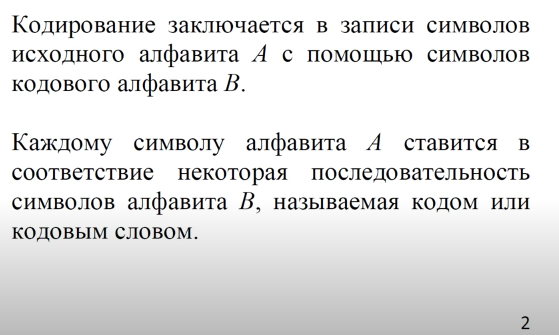
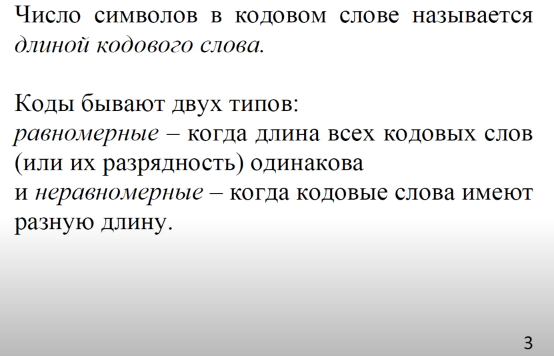
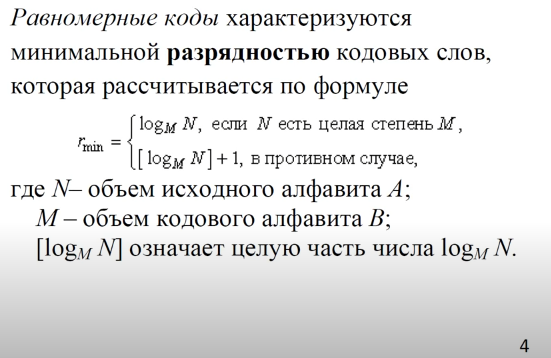
Кодированием называется процесс отображения одного набора знаков в другой, а кодом множество образов при этом отображении. Предположим, у нас есть исходный алфавит А, который содержит N символов, необходимо его закодировать виде символов кодового алфавита В количество этих символов равно M.

Как правило, количество символов в исходном алфавите больше, чем количество символов в кодовым алфавите.

Кодирование заключается в записи символов исходного алфавита А, с помощью символов кодового алфавита В, при этом каждому символу алфавита A ставится в соответствие некоторая последовательность символов алфавита В, именно она и называется кодовым словом.  
 Число символов в кодовом слове, называется длиной кодового слова.

Коды бывает двух типов равномерные, когда длина всех кодовых слов одинакова и неравномерные, когда кодовые слова имеют разную длину.

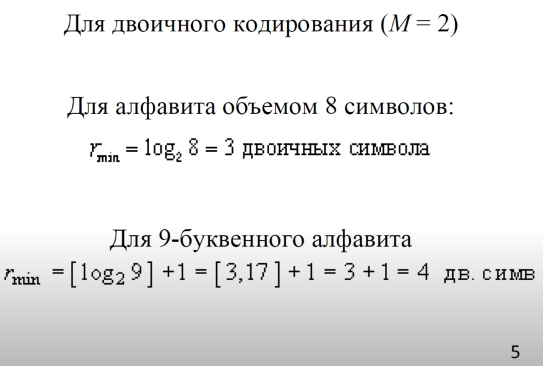
Рассмотрим первый случай. Равномерные коды характеризуется минимальной разрядностью кодовых слов, которая может быть получена по формуле, представленной на слайде.

Минимальная разрядность кодовых слов, равна логарифму от N по основанию M, если N есть целая степень M.

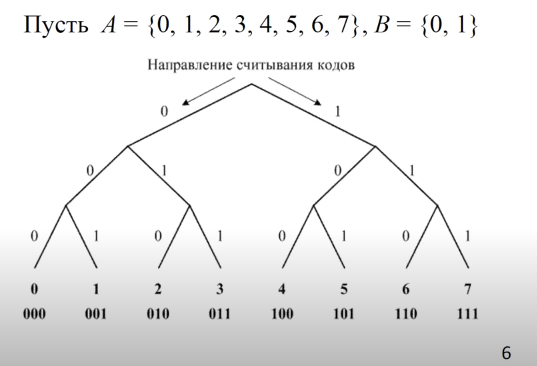
N – это количество символов исходом алфавита, а   
 M – это количество символов кодового алфавита Б.

Например, если у нас есть бинарный кодовый алфавит, двоичный кодовый алфавит, и мы кодируем исходный алфавит А, который содержит 16 символов, в этом случае N у нас целая степень M и тогда минимальный разрядность кодового слова будет логарифм от 16 по основанию 2.

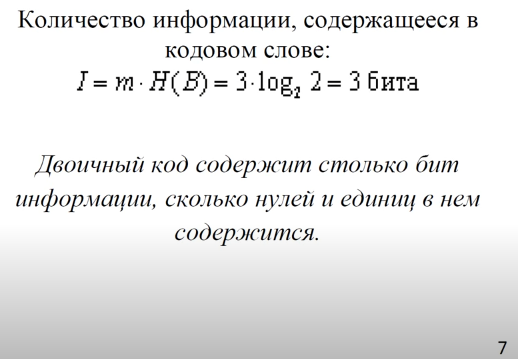
Если же у нас это условие не выполняется, то есть, N не является целой степенью M, тогда мы берем этот логарифм и округляем в большую сторону.

Так для двоичного кодирования M равно 2.

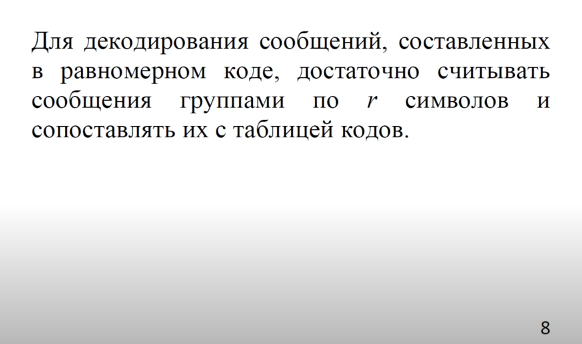
Предположим что у нас, исходная позиция содержит 8 символов, тогда минимальная разрядность кодового слова, будет равна логарифм от 8 по основанию 2, то есть три двоичных символа, однако для 9-символьного исходного алфавита, минимальная разрядность кодового слова будет - логарифм от 9 по основанию 2, мы округляем его в большую сторону и получаем 4-е двоичных символа.

Нетрудно заметить что, минимальная разрядность кодового слова это целое число, и в большинстве случаев она больше энтропии исходного алфавита, за счет округления этого значения, энтропии до целого, это порождает избыточность кода.   
Если кодовый алфавит состоит из двух символов 0 и 1, то такое кодирование называется двоичным.

Двоичный равномерный код легко получить путем построения двоичного дерева, предположим, исходный алфавит у нас содержит 8 символов, тогда двоичное кодирование будет иметь вид, представленный на следующем рисунке (выше).

Код каждого символа исходного алфавита считывается с дерева, двигаясь от корня к листьям, и состоит из трех символов кодового алфавита.   
 Можно легко найти количество информации, которая содержится в кодовом слове, это количество символов (сообщений), то есть в кодовом слове, умножить на энтропию кодового алфавита H(B), то есть 3 умножить на логарифм от 2 по основанию 2. Количество информации, которое содержится в кодовом слове, равно трем битам.

Для двоичного кодирования справедливо следующее утверждение: *двоичный код содержит столько бит информации, сколько нулей и единиц содержится в двоичном коде.*



После того, как мы закодировали сообщение, передали это сообщение, на стороне приемника информации, необходимо это сообщение декодировать.

Для этого достаточно считывать сообщение группами по r символов, то есть группами по количеству символов, равным минимальной разрядности кодового слова, и сопоставлять их с таблицей кодов, таким образом, очень легко можем декодировать полученное сообщение.