

Информатика

Для всех специальностей

Институт ИИИБС, кафедра ИИКГ

Колмыкова Оксана Владимировна

Тема 2. Математические основы информатики

МЕТОДЫ И МОДЕЛИ ОЦЕНКИ КОЛИЧЕСТВА ИНФОРМАЦИИ

способы измерения информации:

- объемный,
- энтропийный,
- алгоритмический.

Объемный метод измерения информации

- *Объемный* является самым простым и грубым способом измерения информации.
- Соответствующую количественную оценку информации естественно назвать объемом информации.
- *Объем информации в сообщении* — это количество символов в сообщении.
 - Поскольку, например, одно и то же число может быть записано многими разными способами (с использованием разных алфавитов): «двадцать один», 21, 11001, XXI, то этот способ чувствителен к форме представления (записи) сообщения.
 - В вычислительной технике вся обрабатываемая и хранимая информация вне зависимости от ее природы (число, текст, изображение) представлена в двоичной форме (с использованием алфавита, состоящего всего из двух символов 0 и 1). Такая стандартизация позволила ввести две стандартные единицы измерения: **бит и байт**.
 - Байт — это восемь бит.

Энтропийный подход к измерению информации

- Получатель информации (сообщения) имеет определенные представления о возможных наступлениях некоторых событий.
- Эти представления в общем случае недостоверны и выражаются вероятностями, с которыми он ожидает то или иное событие.
- Общая мера неопределенности (энтропия) характеризуется некоторой математической зависимостью от совокупности этих вероятностей.
- Количество информации в сообщении определяется тем, насколько уменьшится эта мера после получения сообщения.

Алгоритмический метод оценки информации

- Любому сообщению можно приписать количественную характеристику, отражающую сложность (размер) программы, которая позволяет ее произвести.
 - Каждый согласится, что слово $0101\dots01$ сложнее слова $00\dots0$, а слово, где 0 и 1 выбираются из эксперимента — бросания монеты (где 0 — герб, 1 — решка), сложнее обоих предыдущих.
 - Компьютерная программа, производящая слово из одних нулей, крайне проста: печатать один и тот же символ.
 - Для получения $0101\dots01$ нужна чуть более сложная программа, печатающая символ, противоположный только что напечатанному.
 - Случайная, не обладающая ни какими закономерностями последовательность не может быть произведена никакой «короткой» программой. Длина программы, производящей хаотичную последовательность, должна быть близка к длине последней.
- Для определенности задаются некоторой конкретной вычислительной машиной, например машиной Тьюринга, а предполагаемая количественная характеристика — сложность слова (сообщения) определяется как минимальное число внутренних состояний машины Тьюринга, требующиеся для его воспроизведения.

Для обработки данных с помощью средств вычислительной техники они должны быть преобразованы в понятную для ЭВМ форму.

ФОРМЫ ПРЕДСТАВЛЕНИЯ И ПРЕОБРАЗОВАНИЯ

- Формирование представления информации называется ее **кодированием**.
- В более узком смысле под **кодированием** понимается переход от исходного представления информации, удобного для восприятия человеком, к представлению, удобному для хранения, передачи и обработки.
- Обратный переход к исходному представлению называется **декодированием**.

При кодировании информации ставятся следующие цели:

- 1) удобство физической реализации;
- 2) удобство восприятия;
- 3) высокая скорость передачи и обработки;
- 4) экономичность, т.е. уменьшение избыточности сообщения;
- 5) надежность, т.е. защита от случайных искажений;
- 6) сохранность, т.е. защита от нежелательного доступа к информации.

- В отличие от обычной словесной формы, принятой в письменной речи, *информация в памяти ЭВМ записывается в форме цифрового двоичного кода.*
 - Это объясняется тем, что электронные элементы, из которых строится оперативная память, могут находиться только в одном из двух устойчивых состояний, которые можно интерпретировать как 0 или 1.
- Количество информации, которое может помещаться в один элемент памяти (0 или 1), называемое *битом*, очень мало и не несет никакой смысловой нагрузки. Однако если соединить несколько таких элементов в ячейку, то тогда можно сохранить в запоминающем устройстве столько информации, сколько потребуется.
 - Последовательность битов, рассматриваемых аппаратной частью ЭВМ как единое целое, называется *машинным словом*.
- Так как оперативная память ЭВМ состоит из конечной последовательности слов, а слова - из конечной последовательности битов, то объем представляемой в ЭВМ информации ограничен *емкостью памяти*, а числовая информация может быть представлена только с определенной точностью, зависящей от архитектуры памяти данной ЭВМ.

- Для автоматизации работы с данными унифицируют форму представления данных – применяют кодирование
- Кодирование- выражение данных одного типа через данные другого типа.

- В более узком смысле под **кодированием** понимается переход от исходного представления информации, удобного для восприятия человеком, к представлению, удобному для хранения, передачи и обработки.
- Обратный переход к исходному представлению называется ***декодированием***.

При кодировании информации ставятся следующие цели:

- 1) удобство физической реализации;
- 2) удобство восприятия;
- 3) высокая скорость передачи и обработки;
- 4) экономичность, т.е. уменьшение избыточности сообщения;
- 5) надежность, т.е. защита от случайных искажений;
- 6) сохранность, т.е. защита от нежелательного доступа к информации.

Кодирование данных двоичным кодом

- Двоичное кодирование- представление данных последовательностью двух знаков : 0 и 1.
- Двоичные цифры – binary digit – bit (бит)
- Один бит выражает два понятия: 0 и 1 (да и нет, черное и белое)

В качестве эталона меры измерения информации выбран абстрактный объект, который может находиться в одном из 2-х состояний:

ДА - НЕТ, ВКЛ.- ВЫКЛ..

Такой объект содержит информацию в 1 БИТ

8 бит = 1 байт

2^{10} байт = 1 кб

Единицы измерения данных

- 1 Мбайт = 1024 Кбайт = 2^{10} байт
- 1 Гбайт = 1024 Мбайт = 2^{30} байт
- 1 Тбайт = 1024 Гбайт = 2^{40} байт

Кодирование данных двоичным кодом

- 00 01 10 11
- 000 001 010 011 100 101 110 111
- Для кодирования целых чисел от 0 255 достаточно иметь 8 разрядов двоичного кода (8 бит)
- 0000 0000 = 0
- 0000 0001 = 1
- 1111 1111 = 255

СИСТЕМЫ СЧИСЛЕНИЯ

- **Системой счисления** называется совокупность приемов наименования и записи чисел.
- В любой системе счисления для представления чисел выбираются некоторые символы (слова или знаки), называемые **базисными числами**, а все остальные числа получаются в результате каких-либо операций из базисных чисел данной системы исчисления.
- Символы, используемые для записи чисел, могут быть любыми, только они должны быть разными и значение каждого из них должно быть известно.

ПОЗИЦИОННЫЕ СИСТЕМЫ СЧИСЛЕНИЯ

- Система счисления называется ***позиционной***, если значение каждой цифры (ее вес) изменяется в зависимости от ее положения (позиции) в последовательности цифр, изображающих число.

Десятичная позиционная система счисления

- основана на том, что десять единиц каждого разряда объединяются в одну единицу соседнего старшего разряда.
- Таким образом, каждый разряд имеет вес, равный степени 10.
 - Например, в записи числа 343.32 цифра 3 повторена три раза, при этом самая левая цифра 3 означает количество сотен (ее вес равен 10^2); цифра 3, стоящая перед точкой, означает количество единиц (ее вес равен 100), а самая правая цифра 3 — количество десятых долей единицы (ее вес равен 10^{-1}), так что последовательность цифр 343.32 представляет собой сокращенную запись выражения :

$$3 \times 10^2 + 4 \times 10^1 + 3 \times 10^0 + 3 \times 10^{-1} + 2 \times 10^{-2}.$$

- Десятичная запись любого числа X в виде последовательности цифр:

$$a_n a_{n-1} \dots a_1 a_0 a_{-1} \dots a_m \dots$$

основана на представлении этого числа в виде полинома:

$$X = a_n 10^n + a_{n-1} 10^{n-1} + \dots + a_1 10^1 + a_0 10^0 + a_{-1} 10^{-1} + \dots + a_{-m} 10^{-m} \dots,$$

- Число K единиц какого-либо разряда, объединяемых в единицу более старшего разряда, называют **основанием позиционной системы счисления**, а сама система счисления называется **K -ичной**.
 - Например, основанием десятичной системы счисления является число 10;
 - двоичной — число 2;
 - троичной — число 3 и т.д.
- Для записи произвольного числа в K -ичной системе счисления достаточно иметь K разных цифр $a_i, i=1, \dots, K$.
 - Например, в троичной системе счисления любое число может быть выражено посредством цифр 0, 1, 2. Эти цифры служат для обозначения некоторых различных целых чисел, называемых базисными.

СИСТЕМЫ СЧИСЛЕНИЯ:

10 (десятичная): 0,1,2,3,4,5.....9

2 (двоичная): 0,1

16 (шестнадцатиричная): 0...9ABCDEF

10

0

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

2

0000

0001

0010

0011

0100

0101

0110

0111

1000

1001

1010

1011

1100

1101

1110

1111

16

0

1

2

3

4

5

6

7

8

9

A

B

C

D

E

F

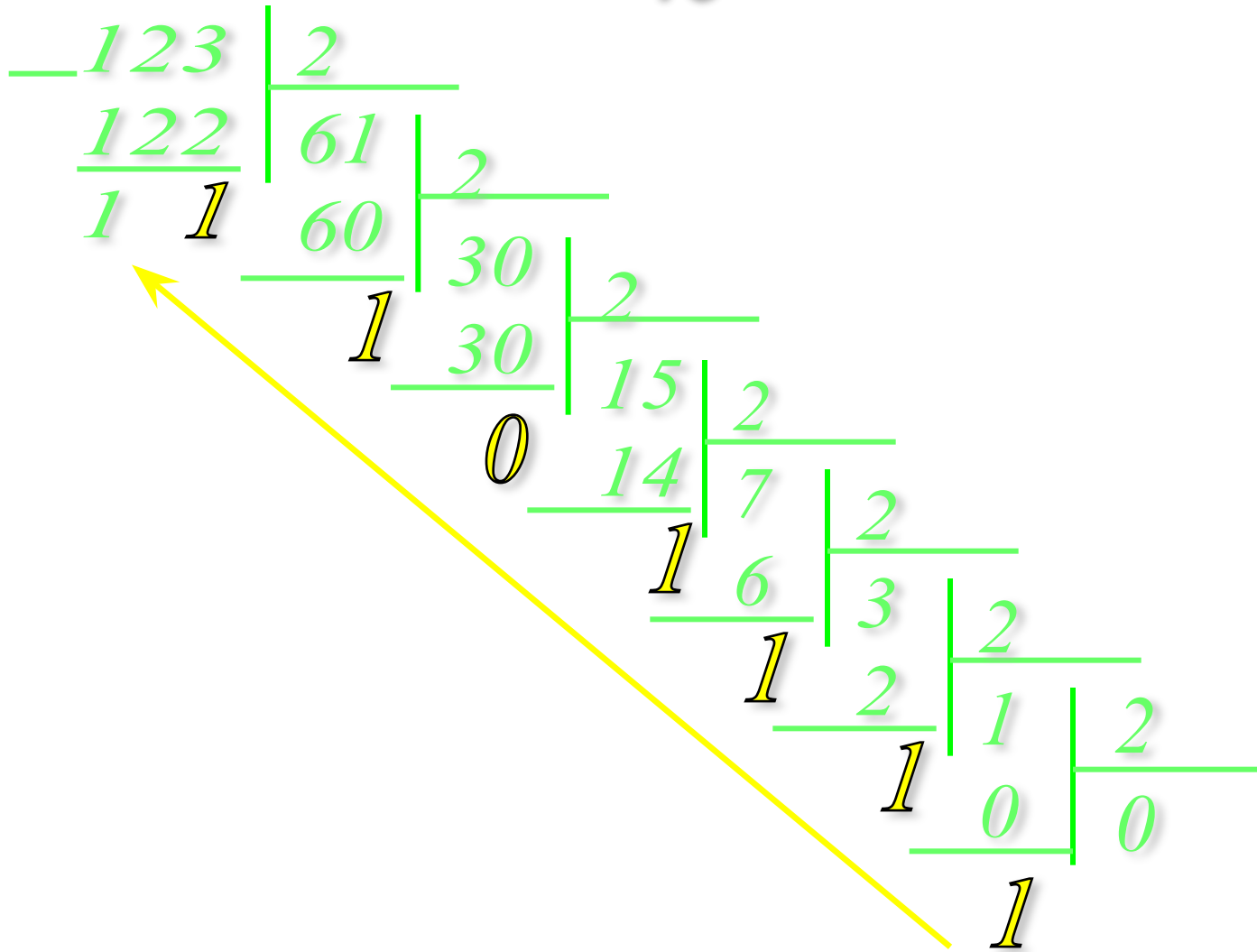


*Правила перевода из
одной системы счисления
в другую*

1) Из 10 в 2

Исходное число в 10 С/С подвергается делению на основание той С/С, в которую осуществляется перевод.

$$123_{10} \longrightarrow 1111011_2$$



$$123_{10} \longrightarrow 7B_{16}$$

$$\begin{array}{r|l} 123 & 16 \\ \hline 112 & 7 & 16 \\ \hline 111 & 0 & 0 \\ \hline & 7 & \end{array}$$

2) 2, 16 в 10

Исходное число раскрывается как
сумма n *-ий соответствующих цифр
исходного числа на основание
исходной С/С в нужной степени.

$$1111011_2 = 1*2^6 + 1*2^5 + 1*2^4 + 1*2^3 + 0*2^2 + 1*2^1 + 1*2^0$$
$$= 123_{10}$$

$$7B_{16} = 7*16^1 + 11*16^0 = 112 + 11 = 123_{10}$$

3) 2 в 16 и 16 в 2

При переводе 2 в 16 исходное число делится на группы по 4 цифры в каждой, и к каждой группе в соответствие ставится символ 16 системы.

10

0

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

2

0000

0001

0010

0011

0100

0101

0110

0111

1000

1001

1010

1011

1100

1101

1110

1111

16

0

1

2

3

4

5

6

7

8

9

A

B

C

D

E

F



$$7B_{16} = 01111011_2$$

$$\boxed{0001} \boxed{1011} \boxed{1011} = 1BB_{16}$$

1 B B

- *Арифметические действия над числами в любой позиционной системе счисления производятся по тем же правилам, что и в десятичной системе, так как все они основываются на правилах выполнения действий над соответствующими полиномами.*
- *При этом нужно только пользоваться теми таблицами сложения и умножения, которые имеют место при данном основании K системы счисления.*

ДВОИЧНАЯ АРИФМЕТИКА





A

B

=

0

0

0

0

1

1

1

0

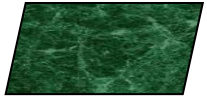
1

1

1

1 0





| A | B | = |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 1 |
| 1 | 0 | 1 |
| 1 | 1 | 0 |



| A | B | = |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

A

B

=

0

0

0

0

1

0

1

0

1

1

1

~~1011~~ | 11
~~11~~ | 11,101...
~~101~~
~~11~~
 100
 11...

Кодирование целых чисел:

От 0 до 255 - 8 бит

от 0 до 65535 - 16 бит

от 0 до 16,5млн - 24 бита

Для кодирования
действительных чисел
используют **80**-разрядное
кодирование.

При этом число предварительно
преобразуется в
НОРМАЛИЗОВАННУЮ ФОРМУ:

$$3,1415926 = 0,31415926 * 10^1$$

$$300000 = 0,3 * 10^6$$

$$123456789 = 0,123456789 * 10^{10}$$

0, 1 2 3 4 5 6 7 8 9 * 10

Мантисса

Характеристика

Кодирование текстовых данных

- Двоичных код используют при кодировании текста, когда каждому символу алфавита сопоставляется определенное число.

Кодирование текстовых данных:

При кодировании текстовой информации
каждому символу соответствует
определенный код

1 буква - **8** разрядов

0 - 31 (первые **32**) - управляющие коды

32 - 127 - английский алфавит, цифры, знаки препинания. (**ASCII**)

128 - 255 - расширенная часть (национальная система кодирования)

UNICODE - универсальная система кодирования,
основанная на **16**-разрядном кодировании
символов, позволяющих обеспечить
уникальные коды для **65 536** различных символов.

Кодирование графических данных:

Для кодирования графических данных применяется принцип декомпозиции; производные цвета зеленый, красный, синий **(RGB)**

Для каждого составляющего - 8разр.=>
=> для каждой точки - **24** разр.

Обеспечивается **16,5** млн. цветов
(TRUE COLOR) ПОЛНОЦВЕТНЫЙ

Дополнительные цвета:

Голубой (С), пурпурный (М), желтый (Y), черный (К) **(CMYK)**

Кодирование звуковых данных:



Метод FM



Метод Wave-Table

Использование материалов презентации

Использование данной презентации, может осуществляться только при условии соблюдения требований законов РФ об авторском праве и интеллектуальной собственности, а также с учетом требований настоящего Заявления.

Презентация является собственностью авторов. Разрешается распечатывать копию любой части презентации для личного некоммерческого использования, однако не допускается распечатывать какую-либо часть презентации с любой иной целью или по каким-либо причинам вносить изменения в любую часть презентации. Использование любой части презентации в другом произведении, как в печатной, электронной, так и иной форме, а также использование любой части презентации в другой презентации посредством ссылки или иным образом допускается только после получения письменного согласия авторов.