с.