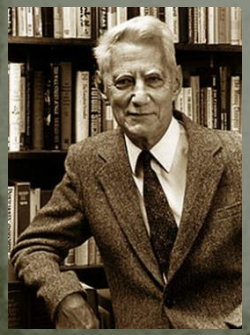
Лекція 3.

Умовна ентропія та інформаційні характеристики каналу передачі інформації

1. Сутність інформаційних характеристик
2. Математичний опис
3. Технічні та інформаційні характеристики каналу зв'язку





Ю. Бродський (кафедра КБ та ЗІ)

1. Сутність інформаційних характеристик

При оцінюванні невизначеності часто необхідно врахувати статистичні зв'язки, які в більшості випадків мають місце між станами двох або більше джерел, об'єднаних в рамках однієї системи, або між станами, які послідовно обираються одним джерелом.

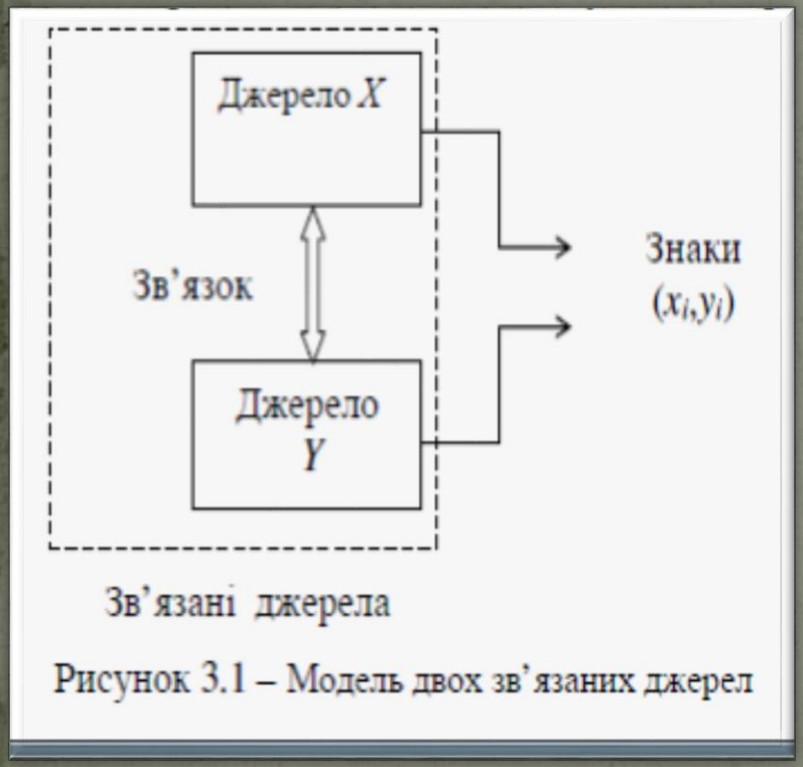
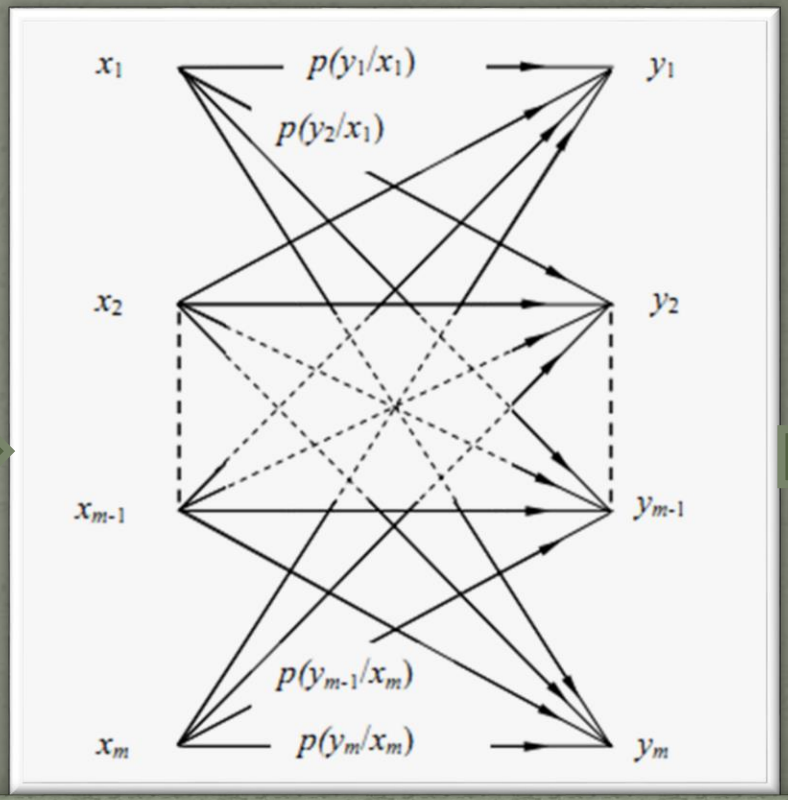


Рисунок 3.1 – Модель двох зв'язаних джерел

X



Y

Рис. 3.2 – модель каналу «X – Y»

Основні поняття:

спільна інформація (ентропія***)***– загальна кількість інформації від різних джерел;

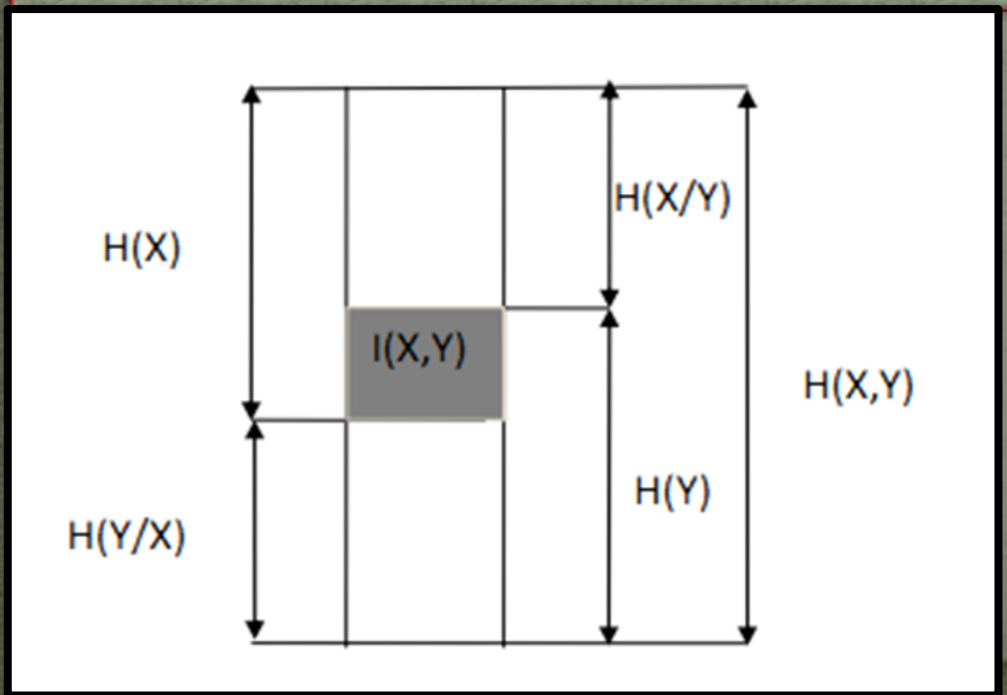
взаємна інформація (ентропія***)***– кількість інформації, яку можна отримати про джерело *X*, спостерігаючи джерело *Y*, або навпаки.

Спільна ентропія розраховується

$H(X, Y) = H(X) + H(Y/X) = H(Y) + H(X/Y)$

Взаємна інформація розраховується

$I(X,Y) = H(X) – H(X/Y) = H(Y) – H(Y/X)$



$$I(X,Y) = H(X) - H(X/Y),$$

$$I(X,Y) = I(Y,X) = H(Y) - H(Y/X)$$

$$I(X,Y) = H(X) + H(Y) - H(X,Y)$$

$$0 \le I(X,Y) \le H(X), \quad 0 \le I(X,Y) \le H(Y).$$

Поняття взаємної інформації широко використовується в теорії передачі інформації. Вимоги до взаємної інформації різні в залежності від того, з якою інформацією працює споживач.

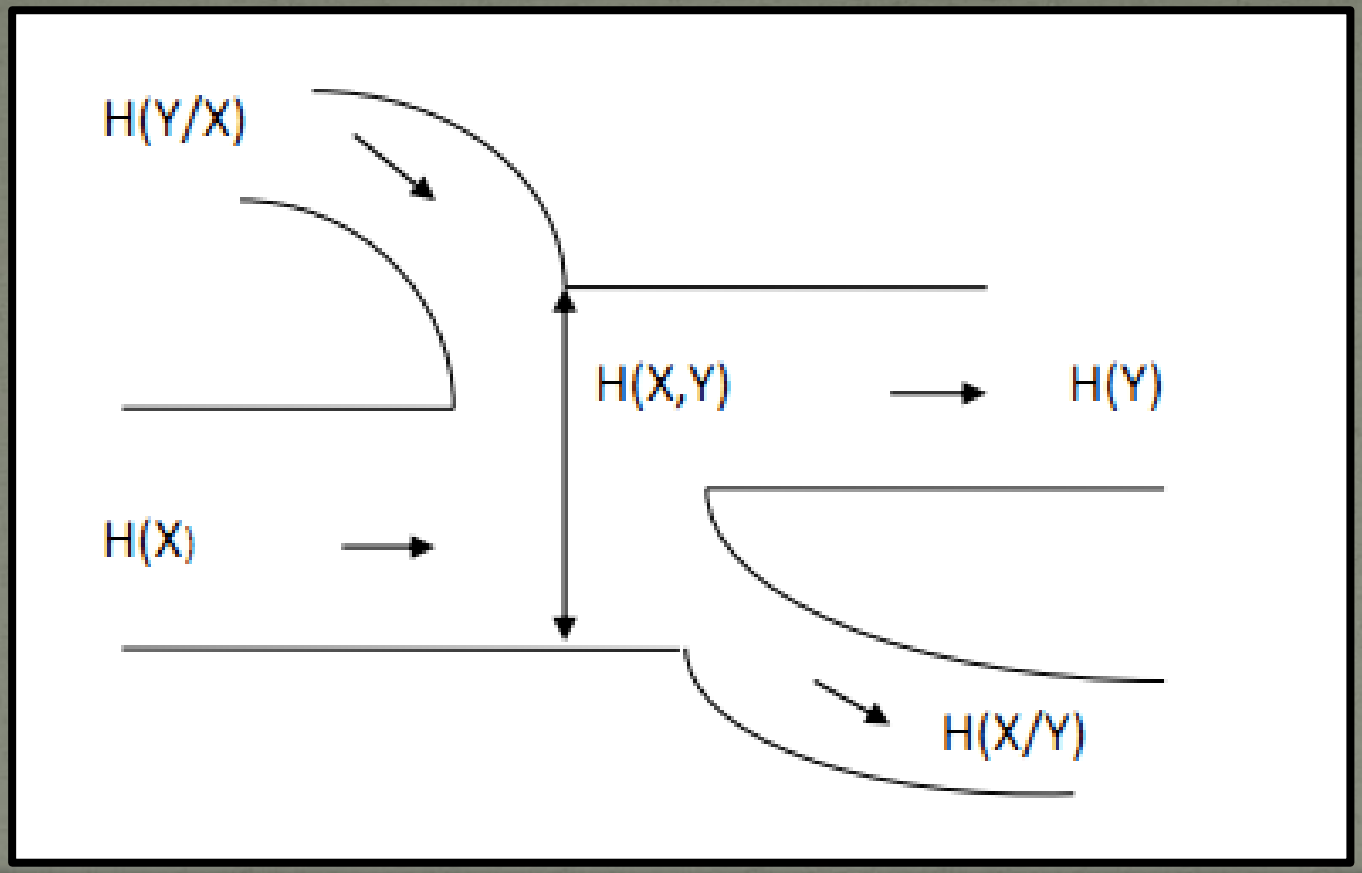
Якщо X і Y - це повідомлення, що передаються різними джерелами, то для отримання якомога більшої сумарної (спільної) інформації, взаємна (тобто однакова в даному випадку) інформація повинна бути мінімальною.

Якщо X і Y - це повідомлення на вході і на виході каналу зв'язку з перешкодами, то для отримання якомога більшої інформації її одержувачем необхідно, щоб взаємна інформація була найбільшою.

Взаємна інформація - це нове поняття теорії інформації, що відіграє центральну роль в інформаційній техніці і описує передачу інформації каналом зв'язку.

Взаємна інформація зв'язує канал і можливість передачі інформації через нього. Тому, взаємна інформація (як і спільна) визначається умовною ентропією.

Сутність умовної ентропії



Умовна ентропія *H(X/Y)* – це втрати інформації в каналі зв'язку (ненадійність каналу).

Умовна ентропія *H(Y/X)* – це інформація про перешкоди (ентропія джерела перешкод), що подається в канал ззовні або утворюється внутрішніми перешкодами в каналі.



Рисунок 6.5 – Діаграма інформаційних потоків

$H(X \mid Y)$ – визначає середню міру невизначеності посланого символу джерела X у тому випадку, коли символи приймача джерела Y відомі, тобто невизначеність приймача, яка залишилася.

Величину $H(X \mid Y)$ часто називають також «витоком» інформації, тому що ентропію $H(X)$ можна інтерпретувати як власну інформацію джерела X, $H(X) = I(X;X)$.

У безшумному каналі інформація передається від джерела X до джерела Y і назад без втрат (без «витоку»). Якщо канал цілком зашумлений, то $H(X \mid Y) = H(X)$ і ніякої передачі інформації не відбувається («витікає» вся інформація).

$H(Y \mid X)$ – визначає середню невизначеність прийнятого символу при відомих посланих символах, тому її називають «сторонньою» шумовою інформацією.

Передачу інформації зашумленим каналом можна розглядати як серію випадкових експериментів, які сприяють зниженню невизначеності. З точки зору теорії інформації, канал є джерелом шумів.

2. Математичний опис

Таблиця 6.2 – Дискретні джерела без пам'яті X і Y із символами

$x \in X = \{x_1, x_2, \ldots, x_m\}$ і $y \in Y = \{y_1, y_2, \ldots, y_N\}$

Теорія ймовірностей	Теорія інформації
Імовірність окремого символу (апріорна імовірність) $p(x)$	Інформація окремого символу $I(x) = -\log_2 p(x)$, біт
Спільна імовірність двох символів $p(x,y)$	Інформація пари символів $I(x,y) = -\log_2 p(x,y)$, біт
Умовна імовірність (апостеріорна імовірність) $p(x\|y) = \frac{p(x,y)}{p(y)}$; $p(x\|y) = \frac{p(x,y)}{p(x)}$	Умовна інформація $I(x\|y) = -\log_2 p(x\|y)$, біт $I(y\|x) = -\log_2 p(y\|x)$, біт
Немає аналогів	Взаємна інформація $I(x;y) = \log_2 \frac{\text{апостеріор. ймовірн.}}{\text{апріорная ймовірн.}} =$ $= \log_2 \frac{p(x\|y)}{p(x)} = \log_2 \frac{p(y\|x)}{p(y)}$, біт Завжди додатна.

Теорія ймовірностей	Теорія інформації
Позначення: $P(x) = P_X(x_i)\ x_i$ для $x_i \in X$ $P(x\|y) = P_{X\|Y}(x_i\|y_j)$ для $x_i \in X,\ y_j \in Y$ $\sum_X p(x) = \sum_{i=1}^{M} P_X(x_i)$ $\sum_Y p(y) = \sum_{j=1}^{N} P_Y(y_j)$	Деякі важливі співвідношення, що використовуються для розрахунків $\sum_X p(x) = 1;\ \sum_Y p(y) = 1;$ $\sum_X \sum_Y p(x,y) = 1;$ $\sum_Y p(x,y) = p(x);\ \sum_X p(x,y) = p(y);$ $\sum_Y p(x\|y) = 1;\ \sum_X p(x\|y) = 1;$

Таблиця 6.3 – Опис “у середньому” дискретних джерел без пам'яті із символами $x \in X = \{x_1, x_2, \ldots, x_m\}$ і $y \in Y = \{y_1, y_2, \ldots, y_N\}$

Ентропія	$H(X) = -\sum_X p(x)\log_2 p(x)$, біт $H(Y) = -\sum_Y p(y)\log_2 p(y)$, біт
Спільна ентропія	$H(X,Y) = -\sum_X \sum_Y p(x,y)\log_2 p(x,y)$, біт
Умовна ентропія	$H(X\|Y) = -\sum_X \sum_Y p(x,y)\log_2 p(x\|y)$, біт $H(Y\|X) = -\sum_X \sum_Y p(x,y)\log_2 p(y\|x)$, біт
Передавання інформації	$I(X;Y) = \sum_X \sum_Y p(x,y)\log_2 \frac{\text{апостеріор. ймовірн.}}{\text{апріорна ймовірн.}} =$ $= \sum_X \sum_Y p(x,y)\log_2 \frac{p(y\|x)}{p(y)}$, біт = $= \sum_X \sum_Y p(x,y)\log_2 \frac{p(x\|y)}{p(x)}$, біт ≥ 0.
Найважливіші залежності (знак рівності справедливий для незалежних джерел)	$H(X) \geq H(X\|Y)$ і $H(Y) \geq H(Y\|X)$ $H(X,Y) = H(X) + H(Y\|X) = H(Y) + H(X\|Y)$ $H(X,Y) = H(X) + H(Y) - I(X;Y)$; $I(X;Y) \geq 0$. $I(X;Y) = H(X) + H(Y) - H(X,Y)$.

3. Технічні та інформаційні характеристики каналу зв'язку

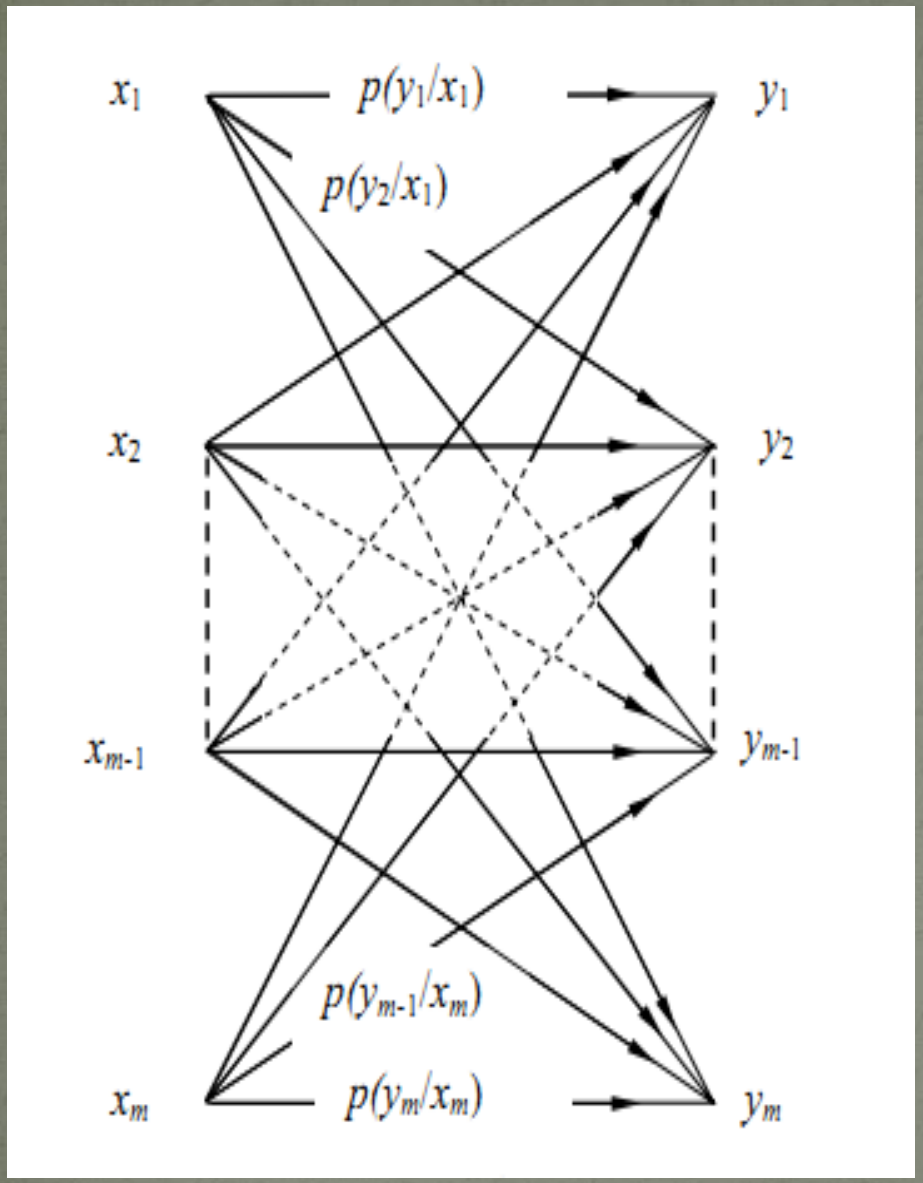
Однією із характеристик, яка визначає швидкість $\overline{I}$ передачі інформації по дискретному каналу, є ***середня швидкість v_0 передачі символів*** (не треба плутати ці дві характеристики). Чисельно v_0 дорівнює середній кількості символів, що передаються по каналу за одиницю часу, і вимірюється в *Бодах*. 1 *Бод* відповідає передачі 1 символу за секунду. В технічних системах, як правило, тривалості всіх символів, тобто час їх перебування на вході та на виході каналу є однаковим; позначимо цю тривалість через T. Тоді швидкість передачі символів по дискретному каналу (її ще називають технічною швидкістю або швидкістю модуляції) $v_0 = 1/T$.

Дискретний канал вважають цілком заданим, якщо можна визначити ймовірність переходу будь-якої послідовності символів на вході каналу для будь-яких фіксованих, кратних T моментів часу у будь-яку послідовність символів на виході каналу для тих же моментів часу.

Для деяких каналів достатньо мати ***матрицю перехідних ймовірностей***

$$\begin{bmatrix} p(y_1/x_1) & p(y_2/x_1) & \dots & p(y_N/x_1) \\ p(y_1/x_2) & p(y_2/x_2) & \dots & p(y_N/x_2) \\ \vdots & \vdots & \vdots & \vdots \\ p(y_1/x_M) & p(y_2/x_M) & \dots & p(y_N/x_M) \end{bmatrix}. \qquad (3.1)$$

Тут $p(y_k/x_i)$ – умовна ймовірність появи на виході каналу в деякий момент часу символу y_k при умові, що на вході каналу в цей же момент мав місце символ x_i. Зрозуміло, що сума кожного рядка матриці (3.1) дорівнює одиниці.



Дія перешкод приводить до того, що передаючи сигнал *xi*, у послідовності після нього будемо одержувати різні сигнали *yj*, що й відображається графом і канальною матрицею.

Вибором якісних ознак сигналів (частоти, фази та ін.) з урахуванням характеристик перешкод забезпечуються найбільші значення умовних ймовірностей на головній діагоналі матриці, чим і досягається найбільш достовірна передача сигналів .

Дискретні канали класифікують в залежності від властивостей перехідних ймовірностей.

Якщо перехідні ймовірності не залежать від часу, канал називають ***стаціонарним*** або ***однорідним;*** в протилежному випадку – ***нестаціонарним, неоднорідним.***

Канал називають каналом без пам'яті, якщо перехідні ймовірності не залежать від передісторії. Термін "без пам'яті" підкреслює той факт, що при передачі чергового символу канал нібито "не пам'ятає" результатів попередніх передач. Канал, для якого перехідні ймовірності залежать від передісторії називають, ***каналом із пам'яттю.*** Для каналу без пам'яті достатньо мати матрицю перехідних ймовірностей типу (3.1). Канали ж із пам'яттю потребують значно складнішого опису. Це наприклад, може бути сукупність матриць перехідних ймовірностей, кожна з яких відповідає одному із станів каналу.

Швидкістю передачі інформації по каналу називають кількість інформації, що передається по каналу за одиницю часу (біт/с).

$$\tilde{I} = \nu_0 I(Y,X) = \nu_0 \left(H(X) - H(X/Y)\right) = \nu_0 (H(Y) - H(Y/X))$$

$$\nu_0 = \frac{1}{T}$$

Ю. Бродський (кафедра КБ та ЗІ)

Пропускна здатність C каналу – це максимально можлива швидкість передачі інформації через цей канал

$$C = \max_{\{p(x_i)\}} \tilde{I}$$

Зрозуміло, що пропускна здатність вимірюється, як і швидкість передачі інформації, в *біт/с*.

Слід зауважити, що при визначенні пропускної здатності каналу всі його характеристики (швидкість v_0 передачі символів, перехідні ймовірності) є фіксованими, а максимум відшукується по усім розподілам ймовірностей появи символів на вході каналу. Тобто знаходиться такий набір значень ймовірностей появи символів на вході каналу, який забезпечує максимум C. Звідси виходить, що пропускна здатність визначається тільки характеристиками каналу.

$$C = v_0 \max_{\{p(x_i)\}} \left(H(X) - H(X/Y) \right)$$

$$C = v_0 \max_{\{p(x_i)\}} \left(H(Y) - H(Y/X) \right)$$

Висновки:

- Взаємна інформація - відіграє центральну роль в інформаційній техніці і описує передачу інформації каналом зв'язку і визначається умовною ентропією.
- Умовна ентропія указує на втрати інформації в каналі зв'язку, надає інформацію про перешкоди.
- Швидкість передачі інформації – це кількість інформації, що передається по каналу за одиницю часу (біт/с).
- Пропускна здатність каналу – це максимально можлива швидкість передачі інформації.

Ю. Бродський (кафедра КБ та ЗІ)