Лекция № 2

«КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ.

ЕДИНИЦЫ КОЛИЧЕСТВА ИНФОРМАЦИИ»

2 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Определение

Под ***информацией*** понимают совокупность сведений о явлениях, процессах, событиях, фактах и т.д., которые принимает человек в процессе жизнедеятельности.

Сообщение, получаемое на приемной стороне, несет полезную информацию лишь в том случае, если имеется ***неопределенность*** относительно состояния источника.

Информацию несут в себе те сообщения, которые снимают неопределенность, существующую до их поступления.

3 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Дискретные сообщения

Количество возможных дискретных сообщений является простейшей характеристикой источника дискретных сообщений.

Опыт с одним исходом

Опыт с двумя исходами

Опыт с множеством

$$V_i \in [0..60]\, Вольт$$

4 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Общий случай:

m Количество символов сообщения

n Количество символов в каждом сообщении

$$N = m^n \qquad (2.1)$$

Количество различных сообщений длиной n из символом m

Ансамбль **A** $\{a_i\}, i = \overline{1, m}$

$$A = \begin{pmatrix} a_1 & a_2 & \dots & a_m \\ p(a_1) & p(a_2) & \dots & p(a_m) \end{pmatrix} \qquad \sum_{i=1}^{m} p(a_i) = 1 \qquad (2.2)$$

5 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Мера Хартли

Ральф Винтон Лайон Хартли

30 ноября 1888,
Спрус, Невада — 1
мая 1970, Нью-Джерси) —
американский учёный-
электронщик

В 1928 г. Р. Хартли предложил использовать логарифмическую функцию от N в качестве количественной меры информации.

$$I = \log N = n \log m \qquad \textbf{(2.3)}$$

Единицы измерения – бит.

Для равновероятных символов можно записать

$$p(a_1) = \frac{1}{m} \qquad \textbf{(2.4)}$$

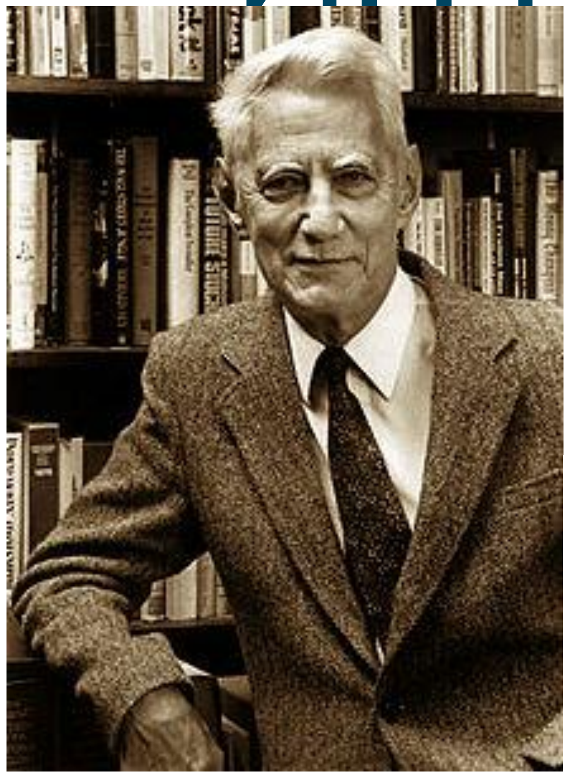
(2.3) → (2.4)

$$I_1 = \log \frac{1}{p(a_1)} = -\log p(a_1) = \log_2 N \qquad \textbf{(2.5)}$$

(2.5) – Количество информации на символ алфавита

6 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Мера Шеннона



Клод Э́лвуд Ше́ннон

30 апреля 1916 Мичиган, США — 24 февраля 2001, Медфорд, Массачусетс, США) — американский инженер и математик

Мера Хартли не отражает вероятностный характер сообщений. Эту задачу решил в 1946 г. К. Шеннон.

В общем случае, для n несовместных и независимых сообщений с вероятностями появления равными $p(a_1)$ $p(a_2)$ ... $p(a_n)$. Чем меньше априорная вероятность события, тем большее количество информации оно несет. Поэтому можно предположить, что количественной мерой неопределенности отдельного сообщения, а также передаваемой им информации, может быть величина, обратная его априорной вероятности

$$\frac{1}{p(a_i)}$$

Сложное событие: a_i a_j

Вероятность события: $p(a_i)p(a_j)$

Количество информации в сложном сообщении: $\dfrac{1}{p(a_i)p(a_j)}$

Такая мера не обладает свойством аддитивности

Мера Хартли

Удобнее воспользоваться логарифмической мерой количества информации

$$I(a_i) = \log \frac{1}{p(a_i)} = -\log p(a_i) \quad [\text{бит}] \qquad \textbf{(2.6)}$$

Количество информации в сложном сообщении:

$$I(a_i, a_j) = \log \frac{1}{p(a_i)p(a_j)} = \log \frac{1}{p(a_i)} + \log \frac{1}{p(a_j)} = I(a_i) + I(a_j)$$

Количество информации для единичного исхода:

$$I(a_1) = \log \frac{1}{1} = 0$$

Прикладная теория цифровых автоматов / К. Г. Самофалов, А. М. Ромлинкевич, В. Н. Валуйский, Ю. С. Каневский, М. М. Пиневич.— К.: Вища шк. Головное изд-во, 1987. — 375 с.

8 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Частная энтропия

$$H(a_i) = \log\frac{1}{p(a_i)} = -\log(p(a_i)) \text{ [бит/символ]} \qquad \mathbf{(2.7)}$$

Количество информации и неопределенность для всей совокупности случайных сообщений можно получить усреднением по всем событиям

$$I(A) = \sum_{i=1}^{n} p(a_i)\log\frac{1}{p(a_i)} = -\sum_{i=1}^{n} p(a_i)\log p(a_i) \qquad \mathbf{(2.8)}$$

$$H(A) = -\sum_{i=1}^{n} p(a_i)\log p(a_i) \qquad \mathbf{(2.9)}$$

(2.9) – формула Шеннона для энтропии источника дискретных сообщений

Энтропия − это среднестатистическая мера неопределенности знаний получателя информации относительно состояния наблюдаемого объекта (источник сообщений)

Кузьмин И.В., Кедрус В.А. Основы теории информации и кодирования с. 115

9 КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Свойства энтропии:

Чем выше энтропия источника, тем большее количество информации в среднем закладывается в каждое сообщение, тем сложнее запомнить, передать или сохранить такое сообщение по каналу связи.

Смысл энтропии Шеннона:

Энтропия дискретной случайной величины − это минимум среднего количества бит, которое нужно передавать по каналу связи о текущем значении данной случайной величины.

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Свойства энтропии

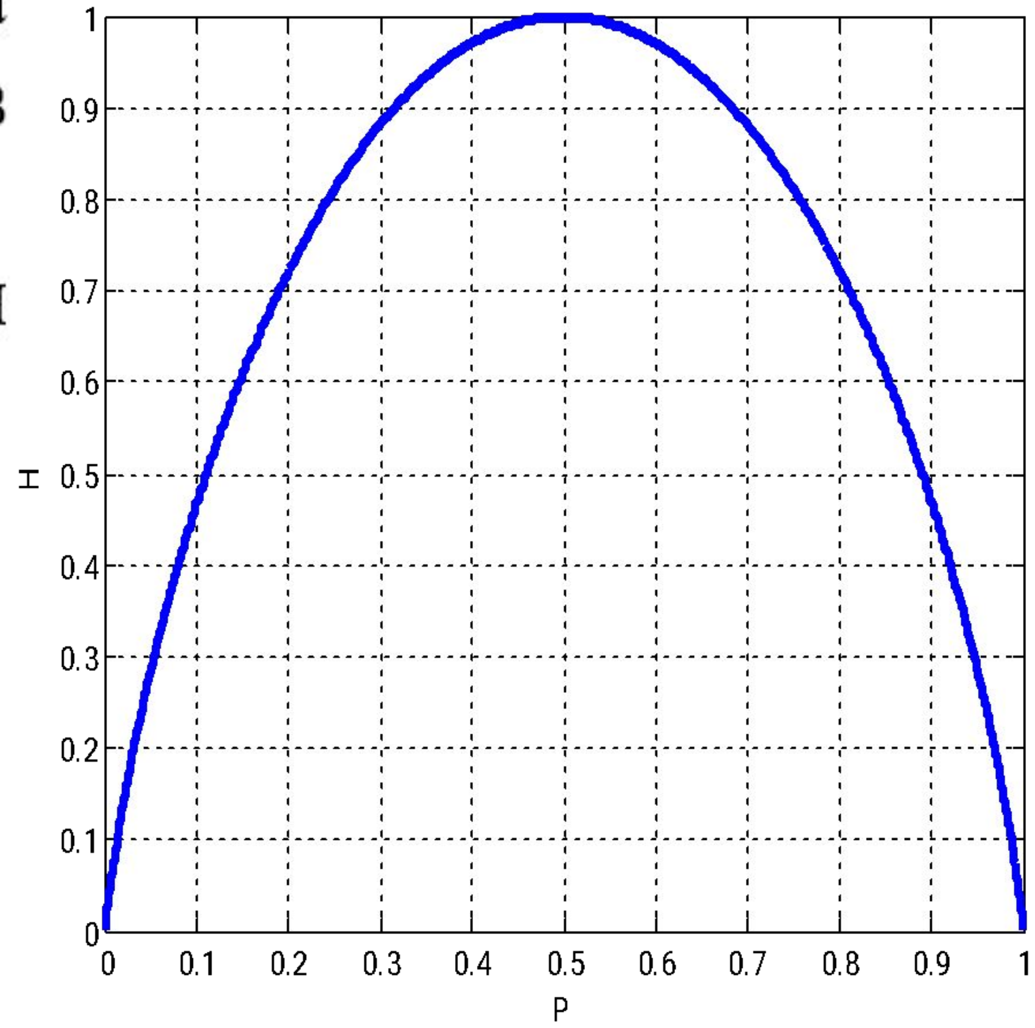
1. Энтропия источника дискретных сообщений есть величина вещественная, ограниченная и положительно определенная;
2. Энтропия равна нулю, если с вероятностью единица всегда выбирается один и тот же символ (в этом случае неопределённость в поведении источника сообщений отсутствует).
3. Энтропия максимальна, если все символы источника появляются независимо и с одинаковой вероятностью, в этом случае она равна

$$H_{\max} = -mp(a_i)\log p(a_i) = -m\frac{1}{m}\log\frac{1}{m} = \log m$$

Пусть $m = 2$, тогда

$$H(A) = -\sum_{i=1}^{2} p(a_i)\log p(a_i) = -p(a_1)\log p(a_1) - (1 - p(a_1))\log(1 - p(a_1))$$

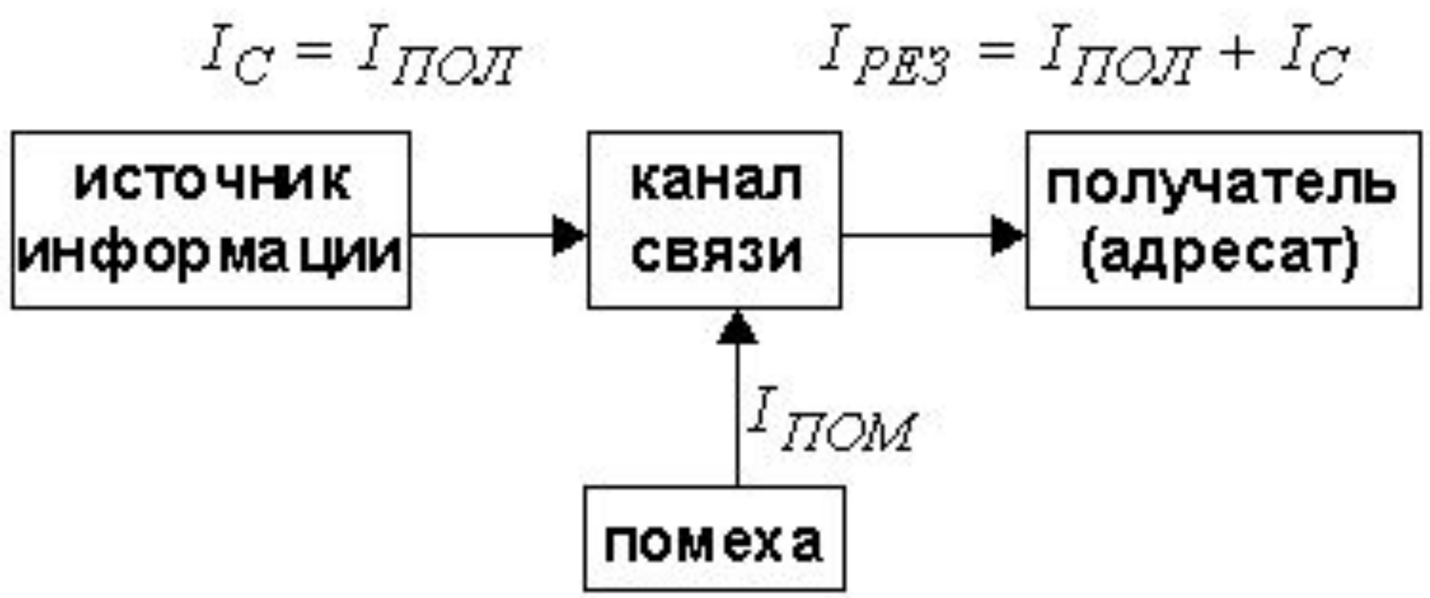
В этом случае $H_{\max} = \log_2 2 = 1$, графическая зависимость энтропии от вероятности будет иметь вид:

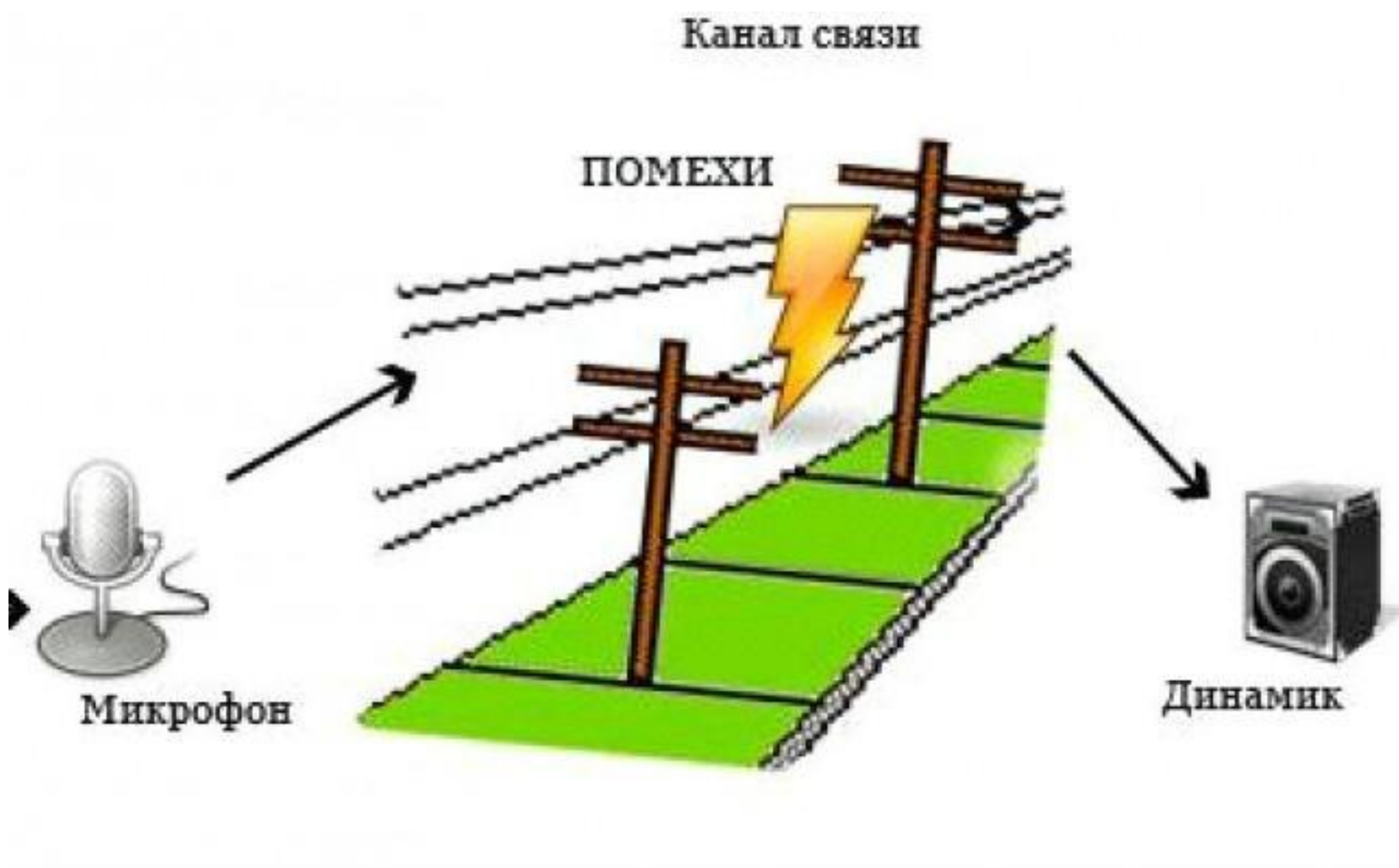


КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Условная энтропия

Процесс передачи информации носит случайный характер





$P\left(X/Y\right)$ вероятность появления события X при условии наступления события Y

Условная энтропия

Пусть при передаче n сообщений символ A появляется m раз, символ B появляется l раз, а символ A вместе с символом B – k раз, то вероятность появления символа A определяется соотношением $P(A)=\frac{m}{n}$, вероятность появления символа B соотношением $P(B)=\frac{l}{n}$, вероятность совместного появления символов A и B соотношением $P(AB)=\frac{k}{n}$

$$P\left(A/B\right)=\frac{P(AB)}{P(B)}=\frac{k/n}{l/n}=\frac{k}{l}$$

$$P\left(B/A\right)=\frac{P(AB)}{P(A)}=\frac{k/n}{m/n}=\frac{k}{m}$$

(2.10)

Из (2.10) $\Rightarrow$ $$P(AB)=P\left(A/B\right)P(B)=P\left(B/A\right)P(A) \quad \textbf{(2.11)}$$

Частная условная энтропия

Различают понятия **частной** и **общей условной энтропии**

Из (2.9) $\Rightarrow$

$$H\left(b_j / a_i\right) = -\sum_{j=1}^{n} P\left(b_j / a_i\right) \times \log\left(P\left(b_j / a_i\right)\right)$$

$$H\left(a_i / b_j\right) = -\sum_{i=1}^{n} P\left(a_i / b_j\right) \times \log\left(P\left(a_i / b_j\right)\right) \qquad \textbf{(2.12)}$$

i – характеристика состояния источника **А**

j – характеристика состояния источника **В**

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Общая условная энтропия

Общая условная энтропия сообщения **B** относительно сообщения **A** характеризует количество информации, содержащейся в любом символе алфавита, и определяется усреднением по всем символам, т.е. по всем состоянием с учетом вероятности появления каждого из состояний, и равна сумме вероятностей появления символов алфавита на неопределенность, которая остается после того как адресат принял сигнал.

$$H\left(B/A\right) = -\sum_{i=1}^{n} P(a_i) H\left(b_j/a_i\right) = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(a_i) P\left(b_j/a_i\right) \times \log\left(P\left(b_j/a_i\right)\right) =$$

$$= -\sum_{i=1}^{n} P(a_i) \sum_{j=1}^{n} P\left(b_j/a_i\right) \times \log\left(P\left(b_j/a_i\right)\right) \qquad \textbf{(2.13)}$$

(2.13) − общее выражение для определения количества информации на один символ сообщения для случая неравновероятных символов

Общая условная энтропия

со стороны источника

$$H\left(B/A\right) = -\sum_{i=1}^{n} P(a_i) \sum_{j=1}^{n} P\left(b_j/a_i\right) \times \log\left(P\left(b_j/a_i\right)\right)$$

со стороны приемника **(2.13')**

$$H\left(A/B\right) = -\sum_{j=1}^{n} P(b_j) \sum_{i=1}^{n} P\left(a_i/b_j\right) \times \log\left(P\left(a_i/b_j\right)\right)$$

Так как $P(a_i, b_j) = P(a_i) P\left(b_j/a_i\right) = P(b_j) P\left(a_i/b_j\right)$ то из (2.13) $\Rightarrow$

$$H\left(B/A\right) = -\sum_{i=1}^{n} \sum_{j=1}^{n} P(a_i, b_j) \log P\left(b_j/a_i\right)$$

(2.14)

$$H\left(A/B\right) = -\sum_{i=1}^{n} \sum_{j=1}^{n} P(a_i, b_j) \log P\left(a_i/b_j\right)$$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Каналы связи с шумами

Для идеальных условий

$$I = kH$$

Для равновероятных независимых сообщений

$$H_1 = \log_2 N = n \log_2 m$$

Для неравновероятных независимых сообщений

$$H_2(A) = -\sum_{i=1}^{m} p(a_i) \log p(a_i)$$

Со стороны источника

$$H_3\left(B/A\right) = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(a_i, b_j) \log P\left(b_j/a_i\right)$$

Со стороны приемника

$$H_4\left(A/B\right) = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(b_j, a_i) \log P\left(a_i/b_j\right)$$

$$H_5(A, B) = H_5(B, A) = -\sum_{i=1}^{n}\sum_{j=1}^{n} P(a_i, b_j) \log P(a_i, b_j)$$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Каналы связи с шумами

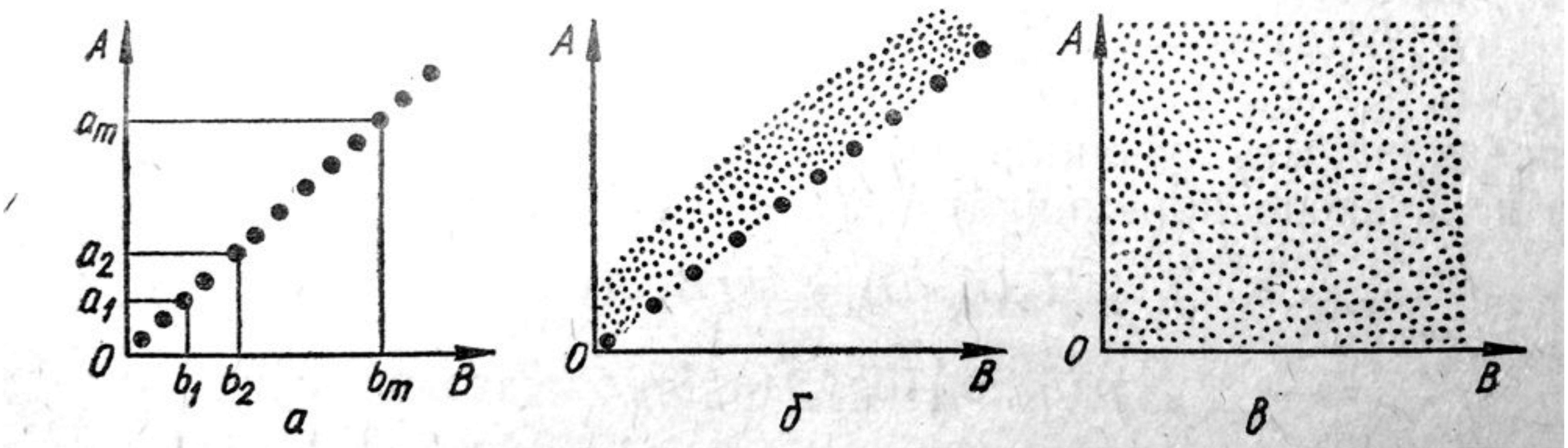


Рис. 21. Графическое представление различных уровней помех в каналах связи.

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Каналы связи с шумами

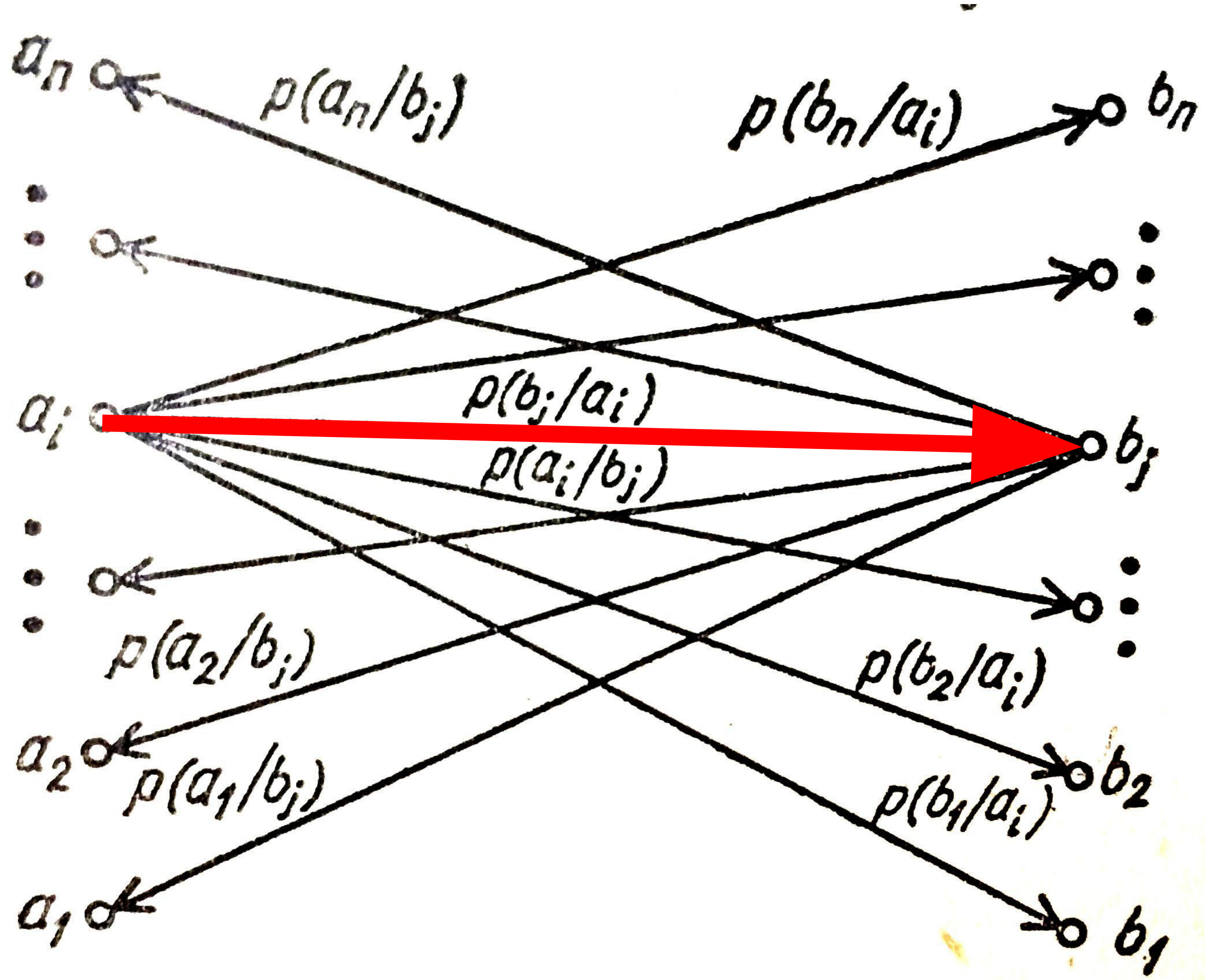


Рис. 12. Иллюстрация неопределенности получения сигнала b_j при передаче a_i.

Канальная матрица

Канальная матрица − матрица, статистически описывающая канал связи.

В общем случае, если передается n сигналов от источника A, и ожидается получение n сигналов на стороне получателя B, влияние помех в канале связи полностью описывается матрицей вида

$A \setminus B$	b_1	b_2	$\dots$	b_j	$\dots$	b_n
a_1	$p(b_1/a_1)$	$p(b_2/a_1)$	$\dots$	$p(b_j/a_1)$	$\dots$	$p(b_n/a_1)$
a_2	$p(b_1/a_2)$	$p(b_2/a_2)$	$\dots$	$p(b_j/a_2)$	$\dots$	$p(b_n/a_2)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_i	$p(b_1/a_i)$	$p(b_2/a_i)$	$\dots$	$p(b_j/a_i)$	$\dots$	$p(b_n/a_i)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_n	$p(b_1/a_n)$	$p(b_2/a_n)$	$\dots$	$p(b_j/a_n)$	$\dots$	$p(b_n/a_n)$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Канальная матрица (со стороны источника)

Канальная матрица – матрица, статистически описывающая канал связи.

В общем случае, если передается n сигналов от источника A, и ожидается получение n сигналов на стороне получателя B, влияние помех в канале связи полностью описывается матрицей вида

$A \setminus B$	b_1	b_2	…	b_j	…	b_n
a_1	$p(b_1/a_1)$	$p(b_2/a_1)$	…	$p(b_j/a_1)$	…	$p(b_n/a_1)$
a_2	$p(b_1/a_2)$	$p(b_2/a_2)$	…	$p(b_j/a_2)$	…	$p(b_n/a_2)$
…	…	…	…	…	…	…
a_i	$p(b_1/a_i)$	$p(b_2/a_i)$	…	$p(b_j/a_i)$	…	$p(b_n/a_i)$
…	…	…	…	…	…	…
a_n	$p(b_1/a_n)$	$p(b_2/a_n)$	…	$p(b_j/a_n)$	…	$p(b_n/a_n)$

$$p(b_1/a_1) + p(b_2/a_1) + \ldots + p(b_j/a_1) + \ldots + p(b_n/a_1) = 1 \qquad \textbf{(2.15)}$$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Канальная матрица (со стороны источника)

Потери информации, приходящиеся на долю сигнала a_i описываются при помощи *частной условной энтропии*.

$$H\left(B/a_1\right) = -\sum_{j=1}^{n} P\left(b_j/a_1\right)\log P\left(b_j/a_1\right) \qquad \textbf{(2.16)}$$

Общая условная энтропия

$$H\left(B/A\right) = -\frac{1}{n}\sum_{j=1}^{n}\sum_{i=1}^{n} P\left(b_j/a_i\right)\log P\left(b_j/a_i\right) \qquad \textbf{(2.17)}$$

$$H\left(B/A\right) = -\sum_{j=1}^{n}\sum_{i=1}^{n} P(a_i)P\left(b_j/a_i\right)\log P\left(b_j/a_i\right) \qquad \textbf{(2.18)}$$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Канальная матрица (со стороны приемника)

$A \setminus B$	b_1	b_2	$\dots$	b_j	$\dots$	b_n
a_1	$p(a_1/b_1)$	$p(a_1/b_2)$	$\dots$	$p(a_1/b_j)$	$\dots$	$p(a_1/b_n)$
a_2	$p(a_2/b_1)$	$p(a_2/b_2)$	$\dots$	$p(a_2/b_j)$	$\dots$	$p(a_2/b_n)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_i	$p(a_i/b_1)$	$p(a_i/b_2)$	$\dots$	$p(a_i/b_j)$	$\dots$	$p(a_i/b_n)$
$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$	$\dots$
a_n	$p(a_n/b_1)$	$p(a_n/b_2)$	$\dots$	$p(a_n/b_j)$	$\dots$	$p(a_n/b_n)$

$$p(a_1/b_j)+p(a_2/b_j)+\dots+p(a_i/b_j)+\dots+p(a_n/b_j)=1 \qquad \textbf{(2.19)}$$

Частная условная энтропия

$$H\left(A/b_j\right)=-\sum_{i=1}^{n}\left(a_i/b_j\right)\log P\left(a_i/b_j\right) \qquad \textbf{(2.20)}$$

Общая условная энтропия

$$H\left(A/B\right)=-\sum_{j=1}^{n}\sum_{i=1}^{n}P(b_j)P\left(a_i/b_j\right)\log P\left(a_i/b_j\right) \qquad \textbf{(2.21)}$$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Канальная матрица

Безусловные вероятности приемника

$$p(b_j) = \sum_{i=1}^{n} p(a_i) p(b_j / a_i) \quad \textbf{(2.22)}$$

Энтропия приемника

$$H(B) = -\sum_{j=1}^{n} p(b_j) \log p(b_j) \quad \textbf{(2.23)}$$

Безусловные вероятности источника

$$p(a_i) = \sum_{j=1}^{n} p(b_j) p(a_i / b_j) \quad \textbf{(2.24)}$$

Энтропия источника

$$H(A) = -\sum_{i=1}^{n} p(a_i) \log p(a_i) \quad \textbf{(2.25)}$$

Энтропия объединения

Энтропия объединения используется для вычисления энтропии совместного появления статистически зависимых сообщений либо энтропии взаимосвязанных систем.

$$H(A,B)=H(B,A)=-\sum_{i=1}^{n}\sum_{j=1}^{n}P(a_i,b_j)\log P(a_i,b_j)\left[\frac{\textit{бит}}{\textit{два символа}}\right] \qquad \textbf{(2.26)}$$

Связь с условной энтропией

$$H(A,B)=H(A)+H\left(B/A\right)=H(B)+H\left(A/B\right) \qquad \textbf{(2.27)}$$

$$p(a_i,b_j)=\begin{vmatrix} p(a_1,b_1) & p(a_1,b_2) & \cdots & p(a_1,b_n) \\ p(a_2,b_1) & p(a_2,b_2) & \cdots & p(a_2,b_n) \\ \vdots & \vdots & \vdots & \vdots \\ p(a_n,b_1) & p(a_n,b_2) & \cdots & p(a_n,b_n) \end{vmatrix} \qquad \begin{array}{l} \sum_{i=1}^{n} p(a_i,b_j)=p(b_j) \\ \sum_{j=1}^{n} p(a_i,b_j)=p(a_i) \end{array} \qquad \textbf{(2.28)}$$

КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

Энтропия объединения

$$H(A) = -\sum_{i=1}^{n} p(a_i)\log p(a_i) = -\sum_{i=1}^{n}\sum_{j=1}^{n} p(a_i, b_j)\log\sum_{j=1}^{n} p(a_i, b_j)$$

(2.29)

$$H(B) = -\sum_{j=1}^{n} p(b_j)\log p(b_j) = -\sum_{j=1}^{n}\sum_{i=1}^{n} p(a_i, b_j)\log\sum_{i=1}^{n} p(a_i, b_j)$$

$$p\left({a_i}/{b_j}\right) = \frac{p(a_i, b_j)}{p(b_j)} = \frac{p(a_i, b_j)}{\sum_{i=1}^{n} p(a_i, b_j)}$$

$$p\left({b_j}/{a_i}\right) = \frac{p(a_i, b_j)}{p(a_i)} = \frac{p(a_i, b_j)}{\sum_{j=1}^{n} p(a_i, b_j)}$$

(2.30)

$$I(A, B) = H(A) + H(B) - H(B, A)$$

$$I(A, B) = H(A) - H(A/B)$$

$$I(A, B) = H(B) - H(B/A)$$

(2.31)