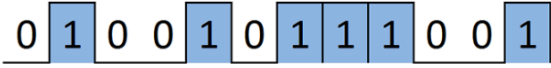


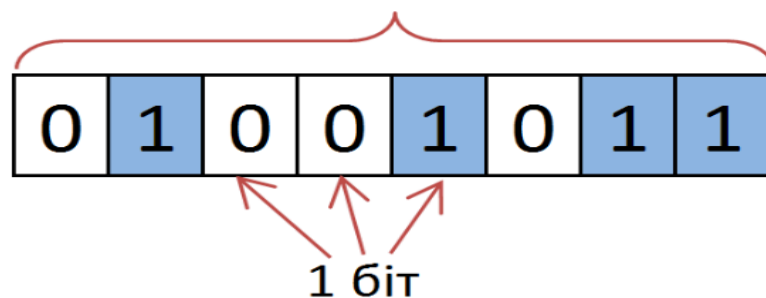
## Урок 2 Двійкове кодування

Подання чисел у двійковій СЧ називають **двійковим кодом**

|                                                                                   |                                                          |
|-----------------------------------------------------------------------------------|----------------------------------------------------------|
|  | 1- Наявність електричного сигналу<br><br>0 – відсутність |
| <b>Двійковий код</b>                                                              | <b>01001011001</b>                                       |

|                                                                                 |                                                                  |
|---------------------------------------------------------------------------------|------------------------------------------------------------------|
| Кожна цифра у двійковому коді – <b>біт</b><br><b>1 біт – це цифра (0 або 1)</b> | Послідовність із 8 бітів – <b>байт</b><br><b>1байт = 8 бітів</b> |
|---------------------------------------------------------------------------------|------------------------------------------------------------------|

1 байт=8 бітів

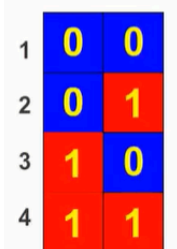


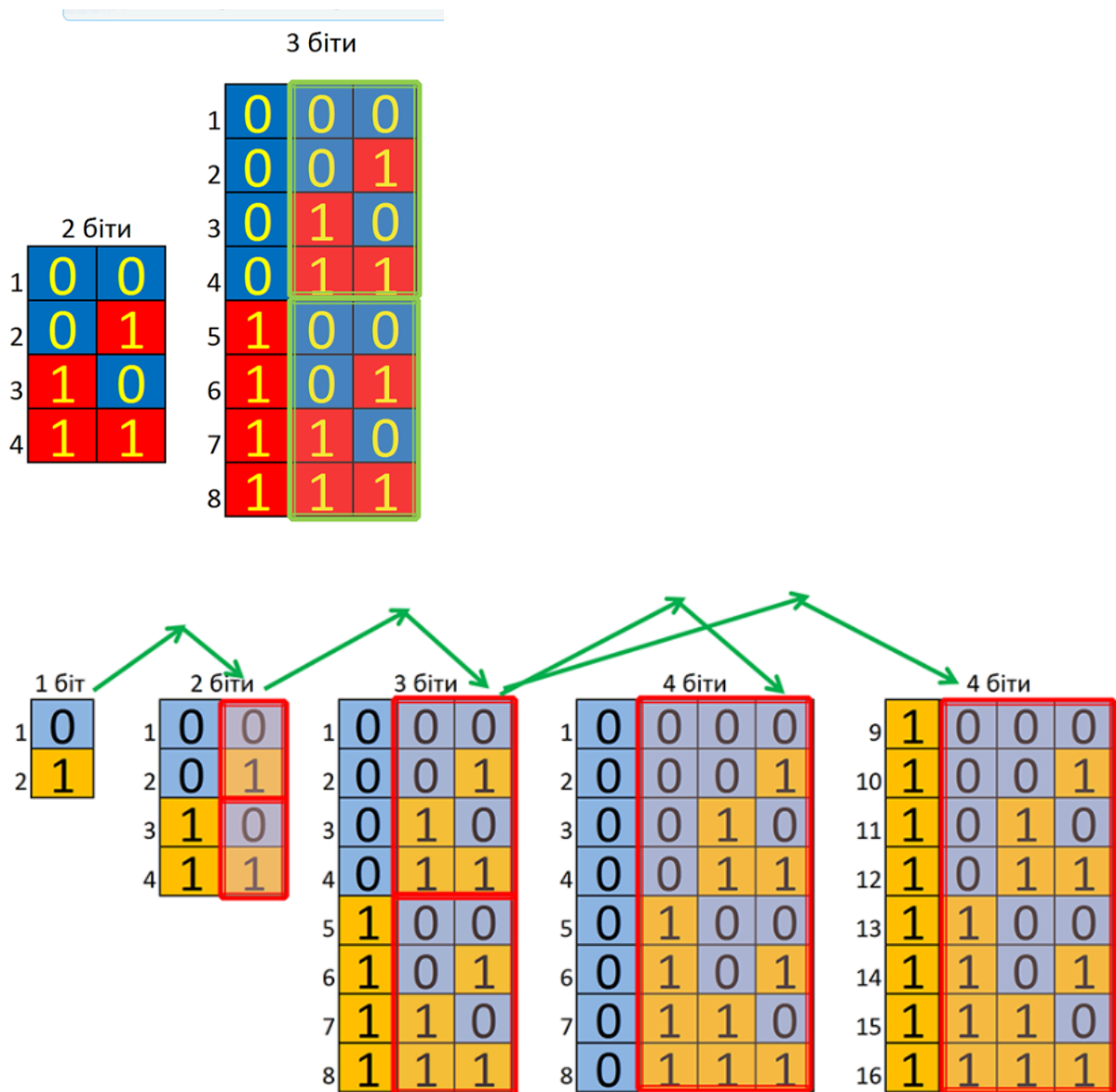
| Довжина в бітах | Довжина в байтах  |
|-----------------|-------------------|
| 24              | 3                 |
|                 | 128               |
|                 | 6 байтів і 4 біти |

За допомогою **1 біта** можна закодувати **2** повідомлення.

Присвоїмо різним повідомленням два коди 0 та 1, які можна отримати з 1 біта.

|                                                           |                                                     |                                                   |
|-----------------------------------------------------------|-----------------------------------------------------|---------------------------------------------------|
| біт – це цифра<br>двійкового коду<br>отримує код <b>0</b> | бібліотека<br>працює з 9-00<br>отримує код <b>1</b> | завтра урок<br>інформатики<br>коди<br>закінчилися |
|-----------------------------------------------------------|-----------------------------------------------------|---------------------------------------------------|

|                                                         |                                                                                      |
|---------------------------------------------------------|--------------------------------------------------------------------------------------|
| Які значення кодів можна утворити за допомогою 2 бітів? |                                                                                      |
| <b>00, 01, 10, 11</b>                                   |  |



Яку кількість повідомлень можна закодувати з допомогою 1 байта?



$$2*2*2*2*2*2*2*2=$$

## Закономірність

|   | 4 біти |   |   |   |    | 4 біти |   |   |   |
|---|--------|---|---|---|----|--------|---|---|---|
| 1 | 0      | 0 | 0 | 0 | 9  | 1      | 0 | 0 | 0 |
| 2 | 0      | 0 | 0 | 1 | 10 | 1      | 0 | 0 | 1 |
| 3 | 0      | 0 | 1 | 0 | 11 | 1      | 0 | 1 | 0 |
| 4 | 0      | 0 | 1 | 1 | 12 | 1      | 0 | 1 | 1 |
| 5 | 0      | 1 | 0 | 0 | 13 | 1      | 1 | 0 | 0 |
| 6 | 0      | 1 | 0 | 1 | 14 | 1      | 1 | 0 | 1 |
| 7 | 0      | 1 | 1 | 0 | 15 | 1      | 1 | 1 | 0 |
| 8 | 0      | 1 | 1 | 1 | 16 | 1      | 1 | 1 | 1 |

| кількість бітів | кількість повідомлень, що можна закодувати |
|-----------------|--------------------------------------------|
| 1               | 2                                          |
| 2               | $2*2=2^2=4$                                |
| 3               | $2*2*2=2^3=8$                              |
| 4               | $2*2*2*2=2^4=16$                           |
| 5               | $2*2*2*2*2=2^5=32$                         |
| 6               | $2*2*2*2*2*2=2^6=64$                       |
| 7               | $2*2*2*2*2*2*2=2^7=128$                    |
| 8               | $2*2*2*2*2*2*2*2=2^8=256$                  |

|                                                         |           |
|---------------------------------------------------------|-----------|
| За допомогою n бітів можна закодувати $2^n$ повідомлень |           |
| $2^7=128$                                               | $2^9=512$ |
| $2^{10}=1024$                                           |           |

1 кілобайт (1 Кбайт) = 1024 байти;

1 мегабайт (1 Мбайт) = 1024 Кбайт;

1 гігабайт (1 Гбайт) = 1024 Мбайт;

1 терабайт (1 Тбайт) = 1024 Гбайт.

|                     |                           |
|---------------------|---------------------------|
| 8байт=8*8=64бітів   | 2мб=2*1024*1024=2097152б  |
| 1кб=1026=4б         | 917504б=917504:1024=896кб |
| 2кб=2*1024байт      |                           |
| 16кб=16*1024=16384б |                           |

У двійковій СЧ «круглі величини» - степені двійки, а  $2^{10}=1024$

Те, що кожна наступна одиниця вимірювання інформації в 1024 рази більша від попередньої, пов'язано з тим, що звичною для нас є десяткова СЧ, де всі «круглі» величини є степенями 10,

**а у двійковому кодуванні застосовують двійкову СЧ, де «круглі» величини – степені двійки.**

Давайте підрахуємо, скільки повідомлень можна закодувати за допомогою 1Кб.

**1 Кб = 1024 байти = 1024\*8 бітів = 8192 біти.**

Отже, така кількість бітів дозволяє закодувати

$2^{8192}$  повідомлень

Це число надзвичайно велике.

Уявіть собі: двійку множимо саму на себе 8192 рази!

Насправді це незрівнянно більше, ніж кількість найелементарніших часток у Всесвіті.

## Азбука Морзе

Символи алфавіту кодувалися певним набором крапок та тире.

На початку 20 століття завдяки появі радіозв'язку популярною стала азбука, яка була запропонована ще в 1838 і названа на честь американського винахідника Самуеля **Морзе**.

Її можна вважати двійковим кодом, оскільки для передачі повідомлень використовувалися два види звукових сигналів: короткий (на папері його позначали крапкою) і довгий (тире).

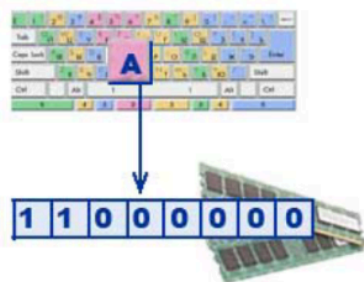


| буква | код     |
|-------|---------|
| А     | . -     |
| Б     | - ...   |
| В     | . - -   |
| Г     | ....    |
| Ґ     | - - .   |
| Д     | - ..    |
| Е     | .       |
| Є     | .. - .. |
| Ж     | ... -   |

|   |       |                                  |
|---|-------|----------------------------------|
| і | ..    | — — . — — — ... — . — — . — .    |
| м | — —   |                                  |
| н | — .   | — — — . — .. — . — — . — .       |
| о | . — — |                                  |
| р | . — . | — — . — — — . — .. — . — — . — . |
| т | —     |                                  |
| б | — ... | — — — — . — . .. — — — . — .     |

Монітор

Натискання клавіші на клавіатурі надсилає на *контролер* (спеціальну мікросхему) електричний сигнал, який перетворюється на двійковий код і зіставляється з певним символом **таблиці кодів**.



Ви вже знаєте, що за допомогою **1 байта** можна закодувати 256 повідомлень, тобто створити 256 двійкових кодів. Така кількість дає змогу закодувати латинські символи, символи кирилиці, цифри та спеціальні знаки.

## Система кодування ASCII

(American Standard Code for Information Interchange – стандартний код інформаційного обміну)

Створена інститутом стандартизації США.  
Для запису одного символу виділено 1 байт.  
Символам присвоєно **256** кодів від 00000000 до 11111111.

Таблиця кодів поділена на дві частини.

## Структура таблиці ASCII

|                                                                | номери  | коди                   | призначення                                                                                             |
|----------------------------------------------------------------|---------|------------------------|---------------------------------------------------------------------------------------------------------|
| Стандартна<br>(інтернаціональна)<br>частина.<br>(128 символів) | 0-31    | 00000000 -<br>00011111 | Керуючі символи.                                                                                        |
|                                                                | 32-127  | 00100000 -<br>01111111 | Цифри, букви латинського алфавіту<br>і специфічні символи, що<br>використовуються найчастіше.           |
| Альтернативна<br>частина.<br>(128 символів)                    | 128-255 | 10000000 -<br>11111111 | Національні алфавіти, написання<br>яких відмінно від латинського.<br><b>(Може мати різні варіанти).</b> |

Фрагмент таблиці ASCII-  
кодів для кодування цифр

| Код у десятковій системі | Код у двійковій системі | Символ |
|--------------------------|-------------------------|--------|
| 48                       | 110000                  | 0      |
| 49                       | 110001                  | 1      |
| 50                       | 110010                  | 2      |
| 51                       | 110011                  | 3      |
| 52                       | 110100                  | 4      |
| 53                       | 110101                  | 5      |
| 54                       | 110110                  | 6      |
| 55                       | 110111                  | 7      |
| 56                       | 111000                  | 8      |
| 57                       | 111001                  | 9      |

Фрагмент таблиці ASCII-  
кодів для кодування  
символів латиниці

| Код у десятковій системі | Код у двійковій системі | Символ |
|--------------------------|-------------------------|--------|
| 65                       | 1000001                 | A      |
| 66                       | 1000010                 | B      |
| 67                       | 1000011                 | C      |
| 68                       | 1000100                 | D      |
| 69                       | 1000101                 | E      |
| 70                       | 1000110                 | F      |
| 71                       | 1000111                 | G      |
| 72                       | 1001000                 | H      |
| 73                       | 1001001                 | I      |
| 74                       | 1001010                 | J      |
| 75                       | 1001011                 | K      |
| 76                       | 1001100                 | L      |
| 77                       | 1001101                 | M      |
| 78                       | 1001110                 | N      |
| 79                       | 1001111                 | O      |

Фрагмент таблиці ASCII-  
кодів для кодування  
символів кирилиці

| Код у десятковій системі | Код у двійковій системі | Символ |
|--------------------------|-------------------------|--------|
| 192                      | 11000000                | A      |
| 193                      | 11000001                | Б      |
| 194                      | 11000010                | В      |
| 195                      | 11000011                | Г      |
| 196                      | 11000100                | Д      |
| 197                      | 11000101                | Е      |
| 198                      | 11000110                | Ж      |
| 199                      | 11000111                | З      |
| 200                      | 11001000                | И      |
| 201                      | 11001001                | Й      |
| 202                      | 11001010                | К      |
| 203                      | 11001011                | Л      |
| 204                      | 11001100                | М      |
| 205                      | 11001101                | Н      |
| 206                      | 11001110                | О      |

| Код у десятковій системі | Код у двійковій системі | Символ |
|--------------------------|-------------------------|--------|
| 97                       | 1100001                 | a      |
| 98                       | 1100010                 | b      |
| 99                       | 1100011                 | c      |
| 100                      | 1100100                 | d      |
| 101                      | 1100101                 | e      |
| 102                      | 1100110                 | f      |
| 103                      | 1100111                 | g      |
| 104                      | 1101000                 | h      |
| 105                      | 1101001                 | i      |
| 106                      | 1101010                 | j      |
| 107                      | 1101011                 | k      |
| 108                      | 1101100                 | l      |
| 109                      | 1101101                 | m      |
| 110                      | 1101110                 | n      |
| 111                      | 1101111                 | o      |
| 112                      | 1110000                 | p      |
| 113                      | 1110001                 | q      |
| 114                      | 1110010                 | r      |
| 115                      | 1110011                 | s      |
| 116                      | 1110100                 | t      |
| 117                      | 1110101                 | u      |
| 118                      | 1110110                 | v      |
| 119                      | 1110111                 | w      |
| 120                      | 1111000                 | x      |

Виберіть фрагмент,  
на якому подано слово **byte** у  
двійковому форматі

|         |
|---------|
| 1110011 |
| 1111001 |
| 1110100 |
| 1100101 |

|         |
|---------|
| 1100010 |
| 1111001 |
| 1110100 |
| 1100101 |

|         |
|---------|
| 1100010 |
| 1111001 |
| 1100010 |
| 1100101 |

|         |
|---------|
| 1100010 |
| 1100001 |
| 1110100 |
| 1100101 |

## Система кодування Unicode (уніфіковане кодування)



- З поширенням Інтернету та розвитком інформаційних технологій виникла необхідність розробки **єдиної системи** кодування текстової інформації, яка б охоплювала всі алфавіти та галузі знань.
- На зміну системі ASCII, у якій виділено для коду коду 1 байт і можна кодувати лише 256 різних символів, в 1991 році було запропоновано систему багатобайтового кодування – **Unicode (Юнікод)**.

|   | A64 | A65 | A66 | A67 | A68 | A69 |
|---|-----|-----|-----|-----|-----|-----|
| 0 | Ї   | Ы   |     | ☼   | Д   | Т   |
| 1 | Ѹ   | ы   |     | ☼   | д   | т   |
| 2 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 3 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 4 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 5 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 6 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 7 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 8 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| 9 | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| A | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| B | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| C | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| D | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| E | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |
| F | Ѹ   | Ѹ   | Д   | ☼   | Ї   | Ї   |

### Що представлено в Unicode?

- Практично всі сучасні писемності.
- Багато історичних писемностей, таких як давньогрецька, єгипетські ієрогліфи, клинопис, писемність майя та інші.
- Широкий набір піктограм, математичних, технічних, комерційних і музичних символів.

В **Unicode (Юнікод)** представлені практично всі письмові мови світу — як сучасні, так і мертві, а також величезна кількість технічних, математичних та графічних символів.

Головна мета Unicode — надати кожному канонічному символу унікальний цифровий код (номер), який не залежить від комп'ютерної платформи, операційної системи чи мови програмування.

У сучасному стандарті Unicode представлені такі основні категорії:

#### 1. Письмо сучасних народів світу

- **Європейські алфавіти:** кирилиця (зокрема, українська абетка з літерами І, Є, І, Ї), латиниця, грецька, вірменська, грузинська мови тощо.
- **Азійські системи письма:** китайські ієрогліфи (ханьзи), японські (кана та кандзі), корейський хангиль, арабська в'язь, іврит, деванагарі (індійські мови) тощо.
- **Рідкісні письмена:** алфавіти корінних народів Африки, Америки та Австралії.

## 2. Історичні та мертві мови

- Давньоєгипетські ієрогліфи, шумерський та вавилонський клинопис.
- Скандинавські руни, давньогрецькі знаки, глаголиця та старослов'янська кирилиця.
- Письмо майя, ронго-ронго (острова Пасхи) та багато інших.

## 3. Спеціальні та технічні символи

- **Математичні та наукові знаки:** інтеграли, суми, квадратні корені, нескінченність ( $\int$ ,  $\sum$ ,  $\sqrt{\quad}$ ,  $\infty$ ), грецькі літери для формул ( $\pi$ ,  $\alpha$ ,  $\beta$ ).
- **Валюти:** символи гривні (₴), долара (\$), євро (€), єни (¥) тощо.
- **Технічні та типографські знаки**  
: стрілки, лапки, маркери списків, дробі, шахматні фігури, картярські масті, ноти.

## 4. Смайли та Емоджі (Emoji)

- Unicode стандартизує графічні емоджі — обличчя, жести, тварин, їжу, прапори країн, транспорт та побутові предмети. Завдяки цьому емоджі, надіслане з iPhone, коректно відображається на Android чи Windows.

### Як це працює технічно?

Кожному символу присвоюється код у шістнадцятковій системі числення, який записується як U+XXXX. Наприклад: 0

- U+0407 — українська велика літера І
- U+0041 — латинська велика літера A
- U+1F600 — смайлик 😊

Сам стандарт Unicode містить понад **150 000 символів**, і цей список постійно доповнюється щороку.

## Найпоширеніші формати перетворення Unicode

### UTF-8

- загальна кількість кодів **65 536**;
- сумісна з 8-бітовим (ASCII) кодуванням тексту;
- використовується в операційних системах та веб-просторі.

### UTF-16

- загальна кількість кодів **1 112 064**;
- сумісна з UTF-8;
- використовується в сучасних операційних системах, веб-просторі.

### UTF-32

- загальна кількість кодів **2 147 483 648**;
- сумісна з UTF-16;
- використовується для зберігання рядків в ОС UNIX.

<https://www.unicode.org/charts/>

| Символ | Код Юнікод (HEX) | Назва символу                          |
|--------|------------------|----------------------------------------|
| Ґ      | U+0490           | Кирилична велика літера «Ґ з підйомом» |
| ґ      | U+0491           | Кирилична мала літера «ґ з підйомом»   |
| Є      | U+0404           | Кирилична велика літера «Є»            |
| є      | U+0454           | Кирилична мала літера «є»              |
| І      | U+0406           | Кирилична велика літера «І»            |
| і      | U+0456           | Кирилична мала літера «і»              |
| Ї      | U+0407           | Кирилична велика літера «Ї»            |

Головна різниця між UTF-8 та UTF-16 полягає у довжині коду, який використовується для запису одного символу, та в ефективності розподілу пам'яті для різних мов. Обидва кодування є частиною стандарту Unicode та здатні закодувати будь-який його символ, але роблять це по-різному.

#### Порівняльна таблиця та ключові відмінності:

| Критерій                              | UTF-8                                                      | UTF-16                                          |
|---------------------------------------|------------------------------------------------------------|-------------------------------------------------|
| <b>Розмір символу</b>                 | Змінний: <b>від 1 до 4 байт</b>                            | Змінний: <b>2 або 4 байти</b>                   |
| <b>Мінімум для символу</b>            | <b>1 байт (8 біт)</b>                                      | <b>2 байти (16 біт)</b>                         |
| <b>Ефективність для латиниці</b>      | <b>Дуже висока</b> (1 байт на символ)                      | <b>Низька</b> (файли збільшуються вдвічі)       |
| <b>Ефективність для кирилиці</b>      | <b>Середня</b> (2 байти на символ)                         | <b>Середня</b> (2 байти на символ)              |
| <b>Ефективність для азійських мов</b> | <b>Низька</b> (часто 3 байти на символ)                    | <b>Висока</b> (переважно 2 байти)               |
| <b>Сумісність з ASCII</b>             | <b>Повна</b> (коди 0-127 ідентичні)                        | <b>Несумісна</b>                                |
| <b>Де використовується</b>            | <b>Веб (HTML, XML), ОС Linux/macOS, файли конфігурацій</b> | Внутрішня пам'ять <b>Windows, Java, .NET/C#</b> |