

## § 3. Інформація й повідомлення.

### Кодування повідомлень

*Вивчивши цей параграф, ми:*

*дізнаємося, який взаємозв'язок існує між інформацією й повідомленням;*

*зрозуміємо, що таке кодування і як кодуються в комп'ютері тексти, зображення, звуки й числа;*

*довідаємося, як вимірюється обсяг інформації в комп'ютері.*

#### ====3.1. Інформація й повідомлення=====

Інформація подається, передається і зберігається у вигляді **повідомлень**. Для створення повідомлень застосовують певні виразні засоби, які називають мовою.

Кожна мова має певний набір знаків, який складає її основу, і правила утворення припустимих конструкцій мови. За історію свого існування людство виробило багато різних мов: усні й писемні мови, якими користуються різні народи, мову музики й математики, креслень і умовних знаків тощо. При виборі мови для створення повідомлення виходять із того, щоб вона була зрозумілою для адресата повідомлення і зручною для подання інформації за даних обставин.

Між інформацією та повідомленням немає однозначної відповідності. Одну й ту саму інформацію можна відтворити в різних повідомленнях, застосовуючи різні мови. Одне й те саме повідомлення може нести різну за змістом інформацію. Наприклад, похвалити учня за правильну відповідь можна і словами, і жестом, і високою оцінкою в журналі. А зміст повідомлення «Он як ти постарався!» цілком залежить від обставин, в яких ми його почули, і може означати і схвалення, і догану.

На розуміння змісту повідомлення впливає власний досвід того, хто його сприймає. Так, написи на стіні стародавнього храму розкривають досвідченому досліднику певну інформацію, а для новачка зберігають свою таємницю.

Процес інформаційного обміну завжди відбувається як перетворення інформації на повідомлення і навпаки — повідомлення на інформацію. Той, хто постає джерелом інформації, складає на її основі повідомлення, а той, хто одержує це повідомлення, здійснює зворотний процес: він сприймає повідомлення, осмислює його і дістає в такий спосіб інформацію.

#### ====3.2. Кодування повідомлень=====

У процесі передавання повідомлень, як ми вже знаємо, можуть брати участь технічні системи. У технічних системах повідомлення подаються за допомогою сигналів. Наприклад, якщо ви надсилаєте телеграму, то телеграфний апарат перетворює текст повідомлення на послідовність електричних сигналів. На приймальному телеграфному пункті електричні сигнали перетворюються на послідовність літер, яку і прочитає адресат. Обидва перетворення повідомлення — з текстової форми на сигнальну й із сигнальної форми на текстову — здійснювалися формально, за певними однозначними правилами, які складають основу функціонування телеграфних апаратів.

**Перетворення повідомлення** за певними правилами з однієї форми його подання на іншу називається кодуванням (від лат. *codex* — звід законів), а результат цього перетворення — кодом.

У багатьох сучасних технічних системах використовується цифровий спосіб кодування повідомлень. До таких систем належить і комп'ютер.

### ====3.3. Двійкове кодування — основа роботи комп'ютера =====

Фізичною основою роботи комп'ютера є впевнене розрізнення двох рівнів (низького і високого) електричної напруги, намагніченості тощо. Низький рівень можна співставити із значенням 0, високий — із значенням 1. Це дозволяє стверджувати, що комп'ютер сприймає і опрацьовує тільки два знаки — «0» і «1». Іншими словами, алфавіт комп'ютерної мови є двійковим.

Зазначимо, що вибір знаків «0» і «1» є умовним. Із тим самим успіхом можна було б вибрати позначення «+» і «-» або «<» і «>», «○» і «×» тощо.

Знаки 0 та 1 двійкового алфавіту, на відміну від звичайних десяткових цифр, називають **двійковими знаками, або бітами**. Слово «біт» є скороченням від англ. binary digit — двійковий знак, двійкова цифра. Бітом називають також двійковий розряд.

Подання повідомлень у комп'ютері за допомогою двійкових знаків називається **двійковим кодуванням**.

Завдяки двійковому кодуванню всі дії з опрацювання повідомлень, які виконує комп'ютер, зводяться до сукупності простих дій над 0 та 1. Спроможність комп'ютера надзвичайно швидко і точно виконувати ці дії зумовлює його потужність як універсального інструменту для роботи з інформацією.

Для кодування текстових, числових, графічних і звукових повідомлень в комп'ютері застосовуються відповідні системи правил.

### ====3.4. Кодування текстів і чисел=====

Розглянемо спочатку кодування текстових повідомлень в комп'ютері. Для кодування тексту необхідно призначити певний індивідуальний двійковий код кожному символу, який може зустрітися в тексті, у тому числі й пропуску. Усім символам призначаються коди однакової довжини. Це дозволяє подавати текст як суцільну послідовність бітів, а для відокремлення символів просто поділяти її на рівні за довжиною порції.

Щоб утворити потрібну кількість різних між собою кодів, необхідно вибрати відповідну довжину коду. Позначимо довжину коду через  $n$ . Якщо код складається з одного біта ( $n = 1$ ), то зрозуміло, що можна мати тільки два різні коди — 0 і 1. Цього замало: адже в тексті не дві літери.

Для збільшення кількості кодів будемо нарощувати його довжину. Для  $n = 2$  кількість різних кодів становитиме 4: 00, 01, 10, 11. Для  $n = 3$  отримаємо 8 кодів: 000, 001, 010, 011, 100, 101, 110, 111.

Як бачимо, кожне збільшення довжини коду на 1 біт при-водить до подвоєння кількості різних кодів: до тих кодів, що були отримані на попередньому етапі, один раз додаємо 0, другий раз — 1. Отже, кількість різних між собою двійкових кодів  $k$  пов'язана із довжиною коду  $n$  таким співвідношенням:

$$k = 2^n$$

Призначення індивідуальних кодів текстовим символам закріплюється стандартами. Стандарти кодування забезпечують можливість здійснювати обмін електронними текстовими документами між окремими особами, організаціями, країнами.

Першоосною для всіх стандартів є ASCII — American Standard Code for Information Interchange (американський стандартний код для обміну інформацією). Коди ASCII мають довжину 8 біт.

Вісімка бітів називається **байтом** (від англ. byte). Отже, 1 байт = 8 біт.

Довжина у 8 біт дозволяє утворити  $2^8 = 256$  різних кодів — від 00000000 до 11111111. Це дозволяє призначити індивідуальні коди великим і маленьким літерам латинського й

українського алфавіту, розділовим знакам та цифрам, знакам математичних операцій і операцій порівняння, лапкам і дужкам і т. ін.

Таблиця *кодів ASCII* складається з двох рівних частин. Одна частина (з кодами від 00000000 до 01111111) є базовою, вона містить коди великих і маленьких літер латинського алфавіту, цифр, розділових і математичних знаків тощо. Друга частина — розширена (з кодами від 10000000 до 11111111) — містить коди великих і маленьких літер національної мови та інші символи, які застосовуються в країні для внутрішнього обміну інформацією.

Для того щоб включити в кодову таблицю символи всіх національних алфавітів, у тому числі й основні ієрогліфи, було розроблено стандарт кодування Unicode, який спирається на двохбайтне кодування символів, тобто на 16-бітні коди. Це дозволяє утворити  $2^{16} = 65\,536$  кодів і закріпити їх за різними символами. Система Unicode вбирає в себе ASCII-коди, тобто зберігає базову таблицю ASCII.

Якщо текст містить число, то кожна його цифра кодується так само, як і літера слова. Проте над числами в тексті не можна здійснювати математичних операцій. Для того щоб комп'ютер сприймав числове значення і застосовував його так, як це прийнято в математиці, числа кодуються іншим чином.

Спосіб кодування чисел залежить від знаку числа (додатне чи від'ємне), його типу (ціле чи дробове) тощо. Довжина коду числа може бути різною: 1, 2, 4, 8 байт. Різні способи кодування чисел дозволяють оптимально використовувати ресурси комп'ютера при виконанні обчислень і забезпечувати їх високу точність.

Сучасний комп'ютер відтворює множину цілих чисел в діапазоні від  $-10^{20}$  до  $+10^{20}$  і працює з багатозначними дробовими числами в діапазоні від  $\pm 10^{-308}$  до  $\pm 10^{+308}$ .

### ====3.5. Кодування зображень=====

Найпростішим способом побудови зображення на екрані комп'ютера є його формування мозаїчним способом з окремих елементів — *пікселів* (від англ. pixel (picture element) — елемент зображення). Надаючи кожному пікселю певного кольору, можна створити будь-яке зображення на екрані. Створене зображення зберігається й обробляється в комп'ютері як таб-лиця значень кольорів усіх пікселів тієї прямокутної частини, яку займає зображення на екрані.

Якість зображення визначається двома факторами: кількістю пікселів на екрані й кількістю різних кольорів, які можна застосовувати. Чим більше пікселів, тим краще передається зображення. Часто на екрані налічується 1024 пікселів вздовж горизонталі й 768 вздовж вертикалі, тобто 786 432.

Кількість кольорів, які можна застосовувати для побудови зображення, залежить від довжини двійкового коду кольору. Чим більша довжина коду, тим більше різних кодів можна утворити й позначити ними різні кольори. Отже, багатство кольорової палітри визначається кількістю бітів у коді кольору.

У сучасних комп'ютерах для подання кольору звичайно використовуються від 2 до 4 байт. Два байти (16 біт) дозволяють розрізняти  $2^{16}$ , тобто 65 536, кольорів і відтінків. Такий режим подання зображень називається High Color — висококольоровий. Чотири байти (32 біти) забезпечують кольорову гаму в  $2^{32}$ , тобто 4 294 967 296, кольорів і відтінків (приблизно 4,3 мільярда). Такий режим називається True Color — справжній (істинний) колір.

### ====3.6. Кодування звуків=====

Звук у комп'ютері спочатку перетворюється на електричний сигнал. Цей сигнал є неперервним, і для кодування здійснюють його вимірювання певну кількість разів за кожен секунду, тобто з певною частотою.

Подання неперервного сигналу сукупністю його окремих значень називається **дискретизацією** (від лат. *discretus* — розділений, переривчастий), а кількість значень, які відповідають тривалості сигналу в одну секунду, — частотою дискретизації.

Таким чином, як графічний образ складається із сукупності окремих різнокольорових точок, так і електричний сигнал подається як послідовність його окремих значень у певні моменти часу.

Якість подання звуку в комп'ютері визначається частотою дискретизації й точністю подання вимірених значень. Чим вища частота дискретизації і чим більша кількість різних значень, якими можна характеризувати сигнал, тим краща якість відображення звуку.

У сучасних комп'ютерах звичайно застосовується частота дискретизації в 22 КГц або 44,1 КГц (1 КГц (кілогерц) — це тисяча вимірювань за 1 секунду), а для подання значення сигналу виділяються 2 байти (16 біт), що дозволяє розрізнити  $2^{16}$ , тобто 65 536, значень.

### ====3.7. Вимірювання обсягу повідомлень і ємності носіїв =====

Сукупна довжина двійкового коду повідомлення є мірою його обсягу в комп'ютері або на носії. Така міра є зручною, тому що вона не пов'язана ні з видом повідомлення, ні з його смисловою значимістю. Текстове чи звукове, надзвичайно важливе чи ні, повідомлення однаково характеризується тією кількістю двійкових знаків, яку воно являє собою у вибраній системі кодування.

Обсяг повідомлення вимірюється в бітах і байтах. Таким чином, байт і біт виконують ще одну роль — роль одиниць вимірювання обсягу повідомлень.

Якщо на ваш запит про погоду на завтра комп'ютер надіслав повідомлення: «Завтра очікується ясна і тепла погода.» або «Завтра очікується землетрус у 8 балів!», то в обох випадках ви однаково отримали 38 байт — за умови, що комп'ютер працює в кодах ASCII. Якщо ж комп'ютер застосовує, наприклад, систему кодування Unicode, де символи кодуються двома байтами, то обсяг кожного з повідомлень виявиться вдвічі більшим і складатиме 76 байт, або 608 біт.

Обсяг двійкового коду, який передається за 1 секунду, є характеристикою швидкодії пристроїв передавання в комп'ютерних мережах і каналах зв'язку. Швидкість передавання повідомлень звичайно вимірюється в бітах за секунду.

Максимальний обсяг двійкового коду, який може бути розміщений на носії, є показником його ємності. Ємність носіїв вимірюється в байтах.

На практиці, щоб уникнути великих чисел, застосовують більш ємні, ніж байт, одиниці вимірювання обсягу повідомлень і ємності носіїв — **кілобайти**, **мегабайти**, **гігабайти** і **терабайти**. Множником при переході до наступної одиниці вимірювання виступає число 1024, яке дорівнює  $2^{10}$ .

#### **Одиниці вимірювання обсягу повідомлень і ємності носіїв**

Назва	Умовне позначення	Співвідношення з іншими одиницями
Кілобайт	Кбайт (Кб)	1 Кбайт = $2^{10}$ байт = 1024 байт
Мегабайт	Мбайт (Мб)	1 Мбайт = $2^{20}$ байт = 1024 Кбайт
Гігабайт	Гбайт (Гб)	1 Гбайт = $2^{30}$ байт = 1024 Мбайт
Терабайт	Тбайт (Тб)	1 Тбайт = $2^{40}$ байт = 1024 Гбайт

Визначимо, наприклад, чи можна записати на дискету ємністю в 1,44 Мб створене на екрані комп'ютера зображення, якщо відомо, що його побудовано в режимі True Color і воно займає весь екран у  $1024 \times 768$  пікселів.

Знайдемо обсяг зображення як добуток кількості пікселів, з яких воно складається, і кількості байтів у коді кольору:

$$1024 \times 768 \times 4 \text{ байта} = 768 \times 4 \text{ Кб} = 3072 \text{ Кб} = 3 \text{ Мб.}$$

Обсяг зображення перевищує ємність дискети, і тому безпосередньо його записати на дискету не можна.

## ВИСНОВКИ

Інформація подається, передається і зберігається у вигляді повідомлень. Застосовуючи різні мови, можна відтворювати одну й ту саму інформацію в різних повідомленнях. У комп'ютері повідомлення всіх видів — текстові, графічні, числові, звукові — за певними правилами перетворюються на послідовність двійкових кодів. Двійкові коди складаються тільки із двох знаків: 0 і 1, які називають бітами. Коди, що застосовуються в комп'ютері, кратні 8 бітам. Вісімку бітів називають байтом. Обсяг інформації в комп'ютері вимірюють сумарною кількістю байтів, якими вона подається. Більш ємними одиницями вимірювання обсягу інформації є кіло-, мега-, гіга- і терабайти. Кожна наступна одиниця більша за попередню в  $2^{10}$ , тобто 1024, разів.

---



---

### Контрольні питання та вправи

1. Біт — це:
  - а) двійкова цифра;
  - б) двійкове число;
  - в) двійковий знак;
  - г) двійковий код.
2. Збільшення довжини коду на один біт приводить до збільшення множини різних кодів, які можна утворити:
  - а) на один код;
  - б) на два коди;
  - в) у два рази;
  - г) у 10 разів.
3. Двійкове кодування тексту на комп'ютері означає подання його у вигляді:
  - а) сукупності кодів різної довжини;
  - б) послідовності кодів однакової довжини;
  - в) послідовності кодів довільної довжини.
4. Якість подання зображення на екрані комп'ютера визначається:
  - а) кількістю пікселів на екрані;
  - б) розміром екрана;
  - в) кількістю припустимих кольорів;
  - г) кількістю бітів у байті.
5. Якість подання звуку в комп'ютері тим краща, чим:
  - а) вища частота дискретизації;
  - б) нижча частота дискретизації;
  - в) вища частота звуку;
  - г) нижча частота звуку.

6. 4096 біт — це:

- а) 0,5 Кб;
- б) 4 Кб;
- в) 512 байт;
- г) 4,096 Кб.

7. Якщо на умовній шкалі відобразити 1 байт відрізком в 1 см, то як у такому масштабі буде виглядати: 1 Кб; 1 Мб; 1 Гб; 1 біт? Наведіть приклади таких відстаней із реального життя.

8. Обсяг текстового повідомлення в ASCII-кодуванні складає 640 біт. Скільки символів містить повідомлення?

9. Для кодування кольору пікселя виділено 2 байти. Скільки різних кольорів може мати зображення на екрані комп'ютера?

10. Кольорове зображення розміром  $64 \times 128$  пікселів, створене на саморобному комп'ютері, має обсяг 3 Кб. Яку кількість різних кольорів може мати зображення?

11. Алфавіт тарабарської мови складається з 28 літер. Якою мінімальною кількістю бітів можна кодувати тарабарські тексти, якщо в них, окрім літер, застосовуються арабські цифри і 7 розділових знаків?

12. Перехоплено таємне повідомлення:

▲♠▼☀♀♠◀ — ☀Ω▼♣Ω▶: ▶ ♂▶♠♣▼▶ 30 ♂♠†Ω△ 2010  
 △♣◀▶ ▲♠♥♠▶♥■▼•♂△◀△♣♂ ▶†Ω♠▲ 10 ◀♀☀♂♠▲ Ω☀  
 ◀▶♠♣◀ ◀♀▶♠▶ «♂▲♠▼♣†».

Чи можна його розшифрувати без ключа?

байт, біт, гігабайт, двійкове кодування, двійковий знак, дискретизація, кілобайт, кодування, код, коди ASCII, мегабайт, піксель, повідомлення, терабайт, частота дискретизації