ВВЕДЕНИЕ

1) Число C определяет пропускную способность канала и измеряется в битах в секунду.

2) Ri – скорость передачи информации, измеряется в битах в секунду.

3) Обобщенная модель канала передачи информации: сообщение – кодер первичного кода – кодер источника – криптокодер – корректирующий кодер – кодер канала – модулятор – канал (- помехи) – демодулятор – декодер канала – корректирующий декодер – дешифратор – декодер источника – ЦАП – получатель.

4) y(t) = x(t)u(t) + n(t) – формула смеси сигнала и помех, x(t) – передаваемый сигнал, u(t) – мультипликативная помеха, n(t) – аддитивная помеха.

КАЧЕСТВЕННАЯ И КОЛИЧЕСТВЕННАЯ ОЦЕНКА ИНФОРМАЦИИ

5) p – вероятность события

6) I – количество информации. I ~ -log p (чем выше вероятность события, тем меньше в нем информации, и наоборот).

7) i – дискретный момент времени, xi – символ, который формируется в этот момент. X – множество исходных символов. Все вероятности появления символов в сумме должны давать 1. m – размерность X. Если источника два (X и T), то они независимы, если px,t = pxpt.

8) Xn – блоковый источник, n – количество символов в блоке, mn – количество символов на выходе.

9) p(y|x) – вероятность того, что на выходе канала будет символ y, хотя посылался символ x (вероятность ошибки). Дискретный канал без памяти – такой канал, в котором количество входов и выходов является конечным, и в котором p(y|x) не зависит от вероятностей предыдущих символов. Чем выше p(y|x), тем хуже (больше ошибок при передаче). P{xi|yj} = (p(yj|xi)p(xi))/p(yj) – вероятность того, что при передаче символ не изменится. p(yi) = Em p(yi|xj)p(xj) – вероятность появления yi (иными словами, можно вычислить вероятность символа на выходе, при известной вероятности исходного символа и известной возможности изменения символа). S/N – отношение мощности сигнала к мощности шума. S/N = (10log10S/N) dB.

10) P(xi.xj) = pipj – вероятность совместного события. I(pi,j) = I(pi) + I(pj) – совместная информация. I = log 1/p = -log p. Если выбрать в качестве основания логарифма 10 – то результат будет измеряться в единицах Хартли. Если выбрать 2 – то в битах.

11) Энтропия – среднее количество информации на один символ. E(I) = Em piI(xi) = Em –pilog pi. Энтропией называют также неопределенностью источника. H = E(I). H0 – емкость дискретного источника. H0 = log m. H’ = nH – энтропия блокового источника, где n – множитель источника. Энтропия блокового источника в n раз больше энтропии источника без множителя.

12) R = H0 – H – избыточность дискретного источника. r = R/H0 = 1 – H/H0 – относительная избыточность источника.

ЭФФЕКТИВНОЕ КОДИРОВАНИЕ ДЛЯ ДИСКРЕТНОГО ИСТОЧНИКА БЕЗ ПАМЯТИ

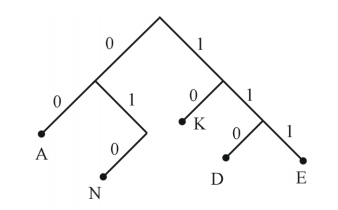
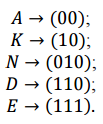
13) Система передачи информации с кодированием: источник информации – кодер – канал – декодер – получатель. Код источника – последовательность событий в виде символов алфавита кодирования. Кодовое слово или кодовый вектор – последовательность символов (состоящая, обычно, из 0 и 1, или из 1 и -1). В результате кодирования каждому уникальному символу присваивается свое кодовое слово. n – число символов кодового слова (длина кода).

14) Дано сообщение, кодовые слова для символов сообщения и вероятность/вероятности символов алфавита кодирования? В первую очередь, запиши сообщение в виде соединенных кодовых слов. Теперь представь, что каждый символ этого закодированного сообщения – это последовательность символов источника X (тоже самое, что мы рассматривали в пункте про качественную и количественную оценку информации). Теперь, можно найти что угодно, т.к. все нужные переменные нам известны.

14) Чем меньше длина кодовых слов, тем эффективнее кодирование. Для событий с высокой вероятностью короткие кодовые слова нужно использовать в первую очередь.

15) Префиксный код – код, в котором менее длинное кодовое слово не совпадает с началом более длинного кодового слова. Моментальный код основывается на префиксном коде, т.к. кодовые слова можно декодировать сразу же после их получения, ибо начала кодовых слов не совпадают. Код с запятой – это когда в конце кодовых слов стоит 0, который является своеобразной «запятой», разделяющей кодовые слова.

16) Даны кодовые слова, нужно построить кодовое дерево? Выбери точку отсчета, из нее нарисуй два отрезка, один – влево (будет обозначать 0) и один – вправо (будет обозначать 1). Окончания отрезков – новые точки отсчета. Повторяем аналогичную процедуру с каждой из этих точек отсчета до тех пор, пока не будут «нарисованы» все кодовые слова. Т.е. нужно сделать так, чтоб каждое кодовое слово можно было «прочитать» по отрезкам, означающим 0 или 1. Кодовое слово «читается», начиная с первой точки отсчета (с которой начали строить дерево). Если в коде идут несколько одинаковых символов подряд, то делать два отрезка из точки отсчета не нужно (но нужно помечать, что этот отрезок подразумевает именно 2 символа кода, а не 1).



17) Неравенство Крафта нужно для того, чтобы узнать, является ли код однозначно декодируемым. Если Em q-n меньше или равно 1, где m – количество кодовых слов, q – число символов кодового алфавита, n – длина кодового слова, то код однозначно декодируемый.

18) Ln = Em PiLi – средняя длина кодовых слов, где m – число символов источника, P – вероятности символов источника, L – длина кодовых слов. Средняя длина на один символ – Ln/n, где n – количество символов источника.

ТЕОРЕМА ШЕННОНА О КОДИРОВАНИИ ДЛЯ КАНАЛА БЕЗ ШУМА

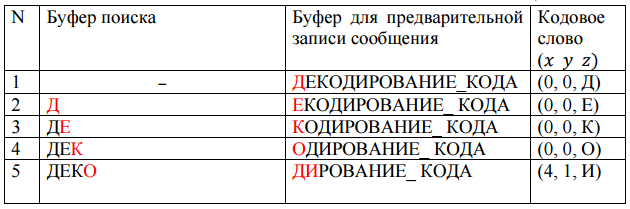
19) Эффективность сжатия оценивается по формуле K = (N-M)/N, где N – затраты на передачу без сжатия, M – затраты на передачу со сжатием. N/M = (KN + M)/M– выигрыш в записи данных. N = R = 2fdn, где R – скорость передачи данных, fd – частота дискретизации, n – длина кодовых слов.

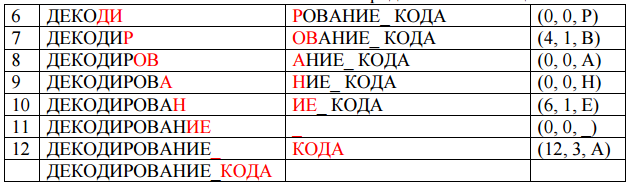
20) Алгоритм Хаффмена: расставляем символы источника в порядке убывания их вероятностей, объединяем 2 символа с наименьшими вероятностями в 1 символ, повторяем эти действия до тех пор, пока все символы не объединятся. В случае, если символы имеют одинаковую вероятность, объединяются те из них, которые до этого имели меньшее число объединений. После этого, начиная с последнего объединения, приписываем первой компоненте составного символа 1, а второй – 0. Продолжать до тех пор, пока все символы не будут закодированы (иными словами, составляем кодовое дерево, и по нему уже записываем коды символов). С блоковыми источниками – тоже самое.

21) Эффективность кода n = H/Ln, где H – энтропия, Ln – средняя длина кода.

22) Адаптивный алгоритм Хаффмена – это когда предполагается, что все вероятности равны. Все остальное – тоже самое.

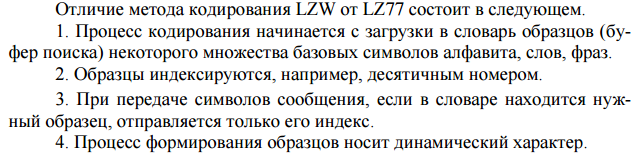
23) LZ77 – алгоритм сжатия. Кодовое слово записывается в формате (x,y,z), где x указывает адрес образца в буфере поиска (адрес определяется числом символов в обратном направлении в буфере, где начинается образец), y – число совпадающих символов, z – следующая буква в буфере.

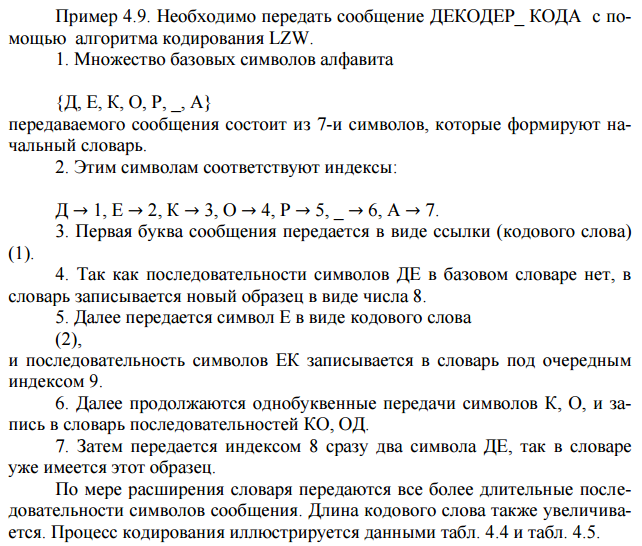


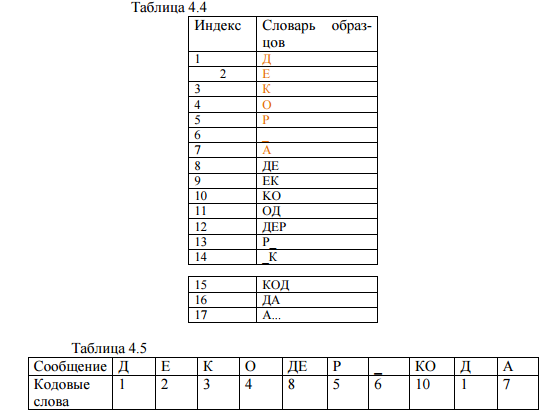


Результат записывается в виде последовательности кодовых слов.

24) LZ78 – модификация LZ77, буферы не имеют конечный раздел и изменена структура кодового слова. LZW – модификация LZ78.

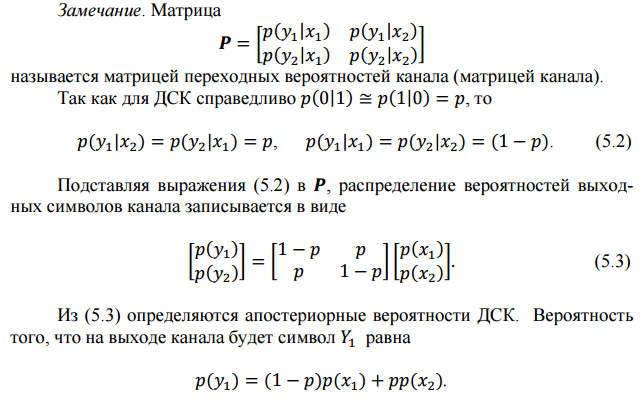






КАНАЛЫ БЕЗ ПАМЯТИ И ПЕРЕДАЧА ИНФОРМАЦИИ

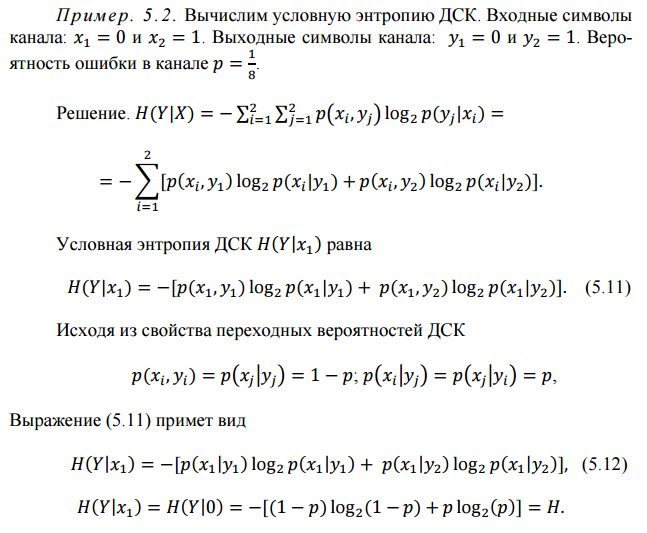
25) Если при передаче информации входные символы (0 и 1) совпадают с выходными, то пространство событий описывается как {00,11}. Если не совпадают – {01,10}. Значит, все возможные события можно описать множеством {00,01,10,11}. P(0|0) => {00}, p(0|1) => {01} и т.д. Если p(0|1) = p (1|0), то канал называется двоичным симметричным. p – вероятность ошибки, (1-p) – вероятность достоверной передачи.

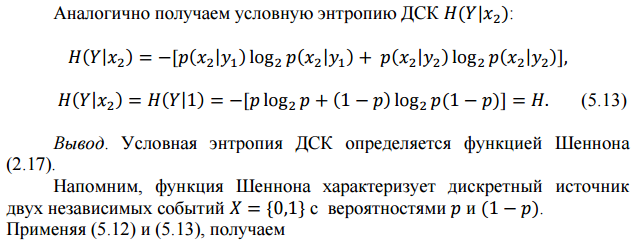


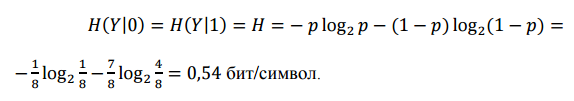


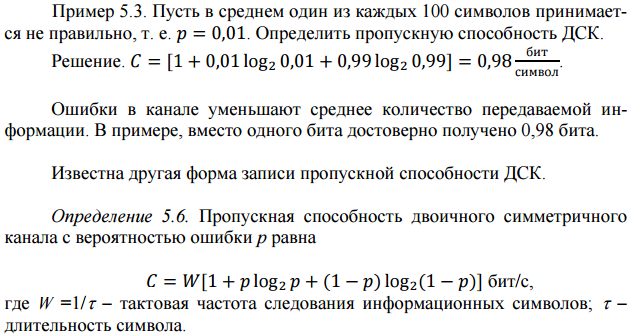
26) l – количество источников. Сумма вероятностей в комбинированном источнике равно 1. H(X,Y) = H(Y) + H(Y) – если источники независимы. H(X,Y) < H(X) + H(Y) – если источники не независимы.

27) Условная энтропия – среднее количество информации на один символ источника Y, при условии наличия символа источника X (и наоборот). H(Y|xi) = -Ey p(yi|xi)log2p(yi|xi) – условная энтропия источника Y для состояния xi источника X. H(Y|X) = -Ex H(Y|xi)p(xi) – среднее значение H(Y|xi) по всем xi.







28) 

29) 