А.Г.Дьячков

Теория информации

§5. Обратная теорема Шеннона для дискретного канала связи

Содержание

1. Свойства энтропии и информации:
   (a) $H(X \mid Y) \leq H(X) \leq \log \mid X \mid$.
   (b) $J(X;Y) = J(Y;X) = H(Y) - H(Y \mid X) = H(X) - H(X \mid Y) = H(X) + H(Y) - H(X;Y) \geq 0$.
   (c) $J(X;Y|Z) = H(Y|Z) - H(Y|X,Z) \geq 0$.
   (d) $J(X;Y_1,Y_2) = J(X;Y_1) + J(X;Y_2|Y_1)$, $J(X;Y_1^N) = \sum_{n=1}^N J(X;Y_n|Y_1^{n-1})$.
   (e) $J(X;f(Y)) \leq J(X;Y)$.
   (f) Если $p(y|x,z) = p(y|x)$, то $J(Y;X|Z) \leq J(Y;X)$. Если $p(x|z) = p(x)$, то $J(Y;X|Z) \geq J(Y;X)$.
   (g) *Неравенство Фано*: если $\Pr\{X \neq Y\} = \epsilon$, то $H(X|Y) \leq \epsilon \log |X| + h(\epsilon)$.
   (h) $H(Y_1^N) \leq \sum_{n=1}^N H(Y_n)$.
   (i) Если $p(y_1^N|x_1^N) = \prod_{n=1}^N p(y_n|x_n)$, то
   $$H(Y_1^N|X_1^N) = \sum_{n=1}^N H(Y_n|X_n);\ J(X_1^N;Y_1^N) \leq \sum_{n=1}^N J(X_n;Y_n).$$
2. Теорема переработки данных для дискретного канала связи.
3. Пропускная способность дискретного канала без памяти (ДКБП).
4. Вычисление пропускной способности ДСК $C = \log 2 - h(p)$.
5. Нижняя граница ϵ-энтропии сообщения с равновероятными значениями.
6. Обратная теорема Шеннона для ДКБП.
7. Обратная теорема Шеннона для КРШ.

8. Математическая модель передачи дискретного сообщения по ДКБП с обратной связью (ОС).

9. Теорема переработки данных для ДКБП с ОС.

10. Пропускная способность ДКБП с ОС.

5.1 Математическая модель передачи сообщения по каналу связи

5.1.1 Обозначения

Мы будем придерживаться следующей системы обозначений. Буквы $X, Y, Z, \ldots$ — случайные величины, а значения этих величин — буквы $x, y, z, \ldots$ соответственно. Распределение вероятностей и условные распределения будем, для краткости, обозначать символами $p(x), p(x|z), \ldots$, под которыми подразумевается более подробная запись, а именно:

$$p(x) = p_X(x) := \Pr\{X = x\}\,; \quad p(x|z) = p_{X|Z}(x|z) := \Pr\{X = x|Z = z\}.$$

Для последовательностей длины N из случайных величин и их значений используем стандартные символы $X_1^N = (X_1, X_2, \ldots, X_N)$, $x_1^N = (x_1, x_2, \ldots, x_N)$. Например,

$$p(y_1^N|x_1^N) := \Pr\{Y_1^N = y_1^N|X_1^N = x_1^N\}$$

обозначает условное распределение случайной величины Y_1^N при условии X_1^N.

5.1.2 Схема передачи: $\xi \to X_1^N \to Y_1^N \to \tilde{\xi}$

$X_1^N = (X_1, X_2, \ldots, X_N)$ — сигнал на входе канала, при $n = \overline{1, N}$, $X_n = \overline{0, K-1}$, $K \geq 2$,

$Y_1^N = (Y_1, Y_2, \ldots, Y_N)$ — сигнал на выходе канала, при $n = \overline{1, N}$, $Y_n = \overline{0, J-1}$, $J \geq 2$,

$N = 1, 2, \ldots$ — время передачи по каналу.

$\xi \in [M] = \{1, 2, \ldots, M\}$ — сообщение источника.

$\tilde{\xi} \in [\tilde{M}] = \{1, 2, \ldots, \tilde{M}\}$, $\tilde{M} \geq M$, — сообщение адресата.

5.1.3 Метод передачи

Возможность передать сообщение ξ, по каналу связи за время N с вероятностью ошибки $\leq \epsilon$ означает существование совместного распределения $p(u,x_1^N, y_1^N,\tilde{u})$ четверки случайных величин $(\xi, X_1^N, Y_1^N, \tilde{\xi})$, которое удовлетворяет следующим условиям (ограничениям):

У1). Задано (фиксированно) распределение $p(u)$ *источника* ξ, $u \in [M]$, а совместное распределение $p(u, \tilde{u})$ пары $(\xi, \tilde{\xi})$ удовлетворяет условию $\Pr\{\xi \neq \tilde{\xi}\} \leq \epsilon$. Указанному условию фактически удовлетворяет условное распределение $p(\tilde{u}|u)$, $u \in [M]$, $\tilde{u} \in [\tilde{M}]$.

У2). В совместном распределении $p(x_1^N; y_1^N)$ пары $(X_1^N; Y_1^N)$, описывающем *дискретный канал связи*, задано условное распределение, т. е. $(K^N \times J^N)$ — матрица

$$P_N := \|p(y_1^N | x_1^N)\| \,.$$

Это означает, что все допустимые совместные распределения пары $(X_1^N; Y_1^N)$ однозначно задаются входными распределениями $p(x_1^N)$.

По определению *отсутствие памяти* в канале связи означает, что для *дискретного канала без памяти (ДКБП)* условное распределение P_N вычисляется с помощью $(K \times J)$–матрицы $\mathbf{P} = P_1 = \|p(y|x)\|$, $x = \overline{0, K-1}$, $y = \overline{0, J-1}$, следующим образом:

$$p(y_1^N | x_1^N) = \prod_{n=1}^{N} p(y_n | x_n) \,.$$

Пример ДКБП. Рассмотрим двоичный симметричный канал (ДСК), в котором $K = J = 2$, а матрица переходных вероятностей имеет вид

$$\mathbf{P} = \begin{pmatrix} 1-q & q \\ q & 1-q \end{pmatrix}, q = p(0|1) = p(1|0), 1-q = p(0|0) = p(1|1),$$

где $q, 0 < q < 1/2,$ — параметр, называемый *вероятностью ошибки* при одной передаче. Условное распределение для ДСК без памяти имеет вид

$$p(y_1^N | x_1^N) = q^{\rho(x_1^N; y_1^N)} (1-q)^{N - \rho(x_1^N; y_1^N)},$$

где $\rho(x_1^N; y_1^N)$ — число несовпадающих позиций (расстояние Хэмминга) в двоичных последовательностях x_1^N и y_1^N.

Нижеприводимое условие **У3)** описывает естественные (с точки зрения здравого смысла) соотношения связи случайных величин в схеме передачи $\xi \to X_1^N \to Y_1^N \to \tilde{\xi}$. Отметим, что свойство цепи Маркова говорит о том, что сообщение источника ξ *не зависит* (имеет другую физическую природу) *от шума* в канале связи.

У3). Задано *кодирование* $f = f(u)$, $u \in [M]$, и *декодирование* $g = g(y_1^N) \in [\tilde{M}]$, так что

$$X_1^N = f(\xi); \tilde{\xi} = g(Y_1^N)$$

и, кроме того, тройка (ξ, X_1^N, Y_1^N) является *цепью Маркова*, т.е. совместное распределение

$$p(y_1^N | x_1^N, u) = p(y_1^N | x_1^N).$$

Определение 1. Совместное распределение $p(u, x_1^N, y_1^N, \tilde{u})$, удовлетворяющее условиям **У1)— У3)** называется *методом передачи* сообщения ξ по дискретному каналу связи за время N с вероятностью ошибки $\leq \epsilon$.

5.1.4 Информационные характеристики источника сообщений и канала связи

Определение 2. Число

$$J(X;Y) := \sum_{xy} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} = \sum_{xy} p(x,y) \log \frac{p(y|x)}{p(y)}$$

называется *информацией* пары случайных величин $(X;Y)$, а число

$$J(X;Y|Z) := \sum_{x,y,z} p(x,y,z) \log \frac{p(x,y|z)}{p(x|z)p(y|z)} = \sum_{x,y,z} p(x,y,z) \log \frac{p(y|x,z)}{p(y|z)}$$

называется *условной информацией* пары величин $(X;Y)$, при условии Z.

Свойство 0. (*Симметрия и неотрицательность*).

$$J(X;Y) = J(Y;X) \geq 0; J(X;Y|Z) = J(Y;X|Z) \geq 0\,.$$

При этом информация $J(X;Y) = 0$ тогда и только тогда, когда величины X и Y — независимы, т.е. $p(y|x) = p(y)$, а условная информация $J(X;Y|Z) = 0$ тогда и только тогда, когда тройка $(X;Z;Y)$ — цепь Маркова, т.е. $p(y|zx) = p(y|z)$.

Доказательство. Пусть $Q_1(\omega)$ и $Q_2(\omega)$, $\omega \in \Omega$, - два произвольных распределения вероятностей на множестве Ω . Применяя неравенство $lnx \leq x - 1$ где равенство лишь при $x = 1$, имеем

$$\begin{aligned} \textstyle\sum_{\omega\in\Omega} Q_1(\omega) \ln \frac{Q_2(\omega)}{Q_1(\omega)} \leq \sum_{\omega} Q_1(\omega) \left[\frac{Q_2(\omega)}{Q_1(\omega)} - 1\right] &= \\ = \textstyle\sum_{\omega} Q_2(\omega) - \sum_{\omega} Q_1(\omega) = 1 - 1 &= 0. \end{aligned}$$

Следовательно, для любых двух распределений вероятностей $Q_i(\omega)$, $i = \overline{1,2}$,

$$\sum_{\omega\in\Omega} Q_1(\omega)\, \log \frac{Q_1(\omega)}{Q_2(\omega)} \geq 0\,, \qquad (*)$$

где знак равенства в $(*)$ тогда и только тогда, когда $Q_1(\omega) = Q_2(\omega)$ при всех $\omega \in \Omega$.

Неравенство $J(X;Y) \geq 0$ вытекает из $(*)$ если положить $\omega = (x,y)$; $Q_1(\omega) = p(x,y)$; $Q_2(\omega) = p(x)p(y)$. Неравенство $J(X;Y|Z) \geq$

0 вытекает из $(*)$ если при фиксированном z положить $Q_1(\omega) = p(x,y|z);\ Q_2(\omega) = p(x|z)p(y|z)$. Свойство симметрии очевидно.

Свойство 0 доказано.

Определение 3. Число

$$H_\xi(\epsilon) := \inf_{\mathbf{1)}} J(\xi, \tilde{\xi}),$$

где нижняя грань берется по всем условным распределениям $p(\tilde{u}|u)$, удовлетворяющим условию **У1)**, называется *ϵ-энтропией источника (сообщения) ξ*.

Определение 4. Число

$$C_N = C_N(P_N) := \sup_{\mathbf{2)}} J(X_1^N; Y_1^N),$$

где верхняя грань берется по всем распределениям $p(x_1^N; y_1^N)$, удовлетворяющим условию **У2)**, называется *пропускной способностью канала за время N*.

5.1.5 Формулировка обратной теоремы Шеннона

В разделе **5.2** будут исследованы свойства информации, условной информации, а также связанного с ними обобщения ранее введенного понятия энтропии. С помощью данных свойств в этом же разделе **5.2** будут доказаны теоремы $1-3$.

Теорема 1. (*Теорема переработки данных*). *Если существует метод передачи сообщения ξ по каналу P_N с вероятностью ошибки $\leq \epsilon$, то необходимо, чтобы $J(\xi;\tilde{\xi}) \leq J(X_1^N; Y_1^N)$, т.е. необходимо выполнение неравенства*

$$H_\xi(\epsilon) \leq C_N(P_N)\,.$$

Теорема 2. *Для дискретного канала без памяти (ДКБП) пропускная способность*

$$C_N = C_N(P_N) = N\,C_1\,,$$

где $C_1 = C_1(P_1)$ — пропускная способность на единицу времени.

Теорема 3. *Если передаваемое сообщение ξ имеет равномерное распределение , т.е. $p(u) = M^{-1}$ при $u \in [M]$ то*

$$H_\xi(\epsilon) \geq (1-\epsilon)\,\log M - \log 2\,.$$

Для ДКБП из теорем $1-3$ следует

Теорема 4 (*Нижняя граница вероятности ошибки*). *Пусть сообщение* ξ *имеет равномерное распределение на* $[M]$ *а* $M = \lceil \exp\{RN\} \rceil$, *где* $R > 0$ *- фиксированный параметр называемый скоростью передачи. Пусть существует метод передачи по ДКБП с пропускной способностью* C_1 *и с вероятностью ошибки* $\leq \epsilon = \epsilon_N$. *Тогда необходимо выполнение неравенства*

$$\epsilon = \epsilon_N \geq 1 - \frac{C_1 + \log 2/N}{R} \,.$$

Из теоремы 4 вытекает важное

Следствие. *Если скорость передачи больше пропускной способности ДКБП, т.е.* $R > C_1$, *то при* $N \to \infty$ *вероятность ошибки любого метода передачи* $\epsilon_N \not\to 0$.

Это следствие называется *обратной теоремой Шеннона*

Доказательство теоремы 4. Из теорем $1-3$ следует, что параметры метода передачи (M, ϵ, C_1, N) удовлетворяют неравенству $(1-\epsilon)\log M - \log 2 \leq NC_1$, которое в силу определения скорости R означает

$$(1-\epsilon)RN - \log 2 \leq NC_1 \,,$$

что равносильно неравенству в теореме 4.

5.2 Свойства информации и энтропии

5.2.1 Условная энтропия и условная информация

Введем следующие понятия:

$$\begin{array}{ll}
H(X) := -\sum_x p(x)\log p(x) & - \text{ энтропия,} \\
H(X;Y) := -\sum_{xy} p(x;y)\log p(x;y) & - \text{ совместная энтропия,} \\
H(Y|X) := -\sum_{xy} p(x;y)\log p(y|x) & - \text{ условная энтропия.}
\end{array}$$

С помощью этих величин можно записать *информацию*:

$$\begin{aligned}
J(X;Y) \;&= H(Y) - H(Y|X) = H(X) - H(X|Y) = \\
&= H(X) + H(Y) - H(X;Y) = \\
&= \textstyle\sum_{xy} p(x,y)\log\frac{p(y|x)}{p(y)} = \sum_{xy} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \,,
\end{aligned} \tag{1}$$

а также *условную информацию*

$$\begin{aligned}
J(X;Y|Z) \;&= H(Y|Z) - H(Y|X,Z) = \\
&= H(X|Z) - H(X|Y,Z) = \textstyle\sum p(x,y,z)\log\frac{p(x|y,z)}{p(x|z)} = \\
&= \textstyle\sum_{xyz} p(x,y,z)\log\frac{p(x,y|z)}{p(x|z)p(y|z)} \,.
\end{aligned} \tag{2}$$

Пусть $|X|$ обозначает число возможных значений величины X. Очевидно, что если X имеет равномерное распределение, т.е. для любого возможного значения x вероятность $\Pr\{X = x\} = 1/|X|$, то энтропия $H(X) = \log|X|$. Имеет место

Свойство 1. (*Экстремальность равномерного распределения*). *Энтропия* $H(X) \leq \log|X|$, *где знак равенства тогда и только тогда, когда* X *имеет равномерное распределение.*

Доказательство. Данное свойство вытекает из неравенства (*) (см. вывод свойства 0), в котором надо положить: $\omega = x$, $Q_1(\omega) = p(x)$, $Q_2(w) = 1/|X|$.

Очевидно (легко проверяется) следующее важное

Свойство 2. *Имеет место формула условной информации*

$$J(X;Y_1,Y_2) = J(X;Y_1) + J(X;Y_2|Y_1)\,, \tag{3}$$

а также обобщенная формула условной информации

$$J(X;Y_1,Y_2|Z) = J(X;Y_1|Z) + J(X;Y_2|Y_1,Z)\,. \tag{4}$$

Из свойства 2 вытекает

Свойство 3. (*"Цепное"правило*).

$$J(X;Y_1^N) = \sum_{n=1}^{N} J(X;Y_n|Y_1^{n-1})\,. \tag{5}$$

5.2.2 Информационные неравенства

Пусть $\tilde{y} = f(y)$ — произвольная функция. Тогда $\tilde{Y} = f(Y)$ — случайная величина и справедливо

Свойство 4. (*Монотонность информации*).

$$J(X;f(Y)) \leq J(X;Y)\,, \tag{6}$$

где равенство тогда и только тогда, когда $(Y;f(Y);X)$ *— цепь Маркова.*

Доказательство. Поскольку распределение пары $(Y;f(Y))$ однозначно задается распределением Y, то $J(X;Y) = J(X;Y,f(Y))$, а следовательно из (3) имеем

$$J(X;Y) = J(X;f(Y)) + J(X;Y|f(Y)) \geq J(X;f(Y))\,.$$

Это неравенство обобщается и на условную информацию, т.е. из (4) вытекает

Свойство 5. (*Монотонность условной информации*).

$$J(X; f(Y)|Z) \le J(X; Y|Z)\,, \tag{7}$$

где равенство тогда и только тогда, когда тройка $(X; (Z, f(Y)); Y)$ *— цепь Маркова.*

Доказательство теоремы 1. Применяя свойство 4, можем написать

$$J(\xi; \tilde{\xi}) \le J(\xi, X_1^N; \tilde{\xi}, Y_1^N) = J(\xi, X_1^N; Y_1^N),$$

где второе равенство вытекает из того, что $\tilde{\xi} = g(Y_1^N)$. Далее, по свойству 2 имеем

$$J(Y_1^N; \xi, X_1^N) = J(X_1^N; Y_1^N) + J(\xi; Y_1^N|X_1^N).$$

Согласно требованию **У3)**, тройка (ξ, X_1^N, Y_1^N) — цепь Маркова. Следовательно, по свойству 0 условная информация $J(\xi; Y_1^N|X_1^N) = 0$. Поэтому для любого совместного распределения $p(u, x_1^N, y_1^N, \tilde{u})$, удовлетворяющего условиям **У1)—У3)**, справедливо неравенство переработки данных

$$J(\xi; \tilde{\xi}) \le J(X_1^N; Y_1^N)\,,$$

из которого очевидным образом следует неравенство между ϵ-энтропией и пропускной способностью, сформулированное в теореме 1.

Из неотрицательности информации и условной информации вытекает

Свойство 6. (*Монотонность энтропии*).

$$H(Y|X) \le H(Y)\,, \quad H(Y|XZ) \le H(Y|X)\,, \tag{8}$$

где равенство тогда и только тогда, когда $p(y|x) = p(y)$ *либо* $p(y|xz) = p(y|x)$*, соответственно.*

Условия для сравнения информации и условной информации дают свойства 7 и 8.

Свойство 7. *Если тройка* (Z, X, Y) *— цепь Маркова, т.е.* $p(y|xz) = p(y|x)$*, то*

$$J(X; Y|Z) \le J(X; Y)\,. \tag{9}$$

Доказательство. Имеем

$$J(X; Y|Z) = H(Y|Z) - H(Y|XZ) \le H(Y) - H(Y|X) = J(X; Y)\,,$$

где применили определение (2), свойство 6, а затем определение (1).

Свойство 8. *Если в тройке* (Z, X, Y) *величины* (X, Z) *— независимы, т.е.* $p(x|z) = p(x)$, *то*

$$J(X;Y|Z) \geq J(X;Y)\,. \tag{10}$$

Доказательство. Имеем

$$J(X;Y|Z) = H(X|Z) - H(X|YZ) \geq H(X) - H(X|Y) = J(X;Y)\,,$$

где применили свойство 6 и равенство $H(X|Z) = H(X)$, вытекающее из независимости X и Z.

Из неотрицательности информации (свойство 0) и представления (1) вытекает, что $H(X;Y) \leq H(X) + H(Y)$, т.е. справедливо

Свойство 9. (*Субаддитивность энтропии*). *Совместная энтропия*

$$H(Y_1^N) \leq \sum_{n=1}^{N} H(Y_n)\,,$$

где знак равенства тогда и только тогда, когда Y_n, $n = \overline{1, N}$, *независимы.*

Свойство 10. *Если* X_1^N *и* Y_1^N *— входной и выходной сигналы ДКБП, то условная энтропия*

$$H(Y_1^N|X_1^N) = \sum_{n=1}^{N} H(Y_n|X_n)$$

Доказательство. Определение условного распределения $p(y_1^N|x_1^N)$ для ДКБП из условия **У2)** означает, что

$$\begin{aligned} H(Y_1^N|X_1^N) \quad &= -\textstyle\sum_{x_1^N\, y_1^N} p(x_1^N, y_1^N) \log \prod_{n=1}^{N} p(y_n|x_n) = \\ &= -\textstyle\sum_{x_1^N\, y_1^N} p(x_1^N, y_1^N) \sum_{n=1}^{N} \log p(y_n|x_n) = \\ &= -\textstyle\sum_{n=1}^{N}\ \sum_{x_1^N\, y_1^N} p(x_1^N, y_1^N) \log p(y_n|x_n) = \\ &= -\textstyle\sum_{n=1}^{N}\ \sum_{x_n\, y_n} p(x_n, y_n) \log p(y_n|x_n) = \sum_{n=1}^{N} H(y_n|x_n)\,. \end{aligned}$$

Доказательство теоремы 2. Используя представление (1) информации через энтропию, имеем

$$J(X_1^N;Y_1^N) = H(Y_N) - H(Y_N|X_N)\,.$$

Поэтому из свойств 9 и 10 вытекает, что для ДКБП

$$J(X_1^N;Y_1^N) \leq \sum_{n=1}^{N} [H(Y_n) - H(Y_n|X_n)] = \sum_{n=1}^{N} J(X_n, Y_n),$$

где знак равенства тогда и только тогда, когда величины Y_n, $n = \overline{1,N}$ – независимы. Следовательно, определение пропускной способности ДК-БП на N единиц времени означает, что

$$C_N = C_N(P_N) \le N\, C_1\,, \tag{11}$$

где C_1 — пропускная способность на единицу времени, т.е.

$$C_1 = C_1(\mathbf{P}) = \max_{p(x)} J(X;Y)\,. \tag{12}$$

Знак равенства в (11) вытекает из того, что в качестве экстремального распределения $\hat{p}(x_1^N)$, на котором достигается максимум в определении 4 величины C_N, можно взять

$$\hat{p}(x_1^N) := \prod_{n=1}^{N} \hat{p}(x_n)\,,$$

где $\hat{p}(x)$, $x = \overline{0, K-1}$, — экстремальное распределение, на котором достигается максимум в (12). Теорема 2 доказана.

Задача. Доказать, что пропускная способность C_1 ДСК без памяти с вероятностью ошибки q достигается на распределении $\hat{p}(0) = \hat{p}(1) = 1/2$ и значение

$$C_1 = \log 2 - h(q)\,,$$

где двоичная энтропия $h(q) = -q\log q - (1-q)\log(1-q)$.

Свойство 11. (*Неравенство Фано*). *Для произвольной пары случайных величин* $(\xi, \tilde{\xi})$, $\xi \in [M]$, $\tilde{\xi} \in [\tilde{M}]$, $\tilde{M} \ge M$, *такой что* $\Pr\{\xi \ne \tilde{\xi}\} = \epsilon$, *условная энтропия*

$$H(\xi|\tilde{\xi}) \le \epsilon \log M + h(\epsilon)\,.$$

Доказательство неравенства Фано. Пусть $p(u,\tilde{u})$ — совместное распределение пары $(\xi,\tilde{\xi})$, т.е.

$$p(u,\tilde{u}) = \Pr\{\xi = u; \tilde{\xi} = \tilde{u}\}\,, \quad u \in [M]\,, \quad \tilde{u} \in [\tilde{M}].$$

Тогда из определения числа ϵ следует, что

$$\epsilon = \sum_{\tilde{u}=1}^{\tilde{M}} \sum_{u \ne \tilde{u}} p(u,\tilde{u})\,,\; 1-\epsilon = \sum_{u=1}^{M} p(u,u),$$

а по определению условной энтропии имеем

$$H(\xi|\tilde{\xi}) = \sum_{\tilde{u}=1}^{\tilde{M}} \sum_{u \ne \tilde{u}} p(u,\tilde{u}) \log\left(\frac{1}{p(u|\tilde{u})}\right) + \sum_{u=1}^{M} p(u,u)\log\left(\frac{1}{p(u|u)}\right)\,.$$

Из этих соотношений вытекает, что разность $\Delta = H(\xi|\tilde{\xi}) - h(\epsilon) - \epsilon \log M = \Delta_1 + \Delta_2$, где

$$\Delta_1 = \sum_{\tilde{u}=1}^{\tilde{M}} \quad \sum_{u \neq \tilde{u}} p(u,\tilde{u}) \log \left[\frac{\epsilon}{p(u|\tilde{u})M} \right], \quad \Delta_2 = \sum_{u=1}^{M} p(u,u) \log \left[\frac{1-\epsilon}{p(u|u)} \right].$$

Используя для оценки неравенство $\log x \leq (x-1)\log e$, имеем

$$\begin{array}{ll} \Delta_1 & \leq \sum_{\tilde{u}} \ \sum_{u \neq \tilde{u}} p(u,\tilde{u}) \left[\frac{\epsilon}{Mp(u|\tilde{u})} - 1 \right] \log e = \left(\sum_{\tilde{u}} \ \sum_{u \neq \tilde{u}} p(\tilde{u}) \frac{\epsilon}{M} - \epsilon \right) \log e \leq (\epsilon \sum_{\tilde{u}} p(\tilde{u}) - \epsilon) \log e = \\ \Delta_2 & \leq \sum_u \ p(u,u) \left[\frac{1-\epsilon}{p(u|u)} - 1 \right] = \left[\sum_u p(u)(1-\epsilon) - (1-\epsilon) \right] \log e = 0\,. \end{array}$$

Следовательно, $\Delta \leq 0$, что равносильно доказываемому неравенству. Свойство 11 доказано.

Доказательство теоремы 3. Пусть $\Pr\{\xi \neq \tilde{\xi}\} = \epsilon$. Поскольку $p(u) = 1/M,\ u \in [M]$, то $H(\xi) = \log M$. Следовательно, информация

$$J(\xi;\tilde{\xi}) = H(\xi) - H(\xi|\tilde{\xi}) = \log M - H(\xi|\tilde{\xi}))\,.$$

Поэтому из неравенства Фано следует, что

$$J(\xi;\tilde{\xi}) \geq \log M - h(\epsilon) - \epsilon \log M \geq (1-\epsilon)\log M - \log 2\,,$$

где мы учли, что двоичная энтропия $h(\epsilon) = -\epsilon \log \epsilon - (1-\epsilon)\log(1-\epsilon) \leq \log 2$. Так, как функция $(1-\epsilon)$ убывает с ростом ϵ, то неравенство

$$J(\xi;\tilde{\xi}) \geq (1-\epsilon)\log M - \log 2\,,$$

верно и для любой пары $(\xi,\tilde{\xi})$, такой, что $\Pr(\xi \neq \tilde{\xi}) \leq \epsilon$. Отсюда, применяя определение ϵ-энтропии, имеем

$$H_\xi(\epsilon) = \inf_{\text{У1}} J(\xi;\tilde{\xi}) \geq (1-\epsilon)\log M - \log 2,$$

если сообщение ξ имеет равномерное распределение $p(u) = 1/M$, $u = \overline{1,M}$.

Теорема 3 доказана.

5.3 Обратная теорема Шеннона для двоичного канала с равновесным шумом (КРШ)

Для двоичного канала с равновесным шумом (КРШ), введенного в §2, условное распределение

$$P(y_1^N|x_1^N) = \begin{cases} \left(C_N^t\right)^{-1}, & \text{если } \rho(x_1^N;y_1^N) = t, \\ 0, & \text{если } \rho(x_1^N;y_1^N) \neq t, 1 \leq t < N, \end{cases}$$

где $\rho(x_1^N; y_1^N)$ — число несовпадающих позиций (расстояние Хэмминга) в двоичных последовательностях y_N и x_N. Данный канал является примером *канала с памятью* и очевидно, что условная энтропия КРШ $H(Y_1^N|X_1^N) = \log C_N^t$. Поэтому информация

$$J(X_1^N; Y_1^N) = H(Y_1^N) - \log C_N^t \le N \log 2 - \log C_N^t,$$

где неравенство вытекает из свойства 1. Легко понять, что знак равенства в этом неравенстве достигается на равномерном входном распределении, т.е. когда

$$p(x_1^N) = \Pr\{X_1^N = x_1^N\} = 2^{-N}, x_1^N \in (0,1)^N.$$

Таким образом, пропускная способность КРШ за время N

$$C_N = N \log 2 - \log C_N^t$$

и теоремы 1 и 3 означают, что вероятность ошибки $\epsilon = \epsilon_N$ при передаче по КРШ сообщения ξ, принимающего M равновероятных значений, удовлетворяет неравенству

$$(1-\epsilon)\log M - \log 2 \le N \log 2 - \log C_N^t\,.$$

Данное неравенство, с учетом определения скорости передачи $R > 0$, можно переписать в виде

$$\epsilon = \epsilon_N \ge 1 - \frac{\log 2 - \frac{\log C_N^t}{N} + \frac{\log 2}{N}}{R}\,.$$

Пусть $R > 0$, $0 < p \le 1/2$, — фиксированные числа. Рассмотрим асимптотические условия

$$N \to \infty\,, \quad M = \lceil \exp\{RN\} \rceil, t = \lceil N\rho \rceil\,.$$

Используя известную (см. §2) логарифмическую асимптотику биномиальных коэффициентов, из границы для вероятности ошибки ϵ_N получаем

$$\underline{\lim}_{N\to\infty}\epsilon_N \ge 1 - \frac{C_1}{R}\,,\ C_1 = \log 2 - h(p)\,.$$

Из полученного неравенства вытекает

Обратная теорема Шеннона для КРШ. *Если* $N \to \infty$, *то при* $R > C_1 = \log 2 - h(p)$ *вероятность ошибки любого метода передачи по* КРШ $\epsilon_N \not\to 0$.

Напомним, что в §2 была доказана прямая теорема Шеннона для КРШ, которая утверждала, что при $R < C_1$ существует метод передачи с вероятностью ошибки $\epsilon_N \to 0$ при $N \to \infty$. Число

$$C_1 = \log 2 - h(p) = \lim_{N\to\infty} \frac{C_N}{N}$$

называется *асимптотической пропускной способностью КРШ на единицу времени.*

5.4 Пропускная способность дискретного канала без памяти (ДКБП) при наличии обратной связи

Схема передачи дискретного сообщения ξ по ДКБП имеет вид

$$\xi \to X_1^N \to Y_1^N \to \tilde{\xi} = g(Y_1^N)\,, \tag{13}$$

где

ξ—передаваемое сообщение (сообщение на входе), $\xi \in [M]$,

$X_1^N = (X_1, X_2, \ldots, X_N)$—входной сигнал, $X_i = \overline{0, K-1}$,

$Y_1^N = (Y_1, Y_2, \ldots, Y_N)$—выходной сигнал, $Y_i = \overline{0, J-1}$,

$N = 1, 2\ldots$—время передачи по ДКБП,

$\tilde{\xi}$—принятое сообщение (сообщение на выхо де), $\tilde{\xi} \in [\tilde{M}]$.

Определение ДКБП означает, что совместное распределение $p(u, x_1^N, y_1^N, \tilde{u})$ величин из (13), называемое *методом передачи*, удовлетворяет условию отсутствия памяти между компонентами пары (X_1^N, Y_1^N):

$$p(y_1^N|x_1^N) = \prod_{i=1}^{N} p(y_i|x_i)\,, \tag{14}$$

где $\mathbf{P} = \|p(y|x)\|$, $x = \overline{0, K-1}$, $y = \overline{0, J-1}$, — матрица переходных вероятностей ДКБП.

Отсутствие обратной связи означает, что кроме (14) выполнены условия: $X_1^N = F_N(\xi)\,;\quad (\xi \to X_1^N \to Y_1^N)$— цепь Маркова, т.е.

$$p(y_1^N|x_1^N, u) = p(y_1^N|x_1^N) = \prod_{i=1}^{N} p(y_i|x_i)\,. \tag{15}$$

По определению, *наличие полной обратной связи* означает, что условие (15) заменяется на следующее условие: при любом $n = \overline{1,N}$ $X_n = f_n(\xi; Y_1^{n-1})$; $\quad (\xi, Y_1^{n-1} \to X_n \to Y_n)$—цепь Маркова,т.е.

$$p(y_n|x_n; y_1^{n-1}, u) = p(y_n|x_n)\,. \tag{16}$$

Лемма 1. *Из условия* (15)*следует* (16), *т.е. метод передачи без обратной связи является частным случаем метода передачи с обратной связью.*

Доказательство. Зафиксируем произвольное $n = \overline{1,N}$ и просуммируем (15) сначала по y_{n+1}^N, а затем по y_n^N. В результате получим два равенства

$$p(y_1^n|x_1^N, u) = \prod_{i=1}^{n} p(y_i|x_i)\,; \quad p(y_1^{n-1}|x_1^N, u) = \prod_{i=1}^{n-1} p(y_i|x_i)\,.$$

Поделив первое равенство на второе и применив определение условной вероятности, имеем

$$\frac{p(y_1^n, x_1^N, u)}{p(y_1^{n-1}, x_1^N, u)} = p(y_n|x_n).$$

Следовательно, для любого фиксированного $n = \overline{1,N}$ справедливо равенство

$$p(y_n; y_1^{n-1}; x_1^N; u) = p(y_n|x_n)\, p(y_1^{n-1}; x_1^N, u).$$

Суммирование последних равенств по x_1^{n-1}, а затем — по x_{n+1}^N дает

$$p(y_n, y_1^{n-1}, x_n, u) = p(y_n|x_n) p(y_1^{n-1}, x_n, u)\,,$$

что равносильно (16).

Аналогично доказывается

Лемма 2. *Из условия* (14) *следует, что в ДКБП тройка* $(Y_1^{n-1} \to X_n \to Y_n)$ — *цепь Маркова, т.е.*

$$P(y_n|x_n, y_n^{n-1}) = p(y_n|x_n)\,.$$

Поэтому, в силу свойства 7, *условная информация*

$$J(X_n; Y_n|Y_1^{n-1}) \leq J(X_n; Y_n)\,.$$

Имеет место

Теорема 5. *Для ДКБП случайные величины в схеме* (13), (14) и (16) *передачи сообщения* ξ *с полной обратной связью удовлетворяют соотношению*

$$J(\xi;\tilde{\xi}) \le \sum_{n=1}^{N} J(X_n;Y_n)\,,$$

называемому неравенством переработки данных при наличии обратной связи.

Из теоремы 5 вытекает, что вероятность ошибки ϵ_N любого метода передачи с полной обратной связью по ДКБП удовлетворяет (как и при отсутствии обратной связи) неравенству

$$H_\xi(\epsilon_N) \le N\,C_1$$

Поэтому из рассуждений, использованных при исследовании метода передачи без обратной связи, следует, что при передаче равновероятных сообщений

$$\epsilon_N \ge 1 - \frac{C_1 + \log 2/N}{R}$$

Таким образом, даже при наличии полной обратной связи при скорости передачи $R > C_1$ вероятность ошибки $\epsilon_N \not\to 0$ при $N \to \infty$.

Доказательство теоремы 5. Как и при выводе теоремы переработки данных при отсутствии обратной связи, можем написать

$$J(\xi;\hat{\xi}) \le J(\xi;\tilde{\xi},Y_1^N) = J(\xi;Y_1^N)\,.$$

Применив "цепное"правило свойства 3, имеем

$$J(\xi;\hat{\xi}) \le \sum_{n=1}^{N} J(\xi;Y_n|Y_1^{n-1})\,.$$

Далее, мы покажем, что при $n=\overline{1,N}$ информация

$$I_n := J(\xi;Y_n|Y_1^{n-1}) \le J(X_n;Y_n)\,, \tag{17}$$

откуда вытекает доказываемое неравенство.

Применяя последовательно свойство 5, свойство 2 и лемму 2, имеем

$$\begin{aligned} I_n \; &\le J(\xi,X_n;Y_n|Y_1^{n-1}) = J(X_n;Y_n|Y_1^{n-1}) + J(\xi;Y_n|Y_1^{n-1},X_n) \le \\ &\le J(X_n;Y_n) + J(\xi;Y_n|X_n,Y_1^{n-1})\,. \end{aligned}$$

Далее рассмотрим

$$\begin{aligned} J(\xi; Y_n | X_n, Y_1^{n-1}) &\overset{(2)}{=} H(Y_n | X_n, Y_1^{n-1}) - H(Y_n | X_n, Y_1^{n-1}, \xi) \overset{(6)}{=} \\ &\overset{(6)}{=} H(Y_n | X_n) - H(Y_n | X_n) = 0\,, \end{aligned}$$

где сначала воспользовались обобщенной формулой условной информации из свойства 2, а затем, учитывая лемму 2 и определение (16) метода передачи с обратной связью, применили свойство 6 о монотонности энтропии. Следовательно, справедиво (17).

Теорема 5 доказана.