



ЖИТОМИРСЬКА  
ПОЛІТЕХНІКА

ДЕРЖАВНИЙ УНІВЕРСИТЕТ

100  
РОКІВ

## Лекція 2

# КІЛЬКІСНІ ОЦІНКИ ВИМІРЮВАЛЬНОЇ ІНФОРМАЦІЇ



ЖИТОМИРСЬКА  
ПОЛІТЕХНІКА

ДЕРЖАВНИЙ УНІВЕРСИТЕТ

100  
РОКІВ

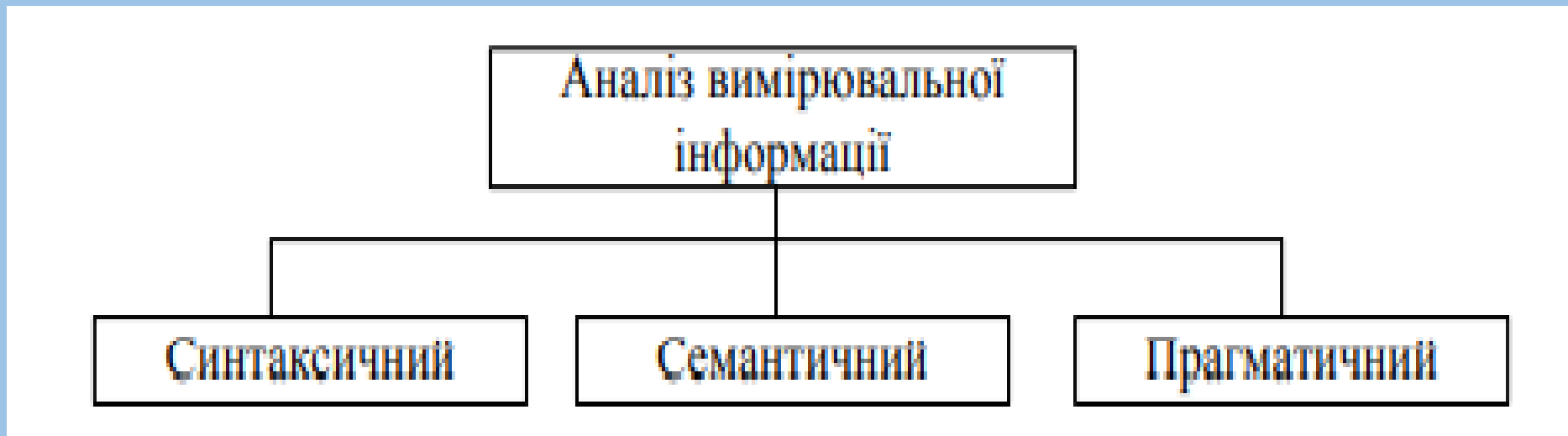
- 2.1. Вимірювальна інформація, її функції та властивості.
- 2.2. Кількісні характеристики інформації.
- 2.3. Міри інформації.
- 2.4. Ентропія та її властивості.
- 2.5. Продуктивність і надмірність джерела інформації.

## 2.1. Вимірювальна інформація, її функції та властивості

- Термін інформація походить від латинського *informatio*, що означає роз'яснення, інформування, викладення.
- **Інформація** - це загальнонаукове поняття, яке охоплює обмін інформацією між людьми, обмін сигналами між живою і неживою природою, людьми і пристроями.
- У цьому сенсі інформація є відображенням реального світу за допомогою відомостей (даних, повідомлень)
- **Вимірювальна інформація** – це сукупність відомостей про навколишній світ, об'єкти і явища навколишнього середовища, а також процеси, які відбуваються в них, відомості про їх параметри, властивості і стан, які зменшують ступінь невизначеності, неповноти знань про них.

- Саме повідомлення – це значення або зміна деякої фізичної величини, яка відображає стан об'єкта (системи або явища).
- *Інформаційно-комунікаційні системами (ІКС)* - шляхи і процеси, які забезпечують передачу повідомлень від джерела інформації до її споживача.
- *Інформаційно-комунікаційні системами (ІКС)* - комплекс організаційно-технічних заходів, інформаційних технологій та інформаційних ресурсів, призначених для забезпечення інформаційних процесів, зокрема створення, поширення, використання, зберігання та знищення інформації.

# Методи аналізу вимірювальної інформації

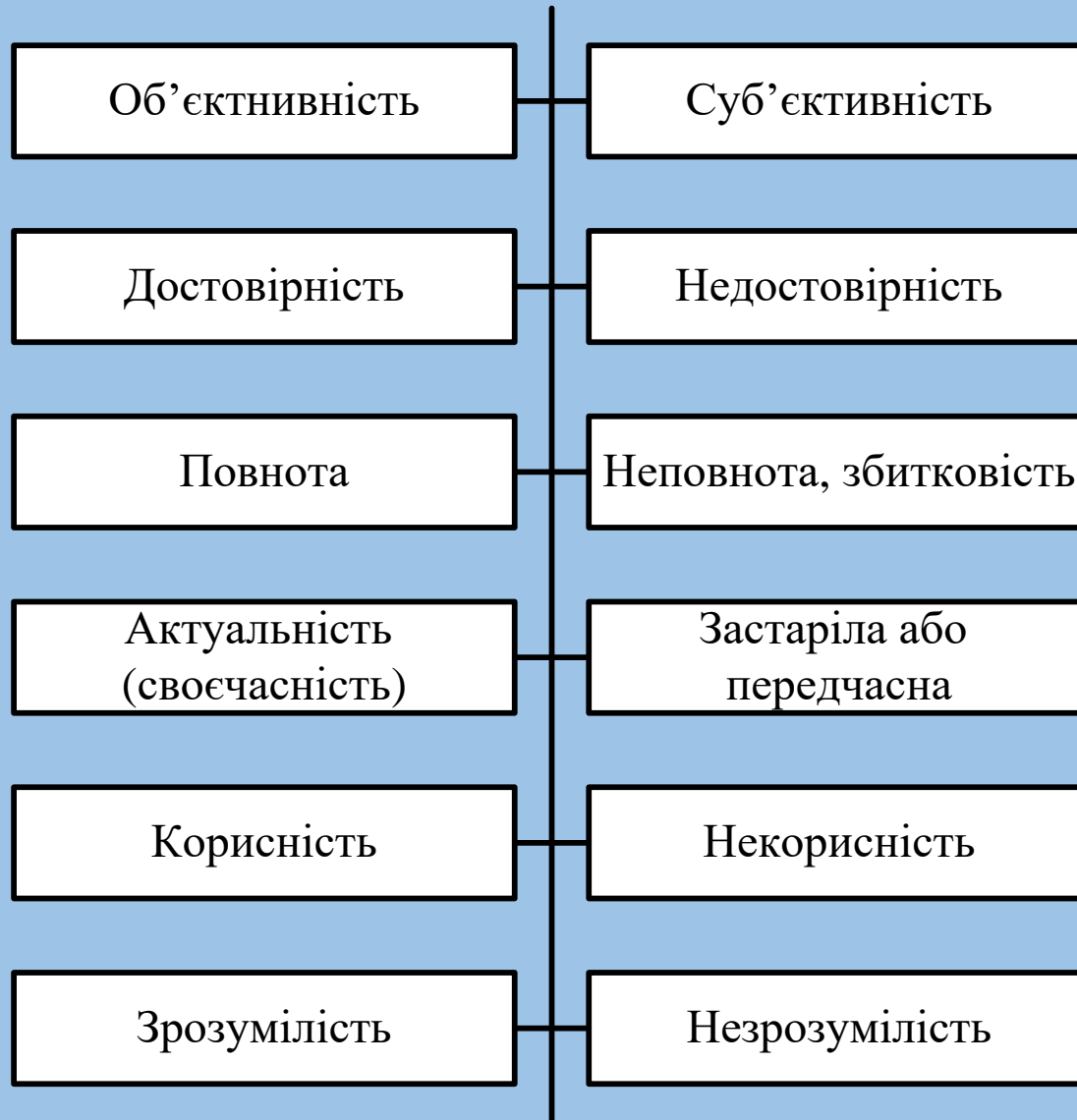


- ***Синтаксичний підхід*** здійснюється структурний аналіз вимірювальної інформації (кількість повідомлень, форма їх подання, статистичні характеристики появи тощо).
- ***Семантичний підхід*** полягає в аналізі змістовності вимірювальної інформації і впливу її на людину.
- ***Прагматичний підхід*** пов'язаний з оцінюванням корисності, цінності отриманої вимірювальної інформації.

# Функцій інформації:

- *Управлінська* у всіх сферах повсякденного життя і покликаної допомогти людині у виборі варіанту своєї поведінки або цілеспрямованого впливу на об'єкти і процеси реального світу;
- *Комунікаційна*, під час обміну інформацією між людьми і спрямованої на організацію взаємодії між ними;
- *Пізнавальна*, зумовлюється потребою в інформації заради загального розвитку, отримання спеціальності і, в загальному, для задоволення потягу до нового безвідносно до його прагматичного значення;
- *Психологічна*, позначається при формуванні певного емоційного настрою за допомогою деяких видів інформації, способів її подання людині.

# Властивості інформації:





## 2.2. Кількісні характеристики інформації

- У теорії інформації під *кількістю інформації* розуміють міру зменшення невизначеності знання.
- *Кількість інформації* - числова характеристика сигналу, яка не залежить від його форми і змісту і характеризує невизначеність, яка зникає після отримання повідомлення у вигляді даного сигналу.
- В цьому випадку кількість інформації залежить від вірогідності отримання повідомлення про ту або іншу подію.
- *Інформаційний об'єм повідомлення* (інформаційна місткість повідомлення) – кількість інформації в повідомленні, виміряна в бітах, байтах або похідних величинах.
- Для абсолютно достовірної події (подія обов'язково відбудеться, тому її імовірність дорівнює 1, кількість інформації в повідомленні про неї дорівнює 0). Чим неймовірніша подія, тим більшу інформацію про нього несе повідомлення. Лише при рівноімовірних відповідях, відповідь “так” чи “ні” несе один біт інформації.



- Г. Хартлі запропонував формулу для обчислення кількості інформації про об'єкт, який може знаходитися в одному з рівноймовірнісних  $N$  станів:

$$I = \log_2 N \quad (2.1)$$

Формулу для обчислення кількості інформації для подій з різною ймовірністю запропонував К. Шеннон в 1948 р.:

$$I = -\sum_{i=1}^N p_i \log_2 p_i \quad (2.2)$$

де  $N$  - кількість можливих подій;  $p_i$  - ймовірність окремих подій.

- **Кількісна міра інформації** (Г. Хартлі і К. Шеннон) враховує тільки кількість  $N$  повідомлень, та ймовірності надходження подій  $\{x_i\}$  на вхід каналу передачі. Сукупність повідомлень подають у наступному вигляді:

$$X = \{(x_1, p_1), \dots, (x_n, p_n)\}, \sum_{i=1}^N p_i(x_i) = 1 \quad (2.3)$$

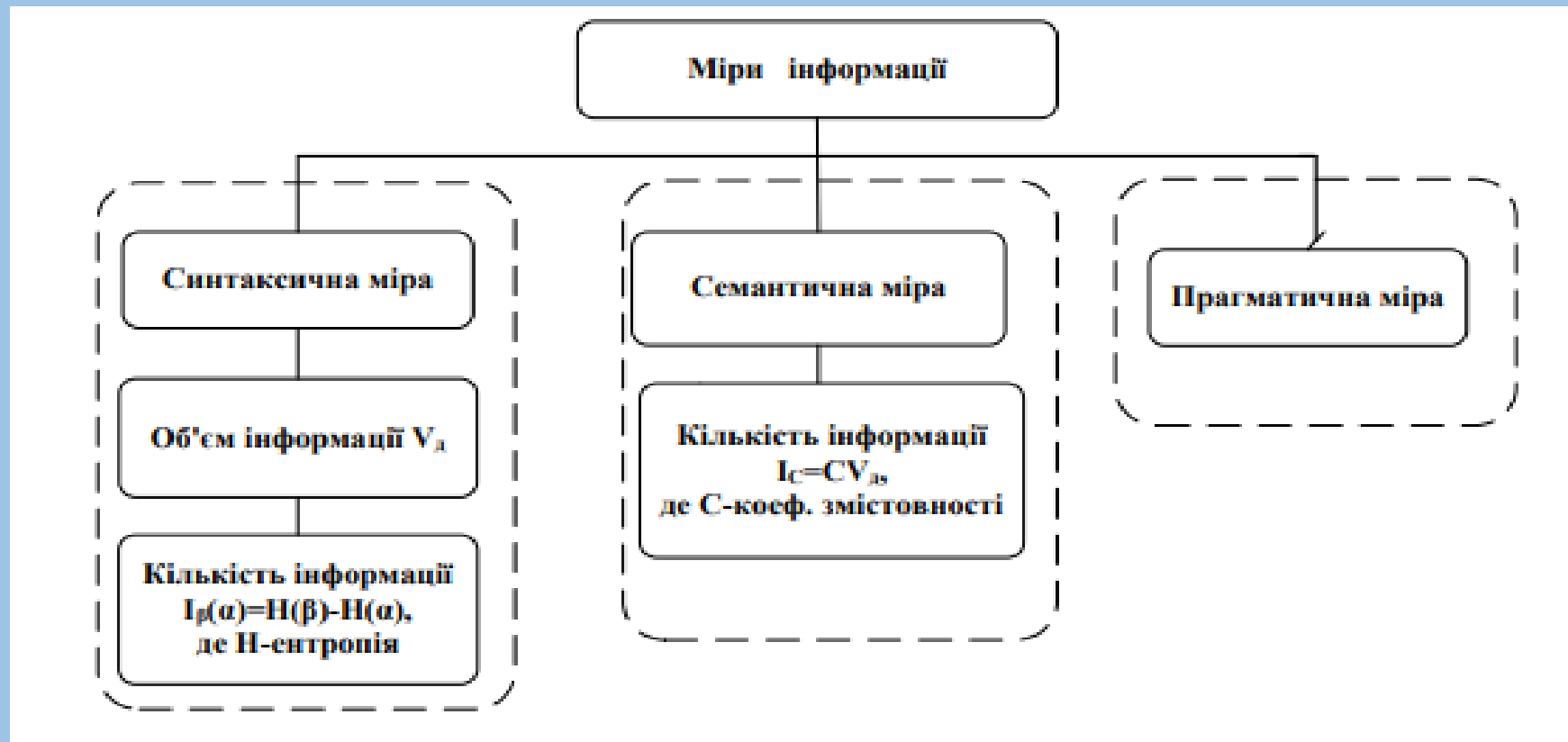
де:  $x_i$  – окремі повідомлення (або їх типи, класи).

В результаті передавання повідомлення  $x_i$  буде отримане повідомлення  $y_j$ . Апостеріорна (післядослідна) ймовірність присутності  $x_i$  в  $y_j$  дорівнює  $p(x_i / y_j)$ .

Кількість інформації, що міститься в події  $y_j$  щодо події  $x_i$  визначається за формулою:

$$I(x_i; y_j) = \log \left[ p(x_i / y_j) / p(x_i) \right] \quad (2.4)$$

## 2.3. Міри інформації



- **Синтаксична міра інформації** - оперує із знеособленою інформацією, що не передає смислового відношення до об'єкту. На синтаксичному рівні враховуються тип носія і спосіб представлення інформації, швидкість передавання та опрацювання, розміри кодів представлення інформації.
- Для визначення такої кількісної міри інформації вводять два параметри: об'єм даних  $V_d$  і кількість інформації  $I$ .
- Об'єм даних ( $V_d$ ) - інформаційний об'єм повідомлення або об'єм пам'яті, необхідний для зберігання повідомлення без яких-небудь змін.
- Об'єм даних  $V_d$  у повідомленні вимірюється кількістю символів (розрядів). Одиниця вимірювання залежить від системи числення.
- **Семантична міра інформації** - слугує для вимірювання змісту інформації. Для цього використовується тезаурусна міра, яка враховує здатність одержувача повідомлення його сприйняти.
- Тезаурусом називають сукупність відомостей, які має в своєму розпорядженні користувач або система.
- **Прагматична міра інформації** – визначає корисність інформації для досягнення користувачем поставлених цілей.



ЖИТОМИРСЬКА  
ПОЛІТЕХНІКА

ДЕРЖАВНИЙ УНІВЕРСИТЕТ

100  
РОКІВ

- Кількість інформації ( $I$ ) в повідомленні про об'єкт, який може знаходитися в одному з рівноімовірних  $N$  станів, визначають, згідно вище зазначеної формули Р. Хартлі (2.1).
- Одиниці вимірювання інформації залежать від вживаної системи числення: у двійковій системі одиниця вимірювання – біт (bit – binary digit – двійковий розряд), у десятковій – діт.
- Біт – кількість інформації, яка потрібна для розрізнення двох рівноімовірних повідомлень.
- Біт – дуже мала одиниця вимірювання. На практиці частіше застосовується одиниця – байт, що дорівнює восьми бітам. Восьми біт достатньо для кодування будь-якого з 256 символів алфавіту клавіатури комп'ютера ( $256 = 2^8$ ).

Широко використовуються наступні похідні одиниці вимірювання інформації:

- 1 Кілобайт (Кбайт) =  $1024$  байт =  $2^{10}$  байт,
- 1 Мегабайт (Мбайт) =  $1024$  Кбайт =  $2^{20}$  байт,
- 1 Гігабайт (Гбайт) =  $1024$  Мбайт =  $2^{30}$  байт.
- 1 Терабайт (Тбайт) =  $1024$  Гбайт =  $2^{40}$  байт,
- 1 Петабайт (Пбайт) =  $1024$  Тбайт =  $2^{50}$  байт.

## 2.4. Ентропія та її властивості

- Кількість інформації ( $I$ ) на синтаксичному рівні визначається через поняття ентропії системи.
- Ентропія визначає міру невизначеності всієї безлічі повідомлень на вході системи і обчислюється як середня кількість власної інформації у всіх повідомленнях:

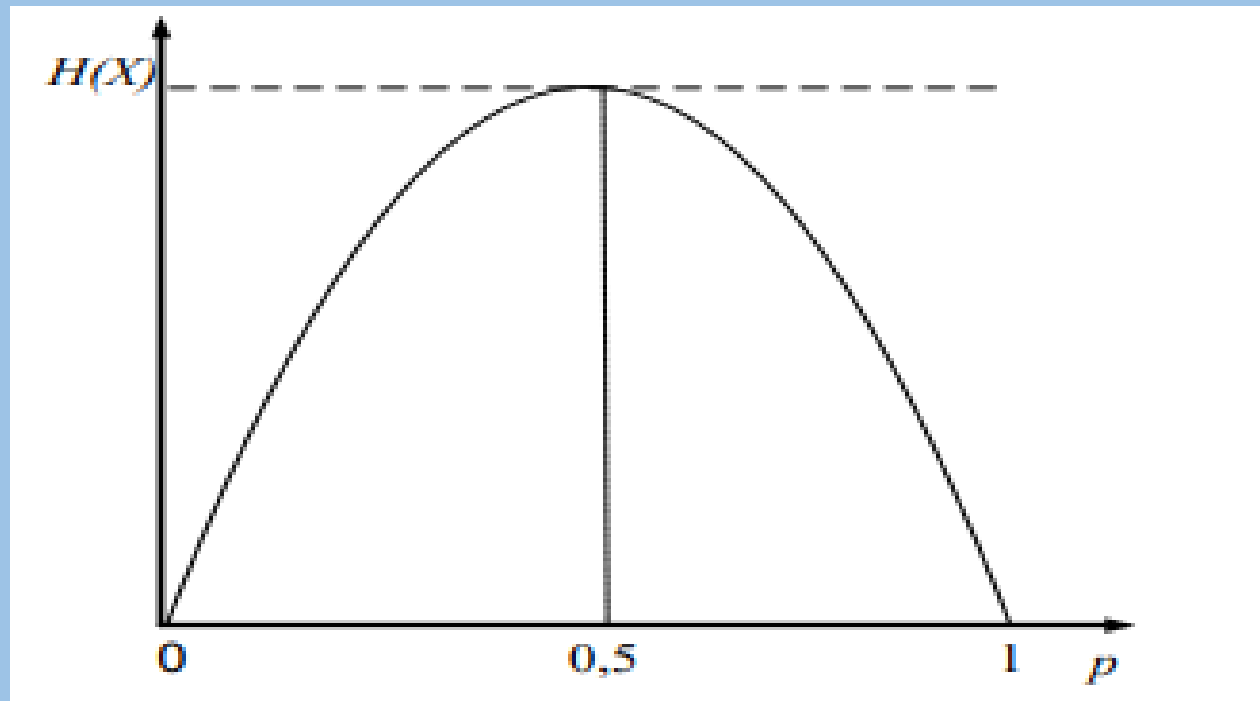
$$I(X) = -\sum p(x_i) \log p(x_i) = H(X) \quad (2.5)$$



# Властивості ентропії:

1. Ентропія  $H(X)$  додатня:  $H(X) > 0$ , де  $X$  – позначення повідомлення джерела інформації.
2. Ентропія  $H(X) < \log N$ .
3. Величина  $\log N = D$  називається інформаційною пропускною спроможністю алфавіту (інформаційною місткістю алфавіту).
4. Якщо  $N = 2$ , то  $p(x_1) = p$ ,  $p(x_2) = 1-p$  і  $H(X) = -p \log_2 p - (1-p) \log_2 (1-p)$ .
5. Максимум  $H(X) = -\log_2 0,5 = \log_2 2 = 1$  – місткість двійкового алфавіту дорівнює 1 біт.

# Залежність $H(X)$ від величини $p$



## 2.5. Продуктивність і надмірність джерела інформації

- Швидкість передавання повідомлення визначається ресурсами її джерела на певному інтервалі часу  $T$ .
- *Продуктивністю джерела*  $H'(X)$  називається сумарна ентропія повідомлень переданих за одиницю часу:

$$H'(X) = H(X)/T \quad (2.6)$$

Продуктивність вимірюється в бітах на секунду.

- Завдання підвищення ефективності передачі зводиться до зменшення надмірності повідомлення й сигналу, точніше до раціонального її використання.
- Надмірність повідомлення обумовлена тим, що елементи повідомлення не є рівноймовірними й між ними є статистичний зв'язок. У випадку кодування можна перерозподілити ймовірності вихідного повідомлення таким чином, щоби розподіл ймовірностей символів коду наближався до оптимального (до рівномірного – у дискретному випадку або до нормального – за умови передавання безперервних повідомлень). Такий перерозподіл дозволяє усунути надмірність, що залежить від розподілу ймовірностей елементів повідомлення.