счёт применения процедур эффективного кодирования. Эффективность кодирования характеризуется величиной $\eta = H/L$, а величина $\mu = 1 - \eta$ наз. избыточностью. Эфф. кодирование, обеспечивающее мин. значение избыточности, можно осуществить с помощью кодов Шеннона, Р. Фано, Д. Хаффмена [1, 3] (в случае $\mu = 0$ код наз. оптимальным).

Код Хаффмена строится след. образом. Сообщения (число к-рых конечно) располагаются в таблице в столбец в порядке убывания их вероятностей. Два последних сообщения объединяются в одно с суммарной вероятностью, и далее по тому же правилу строится след. столбец таблицы. Затем в полученном столбце два последних сообщения вновь объединяются в одно с суммарной вероятностью, строится новый столбец таблицы и т. д. Процесс продолжается до тех пор, пока в последнем построенном столбце не останется двух сообщений. Верхнему из них приписывается кодовое слово 0, нижнему — 1. Далее рассматривается предпоследний столбец, в к-ром для объединявшихся сообщений на втором месте кодового слова ставится 0 для верх. сообщения и 1 — для нижнего. Если сообщения не объединялись, то они сохраняют кодовые слова предыдущего столбца. Процесс кодирования продолжается до тех пор, пока кодовые слова не будут приписаны всем исходным сообщениям в первом столбце.

Пусть, напр., источник создаёт четыре сообщения x_1, x_2, x_3, x_4 с вероятностями $p_1 = 1/2$, $p_2 = 1/4$, $p_3 = p_4 = 1/8$. Процесс построения кода Хаффмена для этого случая — объединение сообщений и кодирование — показан в табл. 1

Табл. 1.

x_i	p_i	p_i	p_i
x_1	$1/2$	$1/2$	$1/2$
x_2	$1/4$	$1/4$ →	$1/2$
x_3	$1/8$ →	$1/4$	
x_4	$1/8$		

Табл. 2.

x_i	Код	Код	Код
x_1	0	0	0
x_2	10	10 ←	1
x_3	110 ←	11	
x_4	111		

и табл. 2 соответственно. Для построенных кодовых слов сообщений 0, 10, 110, 111 ср. длина слова

$$L = 1/2 \cdot 1 + 1/4 \cdot 2 + 1/8 \cdot 3 + 1/8 \cdot 3 = 7/4,$$

что равно значению энтропии для рассматриваемого источника: $H = -1/2\log_2 1/2 - 1/4\log_2 1/4 - 1/8\log_2 1/8 - 1/8\log_2 1/8 = 7/4$. Избыточность кодирования $\mu = 0$, эффективность $\eta = 1$, т. е. построенный код — оптимальный. Использование эфф. кодирования, однако, допустимо только при полной гарантии отсутствия ошибок при кодировании и декодировании сообщений, т. к. в этом случае ошибка в восстановлении одного сообщения может повлечь появление ошибок при восстановлении многих последующих сообщений.

К наиб. важным проблемам Т. и. относится согласование информац. свойств источника сообщений и канала связи. Пропускная способность канала связи C определяется как макс. кол-во информации, к-рое способен передать канал в единицу времени. Единицей измерения пропускной способности канала связи является 1 бит/с.

Пусть источник создаёт сообщения в виде слов, записываемых буквами алфавита $A_m = \{a_1, \ldots, a_m\}$. При вероятностях появления этих букв $p_1, \ldots, p_m$ на одну букву приходится в ср.

$$H = \sum_{i=1}^{m} p_i \log_2 p_i$$

бит информации. Основная теорема Шеннона для канала связи без шума формулируется след. образом.

Пусть источник сообщений характеризуется энтропией H (бит/буква), а канал связи имеет пропускную способность C (бит/с). Тогда можно закодировать сообщения так, чтобы передавать символы по каналу связи со ср. скоростью $C/H - \varepsilon$ (буква/с), где ε — сколь угодно малое число. Передавать буквы со ср. скоростью, превышающей C/H, невозможно. Достижение верх. границы для скорости передачи, указываемой теоремой Шеннона, осуществляется за счёт применения процедур эфф. кодирования.

При передаче сигналов по каналам связи на них возможно действие разл. помех, шумов, к-рые могут привести к искажениям восстанавливаемых сообщений. Пусть, как и ранее, источник сообщений создаёт слова, записываемые буквами алфавита $A_m = \{a_1, \ldots, a_m\}$ при вероятностях их появления $p_1, \ldots, p_m$. Пусть далее вследствие действия помех слова́, восстанавливаемые приёмником, оказываются записанными в алфавите $B_n = \{b_1, \ldots, b_n\}$, к-рый, в частности, может совпадать с алфавитом источника, причём вероятности появления букв алфавита B_n равны $r_1, \ldots, r_n$. Тогда кол-во информации на выходе канала связи относительно его входа, приходящееся на одну передаваемую букву, определяется след. ф-лой:

$$I(B_n, A_m) = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij} \log_2 \frac{p_{ij}}{p_i p_j},$$

где p_{ij} ($i = 1, \ldots, n$; $j = 1, \ldots, m$) — вероятности совместного появления букв входного и выходного алфавитов.

Пропускная способность канала связи с шумами $C_ш$ определяется как макс. кол-во информации $I(B_n, A_m)$, к-рое можно передать по каналу связи за 1 с. Максимум находится для всех возможных источников, к-рые могут быть использованы на входе данного канала связи.

Теорема Шеннона для канала связи с шумами формулируется след. образом.

Пусть H_1 — ср. кол-во информации, создаваемое источником в единицу времени, т. е. производительность источника сообщений, измеряемая в бит/с. Пусть далее $C_ш$ — пропускная способность канала с шумом, тоже измеряемая в бит/с. Тогда если $H_1 \leqslant C_ш$, то такой системы кодирования не существует.

Пропускная способность канала с шумом существенно зависит от действующих на сигналы помех. Рассмотрим двоичный симметричный канал, передающий двоичные буквы 0 и 1 с вероятностью правильной передачи ε и искажающий их с вероятностью $\delta = 1 - \varepsilon$. Пропускная способность такого канала при передаче одной буквы в секунду определяется ф-лой

$$C_ш = 1 + \varepsilon \log_2 \varepsilon + \delta \log_2 \delta.$$

График зависимости $C_ш$ от δ приведён на рис. 4. Если $\varepsilon = \delta = 1/2$, т. е. если вероятность правильной передачи буквы равна вероятности её искажения, то пропускная способность канала с шумом $C_ш = 0$.

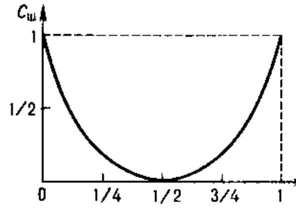

Рис. 4.

Теорема Шеннона для канала с шумом не указывает конкретного способа борьбы с помехами. Простейший способ борьбы с помехами, состоящий в много-