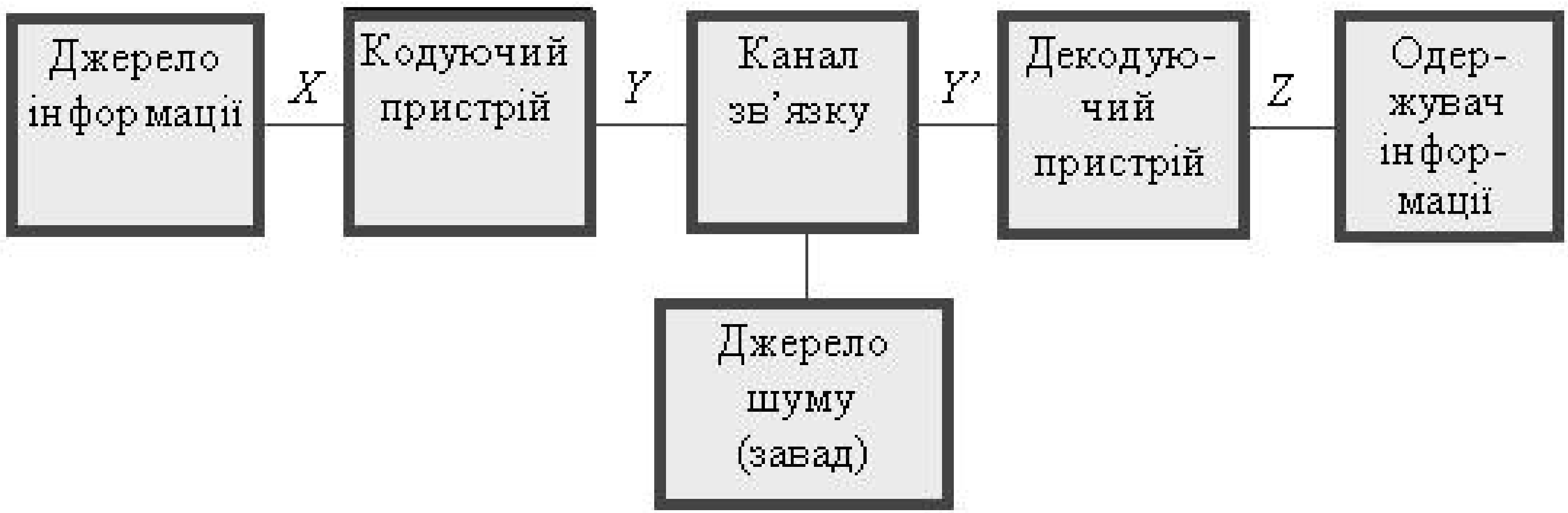
Кількість інформації

Теоретичною базою для оцінки кількості інформації служить теорія інформації Шеннона.



Інформація передається від джерела до одержувача за допомогою повідомлень. Повідомлення зменшує невизначеність знань про джерело інформації.

Повідомлення бувають дискретними і неперервними. **Дискретне повідомлення** складається з *n* окремих елементів (букв, цифр, символів). Елементи повідомлень кодуються.

Використовуються *рівномірні коди,* в яких букви алфавіту представляються комбінаціями з однакового числа елементів, і *нерівномірні коди,* складені з комбінацій різної довжини.

Прикладом рівномірного коду може служити телетайпний код Бодо, а нерівномірного – азбука Морзе.

Буква	**А**	**Б**	**В**	**Г**	**Д**	**Е**
Код Бодо	**10000**	**00110**	**01101**	**01010**	**11110**	**01000**
Азбука Морзе	· —	— · · ·	— —	— — ·	— · ·	·

Неперервне повідомлення*,* на відміну від дискретного, представляється неперервною функцією часу X(t). В КІСУ як неперервні повідомлення використовуються струмові сигнали 4-20 мА.

Згідно Шеннону, кількість інформації визначається як математична міра зменшення невизначеності знань про джерело інформації.

Якщо H_1 - початкова (*апріорна*) невизначеність знання до одержання повідомлення, а H_2 - залишкова (*апостеріорна*) невизначеність, що характеризує стан знання після отримання повідомлення, то кількість інформації, що міститься в цьому повідомленні визначається різницею:

$$I = H_1 - H_2.$$

Якщо дискретне повідомлення X може мати n варіантів x_1, x_2, …, x_n з імовірностями появи цих варіантів $P(x_1)$, $P(x_2)$, …, $P(x_n)$, мірою несподіваності варіанта x_i служить функція

$$\log_2 (1 / P(x_i)) = — \log_2 P(x_i),$$

яку можна розглядати як міру невизначеності $H(x_i)$ стосовно отримання одержувачем варіанта повідомлення x_i.

Для всієї сукупності варіантів повідомлення X Клод Шеннон увів усереднену за ймовірностями $P(x_i)$ міру $H(x_i)$:

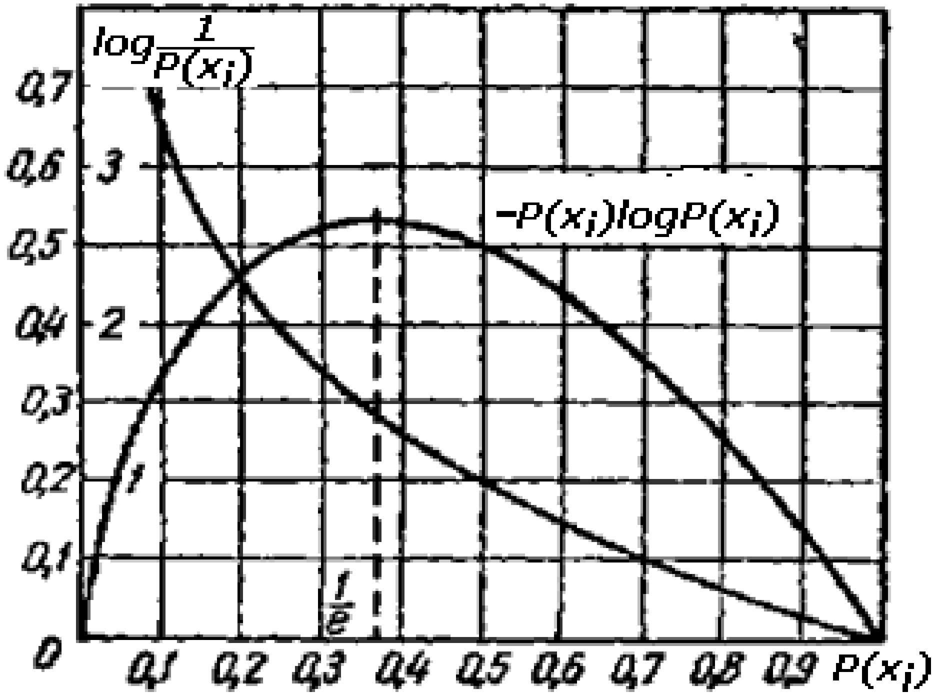
$$H(X) = -\sum_{i=1}^{n} P(x_i) \log P(x_i).$$

Цю величину Шеннон назвав інформаційною ентропією. Інформаційна ентропія оцінює апріорну невизначеність знань про досліджуваний об'єкт.

Якщо всі повідомлення рівноймовірні, тобто $P(x_1)=P(x_2)=\ldots=P(x_n)=1/n$. ентропія максимальна:

$$H(X) = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = \log n.$$

Доданок $-P(x_i)\log P(x_i)$ відображає внесок варіанта повідомлення x_i в ентропію *H.*



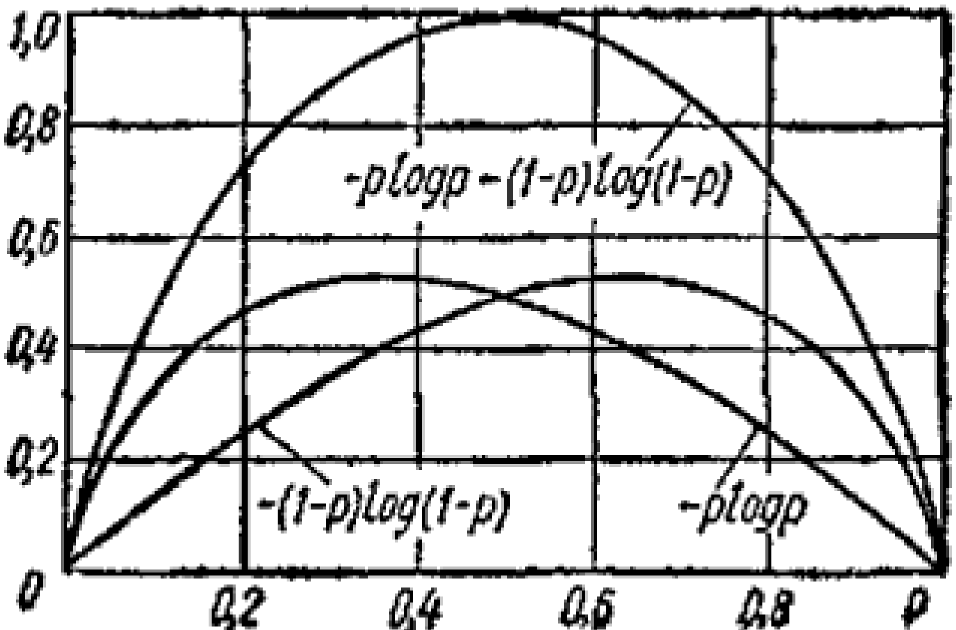
Визначимо точку максимуму цієї функції:

$$0 = \frac{\partial}{\partial P(x_i)}[-P(x_i)\log P(x_i)] = -\log P(x_i) - \log e = \log[P(x_i)e];$$

$$P(x_i)e = 1.$$

При $P(x_i)$ =1/e=0,37 маємо максимум 0,531.

Розглянемо випадок, коли є два тільки варіанти повідомлення, наприклад, x_1=0, x_2=1. Можна покласти $P(x_1) = P$ і $P(x_2) = 1 - P$. Графік $-P(x_2)\log P(x_2)$ буде дзеркальним відображенням графіка $-P(x_1)\log P(x_1)$



Ентропія такого бінарного повідомлення досягає максимуму, рівного 1, при P = 0,5. Це і прийнято за одиницю ентропії **біт**.

В результаті дій шумів і завад у каналі зв'язку повідомлення можуть спотворюватись, і замість переданого X одержувач отримає повідомлення Z, варіанти якого будуть $z_1, z_2, ..., z_n$. Приймаючи повідомлення z_j, неможливо точно встановити, яке зі значень X було передане. Можна лише говорити про умовну ймовірність $P(x_i/z_j)$ передачі значення x_i за умови, що було прийняте значення z_j.

Якщо апріорну невизначеність оцінювати частковою ентропією $H(x_i) = -\log_2 P(x_i)$, а апостеріорну як $H(x_i/z_j) = -\log_2 P(x_i/z_j)$, то кількість інформації в цьому випадку

$$I(z_j / x_i) = H(x_i) - H(x_i / z_j) = \log \frac{P(x_i / z_j)}{P(x_i)}.$$

Для всіх варіантів повідомлення X усереднюємо по ймовірностях $P(x_i/z_j)$:

$$I(z_j/X)=\sum_i P(x_i/z_j)I(z_j/x_i)=\sum_i P(x_i/z_j)\log\frac{P(x_i/z_j)}{P(x_i)}.$$

Для всіх варіантів повідомлення Z треба це усереднити по ймовірностях $P(z_j)$ цих варіантів:

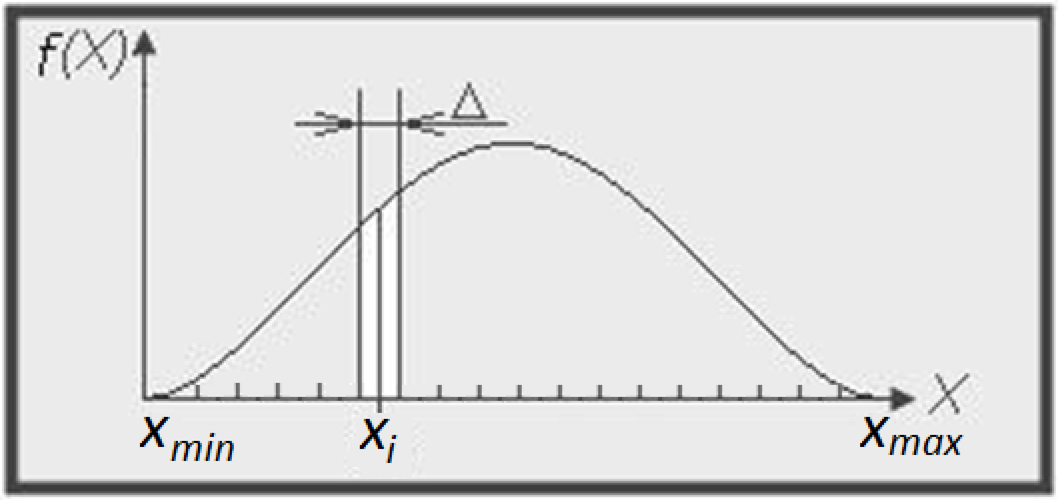
$$I(Z/X)=\sum_j P(z_j)I(z_j/X)=\sum_j\sum_i P(z_j)P(x_i/z_j)\log\frac{P(x_i/z_j)}{P(x_i)}.$$

Після алгебраїчних перетворень цю формулу можна переписати так:

$$I(Z/X)=-\sum_i P(x_i)\log P(x_i)+\sum_j P(z_j)\sum_i P(x_i/z_j)\log P(x_i/z_j)=H(X)-H(X/Z).$$

Тут ентропія $H(X)$ характеризує апріорну невизначеність, а умовна ентропія $H(X/Z)$ – апостеріорну.

Для оцінки ентропії неперервного сигналу діапазон значень сигналу можна розбити на $N=(x_{\max}-x_{\min})/\Delta$ ділянок. Ширину ділянки Δ не має сенсу вибирати меншою, ніж похибка вимірювання датчика сигналу або ступінь квантування q.



Ймовірність того, що значення величини *X* попаде на *i*-ту ділянку, приблизно дорівнює $f(x_j)\times\Delta$. Отримані так ймовірності підставляємо у формулу ентропії:

$$H(X) = -\sum_{i=1}^{N} f(x_i)\Delta \log[f(x_i)\Delta] = -\sum_{i=1}^{N} f(x_i)\Delta \log f(x_i) + \sum_{i=1}^{N} f(x_i)\Delta \log \Delta.$$

Переходячи від суми до інтеграла, отримуємо:

$$H(X) = -\int_{-\infty}^{\infty} f(X)\log f(X)dX - \int_{-\infty}^{\infty} f(X)dX \log\Delta.$$

Враховуючи, що $\int_{-\infty}^{\infty} f(X)dX = 1$, отримаємо:

$$H(X) = -\int_{-\infty}^{\infty} f(X)\log f(X)dX - \log\Delta = h(X) - \log\Delta.$$

Перший доданок *h*(*X*) в отриманому виразенні залежить тільки від розподілу ймовірностей значень *X* і називається **диференціальною ентропією**.

Диференціальна ентропія використовується як оцінка апріорної невизначеності неперервного повідомлення.

Для розрахунку апостеріорної невизначеності виконується аналогічне інтегральне перетворення формули умовної ентропії:

$$h(X/Z) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(Z)f(X/Z)\log f(X/Z)dXdZ.$$

Кількість інформації у неперервному повідомленні:

$$I(Z / X) == h(X) - h(X / Z).$$

При розрахунках найчастіше використовуються два закони розподілу неперервних повідомлень: рівномірний і нормальний.

Стандартний нормальний закон:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{X^2}{2\sigma^2}},$$

де σ — середнє квадратичне відхилення (СКВ). Підставимо у формулу диференціальної ентропії:

$$h(X) = -\int_{-\infty}^{\infty} p(X) \log p(X) dx = -\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{X^2}{2\sigma^2}} \log \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{X^2}{2\sigma^2}}) dX = \log(\sigma\sqrt{2\pi e}).$$

Густина ймовірностей при рівномірному законі величини, що лежить у межах інтервалу значень (*a, b*):

$$f(x)=\begin{cases}0, & при \quad x<a,\\ \dfrac{1}{b-a}, & при \quad a\le x\le b,\\ 0, & при \quad x>b.\end{cases}$$

Диференціальна ентропія:

$$h(X)=-\int_{-\infty}^{\infty}p(X)\log p(X)dX=-\int_{a}^{b}\frac{1}{b-a}\log(\frac{1}{b-a})dX=\log(b-a).$$

Прирівняємо диференціальні ентропії для двох законів:

$$\log(b-a)=\log\sigma\sqrt{2\pi e};\quad b-a=\sigma\sqrt{2\pi e};\quad (b-a)^2=2\pi e\sigma^2.$$

Оскільки дисперсія при рівномірному законі $\sigma_p^2=\dfrac{(b-a)^2}{12},$ то

$$\sigma_p^2=\frac{\pi e\sigma^2}{6}\approx 1{,}42\sigma^2.$$

Отже, при однаковій невизначеності сигналів середня потужність при рівномірному розподілі повинна бути на 42% більше, ніж при нормальному розподілі.