Лекция 4
Эффективные коды

1. Двоичный безизбыточный код
2. Код Грея
3. Код Шеннона-Фано
4. Код Хаффмана
5. Арифметический код

Равномерные коды: двоичный безизбыточный; код Грея (КГ)

- К **равномерным** кодам относят двоичные позиционные коды, любые **m-значные** позиционные коды и коды, построенные на их основе.
- Примером равномерных эффективных кодов является двоичный безизбыточный код ($m_2=2$) в котором разрешены все кодовые комбинации.

1. Двоичный безизбыточный код

Пример

Пусть n=3

- Тогда данный код имеет 8 разрешенных комбинаций: 000, 001, 010, 011, 100, 101, 110, 111.

2. Код Грея

- Код Грея (КГ) — рефлексный* (отраженный) двоичный код, представляющий двоичную систему нумерации, в которой два соседних значения различаются только в одном двоичном разряде.
- КГ был разработан для защиты электромеханических переключателей от ложного срабатывания. В последствии КГ стали применять для выявления и исправления ошибок в системах связи и телеуправления**.
- Для указанных целей простой двоичный код не годится, т.к. в нем имеются соседние значения, для которых предыдущая комбинация отличается от последующей всеми разрядами (например,
- 3 — 4: 011 — 100; 7 — 8: 0111 — 1000;
- 15 — 16: 01111 — 10000 и т.д.).

Код Грея

Соответствие двоичного кода коду Грея (n=3)

Число	Двоичный код	КГ
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

Код Грея

Вывод

- Из таблицы видно, что КГ имеет ось симметрии между комбинациями, соответствующими числам $2^{n-1}-1$ и 2^{n-1} (в нашем примере между числами 3 и 4).

Код Грея

Перевод простого кода в КГ

1. Под двоичным кодом записывают то же число со сдвигом вправо на один разряд (младший разряд при этом теряется.
2. Выполняют поразрядное сложение сдвинутого и не сдвинутого чисел по модулю два.

Пример

$$
\begin{array}{r@{\;}l}
 & 0\,1\,0\,0 \\
\oplus & \;\;\,0\,1\,0\,(0) \\
\hline
 & 0\,1\,1\,0
\end{array}
$$

Код Грея

Обратный перевод (КГ в простой код)

1. Цифра старшего разряда остается без изменений.
2. Каждая последующая цифра инвертируется столько раз, сколько единиц ей предшествует в КГ.

Пример

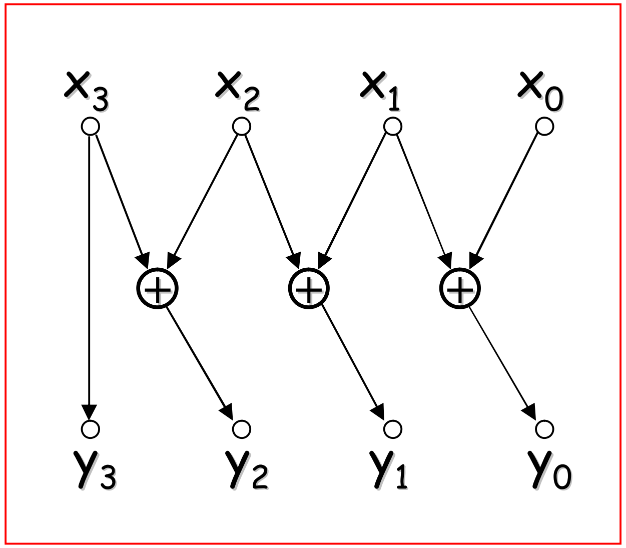
Перевести комбинацию КГ 1110 в простой двоичный код

1. Первая цифра остается без изменений.
2. Вторая цифра инвертируется один раз (1→0), получаем 0.
3. Третья цифра инвертируется два раза (1→0, 0→1), получаем 1.
4. Четвертая цифра инвертируется три раза (0→1,1→0, 0→1), получаем 1.

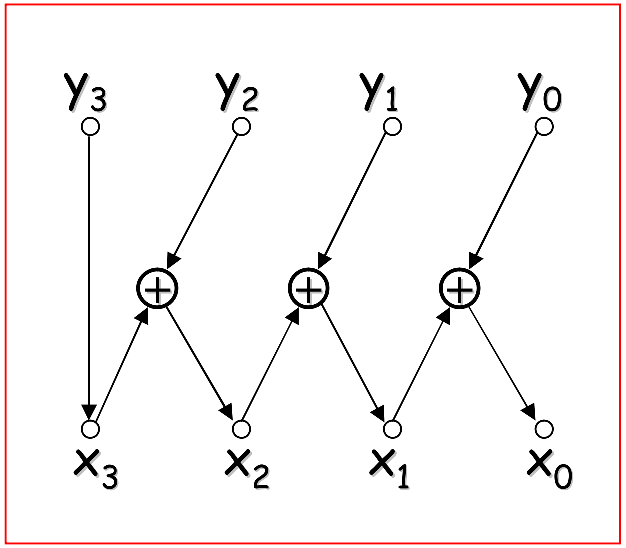
Получили комбинацию двоичного кода 1011.

Код Грея: примеры построения (n=4)

1. Графическое



КГ



ОКГ*

Код Грея: примеры построения (n=4)

2. Аналитическое (см. слайд 9)

КГ	ОКГ	
$y_3 = x_3$		$x_3 = y_3$
$y_2 = x_3 \oplus x_2$	$x_2 = (y_2 - x_3)_2 \equiv$	$x_2 = y_2 \oplus x_3$
$y_1 = x_2 \oplus x_1$		$x_1 = y_1 \oplus x_2$
$y_0 = x_1 \oplus x_0$		$x_0 = y_0 \oplus x_1$

Код Грея: примеры построения (n=4)
3. Матричное

$$
M_{\overrightarrow{\text{КГ}}} = \begin{array}{c|cccc}
 & x_0 & x_1 & x_2 & x_3 \\
\hline
y_0 & 1 & 1 & 0 & 0 \\
y_1 & 0 & 1 & 1 & 0 \\
y_2 & 0 & 0 & 1 & 1 \\
y_3 & 0 & 0 & 0 & 1
\end{array}
$$

$x_3=y_3$

$x_2=(y_2 - x_3)_2$

$x_1=(y_1 - x_2)_2$

$x_0=(y_0 - x_1)_2$

$x_3=y_3$

$x_2=(- y_3 + y_2)_2$

$x_1=(y_3 - y_2 + y_1)_2$

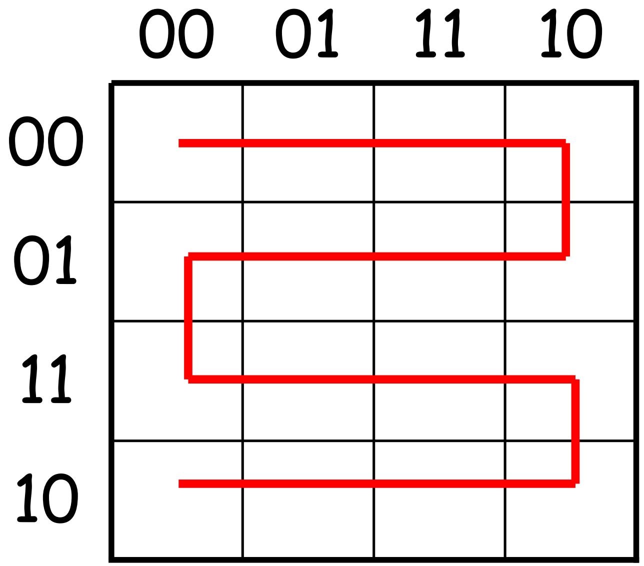
$x_0=(-y_3 + y_2 - y_1+y_0)_2$

$$
M_{\overrightarrow{\text{ОКГ}}} = \begin{array}{cccc|c}
y_0 & y_1 & y_2 & y_3 & \\
\hline
1 & -1 & 1 & -1 & x_0 \\
0 & 1 & -1 & 1 & x_1 \\
0 & 0 & 1 & -1 & x_2 \\
0 & 0 & 0 & 1 & x_3
\end{array}
$$

Код Грея: примеры построения (n=4)

4. По картам Карно

(механизм построения циклического КГ)



Код Грея

Равномерные по Хэммингу (4,3) последовательности

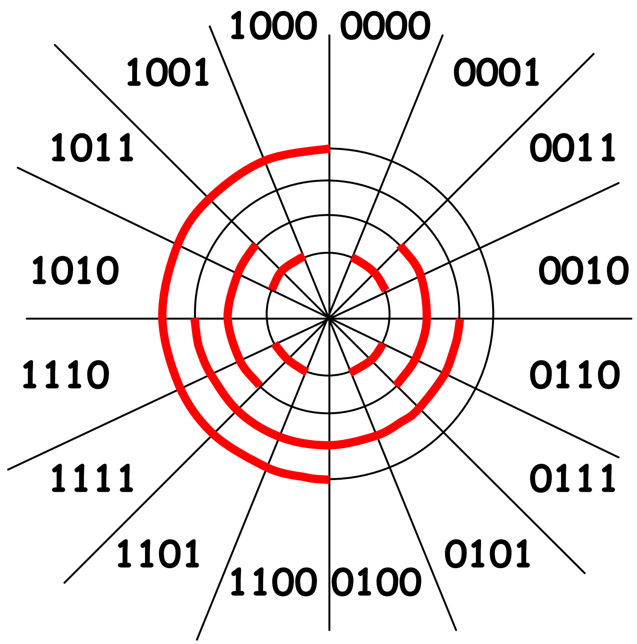
0	0000	4	0101	8	0011	12	0110
1	0111	5	0010	9	0100	13	0001
2	1100	6	1001	10	1111	14	1010
3	1011	7	1110	11	1000	15	1101

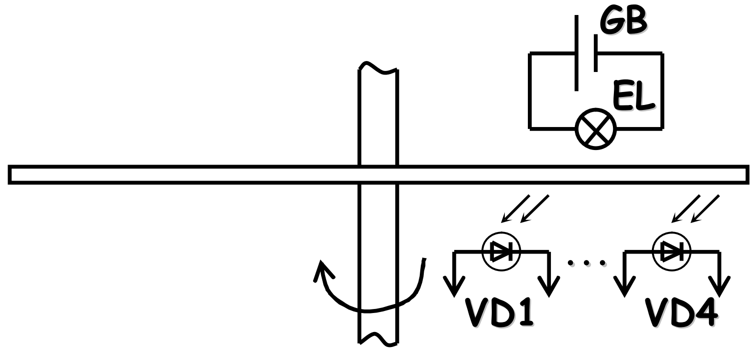
Код Грея: примеры построения (n=4)

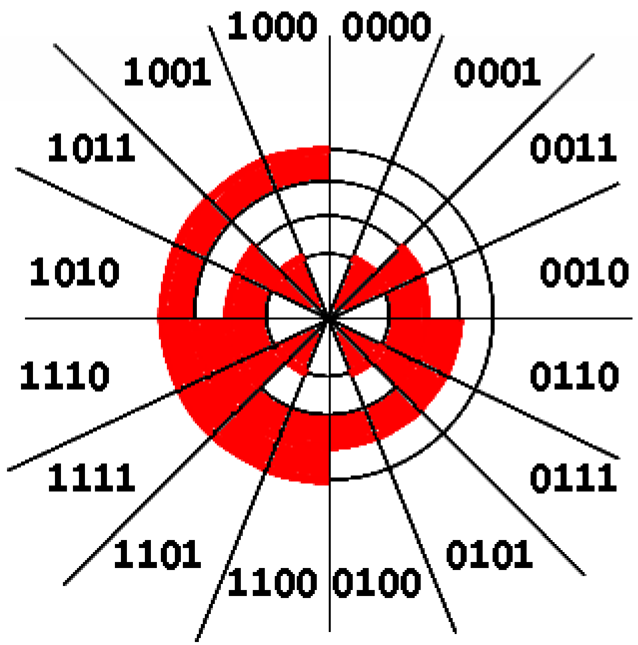
Порядок обхода ячеек карты Карно при построении (4,3)-последовательности

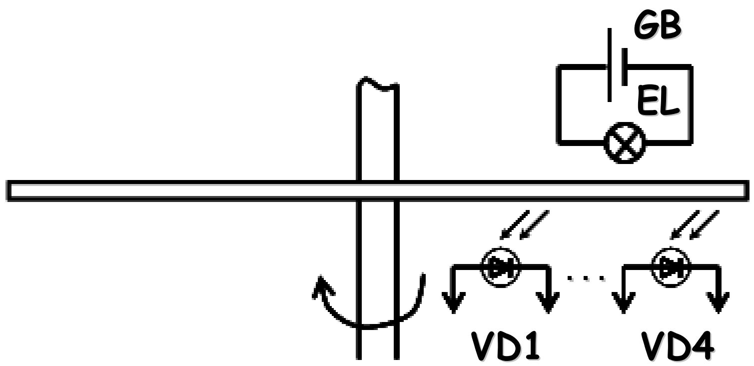
	00	01	11	10
00	0	13	8	5
01	9	4	1	12
11	2	15	10	7
10	11	6	3	14

Код Грея: применение в круговых энкодерах*






1000
0000
0001
0011
0010
0110
0111
0101
0100
1100
1101
1111
1110
1010
1011
1001


GB
EL
VD1
VD4

Неравномерные коды: основные принципы оптимального кодирования

1. В неравномерных кодах наиболее вероятным символам первичного алфавита присваиваются более короткие кодовые комбинации, а символам с меньшей вероятностью — более длинные. Это способствует увеличению скорости передачи информации по КС.
2. Выбор каждого кодового слова необходимо производить так, чтоб содержащееся в нем количество информации было максимальным.

Примерами неравномерных эффективных (оптимальных) кодов являются коды Шеннона-Фано, Хаффмана и арифметический код.

3. Код Шеннона-Фано

Методика построения кодов Шеннона-Фано (КШФ) разработана как для двоичных, так и недвоичных алфавитов.

Построение КШФ для двоичных алфавитов (m=2)

1. Все символы первичного алфавита располагают в ряд сверху вниз в порядке убывания их вероятностей.
2. Упорядоченные символы делят на две группы с примерно равными суммарными вероятностями. Одной из групп присваивают код «1», второй – код «0».
3. Затем каждую из подгрупп снова делят на две группы с примерно равными вероятностями. Им также присваиваются коды «1» и «0». Такое разбиение продолжают до тех пор, пока в каждой подгруппе не останется по одному символу.
4. Далее кодируют символы. Для каждого символа записывают коды подгрупп, в которые он входит, начиная от первого до последнего деления.

Код Шеннона-Фано

Вывод

1. Описанный код является эффективным неоднозначным кодом и называется статическим кодом Шеннона-Фано.

Код Шеннона-Фано

Пример

Закодировать сообщение ABABACAD КШФ и двоичным безизбыточным кодом (ДвК)

Символы	Кодовые слова ДвК	Частоты* символов	1 деление	2 деление	3 деление	Кодовые слова КШФ
A	00	0,5	0	—		0
B	01	0,25	1	0	—	10
C	10	0,125		1	0	110
D	11	0,125			1	111

Код Шеннона-Фано

Результаты кодирования

Исходное сообщение	Закодированное сообщение	Длина сообщения
ABABACAD	00 01 00 01 00 10 00 11 A B A B A C A D	16 бит 11/5 («0»/«1»)
	0 10 0 10 0 110 0 111 A B A B A C A D	14 бит 7/7 («0»/«1»)

— двоичное безизбыточное кодирование

— двоичное кодирование Шеннона-Фано

Код Шеннона-Фано

Расчет средней длины в битах символа сообщения

$$1.\ L_{ср} = \frac{14}{8} = 1{,}75\left[\frac{бит}{симв}\right]$$

$$2.\ L_{ср} = \sum_{i=1}^{4} p(a_i)\,l(a_i) =$$

$$= 0{,}5 \cdot 1 + 0{,}25 \cdot 2 + 0{,}125 \cdot 3 + 0{,}125 \cdot 3 = 1{,}75\left[\frac{бит}{симв}\right]$$

Код Шеннона-Фано

Расчет энтропии сообщения

$$3.\ H(A) = \sum_{i=1}^{4} p(a_i)\log\frac{1}{p(a_i)} =$$

$$= 0{,}5 \cdot \log 2 + 0{,}25 \cdot \log 4 + 0{,}125 \cdot \log 8 + 0{,}125 \cdot \log 8 = 1{,}75 \left[\frac{\text{бит}}{\text{симв}}\right]$$

Код Шеннона-Фано

Построение m-значного КШФ

1. Находим квант разбиения

$$\delta = \frac{1}{m_2}$$

и располагаем символы первичного алфавита в ряд сверху вниз в порядке убывания их вероятностей.

Делим эти символы на m_2 групп с примерно равными суммарными вероятностями и каждой группе присваиваем в определенном порядке по одной из букв вторичного алфавита.

Код Шеннона-Фано

2. На каждом шаге разбиения группы символов первичного алфавита находим остаток разбиения, который равен разности между квантом разбиения и истинной суммарной вероятностью подгруппы

$$\sigma = \delta - \sum_i p_i$$

3. Складывая полученные остатки разбиения по абсолютному значению, получаем среднее отклонение

$$\Delta = \sum_{j=1}^{m_2} \sigma_j \,.$$

Код Шеннона-Фано

4. На каждом следующем этапе разбиения квант разбиения складывается со средним отклонением и делится на m_2 и получается следующий квант разбиения

$$\delta' = \frac{\delta + \Delta}{m_2}$$

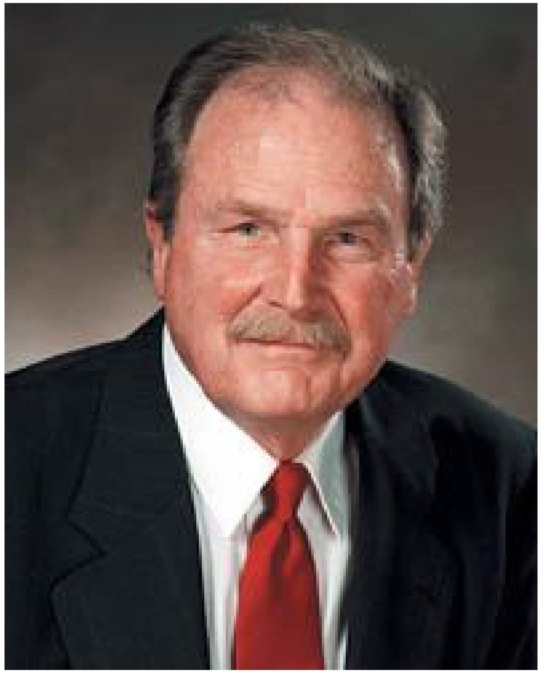
Задача сведения среднего отклонения (Δ) и остатка разбиения (δ) к нулю решается до тех пор, пока в каждой подгруппе не останется m_2 или менее элементов.

5. Решается задача кодирования аналогично ранее рассмотренному случаю.

4. Код Хаффмана* (КХ)

Методика построения КХ является **однозначной**.

1. Все символы первичного алфавита располагают в ряд сверху вниз в порядке убывания их вероятностей.
2. Два последних символа объединяются в один вспомогательный символ, вероятность которого равна сумме вероятностей объединяемых символов.
3. Строится не возрастающий ряд со вспомогательным символом.
4. Повторяются п.п. 2 и 3 данной методики до тех пор, пока не останется один вспомогательный (суммарный) символ с вероятностью 1.



Дэвид Хаффман

Код Хаффмана (КХ)

5. Строится дерево Хаффмана, корнем которого является суммарный символ. Ветвление дерева начинают от суммарной вершины. Ветви с большей вероятностью присваивают код 1 (левой), а ветви с меньшей — код 0 (правой).
6. Тупиковые вершины (листья) — исходные символы.
7. Кодовые комбинации записывают от корня до тупиковой вершины.

Код Хаффмана (КХ)

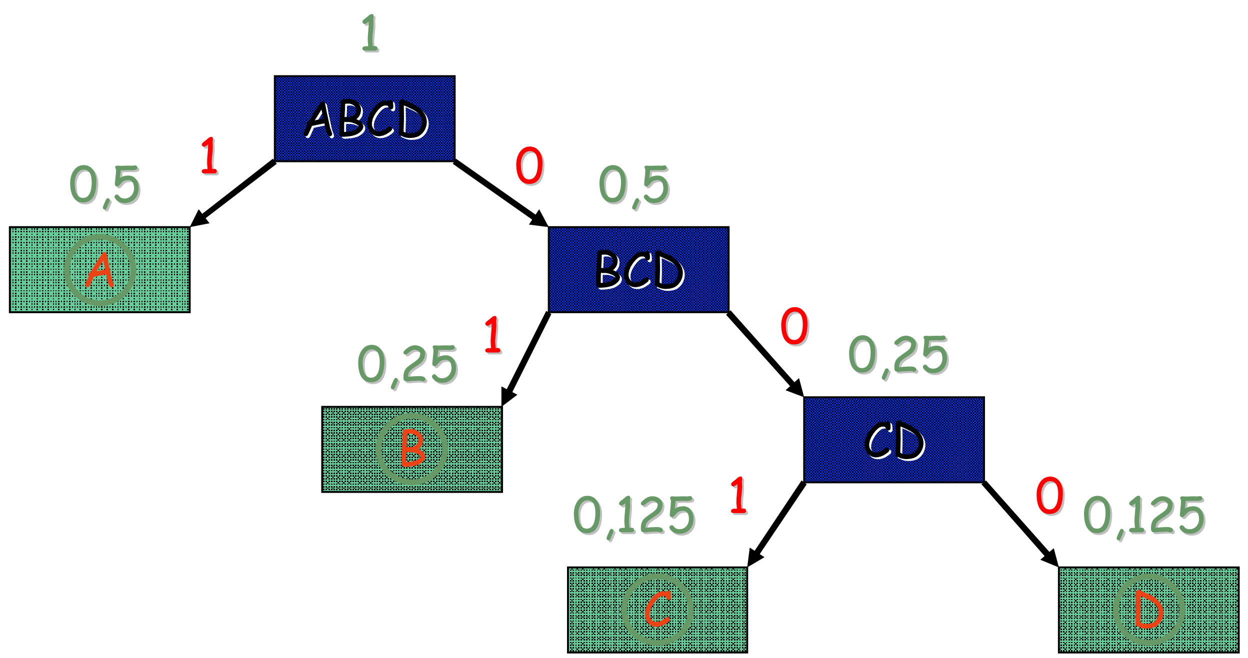
Пример

Закодировать КХ сообщение ABABACAD

Символы	Частоты символов				Кодовые слова КХ*
A	0,5	0,5	0,5	1	1
B	0,25	0,25	0,5		01
C	0,125	0,25			001
D	0,125				000

Код Хаффмана (КХ)

Кодовое дерево



Код Хаффмана (КХ)

Выводы

1. КХ являются однозначными в отличие от КШФ.
2. Все эффективные коды содержат только информационные символы, поэтому нет необходимости отделять информационные символы от проверочных в процессе декодирования.
3. Оптимальные неравномерные коды полностью нагруженные.
4. Кодирование статическими кодами осуществляется в два прохода.

5. Арифметический код

- При арифметическом кодировании (АК) текст представляется вещественными числами в интервале от 0 до 1.
- По мере кодирования текста отображающий его интервал уменьшается, а количество битов для его представления возрастает.

Арифметический код

- Очередные символы текста сокращают величину интервала исходя из значений их вероятностей, определяемых моделью источника. Более вероятные символы в меньшей степени сокращают интервал и, следовательно, добавляют меньше битов к результату, чем менее вероятные.
- Перед началом работы тексту соответствует интервал [0; 1). При обработке очередного символа его ширина сужается за счет выделения этому символу части интервала.

Арифметическое кодирование (пример)

- Дан алфавит {a, e, i, o, u, !} и модель с постоянными вероятностями

Символ	Вероятность	Интервал
a	0,2	[0,0; 0,2)
e	0,3	[0,2; 0,5)
i	0,1	[0,5; 0,6)
o	0,2	[0,6; 0,8)
u	0,1	[0,8; 0,9)
!	0,1	[0,9; 1,0)

- Закодировать текст «eaii!»

Арифметическое кодирование (пример)

Кодирование

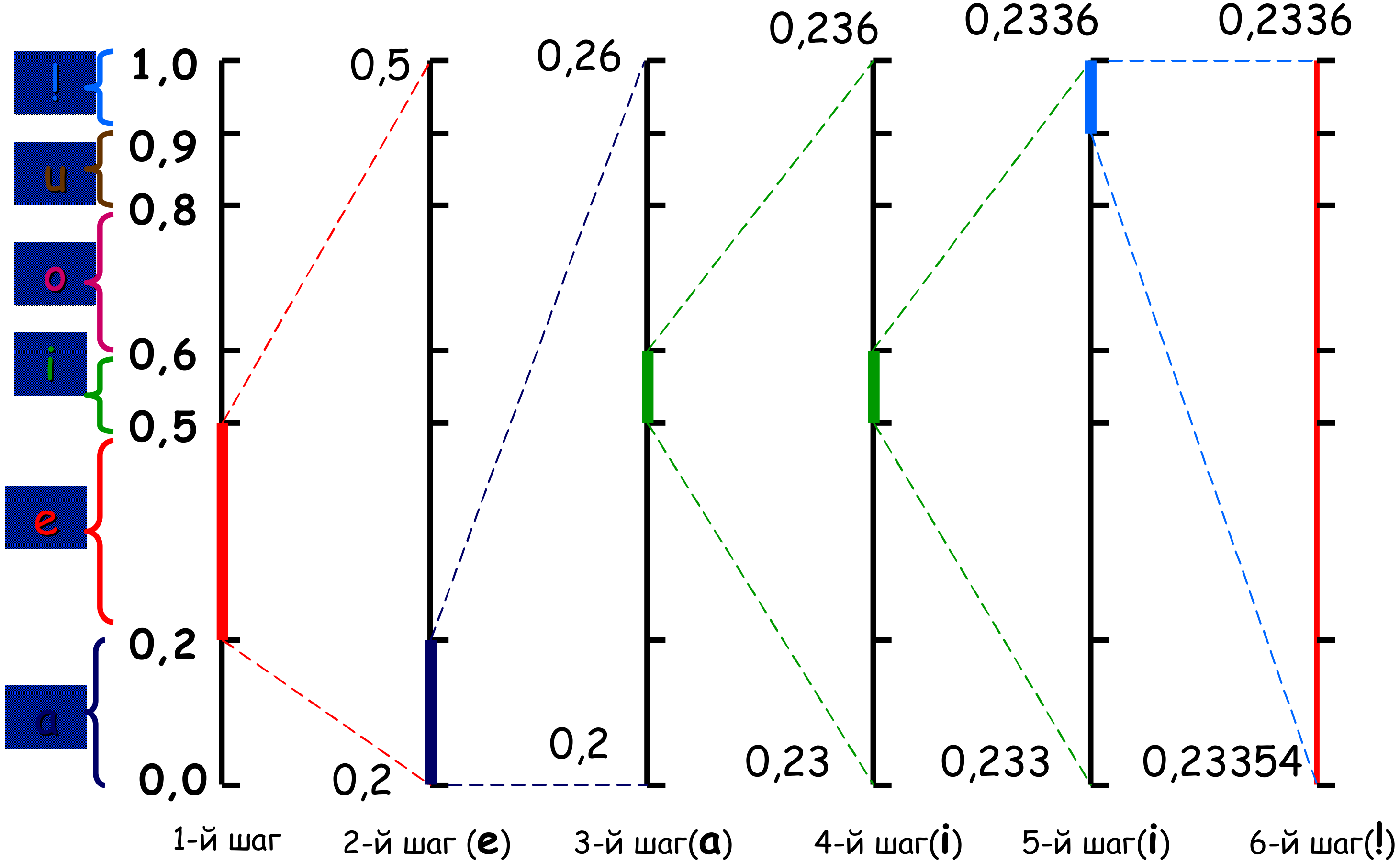
1. Кодеру и декодеру известен начальный интервал [0,1).
2. После просмотра первого символа («е») кодер сужает интервал до [0,2; 0,5), который модель выделяет этому символу.
3. Второй символ («а») сужает этот новый интервал до первой 1/5 его части, т.к. для «а» выделен фиксированный интервал [0,0; 0,2). В результате получается рабочий интервал [0,2; 0,26), т.к. предыдущий интервал имел длину 0,3 единицы и $0{,}3 \times 1/5 = 0{,}06$.
4. Третий символ («i») сужает интервал до [0,23; 0,236), т.к. для «i» выделен фиксированный интервал [0,5; 0,6).
5. Четвертый символ («i») сужает интервал до [0,233; 0,2336).
6. Пятый символ («!») сужает интервал до [0,23354; 0,2336).

Арифметическое кодирование (пример)

- Результаты кодирования можно представить в виде таблицы

Шаг алгоритма кодирования	Границы интервала
1. Начало	[0,0; 1,0)
2. После просмотра «e»	[0,2; 0,5)
3. После просмотра «a»	[0,2; 0,26)
4. После просмотра «i»	[0,23; 0,236)
5. После просмотра «i»	[0,233; 0,2336)
6. После просмотра «!»	[0,23354; 0,2336)

Арифметическое кодирование (пример)*



Арифметическое кодирование (пример)

Декодирование

1. Пусть известно, что [0,23354; 0,2336) — конечный интервал.
2. Тогда первым закодированным символом является «е», т.к. конечный интервал целиком принадлежит интервалу, выделенному этому символу.

Арифметическое кодирование (пример)

Декодирование

1. Для определения остальных символов необходимо повторить действия кодера:

— начальный интервал [0,0; 1,0);

— после декодирования «e» интервал сужается: [0,2; 0,5);

— вторым является символ «a», т.к. «ea» приводит к интервалу [0,2; 0,26), который полностью вмещает интервал [0,23354; 0,2336);

— третьим является символ «i», т.к. «eai» приводит к интервалу [0,23; 0,236);

— четвертым является символ «i», т.к. «eaii» приводит к интервалу [0,233; 0,2336);

— пятый символ — «!», т.к. «eaii!» приводит к интервалу [0,23354; 0,2336).

Арифметическое кодирование (пример)

Выводы

1. Декодер может не знать значения обеих границ итогового интервала. Достаточно знать единственное значение, лежащее внутри него (например, для рассмотренного примера подходят значения 0,23354; 0,23357; 0,2335416 и т.д.).
2. Для завершения процесса декодирования декодер должен вовремя распознать конец текста. Кроме того, одно и то же число 0,0 можно представить как «а», «аа» и т.д. Для устранения неопределенности используют специальные символы (в нашем примере «!»).

Арифметическое кодирование (пример)

Расчет энтропии сообщения

- Для рассмотренного 5-ти символьного текста получим

$$H= -\log0{,}3 - \log0{,}2 - \log0{,}1 - \log0{,}1 - \log0{,}1=$$

$$=- \log0{,}00006 \approx 4{,}22^* \text{ [дит]}.$$

Т.е. в данном примере для кодирования 4-х гласных потребовалось 5 десятичных чисел, что связано с неудачным выбором модели.

Арифметическое кодирование (пример)

Расчет энтропии сообщения

- Если в качестве модели принять {«e» (0,2); «a» (0,2); «i» (0,4); «!» (0,2)}, то получим значение энтропии 2,89(дит), что позволяет закодировать сообщение числом из трех цифр.

Двоичное арифметическое кодирование (Пример)

1. Рассмотрим кодирование дискретной случайной величины (д.с.в.) X, принимающей только два значения 0 и 1 с вероятностями 2/3 и 1/3. Сопоставим значению 0 отрезок [0,0; 2/3), а значению 1 — [2/3; 1,0). Пусть длина блоков д.с.в. X равна трем: n=3. Значения интервалов, их вероятностей и кодов приведем в таблице*.

Двоичное арифметическое кодирование (Пример)

Последовательность и интервал			Вероятность	Код
1→ [2/3; 1.0)	11→ [8/9; 1.0)	111 → [26/27; 1.0)	1/27	31/32=0,11111
		110 → [8/9; 26/27)	2/27	15/16=0,1111
	10→ [2/3; 8/9)	101 → [22/27; 8/9)	2/27	7/8=0,111
		100 → [2/3; 22/27)	4/27	3/4=0,11
0→ [0,0; 2/3)	01→ [4/9; 2/3)	011 → [16/27; 2/3)	2/27	5/8=0,101
		010 → [4/9; 16/27)	4/27	1/2=0,1
	00→ [0,0; 4/9)	001 → [8/27; 4/9)	4/27	3/8=0,011
		000 → [0,0; 8/27)	8/27	1/4=0,01

Двоичное арифметическое кодирование (Пример)

Параметры кода

1. Энтропия

   $H(X)=2/3 \log_2 3/2+1/3 \log_2 3 = \log_2 3-2/3 \approx 0,9183$ [бит/симв.]

2. Средняя длина символа сообщения*

$$L_{cp} = \frac{\sum_{i=1}^{N} n_i p_i}{n}$$

$L_{cp}=(5\cdot 1/27+4\cdot 2/27+3\cdot 2/27+2\cdot 4/27+3\cdot 2/27+1\cdot 4/27+3\cdot 4/27+2\cdot 8/27)/3;$

$L_{cp}=65/81\approx 0,8025$ [бит/симв.]

Двоичное арифметическое кодирование

Замечание

- Среднее количество бит на единицу сообщения получилось меньше энтропии источника. Это связано с тем, что в рассмотренной схеме кодирования не описан код-маркер конца сообщения. Учет дополнительного кода-маркера увеличивает среднее количество бит до значения не меньшего, чем энтропия сообщения.

Двоичное арифметическое кодирование

Алгоритм декодирования

1. По таблице кодирования значений д.с.в. определяется интервал, содержащий текущий код. По этому интервалу однозначно определяется один символ исходного сообщения. Если этот символ — маркер конца сообщения, то конец.
2. Из текущего кода вычитается нижняя граница содержащего его интервала. Полученная разность делится на длину этого же интервала. Результат деления считается новым текущим значением кода. Переход к п. 1.

Двоичное арифметическое кодирование

Декодирование (Пример)

1. Рассмотрим декодирование сообщения 111. Этому сообщению соответствует число $7/8 \in [2/3;\ 1{,}0)$.

   Т.е. первый знак декодируемого сообщения — 1.
2. Из 7/8 вычитается 2/3 и результат делится на 1/3, что дает $5/8 \in [0{,}0;\ 2/3)$.

   Т.е. следующий знак — 0.
3. Вычисляем очередное текущее значение кода $(5/8-0) \times 3/2 = 15/16 \in [2/3;\ 1{,}0)$.

   Т.е. следующий знак —1.

   Все исходное сообщение 101* декодировано.

Двоичное арифметическое кодирование

Замечание

- Т.к. условие остановки вычислений не было определено, то в рассмотренном примере алгоритм декодирования будет работать без остановки. После декодирования последовательности 101 будет получен «следующий символ» 1 и т.д.

Самостоятельная работа

1. Создайте модель дискретного источника текстовых сообщений. Для этого
   - Запишите ваши фамилию, имя и отчество на английском языке (латинскими буквами);
   - Выберите 7 неповторяющихся символов из числа написанных, которые образуют первичный алфавит источника.
   - Каждому символу алфавита источника поставьте в соответствие вероятность, значение которой равно отрицательной степени двойки.
2. Запишите свой пример сообщения, используя полученный алфавит.
3. Закодируйте ваше сообщение двоичным безизбыточным кодом и кодом, соответствующим варианту задания.
4. Рассчитайте энтропию источника, избыточность сообщения и среднюю длину кодового слова (по вариантам п.3).
5. Сделайте выводы по работе.

Самостоятельная работа

Задание

Вариант 1 – код Шеннона-Фано

Вариант 2 – Код Хаффмана

Вариант 3 – Арифметический код