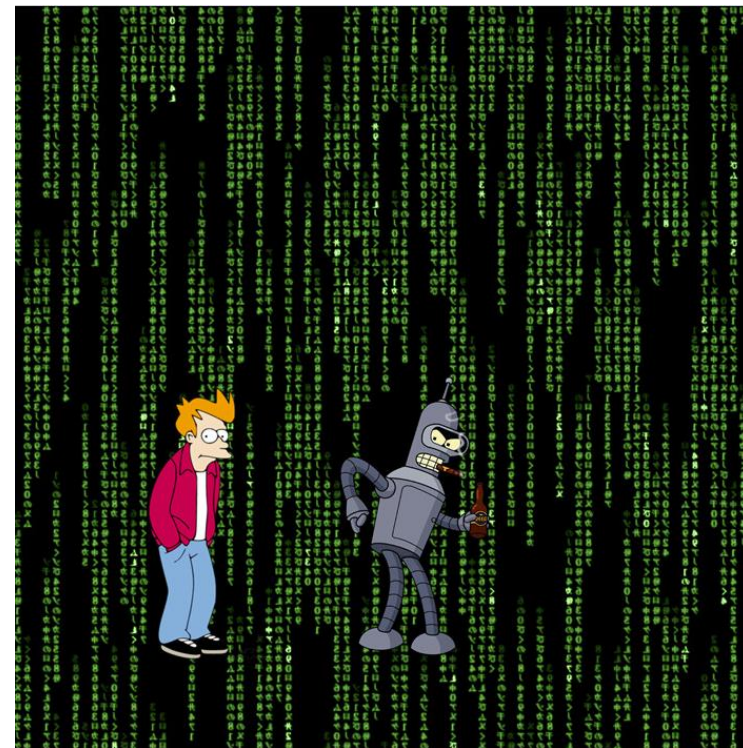


## Лекція 1

# Інформація, сигнали, дані



Термін «інформація» походить від лат. «*informatio*» - пояснення, повідомлення, інформування. Воно використовується в майже в усіх сферах людської діяльності, але строгого, чіткого та загальновизнаного його визначення не існує.

Під поняттям «інформація» часто розуміють структуровану сукупність відомостей про навколишній світ або стан якоїсь матеріальної системи (об'єкта, явища, події, процесу тощо), які призначені для передавання, розподілу, перетворення, зберігання або безпосереднього використання.

В основі інформації лежить взаємозалежність пари об'єктів - джерела і споживача інформації. Джерелами інформації є насамперед природні об'єкти: люди, тварини, рослини, планети тощо. Але у міру розвитку науки і техніки джерелами інформації все частіше стають наукові експерименти, машини, апарати, технологічні процеси. Споживачами інформації є люди, тварини, рідоманітні технологічні пристрої тощо.

# Інформація

```
graph TD; A[Інформація] --> B[Первинна]; A --> C[Вторинна];
```

## Первинна

Дані, що зібрані вперше (наприклад, під час експерименту або спостереження)

## Вторинна

Дані, отримані з існуючих джерел (наприклад, статистичні звіти або інформація з бази даних)

## Інформація

```
graph TD; A[Інформація] --> B[Структурована]; A --> C[Неструктурована];
```

### Структурована

Організована в табличній або ієрархічній формі (наприклад, у файлах певних форматів або базах даних)

### Неструктурована

Дані, які не мають чіткої структури («сирий» сигнал, сукупність байтів)

# Інформація

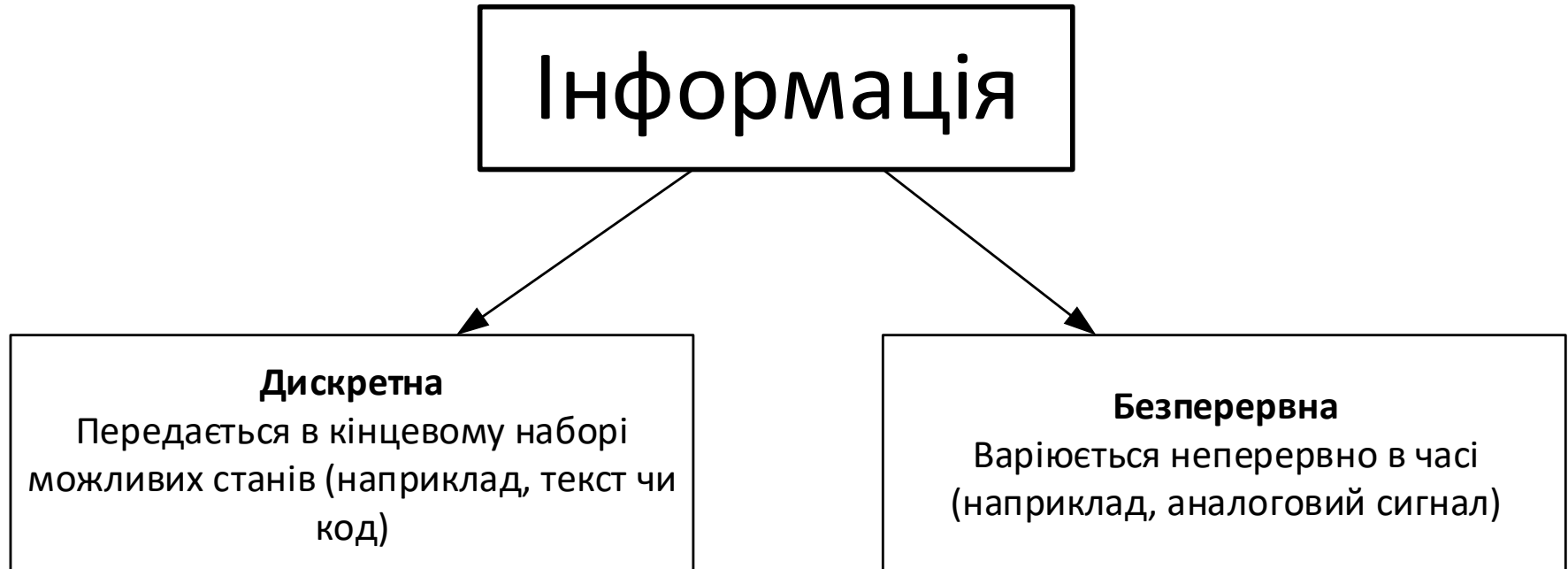
```
graph TD; A[Інформація] --> B[Цифрова]; A --> C[Аналогова];
```

## Цифрова

Дані, представлені в цифровому форматі, придатні для обробки комп'ютером

## Аналогова

Дані, що існують у фізичному світі, наприклад, звукові хвилі або світло



Теорія інформації – це галузь науки, яка займається кількісним вивченням інформації, а також методами її передавання, обробки і зберігання. Вона була заснована в 1948 році Клодом Шенноном, і його робота стала основою для подальших досліджень у телекомунікаціях, криптографії, машинному навчанні та багатьох інших сферах.

Названі вище класифікації інформації застосовується для:

- класифікації текстів у пошукових системах (Google, Yahoo);
- класифікації зображень у системах розпізнавання обличчя;
- організації інформації в базах даних та інформаційних системах.

В теорії інформації інформація розглядається як міра невизначеності системи – під інформацією розуміються всі ті відомості, які зменшують ступінь невизначеності наших знань про конкретний об'єкт або систему (концепція К. Шеннона - чим імовірнішим є повідомлення, тим менше інформації міститься в ньому).

Цей підхід, незважаючи на те, що він не враховує змістового аспекту інформації, став основою для її кількісної оцінки та оптимального кодування повідомлень.

Інформаційна ентропія ( $H$ ) – міра невизначеності або непередбачуваності інформаційного джерела. Вимірюється в бітах або більш загально в натах (натуральний логарифм).

Формула Шеннона для ентропії:

$$H(X) = - \sum_{i=1}^M p_i(x) \log_2(p_i(x))$$

де  $p_i(x)$  - ймовірність того, що подія  $x$  (отримання символу  $x$ ) відбудеться;  
 $M$  - можлива кількість станів системи  $X$ .

Ентропія показує, яка мінімальна кількість біт на символ (в середньому) потрібна для кодування даного повідомлення.

Хоча, як правило, розрахована ентропія буде дробовим числом, його округлюють до найближчого більшого цілого (тобто якщо розраховали  $H(X) = 7,3$  - то реально це буде в 8 біт на символ (при використанні рівномірного коду)).

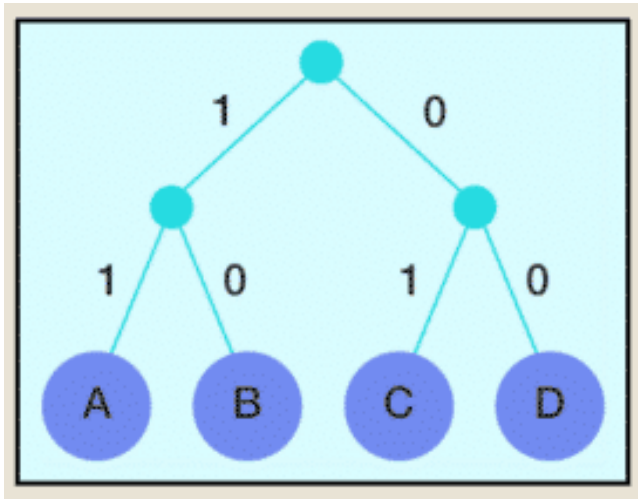


# Основні поняття теорії інформації

## Взаємна інформація, кодування та декодування

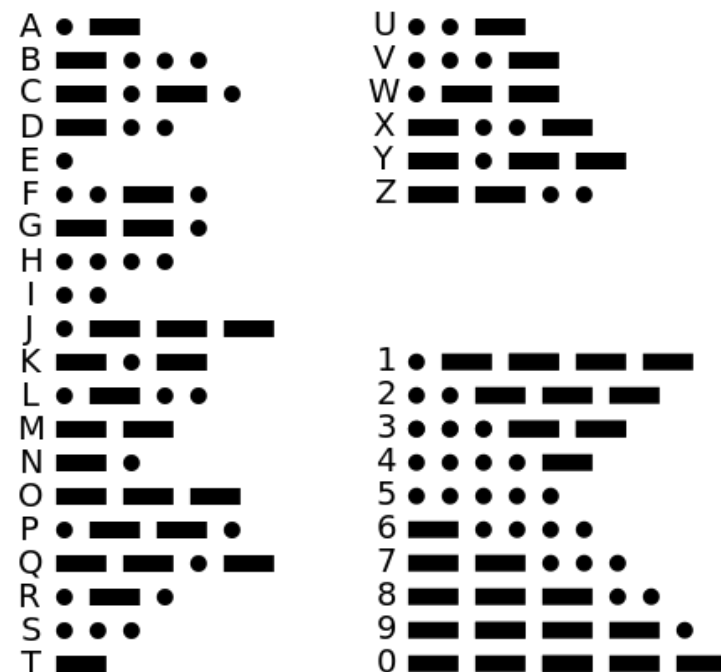
**Взаємна інформація (I)** – міра кількості інформації, яку одне джерело містить про інше. Це показує, наскільки знання про одну змінну зменшує невизначеність щодо іншої.

**Кодування та декодування** – процеси представлення інформації в іншому форматі для більш ефективної передачі (кодування) та відновлення інформації (декодування). Приклад: Хафманівське кодування або кодування Шеннона-Фано.



A = 11  
B = 01  
C = 10  
D = 00

2 біт/символ



# Основні поняття теорії інформації

## Пропускна здатність каналу

Пропускна здатність каналу ( $C$ ) – максимальна кількість інформації, яка може бути передана через комунікаційний канал без помилок. Пропускна здатність каналу визначається за допомогою формули Шеннона:

$$C = B \log_2 \left( 1 + \frac{S}{N} \right)$$

де:

$C$  - це пропускна здатність каналу, тобто максимальна кількість інформації, яку можна передати через канал без помилок (вимірюється в бітах на секунду, біт/с, бод);

$B$  - ширина смуги частот каналу (в Гц);

$S$  - потужність сигналу (середня потужність сигналу на виході передавача);

$N$  - потужність шуму (в середньому за той самий період).

Для «ідеального» каналу без шуму

$$C = B \log_2 S$$

**Інформація** - це властивість (атрибут) матерії, інформацію містять будь-які повідомлення, що сприймають людина, тварина або прилад.

Поява цієї концепції пов'язана з розвитком кібернетики і найяскравіше відображена Віктором Глушковий, який писав, що «інформацію несуть не лише поцятковані буквами листи книги або людська мова, але і сонячне проміння, зморшки гірського хребта, шум водоспаду, шелест трави».

**Інформація** - це діюча, корисна, «працююча» частина знань, яку використовують для орієнтації та спрямування активних дій, управління та самоуправління.

Такі три підходи певною мірою доповнюють один одного, висвітлюють різні сторони сутності поняття інформації.

**Інформація** - це властивість (атрибут) матерії, інформацію містять будь-які повідомлення, що сприймають людина, тварина або прилад.

Поява цієї концепції пов'язана з розвитком кібернетики і найяскравіше відображена Віктором Глушковий, який писав, що «інформацію несуть не лише поцятковані буквами листи книги або людська мова, але і сонячне проміння, зморшки гірського хребта, шум водоспаду, шелест трави».

**Інформація** - це діюча, корисна, «працююча» частина знань, яку використовують для орієнтації та спрямування активних дій, управління та самоуправління.

Такі три підходи певною мірою доповнюють один одного, висвітлюють різні сторони сутності поняття інформації.

**Інформація** - це вибір одного або декількох варіантів з багатьох можливих і рівноправних ( тобто всі варіанти, з яких відбувається вибір - належать одній множині). Три основні варіанти вказаного вибору: якщо вибір підказаний (або вказаний) - то йдеться про *одержання (рецепцію)* інформації; якщо вибір здійснений самостійно та/або випадково - то йдеться про *виникнення* інформації; якщо вибір обумовлений поточною ситуацією або попередніми подіями (тобто вибору немає) - то про інформацію взагалі вести мову не доводиться.



**Адекватність** інформації задає певний рівень відповідності змісту одержаної інформації образу реального об'єкта, який відображається.

З погляду семіотики (науки про знакові системи) адекватність інформації може виражатися у трьох формах: *синтаксичній*, *семантичній* та *прагматичній*.

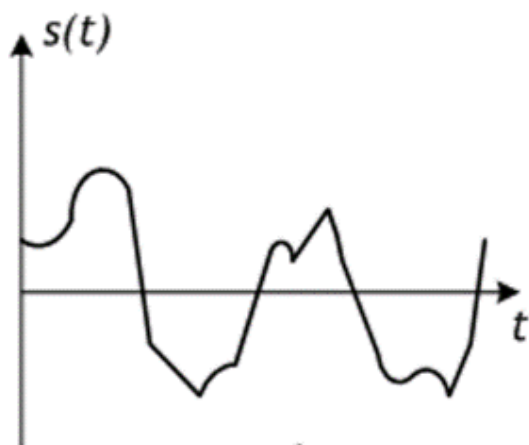
**Точність** визначає ступінь близькості одержаної інформації до реального стану об'єкта, який відображається.

**Достовірність** інформації - це її властивість відображати реальні об'єкти з необхідною точністю. Це ймовірність відсутності помилок.

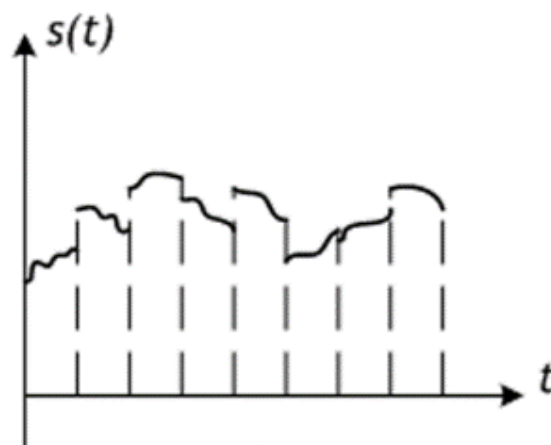
**Достатність** (або повнота) інформації - це наявність мінімальної, але достатньої для прийняття рішення інформації.

**Актуальність** інформації - це її властивість зберігати свою цінність в часі.

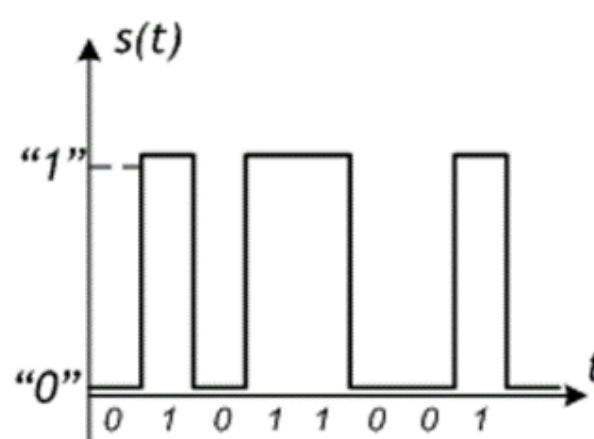
Сигнал - це зміна в часі якоїсь фізичної величини (найчастіше під сигналом розуміється зміна електричної напруги в часі).



Аналоговий

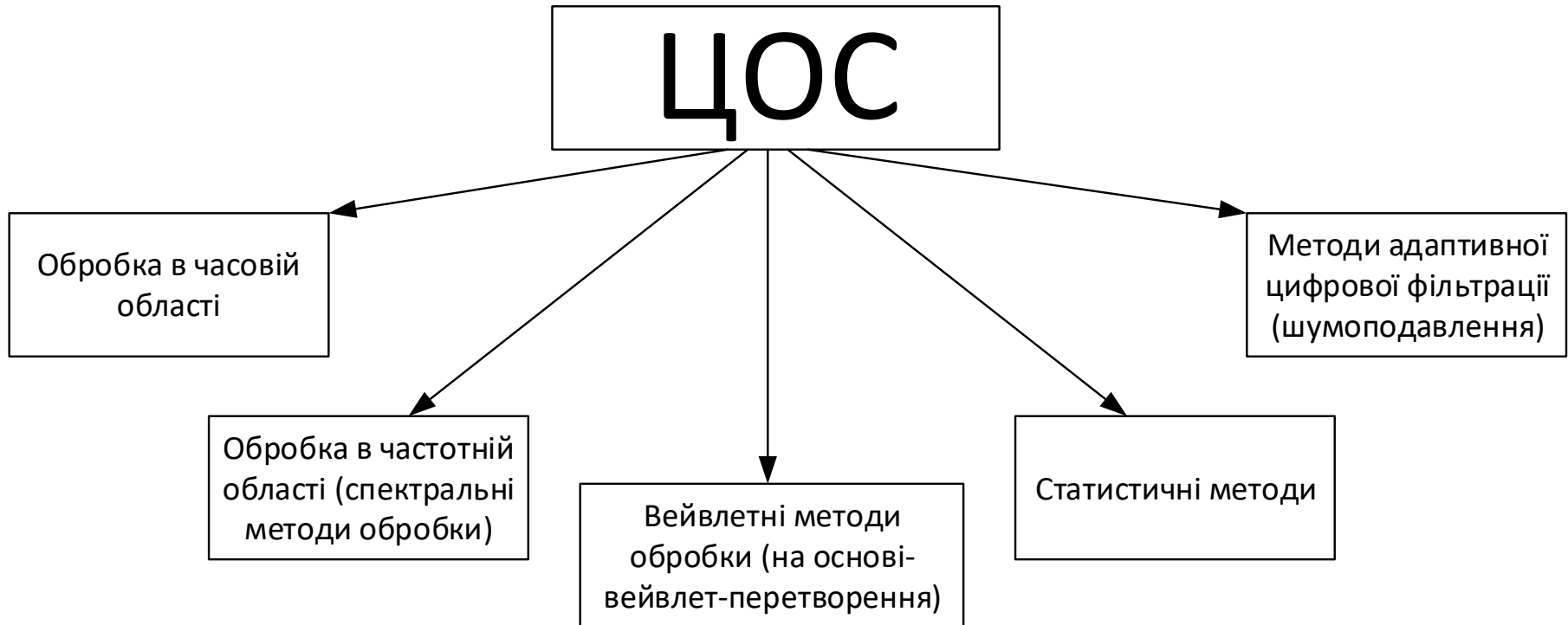


Дискретизований  
по часу



Цифровий

Параметри сигналів: діапазон напруг, смуга частот, скважність (для імпульсних сигналів), спектр, тривалість.





Далі буде...

**...Перетворення аналогових сигналів на цифрові**