

Информация. Двоичное кодирование информации.

1.до20

2.20-36

3.36=>

Презентация создана по учебнику «Информатика и ИКТ» автор Н.Д.Угриновича (10-11 класс)-2002г

Информация. Двоичное кодирование информации

2.1. Понятие «информация» и свойства информации.

Слово «информация» происходит от латинского слова informatio, что в переводе означает сведение, разъяснение, ознакомление.

Понятие «информация» является базовым в курсе информатики, невозможно дать его определение через другие, более «простые» понятия. В геометрии, например, невозможно выразить содержание базовых понятий «точка», «луч», «плоскость» через более простые понятия.

Информация в физике.

- В физике мерой беспорядка, хаоса для термодинамической системы является энтропия системы, тогда как информация (антиэнтропия) является мерой упорядоченности и сложности системы.
- По мере увеличения сложности системы величина энтропии уменьшается, и величина информации увеличивается. Процесс увеличения информации характерен для открытых, обменивающихся веществом и энергией с окружающей средой, саморазвивающихся систем живой природы (белковых молекул, организмов, популяций животных и так далее).
- Таким образом, в физике информация есть антиэнтропия или энтропия с обратным знаком

Информация в биологии.

- В биологии, которая изучает живую природу, понятие «информация» связывается с целесообразным поведением живых организмов. Такое поведение строится на основе получения и использования организмом информации об окружающей среде.
- Понятие «информация» используется также в связи с исследованиями механизмов наследственности. Генетическая информация передается по наследству и хранится во всех клетках живых организмов. Гены представляют собой сложные молекулярные структуры, содержащие информацию о строении живых организмов. Последнее обстоятельство позволило проводить научные эксперименты по клонированию, то есть созданию точных копий организмов из одной клетки.

Информация в кибернетике

- В кибернетике (науке об управлении) понятие «информация» связано с процессами управления в сложных системах (живых организмах или технических устройствах).
- Жизнедеятельность любого организма или нормальное функционирование технического устройства зависит от процессов управления, благодаря которым поддерживаются в необходимых пределах значения их параметров. Процессы управления включают в себя получение, хранение, преобразование и передачу информации.

Свойства информации

- Информация должна быть:
- - понятной;
- - полезной;
- - актуальной;
- - достоверной;
- - полной;
- - точной.

Количество информации как мера Уменьшения неопределенности знаний

- Информация и знания. Человек получает информацию из окружающего мира с помощью органов чувств, анализирует ее и выявляет существенные закономерности с помощью мышления, хранит полученную информацию в памяти.
- Процесс систематического научного познания окружающего мира приводит к накоплению информации в форме знаний (фактов, научных теорий и так далее). Таким образом, с точки зрения процесса познания информация может рассматриваться как знания.

- Процесс познания можно наглядно изобразить в виде расширяющегося круга знания. Вне этого круга лежит область незнания, а окружность является границей между знанием и незнанием. Парадокс состоит в том, что чем бОльшим объемом знания обладает человек (чем шире круг знаний), тем больше он ощущает недостаток знаний (тем больше граница нашего незнания, мерой которого в этой модели является длина окружности) — рис.1.

- Рис.1 НЕЗНАНИЕ



- Так, объем знаний выпускника школы гораздо больше, чем объем знаний первоклассника, однако и граница его незнания существенно больше. (Законы физики)

- Информацию, которую получает человек, можно считать мерой уменьшения неопределенности знаний. Если некоторое сообщение приводит к уменьшению неопределенности наших знаний, то можно говорить, что такое сообщение содержит информацию.
- Например, после сдачи экзамена по информатике вы мучаетесь неопределенностью, вы не знаете какую оценку получили. Наконец, экзаменационная комиссия объявляет результаты экзамена, и вы получаете сообщение, которое приносит полную определенность, теперь вы знаете свою оценку. Происходит переход от незнания к полному знанию, значит, сообщение экзаменационной комиссии содержит информацию. Вопрос только в том, какое кол-во информации содержится в сообщении.

Уменьшение неопределенности знаний

- Пусть у нас имеется монета, которую мы бросаем на ровную поверхность. С равной вероятностью произойдет одно из двух возможных событий — монета окажется в одном из двух положений: «орел» или «решка».
- Можно говорить, что события равновероятны.

- Перед броском существует неопределенность наших знаний (возможны два события), и, как упадет монета, предсказать невозможно.
- После броска наступает полная определенность, так как мы видим (получаем зрительное сообщение), что монета в данный момент находится в определенном положении (например, «орел»).
- Это сообщение приводит уменьшению неопределенности наших знаний в два раза, так как до броска мы имели два вероятных события, а после броска — только одно, то есть в два раза меньше.

Единицы измерения кол-ва информации

- Для количественного выражения любой величины необходимо определить единицу измерения. Так, для измерения длины в качестве единицы выбран метр, для измерения массы - килограмм и так далее.
- Аналогично, для определения количества информации необходимо ввести единицу измерения.
- За единицу количества информации принимается такое количество информации, которое содержит сообщение, уменьшающее неопределенность в два раза. Такая единица названа «бит».

- Если вернуться к опыту с бросанием монеты, то здесь неопределенность как раз уменьшается в два раза и, следовательно, полученное количество информации равно 1 биту.
- Минимальной единицей измерения количества информации является бит, а следующей по величине единицей является байт, причем
 - $1 \text{ байт} = 2^3 \text{ бит} = 8 \text{ бит}$
 - $1 \text{ Кбайт} = 2^{10} \text{ байт} = 1024 \text{ байт};$
 - $1 \text{ Мбайт} = 2^{10} \text{ Кбайт} = 1024 \text{ Кбайт};$
 - $1 \text{ Гбайт} = 2^{10} \text{ Мбайт} = 1024 \text{ Мбайт};$
 - $1 \text{ Тбайт} = 2^{10} \text{ Гбайт} = 1024 \text{ Гбайт};$
 - $1 \text{ Пбайт} = 2^{10} \text{ Тбайт} = 1024 \text{ Тбайт};$

Количество возможных событий и количество информации

- Существует формула, которая связывает между собой количество возможных событий N и количество информации I : $N=2^I$ (1)
- По этой формуле можно легко определить количество возможных событий, если известно количество информации.
- Например, если мы получили 4 бита информации, то количество возможных событий составляло: $N = 2^4 = 16$

- Наоборот, для определения количества информации, если известно количество событий, необходимо решить уравнение относительно I .
- Например, в игре «Крестики-нолики» на поле 8×8 перед первым ходом существует 64 возможных события (64 различных варианта расположения «крестика»), тогда уравнение принимает вид:
 - $64 = 2^I$.
 - Так как $64 = 2^6$,
 - $I = 6$ битов, то есть количество информации, полученное вторым игроком после первого хода первого игрока, составляет 6 битов.

Алфавитный подход к определению количества информации

- при хранении и передаче информации с помощью технических устройств целесообразно отвлекаться от содержания информации и рассматривать ее как последовательность знаков (букв, цифр, кодов цветов точек изображения и так далее).
- Набор символов знаковой системы (алфавит) можно рассматривать как различные возможные состояния (события). Тогда, если считать, что появление символов в сообщении равновероятно, по формуле $N=2^I$ можно рассчитать, какое количество информации несет каждый символ.
- Так, в русском алфавите, если не использовать букву ё, количество событий (букв) будет равно 32. Тогда:
- $32 = 2^I$, откуда $I = 5$ битов.

- Количество информации в сообщении можно подсчитать, умножив количество информации, которое несет один символ, на количество . СИМВОЛОВ.
- **Количество информации, которое содержит сообщение, закодированное с помощью знаковой системы, равно количеству информации, которое несет один знак, умноженному на количество знаков.**

Формула Шеннона

- Существует множество ситуаций, когда возможные события имеют различные вероятности реализации. Например, если монета несимметрична (одна сторона тяжелее другой), то при ее бросании вероятности выпадения «орла» и «решки» будут различаться.
- Формулу для вычисления количества информации в случае различных вероятностей событий предложил К. Шеннон в 1948 году. В этом случае количество информации определяется по формуле:
$$I = -\sum_{i=1}^N P_i \log_2 P_i \quad (2)$$
- где I — количество информации;
- N — количество возможных событий;
- P_i — вероятность i -го события,

- Например, пусть при бросании несимметричной четырехгранной пирамиды вероятности отдельных событий будут равны:
- $p_1=1/2, p_2=1/4, p_3=1/8, p_4=1/8$.
- тогда кол-во информации, которое мы получим после реализации одного из них, можно рассчитать по формуле:
- $I = -(1/2 \log_2 1/2 + 1/4 \log_2 1/4 + 1/8 \log_2 1/8 + 1/8 \log_2 1/8) = 1/2 + 2/4 + 3/8 + 3/8 = 1,75$ бита.
- Этот подход к определению кол-ва информации называется **вероятностным**.
- Для случая, когда все события равновероятны применима формула:
- $I = \log_2 N$ (3)

- Кол-во информации, которое мы получим при бросании симметричной и однородной четырехгранной пирамиды будет равно:
- $I = \log_2 4 = 2$ бита.
- При бросании несимметричной пирамиды мы получаем меньшее кол-во информации (1,75 бита), чем при бросании симметричной пирамиды (2 бита).
- Кол-во информации, которое мы получаем достигает максимального значения, если события равновероятны.

- На получении максимального количества информации строится выбор оптимальной стратегии в игре «Угадай число», в которой первый участник загадывает целое число (например, 3) из заданного интервала (например, от 1 до 16), а второй — должен «угадать» задуманное число. Если рассмотреть эту игру с информационной точки зрения, то начальная неопределенность знаний для второго участника составляет 16 возможных событий (вариантов загаданных чисел).
- При оптимальной стратегии интервал чисел всегда должен делиться пополам, тогда количество возможных событий (чисел) в каждом из полученных интервалов будет одинаково и отгадывание интервалов равновероятно. В этом случае на каждом шаге ответ первого игрока («Да» или «Нет») будет нести максимальное количество информации (1 бит).

- Как видно из таблицы 1 угадывание числа 3 произошло за четыре шага, на каждом из которых неопределенность знаний второго участника уменьшалась в два раза за счет получения сообщения от первого участника, содержащего 1 бит информации. Таким образом, количество информации, необходимое для отгадывания одного из 16 чисел, составило 4 бита. Таб.1

Вопрос	ответ	Неопределенность 16	Полученное кол-во инф.

Кодирование информации.

Язык как знаковая система

- Естественные языки. Русский-33 знака, английский - 26 знаков, китайский - десятки тысяч знаков.
- Формальные языки. Системы счисления, язык алгебры, языки программирования.
- Представление информации может осуществляться с помощью языков, которые являются знаковыми системами. Каждая знаковая система строится на основе определенного алфавита и правил выполнения операций над знаками.
- Каждая цифра машинного двоичного кода несет в себе 1 бит информации.

Представление информации в живых организмах

- Нейрон находится в 2-х состояниях (0,1)
- Молекула ДНК хранит всю информацию о человеке.
- Кодирование — это операция преобразования знаков или групп знаков одной знаковой системы в знаки или группы знаков другой знаковой системы.



Представление числовой информации с помощью систем счисления.

- Система счисления – знаковая система, в которой числа записываются по определенным правилам с помощью символов некоторого алфавита, называемого цифрами.
- Системы счисления бывают 2-х видов - позиционные и непозиционные.
- В позиционных системах счисления количественное значение цифры зависит от ее позиции в числе (десятичная), а непозиционных – не зависит (римская).

- Каждая позиционная система имеет определенный алфавит цифр и основание.
- В позиционных системах счисления основание системы равно кол-ву цифр (знаков в алфавите) и определяет во сколько раз различаются значения одинаковых цифр, стоящих в соседних позициях числа (таб2).

Таб.2

Система	Основание	Алфавит
Десятичная	10	0123456789
Двоичная	2	01
Восьмеричная	8	01234567
Шестнадцатеричная	16	0123456789AB CDEF

- Десятичная система счисления
- Позиция цифры в числе называется **разрядом**. 123. 3-первый разряд, 2-второй, 1-третий.
- Число 123 записано в свернутой форме.
- Мы умножаем в уме цифры числа на различные степени числа 10.
- В развернутой форме умножение записывается в явной форме.
- $123_{10} = 1 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0$

- Для записи десятичных дробей используются отрицательные значения степеней основания.
- $12,35_{10} = 1*10^1 + 2*10^0 + 3*10^{-1} + 5*10^{-2}$
- В десятичной системе счисления запись числа A_{10} , которое содержит N целых разрядов числа и M дробных разрядов числа выглядит так:
- $A_{10} = a_{n-1} * 10^{n-1} + \dots + a_0 * 10^0 + \dots + a_{-m} * 10^{-m}$, где
- a_i – цифры числа.

Двоичная система счисления

- В двоичной системе счисления основание = 2
- Алфавит состоит из 2-х цифр – 0 и 1.
- $A_2 = a_{n-1} * 2^{n-1} + \dots + a_0 * 2^0 + \dots + a_{-m} * 2^{-m}$
- $A_2 = 101,01 = 1 * 2^2 + 0 * 2^1 + 1 * 2^0 + 0 * 2^{-1} + 1 * 2^{-2}$
- Аналогично для 8-ой и 16-ой систем.

Непозиционные системы счисления

- Самая распространенная непозиционная система счисления является римская (таб.3). Таб.3

Десятичная система	Римская система
1	
5	
10	
50	
100	
500	
1000	

- Если меньшая цифра стоит слева от большей, то она вычитается, если справа, прибавляется.

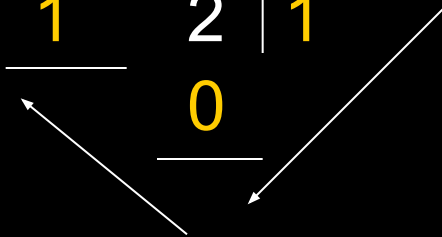
- DCLXVI = + + + +

Перевод чисел в позиционных системах счисления

- Перевод из двоичной в десятичную
- Для перевода из 2-ой, 8-ой, 16-ой системы счисления в 10-ую необходимо записать число в развернутой форме и вычислить его значение
- $101,01_2 = 1*2^2 + 0*2^1 + 1*2^0 + 0*2^{-1} + 1*2^{-2} = 1*4 + 0*2 + 1*1 + 0*1/2 + 1*1/4 = 5,25_{10}$
- Перевод из восьмеричной в десятичную
- $6,5_8 = 6*8^0 + 5*8^{-1} = 6*1 + 5*0,125 = 6,625_{10}$
- Перевод из шестнадцатеричной в десятичную
- $5C_{16} = 5*16^1 + 12*16^0 = 5*16 + 12*1 = 92_{10}$

Перевод чисел из десятичной системы в двоичную, восьмеричную и шестнадцатеричную.

•	5 ₁₀		2	
•	4		2	2
•	1		2	1
•			0	



$$5_{10} = 101_2$$

- Для перевода необходимо число разделить на основание той системы в которую осуществляется перевод, записать полученные остатки в обратной последовательности.

Перевод дробей в двоичную систему.

- Для перевода необходимо умножать дробь до тех пор пока не получится нулевая дробная часть или не будет достигнута требуемая точность. Записать полученные целые части в прямой последовательности.

- $0,75_{10} = 0,11_2$

- $\begin{array}{r} 2 \\ \hline \end{array}$

- $\begin{array}{r} 1,50 \\ \hline \end{array}$

- $\begin{array}{r} 2 \\ \hline \end{array}$

- $\begin{array}{r} 1,00 \\ \hline \end{array}$

- Перевод чисел, содержащих и целую и дробную части осуществляется в 2 этапа. Отдельно переводится целая часть и отдельно – дробная. В итоговой записи целая и дробная часть объединяются (отделяются запятой).
- $5_{10} = 101_2$
- $0,75_{10} = 0,11_2$
- $5,75_{10} = 101,11_2$

Перевод чисел из 2-ой системы счисления в 8-ую, 16-ую и обратно.

- Для перевода целого 2-го числа в 8-ое его нужно разбить на группы по 3 цифры, **справа налево**, а затем преобразовать каждую группу в 8-ую цифру.
- Если в левой группе менее 3 цифр, то необходимо дополнить ее **слева** нулями.
- $11001_2 = 011_001$
- $011 = 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 0 + 2 + 1 = 3$
- $001 = 0 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 0 + 0 + 1 = 1$
- $11001_2 = 31_8$

- Для упрощения перевода используются таблицы преобразования двоичных триад (групп по 3 цифры) в восьмеричные цифры (таб.4).

Таб.4

2-ые триады	000	001	010	011	100	101	110	111
8-ые цифры	0	1	2	3	4	5	6	7

- Для перевода дробного 2-го числа в 8-ое необходимо разбить его на триады **слева направо**. Если в правой группе окажется менее 3 цифр, дополнить ее **справа** нулями.
- $0,1011_2 = 101_100 = 54$

Перевод из 2-ой в 16-ую систему счисления

- Для перевода целого числа необходимо разбить его на группы по 4-е цифры (тетрады), начиная **справа**. Если в левой группе окажется менее 4 цифр, то дополнить ее **слева** нулями.
- Для перевода дробного двоичного числа, необходимо разбить его на тетрады **слева** направо. Если в правой группе окажется менее 4 цифр, дополнить ее **справа** нулями.

- $111011_2 = 0011_1011 = 3B_{16}$

- $0011 = 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 0 + 0 + 2 + 1 = 3$

- $1011 = 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 8 + 0 + 2 + 1 = 11(B)$

- $111011_2 = 3B_{16}$

- Для упрощения перевода используются таблицы преобразования двоичных тетрад (групп по 4 цифры) в шестнадцатеричные цифры таб.5

2-е тетрады	16-е цифры	2-е тетрады	16-е цифры
0000	0	1001	9
0001	1	1010	A
0010	2	1011	B
0011	3	1100	C
0100	4	1101	D
0101	5	1110	E
0110	6	1111	F
0111	7		
1000	8		

Перевод из 8-ой и 16-ой в 2-ую систему.

- Для перевода из 8-ой в 2-ую систему необходимо преобразовать каждую цифру в группу из трех двоичных цифр (триаду), а при преобразовании 16-го числа – в группу из четырех цифр (тетраду).
- $0,47_8 = 0,10111_2$
- $AB_{16} = 10101011_2$

Арифметические операции в позиционных системах счисления

- **Сложение:**

- $0+0=0,$

- $0+1=1,$

- $1+0=1,$

- $1+1=10.$

- $$\begin{array}{r} 111 \\ +11 \\ \hline \end{array}$$

- $$\begin{array}{r} 1010 \\ \hline \end{array}$$

- $$\begin{array}{r} 1010 \\ \hline \end{array}$$

- **Вычитание:**

- $0-0=0$

- $0-1=11$

- $1-0=1$

- $1-1=0$

- $$\begin{array}{r} 101 \\ -11 \\ \hline \end{array}$$

- $$\begin{array}{r} 101 \\ -11 \\ \hline \end{array}$$

- $$\begin{array}{r} 101 \\ -11 \\ \hline \end{array}$$

• Умножение:

- $1*0=0$
- $0*1=0,$
- $0*0=0,$
- $1*1=1.$
- $$\begin{array}{r} 111 \\ *11 \\ \hline 111 \end{array}$$
- $$\begin{array}{r} 111 \\ +111 \\ \hline 10101 \end{array}$$

Деление:

Деление и умножение
аналогично 10-й
системе

$$\begin{array}{r|l} 110 & 11 \\ 11 & 10 \\ \hline & 0 \end{array}$$

Представление чисел в ПК

- Для хранения целых неотрицательных чисел отводится одна ячейка памяти (8 битов).
- Число 10100100_2

1	0	1	0	0	1	0	0
---	---	---	---	---	---	---	---
- Максимальное значение целого неотрицательного числа достигается в случае, когда во всех ячейках 1
- Для n-разрядного представления оно будет =
$$2^n - 1 \quad (4)$$
- Определим диапазон чисел, которые могут храниться в оперативной памяти (целые неотриц).
- $A = 1 \cdot 2^7 + 1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0$
 $= 1 \cdot 2^8 - 1 = 255. \quad (0-255)$ 44

- Для хранения целых чисел со знаком отводится 2 ячейки (16 битов). Левый разряд отводится под знак. (+0,-1). Число $-2006 = -11111010110$

1	0	0	0	0	1	1	1	1	1	0	1	0	1	1	0
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

- Представление положительных чисел в формате «знак-величина» называется **прямым кодом числа**.
- Максимальное положительное число (с учетом выделения одного разряда на знак) для целых чисел со знаком в n-разрядном представлении равно:
$$A = 2^{n-1} - 1 \quad (5)$$
- Для представления отрицательных чисел используется **дополнительный код**.

- Дополнительный код позволяет избежать вычитания, заменив его сложением. Что увеличивает быстродействие процессора.
- Дополнительный код отрицательного числа A , хранящегося в n ячейках, $= 2^n - |A|$ (6)
- Прямой код модуля:
- $-2006 = |-2006_{10}| = 0000011111010110_2$
- Обратный код (инвертирование):
- 1111100000101001 и прибавление единицы
- $+0000000000000001$
- 1111100000101010 — дополнительный код.

- Максимальное значение модуля числа A в n -разрядном представлении = $|A|=2^{n-1}$ (7)
- Минимальное отрицательное число равно
 - $A = -2^{n-1}$ (8)
- Диапазон чисел, которые могут храниться в оперативной памяти в формате длинных целых чисел со знаком (32 бита). 4 ячейки памяти.
- (5). $A=2^{n-1}-1 = A=2^{31}-1 = 2'147'483'647_{10}$
- Минимальное отрицательное целое число =
- (8) $A = -2^{n-1} = -2^{31} = 2'147'483'648_{10}$
- Плюсы чисел с фиксированной запятой: простота представления чисел и операций над ними.

Представление чисел в формате с плавающей запятой.

- Число A может быть представлено в виде:
 - $A = mq^n$, (9)
- где m – мантисса числа;
- q – основание системы счисления;
- n – порядок числа.
- Для единообразия представления чисел с плав. запятой используется нормализованная форма, при которой мантисса отвечает условию:
 - $1/n \leq |m| < 1$

Мантисса должна быть правильной дробью и иметь после запятой цифру, отличную от нуля.

- Запишем число 13,13 в экспоненциальной форме с нормализованной мантиссой:
- $0,1313 \cdot 10^2$, мантисса=0,1313, порядок=2.
- Число в формате с плав.запятой занимает в памяти ПК 4 (обычной точности) или 8 байтов (двойной точности).



- Определим максимальное число и его точность для формата чисел с обычной точностью, если для хранения порядка и его знака отводится 8 разрядов, а для хранения мантиссы и ее знака 24 разряда

Максимальное значение порядка =
 $1111111_2 = 127_{10} \Rightarrow$

Максимальное значение числа =
 $2^{127} = 1,7014118346 \cdot 10^{38}$

Максимальное значение положительной
мантиссы =
 $2^{23} - 1 = \pm 10^7$

Максимальное число обычной точности =
 $1,701411 \cdot 10^{38}$ (кол-во значащих цифр ограничено
7 разрядами).

Кодирование (Смольникова).

ЗАДАНИЕ

- Чт. с.72-82. Зад.: 2.3аб, 2.4.
 - Чт. с.82-94. стр92-вопрос 1,2,3. Зад.:2.6-2.10.
 - Чт. с.103-121.Зад.: 2.19, 2.20.
-
- Презентация создана по учебнику «Информатика и ИКТ» автор Н.Д.Угриновича (10-11 класс)-2002г