

Средняя длина кода $code$
зависит от источника:

$$|code(X)| = \sum_{x \leftarrow X} p_X(x) \cdot |code(x)|$$

1921, Роналд Фишер: для каждого
источника X есть предел сжатия
 $\exists \inf_{code} |code(X)|$

Для РВ X с $N = 2^r$ символами
предел r бит на символ

1928, Ральф Хартли:
а) назовём этот предел $I(X)$

б) для РВ X с N символами
 $I(X) = \log_2 N = -\log_2 \frac{1}{N}$ бит

Объёмный подход: $I(X) = \inf_{code} |code(X)|$

Вероятностный подход: $I(X)$ зависит от вероятностей X

Первая теорема Шеннона — эквивалентность подходов

Для РВ X с N символами
 $\frac{1}{N} = p_X(a_j)$

1948, Клод Шеннон:

а) пусть в сообщении x : $I_X(x) = -\log_2 p_X(x)$ бит

б) усредняем по X как длину: $I(X) = \sum_{x \leftarrow X} p_X(x) \cdot I_X(x)$

в) такая $I(X)$ действительно является $\inf_{code} |code(X)|$
(первая теорема Шеннона в формулировке для сжатия)