Понятие о кодах

Код –это совокупность всех комбинаций из определенного количества символов, которые избраны для представления информации. Каждая комбинация называется кодовой. Коды : **равномерные** – где все комбинации имеют одинаковое количество знаков и **неравномерные** – с различным количеством знаков (код Морзе).

При помощи n двоичных знаков можно получить 2^n кодовых комбинаций.

В зависимости от степени использования всех возможных кодовых комбинаций, коды делятся на **простые** (используются все комбинации) и **корректирующие** (избыточные), для которых нужна дополнительная информация, которые разделяются на **систематические** (с постоянным количеством m информационных и k=n-m избыточных знаков, где эти знаки занимают одни и те же позиции во всех комбинациях) и **несистематические,** где коды или слова нельзя разделить на информационные и контрольные

Основные характеристики:

Избыточность, определяется по формуле **k=n-m**,

где **n**- общее число знаков в коде; **m**- число информационных знаков. **N** чисел или слов в простом коде определяется:

m=log$_2$N, **k**-число контрольных знаков. m=log$_2$4=2, k=1.

Относительная избыточность – R=k/m

Корректирующая способность определяется вероятностью обнаружения или исправления ошибок различных типов.

Вес W(A) кодовой комбинации А определяется количеством содержащихся в ней двоичных единиц.

Для А=111001, $W(A)=\Sigma a_i=4$

Кодовое расстояние между двумя кодовыми комбинациями А и В определяется числом позиций, в которых их элементы не совпадают. Равно весу комбинации С, полученной поразрядным сложением А и В:

$W(C)=W(A \oplus B)= \Sigma(a_i \oplus b_i)$

Минимальная кодовое расстояние **α**

Простейший корректирующий код –с проверкой на четность, образуется добавлением одного избыточного разряда.

Мин кодовое расстояние с проверкой на четность α=2, обнаруживает одиночные ошибки и групповые нечетной кратности.

Код Хэмминга – для исправления одиночных ошибок (при α=3) или обнаружения без исправления двойных ошибок (α=4). N-значный код Хэмминга имеет m информационных разрядов и k контрольных. Число контрольных разрядов должно удовлетворять $k>=\log_2(n+1)$, откуда $m<=n-\log_2(n+1)$

Пример шестизначный код Хэмминга n=6 $k>=\log_2 7$, k=3,

M = n – k = 3

Цифра	Простой код	Код Хэмминга 6 5 4 3 2 1
0	000	0 0 0 0 0 0
1	001	0 0 0 1 1 1
2	010	0 1 1 0 0 1
3	011	0 1 1 1 1 0
4	100	1 0 1 0 1 0
5	101	1 0 1 1 0 1
6	110	1 1 0 0 1 1
7	111	1 1 0 1 0 0

Принят код: 111100 исправлено 110100 ошибка по корректирующему числу в 4 разряде

111010 исправлено 101010 ошибка по корректирующему числу в 5 разряде

100000 исправлено 000000 ошибка по корректирующему числу в 6 разряде

При проверке информации после приема возможны 3 случая

- отсутствие ошибок, к.ч.=0,
- одиночная ошибка, к.ч.=номер искаженного разряда
- двойная ошибка, к.ч. не равно 0.

Циклический код – разновидность систематических кодов

Неприводимые многочлены (не могут быть представлены в виде произведения многочленов низших степеней, делится на себя или 1)при построении циклических кодов играют роль образующих полиномов.

Двоично-кодированное n - разрядное число представляется полиномом (n-1) степени некоторой переменной x, причем коэффициентами полинома являются двоичные знаки соответствующих разрядов. Запись, чтение и передача комбинаций производятся, начиная со старшего разряда (старший – справа).

Циклическая перестановка разрядов соответствует умноже-нию полинома на x, при котором x^n заменяется 1 и перехо-дит в начало полинома.

Алфавитное кодирование

Пусть задано конечное множество $A\{a_1,a_2,\ldots a_n\}$, называемое алфавитом. Элементы алфавита – буквы, последовательность букв – слово, n –длина слова. Множество слов в алфавите –A*. Если слово $\alpha= \alpha_1, \alpha_2$, то α_1 – это префикс слова, α_2 – это постфикс (конец) слова.

Алфавитное или побуквенное кодирование задается схемой или таблицей кодов $\delta:=< \alpha_1 \rightarrow \beta_1, \alpha_2 \rightarrow \beta_2, \ldots, \alpha_n \rightarrow, \beta_n >$, где $\alpha_k \in A$, $\beta_k \in B^*$

Разделимое кодирование – такое, что любое слово, состоящее из элементарных кодов, единственным образом разлается на элементарные коды. (Двоично-десятичный код – разделимый).

Префиксное кодирование – такое, что элементарный код одной буквы не является префиксом элементарного кода другой буквы.

Для получения разделимой схемы алфавитного кодирования необходимо, чтобы длины элементарных кодов удовлетворяли определенному соотношению – неравенству Макмиллана.

Теорема.

Если числа $l_1, l_2, \ldots, ln$ соответствующие длинам элементарных кодов $\beta 1, \beta 2, \ldots, \beta n$, удовлетворяют неравенству

$$\sum 2^{-li} <= 1,$$

то существует разделимая схема алфавитного кодирования $\delta^{\rightarrow} := < \alpha_1^{\rightarrow} \beta 1, \alpha_2^{\rightarrow} \beta_2, \ldots, \alpha_n^{\rightarrow} \beta_n$

“Квадрат Полибия”

В Древней Греции (II в. до н. э.) был известен шифр, называемый “*квадрат Полибия”* (Полибий (200–120 гг. до н. э.) – древнегреческий историк.) . Это устройство представляло собой квадрат 5×5, столбцы и строки которого нумеровали цифрами от 1 до 5.

В каждую клетку записывалась одна буква (в греческом варианте одна клетка оказывалась пустой, а в латинском – в одну клетку помещали две буквы *I, J*). В результате каждой букве отвечала пара чисел по номеру строки и столбца

A	**B**	**C**	**D**	**E**
F	G	H	I,J	K
L	M	N	O	P
Q	R	S	T	U
V	W	X	Y	Z

“Я мыслю, следовательно, существую”
Р. Декарт
Cogito ergo sum – лат.

13 34 22 24 44 34 15 42 22 34 43 45 32

Код Цезаря

В I в. н. э. Ю. Цезарь во время войны с галлами, переписываясь со своими друзьями в Риме, заменял в сообщении первую букву латинского алфавита *A* на четвертую *D*, вторую *B* – на пятую *E*, наконец последнюю – на третью:

ABCDEFGHIJKLMNOPQRSTUVWXYZ
DEFGHIJKLMNOPQRSTUVWXYZABC

YHQL YLGL YLFL

Veni vidi vici –

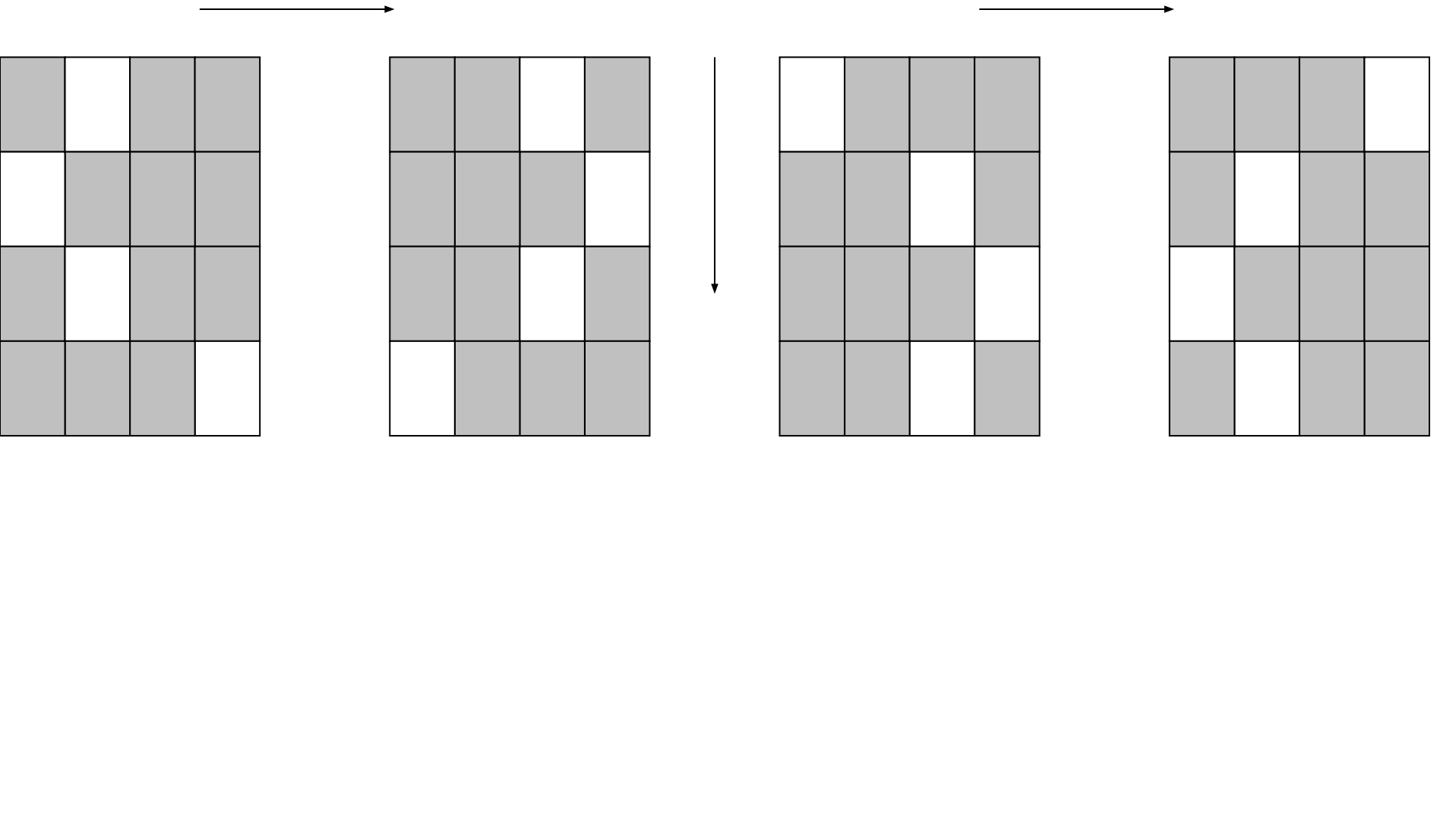
“Пришел, увидел, победил”

Ю. Цезарь. Донесение Сенату о победе
над понтийским царем

“*Решетка Кардано*”

Широко известны шифры, принадлежащие к классу “перестановка”, в частности, “*решетка Кардано*”[2]. Это прямоугольная карточка с отверстиями, чаще всего квадратная, которая при наложении на лист бумаги оставляет открытими лишь некоторые его части. Число строк и столбцов на карточке четное. Карточка сделана так, что при последовательном ее поворачивании каждая клетка лежащего под ней листа окажется занятой. Карточку поворачивают сначала вдоль вертикальной оси симметрии на 180°, а затем, вдоль горизонтальной оси, также на 180°. И вновь повторяют ту же процедуру.

[2] Кардано Джероламо (1501–1576 гг.) – итальянский математик, философ и врач.



“Таблица Виженера”

Неудобство рассмотренных выше шифров очевидно, так как в случае использования стандартного алфавита таблица частот встречаемости букв алфавита позволяет определить один или несколько символов, а этого достаточно для вскрытия шифра, поэтому использовались различные приемы, для того чтобы затруднить дешифрование, например, использование “*таблицы Виженера”*, которая представляет собой квадратную таблицу с числом строк и столбцов, равным количеству букв алфавита. Чтобы получить шифрованный текст, находят очередной знак лозунга, начиная с первого, в вертикальном алфавите, а ему соответствующий знак сообщения в горизонтальном алфавите. На пересечении выделенных столбца и строки находим первую букву шифра. Очевидно, что ключом к такому шифру является используемый лозунг.

р а с к и н у л о с ь м о р е ш и р о к о

э о я к щ а п ы й ю й щ о в ч ф ш л ь ш ы

АБВГДЕЖЗИЙКЛ**М**НОПРСТУФХЦЧШЩЬЫЭЮЯ

БВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯА

ВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБ

ГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБВ

ДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯФБВГ

ЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯФБВГД

ЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯФБВГДЕ

ЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖ

ИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗ

ЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИ

КЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙ

ЛМНОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙК

МНОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛ

НОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМ

ОПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМН

ПРСТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНО

РСТУФХЦЧШЩЬЫ**Э**ЮЯАБВГДЕЖЗИЙКЛМНОП

СТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПР

ТУФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРС

УФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТ

ФХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУ

ХЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФ

ЦЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХ

ЧШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦ

ШЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧ

ЩЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШ

ЬЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩ

ЫЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬ

ЭЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧЩШЬЫ

ЮЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭ

ЯАБВГДЕЖЗИЙКЛМНОПРСТУФХЦЧШЩЬЫЭЮ

Виженер (1523–1596 гг.) – французский посол в Риме, написал большой труд о шифрах. Квадратный шифр Виженера на протяжении почти 400 лет не был дешифрован, считался недешифруемым шифром

“*Одноразовый шифровальный блокнот*”

Примером нераскрываемого шифра может служить так называемый “*одноразовый шифровальный блокнот*” – шифр, в основе которого лежит та же идея, что в шифре Цезаря. Назовем *расширенным алфавитом* совокупность букв алфавита, знаков препинания *{. , : ; ! ? () – “ <пробел>}*, число символов расширенного русского алфавита в данном варианте будет равно *44*. Занумеруем символы расширенного алфавита числами от *0* до *43*. Тогда любой передаваемый текст можно рассматривать как последовательность $\{a_n\}$ чисел множества $A=\{0,1,2,\ldots,43\}$.

Предположим, что имеем случайную последовательность $\{c_n\}$ из чисел множества *A* той же длины, что и передаваемый текст – *ключ*.

Складывая по модулю *44* число из передаваемого текста a_n с соответствующим числом из множества ключа c_n:

$$a_n + c_n = b_n \ (mod\ 44),\ \ 0 <= b_n <= 43,$$

получим последовательность $\{b_n\}$ знаков шифрованного текста.

Чтобы получить передаваемый текст, можно воспользоваться тем же ключом: $a_n = b_n - c_n \ (mod\ 44),\ \ 0 <= a_n <= 43.$

У двух абонентов, находящихся в секретной переписке, имеются два одинаковых блокнота. В каждом из них, на нескольких листах, напечатана случайная последовательность чисел множества *A*. Отправитель свой текст шифрует указанным выше способом при помощи первой страницы блокнота. Зашифровав сообщение, он уничтожает использованную страницу и отправляет текст сообщения второму абоненту; получатель шифрованного текста расшифровывает его и также уничтожает использованный лист блокнота. Нетрудно видеть, что одноразовый шифр нераскрываем в принципе, так как символ в тексте может быть заменен любым другим символом и этот выбор совершенно случаен.