

Прямое и обратное условие Фано

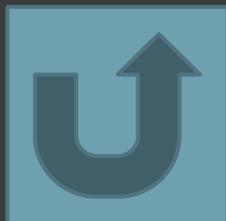
ОДНОЗНАЧНОЕ ДЕКОДИРОВАНИЕ

Учитель информатики и ИКТ
МБОУ СОШ № 7 г. Оха
Сахалинской области
Сергиенко Татьяна Геннадьевна

Задача 1

- Пусть для кодирования фразы «Доброе утро» выбран такой код:

| Д | О | Б | Р | Е | У | Т | Пробел |
|-----|-----|----|---|----|---|----|--------|
| 111 | 000 | 00 | 1 | 01 | 0 | 10 | 11 |



- ⦿ Коды букв «сцепляются» в единую битовую строку и передаются, например, по сети:
- ⦿ Доброе утро →
11100000100001110101000
- ⦿ В пункте назначения возникает проблема – как восстановить исходное сообщение, и возможно ли это.

- ◎ 11100000100001110101000
- ◎ Раскодировать данное сообщение можно разными способами. В том числе предположим, что оно состоит только из букв Р – 1 и У – 0. 🏠 Тогда получим
RRRUUUUUURUUUURRRURURUUU, т.е. бессмысленный набор букв.

- Код называется **однозначно декодируемым**, если любое кодовое сообщение можно расшифровать единственным способом (однозначно).

- ◎ Значит, код



- ◎ не является однозначно декодируемым.

Задача 2

- Равномерные коды. Для той же фразы используем равномерный код:

| Д | О | Б | Р | Е | У | Т | Пробел |
|-----|-----|-----|-----|-----|-----|-----|--------|
| 111 | 000 | 001 | 101 | 011 | 010 | 100 | 110 |

- ◎ Равномерные коды неэкономичны – гораздо длиннее неравномерных. Это приводит к усложнению кодирования, но при этом они раскодируются однозначно, что, естественно, облегчает задачу.

Задача 3

- ⦿ Чтобы сократить длину сообщения, можно попробовать применить неравномерный код, т.е. код, в котором кодовые слова, соответствующие разным символам исходного алфавита, могут иметь разную длину, от одного до нескольких символов.

- Используем следующий код:

| Д | О | Б | Р | Е | У | Т | Пробел |
|----|----|------|-----|------|------|------|--------|
| 01 | 00 | 1011 | 100 | 1010 | 1101 | 1110 | 1111 |

- 010010111000010101111101111010000
- Эта битовая цепочка декодируется однозначно.

- Первая буква - Д (код 01), т.к. ни одно другое кодовое слово не начинается с 01. Вторая буква – О (код 00). Никакое другое слово не начинается с 00. Это же свойство, которое называется условием Фано, выполняется и для кодовых слов других букв.

УСЛОВИЕ ФАНО

- ◎ *Никакое кодовое слово не совпадает с началом другого кодового слова.*
- ◎ *Такие коды называются префиксными (раскодируются с начала сообщения) и декодируются однозначно.*

Задача 4

- Рассмотрим ещё один код:

| Д | О | Б | Р | Е | У | Т | Пробел |
|----|----|------|-----|------|------|------|--------|
| 10 | 00 | 1011 | 001 | 0101 | 1000 | 0111 | 1111 |

- Он не является префиксным, т.к. код буквы Д (10) совпадает с началом кода буквы Б (1011), У(1000) и код буквы О(00) совпадает с началом кода буквы Р (001).

- ⦿ Закодируем наше сообщение:
- ⦿ ДОБРОЕ УТРО → 10 00 1011 001 00
0101 1111 1000 0111 001 00
- ⦿ Начнём декодировать с начала.
Первая – Д, или У, а дальше идут
вообще разные варианты: Р или Б...
Т.е. надо «заглядывать» вперёд, что
очень неудобно.

- ◎ Попробуем раскодировать сообщение с конца – оно однозначно декодируется! Выполняется обратное условие Фано: ***никакое кодовое слово не совпадает с окончанием другого кодового слова.***

- ⦿ Коды, для которых выполняется обратное условие Фано, называются ***постфиксными.***

Сделаем вывод:

- **Сообщение декодируется однозначно, если для используемого кода выполняется прямое или обратное условие Фано.**

Условие Фано - это достаточное, но не необходимое условие однозначной декодируемости

- ⦿ Это значит, что:
- ⦿ - для однозначной декодируемости достаточно выполнения хотя бы одного из двух условий - прямого или обратного.
- ⦿ - могут существовать коды, для которых не выполняется ни прямое, ни обратное условие Фано, но тем не менее обеспечивается однозначное декодирование, т.к. иначе теряется смысл выражения.

Задача 5

- Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г и Д используется неравномерный двоичный код, позволяющий однозначно декодировать полученную двоичную последовательность. Вот этот код: А – 00, Б – 01, В – 100, Г – 101, Д – 110.

- ⦿ Можно ли сократить для одной из букв длину кодового слова так, чтобы код по-прежнему можно было декодировать однозначно? Коды остальных букв меняться не должны. Выберите правильный вариант ответа:
- ⦿ 1) для буквы Д -11 2) это невозможно
- ⦿ 3) для буквы Г - 10 4) для буквы Д -10

РЕШЕНИЕ:

- ⦿ Исходный код – префиксный. Для него выполняется условие Фано – ни один из трёхбитных кодов не начинается ни с 00 (А), ни с 01 (Б). (При этом обратное условие Фано не выполняется – код А (00) совпадает с окончанием В (100), а код Б (01) совпадает с окончанием Г (101).)

- ⦿ Теперь проверим ответы.
- ⦿ Сократим D до 11. Если полученный код нарушит прямое условие Фано, то свойство однозначного декодирования будет нарушено. Но этого не произошло, нет других кодов, начинающихся с 11. Это и есть верное решение. Проверим остальные варианты.

- ⦿ Вариант 2 сразу не рассматриваем – ответ у нас найден. Вариант 3 нарушает прямое условие Фано – с 10 начинается код буквы В (101). Вариант 4 – так же нарушает прямое условие Фано. Т.е. ответ однозначный, других вариантов нет.

Спасибо за внимание!