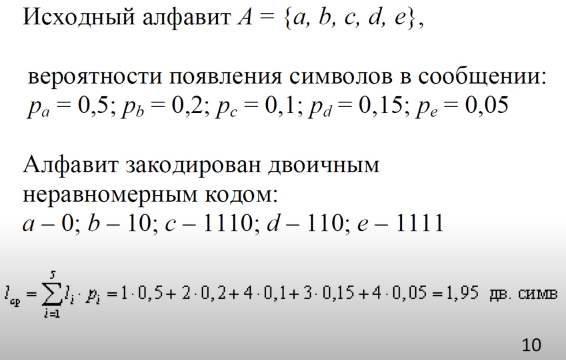


Кодирование — это представление символов исходного алфавита в виде символов кодового алфавита.

Предположим что у нас даны неравномерные коды, то есть каждый символ исходного алфавита А (каждый i-тый символ), закодирован в виде, последовательности символов алфавита В и длина этих последовательностей, для каждого i-тового символа - разная, это дает нам неравномерные коды.

Неравномерные коды характеризуются средней длиной кодового слова, которое может быть рассчитано последующей формуле (выше).

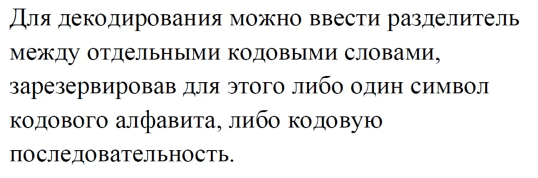
Для того чтобы найти среднюю длину кодового слова, мы берем сумму от единицы до N, то есть пробегаем по всем символам исходного алфавита i - это символ исходного алфавита. Что же у нас под знаком суммы, мы суммируем длину кодового слова для каждого i-тового символа (то есть для каждого символа исходного алфавита) и умножаем на вероятность появления этого символа (символа исходного алфавита), таким образом, средняя длина кодового слова, это сумма длин всех кодовых слов, взятых с весом - равным вероятности появления кодируемого символа.

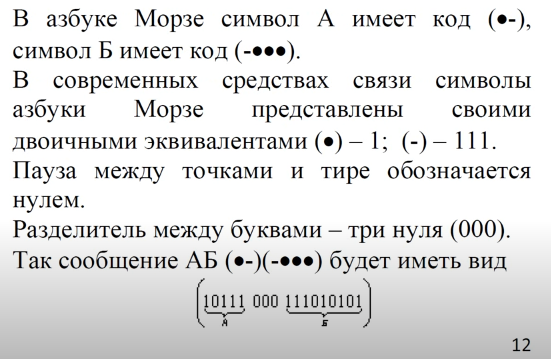
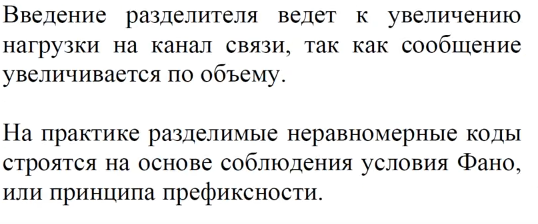


Предположим, у нас есть исходный алфавит А, который содержит 5 символов a, b, c, d, e, мы знаем вероятность появления символов сообщений. Она представлена на слайде.

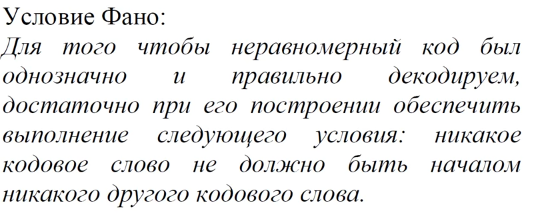
Предположим что у нас, данный алфавит уже закодирован двоичным неравномерным кодом, символ a - закодирован символом 0, b - закодирован как символы 10, c - как последовательность символов 1110, d - 110 и, наконец, e - 1111, тогда мы можем найти легко среднюю длину кодового слова, согласно предыдущей формуле, номер стр.9 выше слайд. То есть, мы берем сумму от единицы до пяти и умножаем длину кодового слова (то есть количество символов кодового алфавита) на вероятность встретить тот или иной символ исходного алфавита, по сколько у нас a закодировано нулем соответственно l первое у нас будет равна единице и вероятность равна 0,5, b - у нас закодировано 10 (2 символа), соответственно следующие слагаемое у нас будет, 2 умножить на вероятность встретить символ b, (0,2), то есть 2 умножить на 0,2, аналогично производя все вычисления, мы получаем, что средняя длина кодового слова равна 1,95 двоичных символов.

Предположим, что мы, закодировали некоторые сообщения в виде символов кодового алфавита, передали это сообщение и теперь на стороне получателя информации (получателя сообщений) нам необходимо декодировать полученные сообщения. Каким образом мы можем это сделать?

Мы можем ввести некоторый разделитель, между отдельными кодовыми словами, зарезервировав для этого один символ кодового алфавита или какую-то кодовую последовательность. Например, предположим у нас есть азбука Морзе.

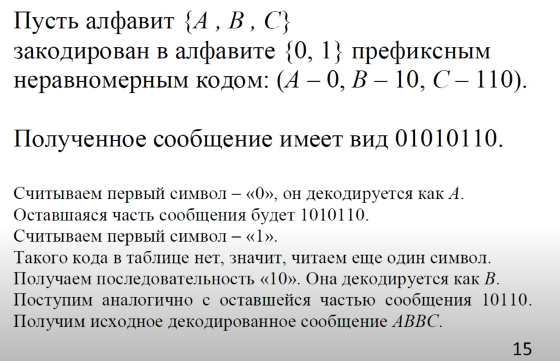


Введение разделителя ведет к увеличению нагрузки на канал, так как сообщение увеличится по объему и для того, чтобы от этого уйти на практике разделимые, неравномерные коды строятся на основе соблюдения условия Фано или иногда его называют принципом префиксной.



Заключается этот принцип (или условие Фано) в следующем: для того, чтобы неравномерный код был однозначно правильно декодируем, достаточно при его построении обеспечить выполнения следующего условия:

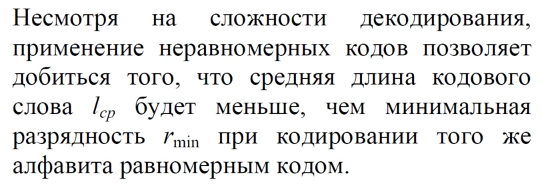
Никакое кодовое слово, не должно быть началом, никакого другого кодового слова.



Предположим у нас есть алфавит который содержит символы и A,B,C, он закодированным двоичным алфавитом, то есть символами 0 и 1, префиксным неравномерным кодом, код уже нам представлен, А у нас закодирован как 0, В как 10 и С как 110, мы видим, что здесь это условие Фано выполняется полностью, никакое кодовое слово не является началом, никакого другого кодового слова.

Тогда если мы получили сообщение вида 0 1 0 1 0 1 1 0, мы считаем символ за символ и стараемся его декодировать.

Например, мы считали 0 мы знаем код, 0 соответствует символу исходного алфавита А, таким образом, первый символ нам дает символ А, мы его декодируем и переходим к считыванию следующего символа. Считали единицу опять же обращаемся к коду, А точно отпадает, поскольку здесь у нас 0 а получили единицу, остались В и С, тому и другому символу подходят единицы, но мы пока не можем сказать однозначно какой это символ. Мы считываем следующий символ из последовательности поступившей нам это 0, 10 у нас соответствует строго символу исходного алфавита В, таким образом, мы декодируем следующий символ как В и переходим к считыванию следующего символа из поступившей последовательности, аналогичным образом мы получаем вторую В, потом получаем единицу опять же на повестке дня В или С получаем единицу, значит В отпала, и наконец получаем 0, однозначно декодируемым как С и таким образом у нас произошло декодирование, мы получили исходное сообщение АВВС.



Несмотря на сложности декодирования при применении неравномерных кодов, средняя длина кодового слова оказывается меньше, чем минимальная разрядность при кодировании этого же алфавита равномерным кодом, таким образом, кодирование неравномерным кодом будет эффективней нежели чем равномерным кодом.