

Техніка безпеки. Інформація та інформаційні процеси

*Кодування
інформації*



Кодування інформації

Щоб передати інформацію, її необхідно закодувати відповідно до природи каналу передачі. Наприклад, розмова по телефону (звукові сигнали) автоматично кодується в електричні сигнали, щоб вони могли по мережі дійти до співрозмовника і там перетворитися у звукові сигнали.

Джерело -> **Кодування** -> Канал передачі -> **Декодування** -> Приймач

Поняття кодування та коду

Кодування - це перетворення інформації без зміни її змісту в інший вигляд за допомогою певного коду.

Код - це набір правил перетворення для кодування.

Прикладом учнівського кодування є такий прийом: виписується алфавіт і всі букви нумеруються за порядком – створюється кодова таблиця. Далі при створенні записки замість букв записуються номери.

Одержувач такої записки повинен користуватися цією ж кодовою таблицею, щоб розшифрувати і прочитати повідомлення.

А – 1	Д – 6	И – 11	Л – 16	Р – 21	Х – 26	Б – 31
Б – 2	Е – 7	І – 12	М – 17	С – 22	Ц – 27	Ю – 32
В – 3	Є – 8	Ї – 13	Н – 18	Т – 23	Ч – 28	Я – 33
Г – 4	Ж – 9	Й – 14	О – 19	У – 24	Ш – 29	
Ґ – 5	З – 10	К – 15	П – 20	Ф – 25	Щ – 30	

Кодування інформації у комп'ютері

У комп'ютері носіями інформації є електричні або магнітні сигнали, які можуть мати лише два значення:

0 - вимкнено (нема струму, розмагнічено) або

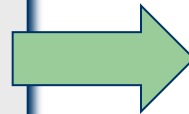
1 - увімкнено (є струм, намагнічено).



Демонстрація перемикача сигналу.
Якщо клікнути лівою кнопкою миші у верхньому квадратику, то там з'явиться прапорець, а під ним цифра **1**, що свідчитиме про наявність сигналу. Наступне клацання знімить прапорець, а унизу з'явиться цифра **0**, що свідчитиме про відсутність сигналу.

Кодування інформації у комп'ютері

За допомогою 0 і 1 кодують будь-яку інформацію, яку обробляє комп'ютер. Якщо застосувати згаданий раніше метод, то цифр 0 і 1 вистачить на позначення лише двох інформаційних одиниць, наприклад символів.



Введемо наступні позначення:

0 – М

1 – А

Тоді код **0101** можна читати як **МАМА**



Введемо наступні позначення:

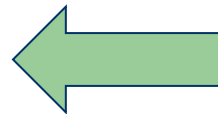
00 – К

01 – Л

10 – О

11 – Я

Тоді **00 10 01 11** можна читати як **КОЛЯ**, а **10 01 11** – як **ОЛЯ**



Можна застосувати кодування комбінацією із кількох 0 і 1. Якщо взяти по два знаки це дозволить кодувати вже 4 інформаційні одиниці (00, 01, 10, 11).

Кодування інформації у комп'ютері

Число комбінацій залежить від того, скількох значний код:

- однозначний – 2 (2^1)
- двозначний – 4 (2^2)
- тризначний – 8 (2^3)
- чотиризначний – 16 (2^4)
- п'ятизначний – 32 (2^5)
- шестизначний – 64 (2^6)
- семизначний – 128 (2^7)
- восьмизначний – **256** (2^8)

Для кодування інформації в комп'ютерах використовується **восьмизначний двійковий код**.

Такої кількості символів вистачає для відображення літер, цифр, розділових, а також графічних елементів. Малюнки і звуки зображуються тими ж комбінаціями, але перед ними стоїть спеціальний знак, який вказує на характер інформації.

Для більшої зрозумілості розглянемо приклади кодування.

Наприклад, кирилична буква **А** має двійковий код **10000000**, а буква **М** – код **10001100**. Тоді слово **МАМА** в комп'ютері матиме представлення **10001100 10000000 10001100 10000000**

Кодова таблиця символів

Усі символи, які використовує комп'ютер, заносяться в таблицю і нумеруються десятковими числами, які перетворюються у двійковий код із 8 знаків 0 і 1 - виходить кодова таблиця символів.

Наприклад, в комп'ютерах застосовується восьмизначний код ASCII, розрахований на 256 символів. Для кожного символу - це послідовність з 8-ми цифр 0 і 1, яка визначає певний символ.

Витяг з кодової таблиці символів ASCII

Номер	Двійковий код	Символ	Номер	Двійковий код	Символ	Номер	Двійковий код	Символ
128	10000000	А	139	10001011	Л	150	10010110	Ц
129	10000001	Б	140	10001100	М	151	10010111	Ч
130	10000010	В	141	10001101	Н	152	10011000	Ш
131	10000011	Г	142	10001110	О	153	10011001	Щ
132	10000100	Д	143	10001111	П	154	10011010	Ї
133	10000101	Е	144	10010000	Р	155	10011011	І
134	10000110	Ж	145	10010001	С	156	10011100	Ь
135	10000111	З	146	10010010	Т	157	10011101	Є
136	10001000	И	147	10010011	У	158	10011110	Ю
137	10001001	Й	148	10010100	Ф	159	10011111	Я
138	10001010	К	149	10010101	Х	160	10100000	а

Кодування байтами

У комп'ютер заноситься та обробляється текстова інформація, цілі і дійсні числа, графічні побудови, малюнки і музика.

І всі вони представлені у вигляді комбінацій із восьми «0» і «1» - **байтами**.

Як б інформація в комп'ютері не була (голос диктора або його зображення), вона завжди кодується байтами.

10001100

10000000

10001100

10000000

1-й байт

2-й байт

3-й байт

4-й байт

Текст у пам'яті комп'ютера

Визначимо головну відмінність між книгою і занесеним в комп'ютер текстом.

Вона полягає в тому, що коли ми відкриваємо книгу, то бачимо зображення символів, сформованих дрібними крихтами друкарської фарби.

А якщо ж «відкриємо» пам'ять комп'ютера, то «побачимо» коди літер, складених із 0 і 1. Звичайний текст зображується в комп'ютері послідовністю кодів. Тобто замість кожної літери тексту зберігається її код за кодовою таблицею.

І тільки при виведенні літер у зовнішнє середовище (на папір або екран монітора) проводиться формування звичних зорових образів літер.

Двобайтне кодування (UNICODE)

В сучасних програмах розповсюджене кодування символів 16-бітним кодом (2 байти)

UNICODE

з номерами від **0** до **65535**, що практично вміщує алфавіти всіх народів світу.

Вимірювання об'єму інформації

Інформація відображається символами (байтами), і її об'єм вимірюється кількістю байт у повідомленні.

Наприклад, текст «*Данило грає у футбол*» має 20 байтів, і це можна порахувати. Об'єм інформації вимірюється кількістю байт у повідомленні (текстовому, графічному тощо).

В тексті розділові знаки і «пропуски» - теж символи.

Нехай на аркуші 56 рядків по 64 символи у рядку. Такий аркуш буде мати $56 \times 64 = 3584$ байт інформації.

Вимірювання великих об'ємів інформації

Для вимірювання великих об'ємів інформації використовують **кіло-, мега-, гігабайти**.

В інформатиці, обчислювальній техніці префікси **кіло-, мега- і гіга-** мають дещо інший зміст, ніж в інших науках.

- 1 КБ(кілобайт) = 2^{10} байт = 1024 байт
- 1 МБ(мегабайт) = 2^{10} КБ = 1024 КБ
- 1 ГБ(гігабайт) = 2^{10} МБ = 1024 МБ

Якщо є необхідність приблизно підрахувати об'єм інформації цілої книжки, треба кількість символів на одній сторінці помножити на кількість сторінок. Нехай книжка має 256 сторінок по 3584 байти на сторінці. Об'єм інформації такої книги:

$3584 \times 256 = 917504$ байт = 896 кілобайт = 0,875 мегабайта.

Питання для самоконтролю:

- 1 Як відбувається кодування-декодування при телефонній розмові?
- 2 Які етапи проходить інформація від джерела до приймача ?
- 3 Як можна найпростіше закодувати інформацію?
- 4 Яким чином кодується інформація в комп'ютері?
- 5 Чому в комп'ютері використовується саме двійковий код?
- 6 У чому відмінність запису тексту в книзі і комп'ютері?
- 7 Яким чином зберігаються літери тексту в комп'ютері?
- 8 Як побудована кодова таблиця символів ASCII?
- 9 Яке слово закодовано кодом ASCII: 145 128 139 142?
- 10 Яке слово закодовано двійковим кодом: 10001010 10001110 10000100?
- 11 Якими одиницями вимірюється об'єм інформації?
- 12 Скільки байт має кілобайт ?
- 13 Скільки кілобайт має мегабайт?
- 14 Скільки байт має мегабайт?
- 15 Як порахувати об'єм інформації на сторінці? У всій книзі?

