



03

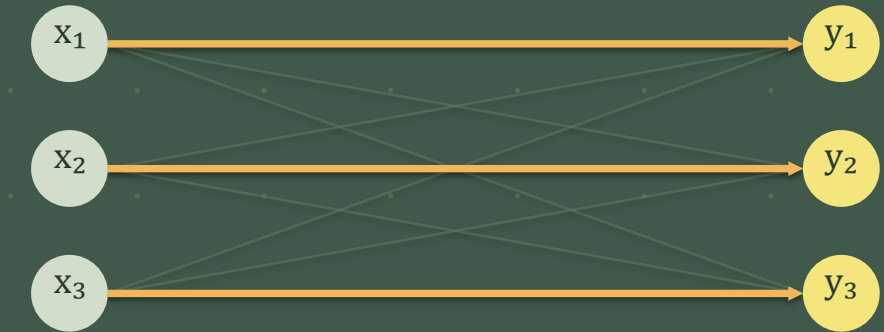
УНІВЕРСИТЕТСКА ЛЕКЦІЯ

Умовна ентропія

та інформаційні характеристики
каналу передачі інформації

- 01 Сутність інформаційних характеристик
- 02 Математичний опис
- 03 Технічні та інформаційні характеристики каналу зв'язку

ІНФОРМАЦІЯ У КАНАЛІ



Вхід X

Вихід Y

 $I(X;Y)$ Взаємна
інформація



Сутність інформаційних характеристик

01

Спільна, взаємна та умовна інформація

Статистичний зв'язок джерел враховують під час оцінювання невизначеності.

$$H(X,Y)$$

Спільна ентропія

Інформація від пари джерел

$$I(X;Y)$$

Взаємна інформація

Відомості про X зі спостереження Y

$$H(X|Y)$$

Умовна ентропія

Невизначеність X , що залишається

Модель пов'язаних джерел X–Y

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.

ДВА ДИСКРЕТНІ ДЖЕРЕЛА

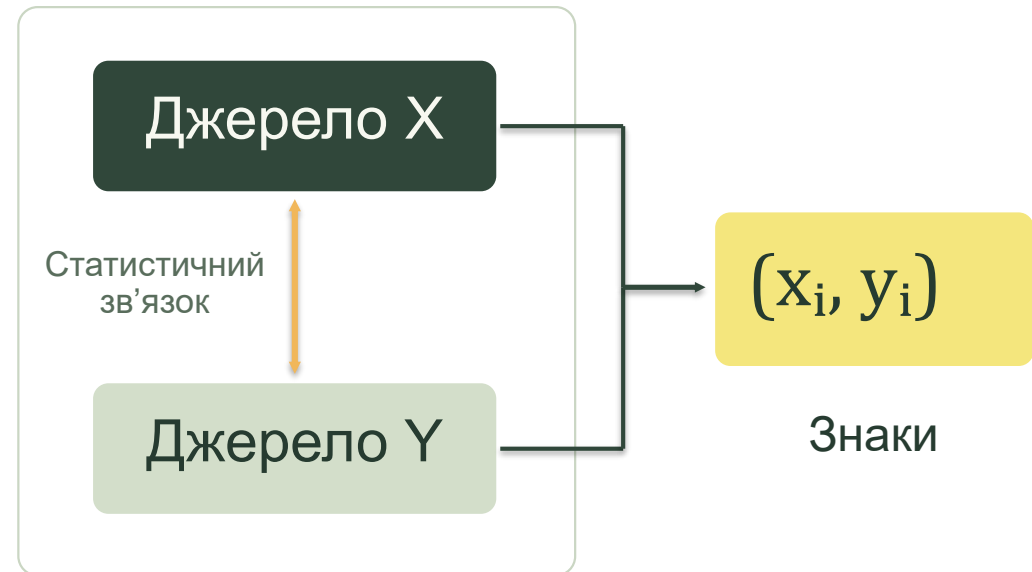


Рис. 3.1 — Модель двох зв'язаних джерел

Модель каналу «X – Y»

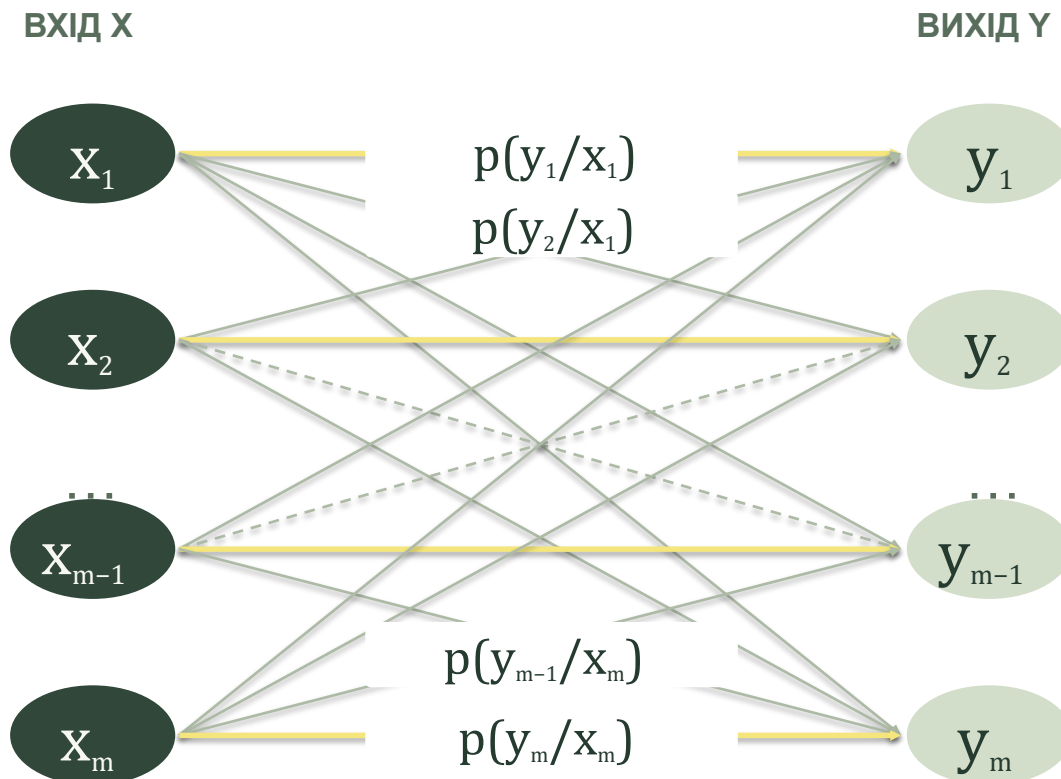


Рис. 3.2 — модель каналу «X – Y»

Статистичні зв'язки між джерелами зображено переходами від вхідних символів X до вихідних символів Y.

Перехідні ймовірності

Позначення зв'язків на графі збережено з рисунка 3.2 початкової лекції.

Спільна та взаємна інформація

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.

ЯК ПОВ'ЯЗАНІ ВЕЛИЧИНИ

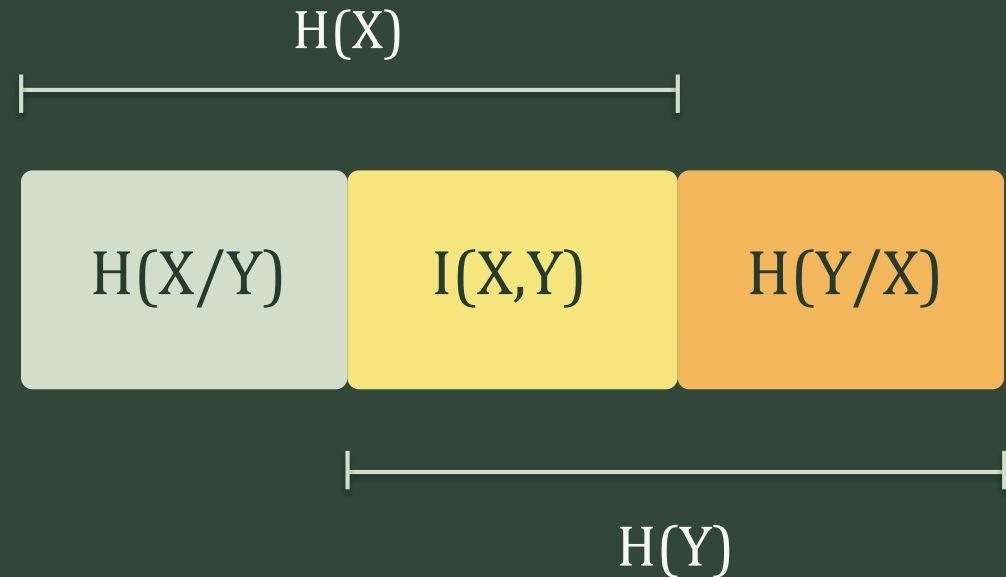


Схема співвідношень, без числового масштабу

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



Співвідношення між інформаційними величинами

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

ТРИ ЗАПИСИ ВЗАЄМНОЇ ІНФОРМАЦІЇ

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

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

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

Незалежні джерела

$$I(X,Y) = 0$$

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

Повна взаємна залежність

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

$$0 \leq I(X,Y) \leq H(X), \quad 0 \leq I(X,Y) \leq H(Y)$$



Коли взаємної інформації потрібно більше?

РІЗНІ ДЖЕРЕЛА

Менше дублювання

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.

$$I(X,Y) \rightarrow \text{мінімум}$$

ВХІД І ВИХІД КАНАЛУ

Більше переданої інформації

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.

$$I(X,Y) \rightarrow \text{максимум}$$

Взаємна інформація пов'язує характеристики каналу з можливістю передавання інформації.



Взаємна інформація та можливості каналу

Теорія передачі

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

Взаємна інформація зв'язує канал і можливість передачі інформації через нього.

Центральна роль

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

Тому, взаємна інформація (як і спільна) визначається умовною ентропією.



Умовна ентропія: втрати та перешкоди

$$H(X/Y)$$



ВТРАТИ ІНФОРМАЦІЇ

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

Характеризує ненадійність каналу

$$H(Y/X)$$

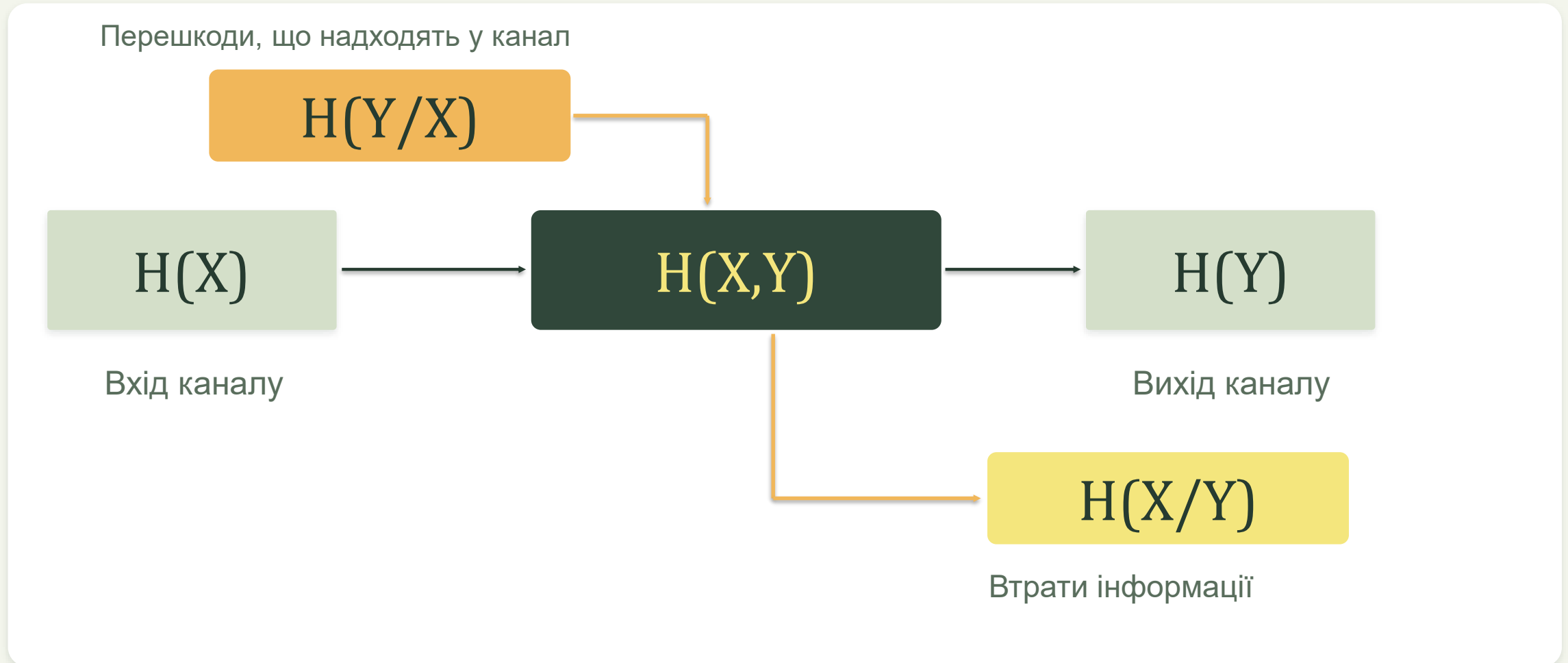


ІНФОРМАЦІЯ ПРО ПЕРЕШКОДИ

Це інформація про перешкоди (ентропія джерела перешкод), що подається в канал ззовні або утворюється внутрішніми перешкодами в каналі.

Перешкоди ззовні або всередині каналу

Сутність умовної ентропії: схема потоків



Діаграма інформаційних потоків

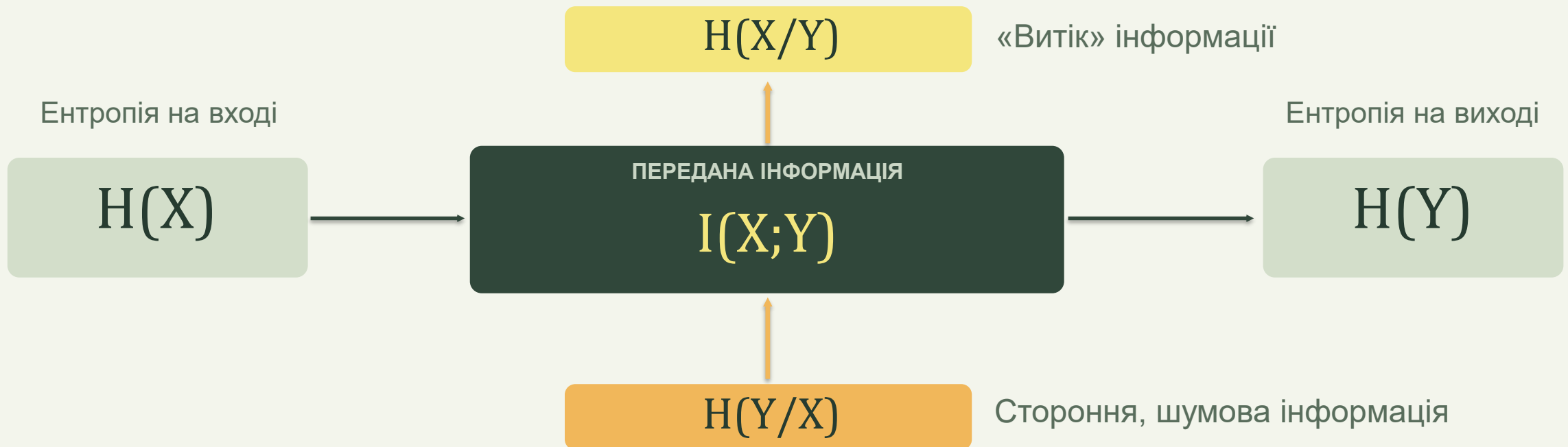


Рисунок 6.5 — Діаграма інформаційних потоків

Безшумний канал

$$H(X/Y) = 0 \rightarrow I(X;Y) = H(X)$$

Цілком зашумлений канал

$$H(X/Y) = H(X) \rightarrow I(X;Y) = 0$$

Що означає умовна невизначеність?

$$H(X|Y)$$



НЕВИЗНАЧЕНІСТЬ ПОСЛАНОВОГО СИМВОЛУ

Визначає середню міру невизначеності посланого символу джерела X у тому випадку, коли символи приймача джерела Y відомі, тобто невизначеність приймача, яка залишилася.

«Вітік» інформації

$$H(Y|X)$$



НЕВИЗНАЧЕНІСТЬ ПРИЙНЯТОГО СИМВОЛУ

Визначає середню невизначеність прийнятого символу при відомих посланих символах, тому її називають «сторонньою» шумовою інформацією.

«Стороння» шумова інформація

Витік інформації та граничні випадки

Власна інформація джерела

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

Без втрат або без передачі

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

Передана інформація та джерело шумів

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

Передачу інформації зашумленим каналом можна розглядати як серію випадкових експериментів, які сприяють зниженню невизначеності.

З точки зору теорії інформації, канал є джерелом шумів.



Математичний опис

02

**Від окремої події
до середніх
характеристик**

Ймовірності, інформаційні
величини та співвідношення
між ними.

 $p(x)$ **Ймовірність**

Опис окремої події

 $I(x)$ **Інформація**

Кількість інформації події

 $H(X)$ **Ентропія**

Усереднений опис джерела

Імовірності та інформаційні величини

Дискретні джерела без пам'яті X і Y із символами:

$$x \in X = \{x_1, x_2, \dots, x_m\}; \quad y \in Y = \{y_1, y_2, \dots, y_N\}$$

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) $p(x y) = \frac{p(x,y)}{p(y)}; \quad p(y x) = \frac{p(x,y)}{p(x)}$	Умовна інформація $I(x y) = -\log_2 p(x y)$, біт $I(y x) = -\log_2 p(y x)$, біт



Взаємна інформація окремої події

ТЕОРІЯ ЙМОВІРНОСТЕЙ

Немає аналогів

Взаємна інформація

ТЕОРІЯ ІНФОРМАЦІЇ

$$I(x;y) = \log_2 \frac{\text{апостеріор. ймовірн.}}{\text{апріорная ймовірн.}}$$

$$I(x;y) = \log_2 \frac{p(x|y)}{p(x)} = \log_2 \frac{p(y|x)}{p(y)}$$

Одиниця кількості інформації — біт.

$$I(X;Y) \geq 0$$



Співвідношення для розрахунків

ТЕОРІЯ ІНФОРМАЦІЇ

Деякі важливі співвідношення, що використовуються для розрахунків

Продовження табл. 6.2

$$\sum_x p(x) = 1; \quad \sum_y p(y) = 1$$

$$\sum_x \sum_y p(x,y) = 1$$

$$\sum_y p(x,y) = p(x); \quad \sum_x p(x,y) = p(y)$$

$$\sum_y p(x|y) = 1; \quad \sum_x p(x|y) = 1$$



Опис дискретних джерел (формули для розрахунків)

Ентропія окремого джерела

$$H(X) = - \sum_x p(x) \log_2 p(x)$$

$$H(Y) = - \sum_y p(y) \log_2 p(y)$$

біт

Спільна ентропія

$$H(X,Y) = - \sum_x \sum_y p(x,y) \log_2 p(x,y)$$

біт

Умовна ентропія

$$H(X|Y) = - \sum_x \sum_y p(x,y) \log_2 p(x|y)$$

$$H(Y|X) = - \sum_x \sum_y p(x,y) \log_2 p(y|x)$$

біт

Взаємна інформація

$$I(X;Y) = \sum_x \sum_y p(x,y) \log_2 \frac{p(y|x)}{p(y)}$$

біт

Передавання інформації

$$I(X;Y) = \sum_X \sum_Y p(x,y) \log_2 \frac{\text{апостеріор. ймовірн.}}{\text{апріорная ймовірн.}}$$

$$= \sum_X \sum_Y p(x,y) \log_2 \frac{p(y|x)}{p(y)}$$

$$= \sum_X \sum_Y p(x,y) \log_2 \frac{p(x|y)}{p(x)}$$

Одиниця кількості інформації — біт.

$$I(X;Y) \geq 0$$

Найважливіші залежності

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

Взаємозв'язок величин

$$I(X;Y) = H(X) - H(X|Y)$$

$$I(X;Y) = H(Y) - H(Y|X)$$

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

Для незалежних джерел

$$I(X;Y) = 0$$

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

Повна взаємна залежність

$$I(X;Y) = H(X) = H(Y)$$

$$I(X;Y) \geq 0$$

$$H(X) \geq H(X|Y), \quad H(Y) \geq H(Y|X)$$



Взаємна інформація через спільні ймовірності

ЕКВІВАЛЕНТНИЙ ЗАПИС ІЗ ІЛЮСТРАЦІЇ ЛЕКЦІЇ

$$I(X,Y) = \sum_{i=1}^N \sum_{j=1}^M p(x_i, y_j) \log_2 \frac{p(x_i, y_j)}{p(x_i) p(y_j)}$$

$$0 \leq I(X,Y) \leq H(X)$$

$$0 \leq I(X,Y) \leq H(Y)$$

Технічні та інформаційні характеристики каналу

03

Символи, інформація та межі передавання

Технічна швидкість і швидкість передачі інформації — різні характеристики.



Технічна швидкість

Скільки символів передано за секунду



Перехідні ймовірності

Як пов'язані символи на вході й виході



Пропускна здатність

Максимальна швидкість інформації

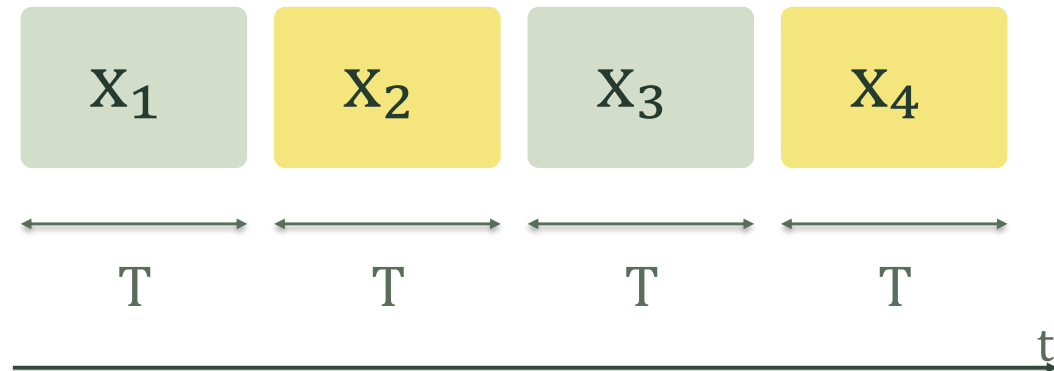
Технічна швидкість: символи за секунду

Однією із характеристик, яка визначає швидкість $\bar{T}$ передачі інформації по дискретному каналу, є середня швидкість v_0 передачі символів (не треба плутати ці дві характеристики).

1 Бод = 1 символ за секунду

Чисельно v_0 дорівнює середній кількості символів, що передаються по каналу за одиницю часу, і вимірюється в Бодах.

ПОСЛІДОВНІСТЬ СИМВОЛІВ



$$v_0 = \frac{1}{T}$$

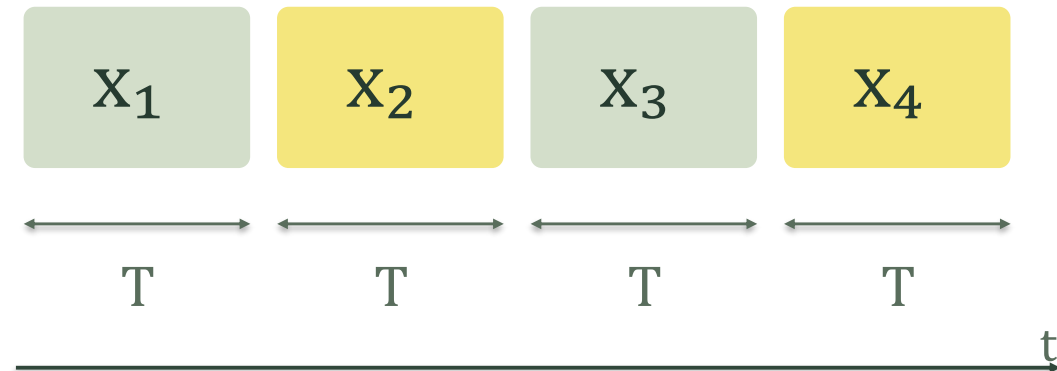
Тривалість символу та швидкість модуляції

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 .

Однакова тривалість — T

Тоді швидкість передачі символів по дискретному каналу (її ще називають технічною швидкістю або швидкістю модуляції) $v_0 = 1/T$.

ПОСЛІДОВНІСТЬ СИМВОЛІВ



$$v_0 = \frac{1}{T}$$

Коли дискретний канал цілком заданий?

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.

Матриця перехідних ймовірностей

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 .





Канальна матриця перехідних імовірностей

Для деяких каналів достатньо мати матрицю перехідних ймовірностей

СТОВПЦІ — СИМВОЛИ Y НА ВИХОДІ

X/Y	y_1	y_2	...	y_N
x_1	$p(y_1/x_1)$	$p(y_2/x_1)$	...	$p(y_N/x_1)$
x_2	$p(y_1/x_2)$	$p(y_2/x_2)$	...	$p(y_N/x_2)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_M	$p(y_1/x_M)$	$p(y_2/x_M)$	...	$p(y_N/x_M)$

Кожен рядок — фіксований вхідний символ x_i .

$$p(y_k/x_i)$$

Умовна ймовірність появи на виході символу y_k за умови, що на вході каналу в цей же момент був символ x_i .

$$\sum_{k=1}^N p(y_k/x_i) = 1$$

Сума кожного рядка — одиниця.

Граф каналу та достовірність передавання

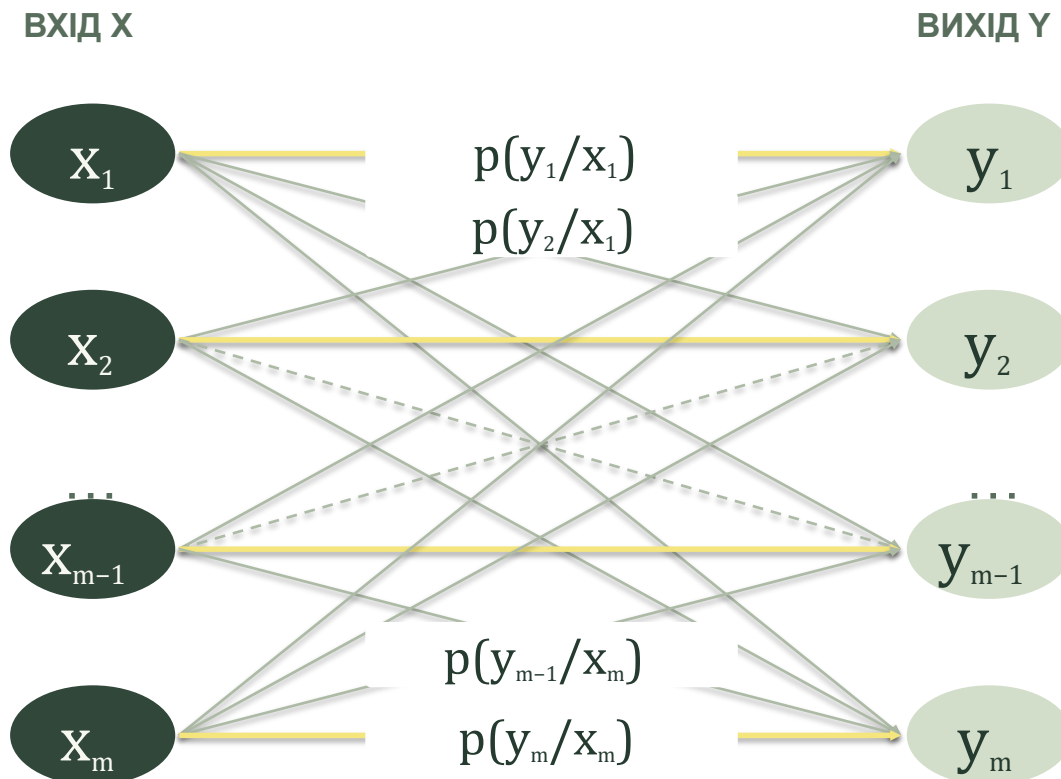


Рис. 3.2 — модель каналу «X – Y»

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

Достовірність передачі досягається

Вибором якісних ознак сигналів — частоти, фази та ін. — та з врахуванням характеристик перешкод.

Якісні ознаки сигналів і достовірність передачі

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.

ГОЛОВНА ДІАГОНАЛЬ МАТРИЦІ

$$p(y_1/x_1)$$

$$p(y_2/x_2)$$

$$p(y_m/x_m)$$

Відповідні символи на вході та на виході



Класифікація дискретних каналів

ЗАЛЕЖНІСТЬ ВІД ЧАСУ

Стаціонарний

Перехідні ймовірності не залежать від часу. Канал однорідний.

Нестационарний

Перехідні ймовірності змінюються з часом. Канал неоднорідний.

ЗАЛЕЖНІСТЬ ВІД ПЕРЕДІСТОРІЇ

Без пам'яті

Перехідні ймовірності не залежать від передісторії.

Із пам'яттю

Перехідні ймовірності залежать від передісторії каналу.

Для каналу без пам'яті достатньо матриці переходів. Для каналу з пам'яттю опис складніший: наприклад, сукупність матриць для різних станів.

Дискретні канали класифікують в залежності від властивостей перехідних ймовірностей.

Канали без пам'яті та канали із пам'яттю

Без пам'яті

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.

Із пам'яттю

Канал, для якого перехідні ймовірності залежать від передісторії називають, каналом із пам'яттю.

Канали ж із пам'яттю потребують значно складнішого опису.

Як описують канали з різними властивостями?

Матриця переходів

Для каналу без пам'яті достатньо мати матрицю перехідних ймовірностей типу.

Для каналу без пам'яті

Сукупність матриць

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.

Для різних станів каналу



Швидкість передачі інформації

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

$$\tilde{I} = v_0 I(Y,X) =$$

$$= v_0 (H(X) - H(X/Y)) =$$

$$= v_0 (H(Y) - H(Y/X))$$

СИМВОЛИ
ЗА СЕКУНДУ

$$v_0 = \frac{1}{T}$$

Технічна
швидкість



Пропускна здатність каналу

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

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

біт/с

ФІКСОВАНО

Технічна швидкість v_0
та перехідні ймовірності каналу.

МАКСИМІЗУЄМО ЗА

Усіма розподілами ймовірностей
появи символів на вході каналу.

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

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

Що залишається фіксованим, а що обирають?

Характеристики каналу

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,

Розподіл на вході

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 .

Пропускна здатність визначається тільки характеристиками каналу.



Висновки: чотири характеристики каналу

01 Взаємна інформація

$$I(X;Y)$$

Відіграє центральну роль в інформаційній техніці, описує передачу інформації каналом зв'язку і визначається умовною ентропією.

02 Умовна ентропія

$$H(X|Y)$$

$$H(Y|X)$$

Указує на втрати інформації в каналі зв'язку, надає інформацію про перешкоди.

03 Швидкість інформації

$$\bar{I}$$

Кількість інформації, що передається по каналу за одиницю часу (біт/с).

04 Пропускна здатність

$$C$$

Максимально можлива швидкість передачі інформації через канал.

[Повернутися до плану](#)