# Самостоятельная работа 2.

# Кодирование и обработка числовой информации

**Цель работы:** ознакомиться со способами кодирования числовой информации в компьютере, понимать связь способа кодирования с диапазоном и точностью данных, уметь выбирать способ кодирования, адекватный задаче обработки данных

[Тема 1. Системы счисления. Перевод из одной системы счисления в другую. 1](#_Toc142777618)

[Тема 2. Кодирование, арифметика целых чисел. 8](#_Toc142777619)

[Тема 3. Числа с плавающей запятой. 17](#_Toc142777620)

[Задание 1. Системы счисления 8](#_Toc142777732)

[Задание 2. Прямой, обратный и дополнительный коды 16](#_Toc142777733)

[Задание 3. Числа с плавающей запятой 25](#_Toc142777734)

## Методические указания по выполнению работы

Работа затрагивает 3 темы: 1) системы счисления, 2)кодирование и обработка целых чисел, 3) стандарт IEEE-754 представления чисел с плавающей запятой. К каждой теме приложен теоретический материал, которого, по нашему мнению, вполне достаточно для успешного выполнения заданий.Более полную информацию можно получить из источников, указанных в библиографическом списке. Внимательно изучите теоретический материал.

Оформление. Задание выполняется в тетради, на которой должны быть указаны ФИО студента, группа, номер варианта.

Критерии оценивания. За каждую правильно решенную задачу начисляется 1 балл. Всего в 3-х заданиях 4+2+1=7 задач. Максимальное число баллов – 7.

## Тема 1. Системы счисления. Перевод из одной системы счисления в другую.

Системы счисления (с/с) бывают позиционными и непозиционными.

Пример непозиционной с/с – римская.

**В *непозиционных* системах вес цифры** (т.е. тот вклад, который она вносит в значение числа) **не зависит от ее позиции** в записи числа. Так, в римской системе счисления в числе ХХХII (тридцать два) вес цифры Х в любой позиции равен просто десяти.

**В *позиционных* системах** счисления **вес каждой цифры изменяется в зависимости от ее положения** (позиции) в последовательности цифр, изображающих число. Например, в числе 757,7 первая семерка означает 7 сотен, вторая – 7 единиц, а третья – 7 десятых долей единицы.

Запись числа 757,7 означает сокращенную запись выражения

700 + 50 + 7 + 0,7 = 7•102 + 5•101 + 7•100 + 7•10-1 = 757,7.

Любая позиционная система счисления характеризуется своим **основанием.**

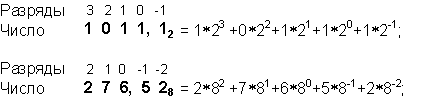
**Определение.** *Основание позиционной системы счисления* — это количество различных знаков или символов, используемых для изображения цифр в данной системе.

Для позиционных систем счисления: 

, (1)

где *a* – разрядная цифра, *q*–основание с/с, *n*–количество целых разрядов, *m* – количество дробных разрядов, 0<=*a* i<=*q*-1

Например:



**Пример 1.** Представить число 124,5378 в виде многочлена.



**Пример 2.** Представить число 21223 в виде многочлена.



**Перевод из одной системы счисления в другую**

Существует несколько способов перевода чисел из одной системы счисления в другую. Например,

1. Перевод целых чисел делением на основание
2. Перевод дробных чисел умножением на основание
3. Использование промежуточной системы счисления

Рассмотрим все варианты подробнее.

# Перевод целых чисел делением на основание

Действия выполняются по правилам исходной системы счисления. Деление происходит на основание новой системы счисления, записанное в исходной системе счисления. Результат записывается перечислением остатков от деления, начиная с последнего.

**Пример.** Перевести  делением на основание.

 Получаем 

**Ответ:**

**Пример.** Перевести  делением на основание.

. Получаем 

**Ответ:**

**Задача 3**. Перевести .

Решение: 10 →2 = 10102. Есть 2 варианта решения;

1) разложением в полином:

10111012=1⋅26+0⋅25+1⋅24+1⋅23+1⋅22+0⋅21+1=64+16+8+4+1=93

2) делением на основание 10 в двоичной системе счисления (10=10102)

|  |  |
| --- | --- |
| 1011101 | 1010 |
| -1010 | 1001=910 |
| 1101  -1010 |  |
| 11=310 |  |

Ответ: 10111012=9310

# Перевод дробных чисел умножением на основание

Действия выполняются по правилам исходной системы счисления. Умножение происходит на основание новой системы счисления, записанное в исходной системе счисления. Результат записывается перечисление целых частей произведений, начиная с первого. Умножение производится либо с заданной точностью, либо до получения нулевого остатка.

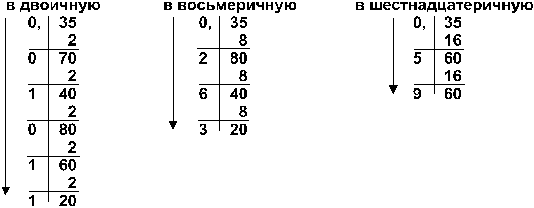
**Пример.** Перевести  умножением на основание.



Стрелка показывает направление записи результата перевода.

**Ответ:**

**Пример:** Перевести число 0,35 из десятичной системы в двоичную, восьмеричную и шестнадцатеричную:



**Ответ:** 0,3510 = 0,010112 = 0,2638 = 0,5916 .

**Пример.** Перевести  умножением на основание.

По окончании деления целые части произведений переводятся в новую с/с и последовательно выписываются, начиная с первого.

**Ответ:**

# Правила округления

При переводе дробной части обычно получаются бесконечные периодические дроби. Округление до *m* знаков после запятой выполняется по обычным правилам округления:

1. вычисления проводятся до *m*+1 цифры после запятой

2. если последняя (отбрасываемая) цифра больше или равна половине основания системы счисления (*q*/2), то к предпоследней цифре добавляется 1, иначе последняя цифра просто отбрасывается.

# Перевод с использованием промежуточной системы счисления

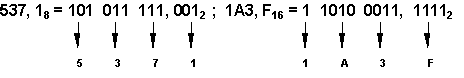
Используется для систем с основанием , k=1,2,3,…, т.е для 2, 8, 16 с/с.

Промежуточной с/с является двоичная.

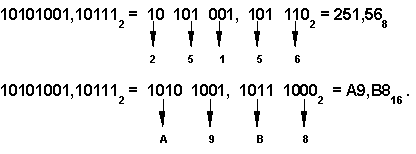
Каждая цифра числа, записанного в исходной с/сс основанием , записывается в двоичной с/с с использованием kколичества разрядов. Для записи числа в новой с/с с основанием  в двоичном представлении выделяется по k1 разрядов, начиная от запятой влево для целой части числа и вправо – для дробной. Затем выделенные части переводятся в числа новой с/с. Запись подряд полученных цифр и будет являться результатом перевода.

Рассмотрим на примере преобразования последовательно.

1. Преобразование числа, записанного в 8 или 16 с/с в двоичную



1. Преобразование двоичного числа в 16 или 8 с/с соответственно.



**Пример.** Перевести 

Исходная с/с - 8, что соответствует , т.е. k=3. Следовательно, цифры исходного числа записываются в двоичной с/с тремя разрядами (триадами).



Новая с/с – 16, что соответствует , т.е. k1=4. В записанном двоичном представлении, начиная от запятой выделяется по 4 разряда (тетрада). Если разрядов для выделения не хватает, то они дописываются нулями. Затем каждая тетрада переводится в цифру 16 с/с. Запись подряд полученных цифр является результатом перевода: 



**Ответ:**.

**Пример.** Перевести 

B E 1 3

BE1316 = 1011 1110 0001 0011 = 001 011 111 000 010 011 = 1370238

1 3 7 0 2 3

**Ответ:**

# Сложение и вычитание чисел в разных системах счисления

Выполняем сложение и вычитание непосредственно в заданной системе счисления.

Существует общий способ – сложение с выделением основания.

Является ручным способом. Для работы неявно используется 10 с/с. Сложение выполняется поразрядно, если при этом число в разряде больше или равно основания целевой системы, то разряд уменьшается на это основание, а 1 переносится старший разряд. Аналогично, если при вычитании получается отрицательное число в разряде, данному разряду прибавляется основание целевой системы, а из старшего разряда делается заем 1.

**Пример 3.** Сложить 59B16+2916= X16

|  |
| --- |
| 5 9 В  + 2 9 |
| (11+9)=20>16  (20-16)=**4**  +1 |
| (9+2+1)=12=**C** |
| 5 C 4 |

*Решение:* Начинаем как обычно, с младших разрядов: B16+916=(11+9)10=2010

Так как 20>16, то выполняем коррекцию: 20-16=4 – окончательное значение,

и добавляем 1 в разряд слева: (9+2+1)10=1210=C16. Старший разряд не изменился.

Проверка: (5⋅162+9⋅16+11)10+(2⋅16+9)10=(1435+41)10=147610

(5⋅162+12⋅16+4)10=147610

*Ответ:* 59816+2916=5C416

|  |
| --- |
| 2 5 4  -1 0 6 |
| (4-6)=-2<0  (-2+7)=**5**  -1 |
| (5-0-1)=**4** |
| (2-1)=**1** |
| 1 4 5 |

**Пример 4.** Вычесть 2547-1067=X7

Начинаем с младших разрядов: 47+67=(4-6)10=-210<0.

Так как получили отрицательное число, то выполняем коррекцию: -2+7=5 – окончательное значение, и занимаем 1 в разряде слева: (5-0-1)10=410=47. В старшем разряде 2-1=1.

Проверка:

2547-1067=(2⋅72+5⋅7+4)10-(72+6)10=(137-55)10=8210

1457=(1⋅72+4⋅7+5)=49+28+5=8210

Ответ: 2547-1067=1147

1. Системы счисления

Вариант выбирается по формуле K=(N-1) modS +1, где N-ваш номер в списке группы, S- число вариантов (10), K-номер варианта (функция A mod B означает целочисленный остаток от деления A на B, например, 4 mod 5=4; 5 mod 5=0; 6 mod 5=1)

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Вариант | | | | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1.Перевести делением на основание | | | | | | | | | |
| 14810→3 | 136910→5 | 10111012  →10 | 11010112  →10 | 11001012  →5 | 100010012  →3 | 138610  →8 | 237610  →4 | 11100102  →12 | 241510  →11 |
| 2.Перевести умножением на основание. Точность 4 знака после запятой(не забывайте про округление) | | | | | | | | | |
| 0,56810  →2 | 0,101112  →10 | 0,024510  →3 | 0,100112  →5 | 0,97110  →6 | 0,110012  →6 | 0,64210  →13 | 0,101112  →12 | 0,74310  →16 | 0,001112  →9 |
| 3.Перевести с использованием промежуточной системы счисления | | | | | | | | | |
| 6DF,E216  →8 | 3AB,7916  →4 | 26,1048  →16 | 577,138  →4 | 3D8,A4216  →4 | 2211,023  →9 | 320,2314  →8 | 863,819  →3 | 754,128  →16 | 1202,23  →9 |
| 4.Сложить два числа, вычесть из первого числа второе в системе счисления (ответ в той же системе счисления) | | | | | | | | | |
| 3EF5,616  1D4,516 | 2301,14  132,24 | B3A8,912  98A,B12 | 2436,327  135,67 | 1042,335  403,245 | 2100,213  12,0223 | 856,479  57,1819 | 9A4,5A11  80A,4411 | 5731,478  237,548 | 3540,056  124,336 |

## Тема 2. Кодирование, арифметика целых чисел.

Представление чисел в компьютере имеет две особенности:

* числа записываются в двоичной системе счисления (в отличие от привычной десятичной);
* для записи и обработки чисел отводится конечное количество разрядов (в привычной арифметике такое ограничение отсутствует).

Если для записи числа в компьютере отводится *N* двоичных разрядов, то, согласно формуле Хартли, число различимых состояний *K* (различных чисел), которые могут быть закодированы *N* разрядами, равно:

|  |  |
| --- | --- |
| *К*=2*N* |  |

Следствие из этой формулы заключается в том, что мы не можем представить в компьютере сколь угодно большие числа со сколь угодно большой точностью.

Способ представления (кодирования) числа определяет диапазон и точность представления чисел в компьютере. Также, способ кодирования влияет на алгоритмы обработки чисел.

Принято различать кодирование чисел:

* целые числа (со знаком и без знака);
* числа с плавающей запятой.

Числа, соответствующие кодам в форме с фиксированной запятой, располагаются на числовой оси с постоянным шагом (см. рис. 1а), а числа в форме с плавающей запятой – с переменным (см. рис. 1б):

X

X

a)

б)

Рис.1. Расположение чисел

а)-с фиксированной запятой; б)-с плавающей запятой

Способ кодирования числа задает соответствие между формой представления числа в памяти компьютера и его значением.

Способ кодирования определяет диапазон и точность представления числа в компьютере.

Далее условимся считать, что *X*(2) – двоичное представление числа *Х* в памяти компьютера, и для кодирования числа *Х* используются *N* двоичных разрядов.

### Беззнаковые целые числа (unsigned)

Форма представления целых чисел без знака соответствует записи числа в двоичной системе счисления.

Значение целого числа без знака определяется формулой:

|  |  |
| --- | --- |
| *X=aN-12N-1+aN-2N-2+…+a0* = *X*(2)*.* |  |

Диапазон представления ***беззнаковых целых чисел*** равен:

|  |  |
| --- | --- |
| 0 *≤ X ≤* (*2N*-1) |  |

Максимальная абсолютная погрешность постоянна во всем диапазоне и равна половине последнего разряда, то есть:

|  |  |
| --- | --- |
| |Δ*X*| *≤ 2-1* |  |

### Целые числа со знаком

Существуют несколько способов кодирования знаковых чисел:

* прямой код;
* обратный;
* дополнительный.

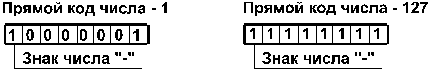
Знак числа в прямом, обратном и дополнительном кодах располагается в старшем разряде двоичного представления числа, и кодируется: 0 – число положительное; 1- отрицательное.

**2.2.1. Положительные числа** в **прямом**, **обратном** и **дополнительном** кодах изображаются ***одинаково***  -  двоичными кодами с цифрой 0 в знаковом разряде. Например:



**Отрицательные числа** в **прямом**, **обратном** и **дополнительном** кодах имеют разное изображение.

2.2.2. **Прямой код (отрицательного числа)**. В знаковый разряд помещается цифра 1, а в разряды цифровой части числа — двоичный код его абсолютной величины. Например:



Для прямого кода справедливы соотношения:

|  |  |
| --- | --- |
|  |  |

И обратно:

|  |  |
| --- | --- |
|  |  |

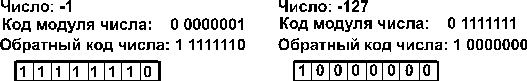
Диапазон представления чисел *в прямом коде* равен:

|  |  |
| --- | --- |
|  |  |

Недостатки прямого кода:

* возможны положительный и отрицательный нули: 00000000 и 10000000;
* алгоритм выполнения операций сложения и умножения зависят от знаков и значений операндов.

2.2.3. **Обратный код (отрицательного числа)**. Получается инвертированием всех цифр двоичного кода абсолютной величины числа, включая разряд знака: нули заменяются единицами, а единицы — нулями. Например:



Для обратного кода справедливы соотношения:

|  |  |
| --- | --- |
|  |  |

Обратное преобразование:

|  |  |
| --- | --- |
|  |  |

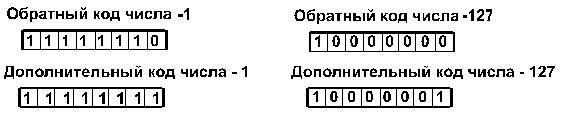
Диапазон представления чисел *в обратном коде* равен:

|  |  |
| --- | --- |
|  |  |

Недостаток обратного кода:

* возможны положительный и отрицательный нули: 00000000 и 11111111;

**2.2.4. Дополнительный код (отрицательного числа)**. Получается из обратного кода прибавлением единицы к его младшему разряду. Например:



Для дополнительного кода справедливы соотношения преобразования:

|  |  |
| --- | --- |
|  |  |
|  |  |

Диапазон представления чисел *в дополнительном коде* равен:

|  |  |
| --- | --- |
|  |  |

Достоинство дополнительного кода:

* операции сложения и вычитания выполняются так же как с беззнаковыми числами, при этом автоматически получается правильный знак результата.

Обычно **отрицательные десятичные числа при вводе в машину автоматически преобразуются в дополнительный двоичный код** и в таком виде хранятся, перемещаются и участвуют в операциях. При выводе таких чисел из машины происходит **обратное преобразование в отрицательные десятичные числа.**

Пример: Представить число +7, -12, -15, -16 в прямом, обратном, дополнительном кодах и со смещением L=32

|  |  |  |  |
| --- | --- | --- | --- |
| **Число** *A* | **Прямой код** *Aпр* | **Обратный код** *Aобр* | **Дополнительный код** *Aдоп* |
| +7 | 0 0000111 | 0 0000111 | 0 0000111 |
| -12 | 1 0001100 | 1 1110011 | 1 1110100 |
| -15 | 1 0001111 | 1 1110000 | 1 1110001 |
| -16 | 1 0010000 | 1 1101111 | 1 1110000 |

Обратите внимание, что прямой, обратный и дополнительный коды положительного числа совпадают, отличаются только коды отрицательного числа. Обратный код получается инвертированием разрядов модуля числа, а дополнительный – получается из обратного добавлением 1.

2.2.6. **Обратное преобразование**, позволяющее получить модуль числа, выполняется по тем же правилам, что и прямое.

При переводе ***из обратного в прямой код*** нужно выполнить *инверсию цифр числа*.

При переводе ***из дополнительного кода в прямой*** выполняем:

1) инверсию цифр числа,

2) добавляем +1 в младший разряд инвертированного числа.

Пример: Представить число -12 в обратном и дополнительном кодах, выполнить обратное преобразование:

|  |  |
| --- | --- |
| Код модуля числа -12 : 0 0001100 (12) | |
| Преобразование в обратный код:  0 0001100 (12)  1 1110011 (-12*обр*) | Преобразование из обратного в прямой код  1 1110011 (-12*обр*)  0 0001100 (12) |
| Преобразование в дополнительный код  0 0001100 (12)  1 1110011 (инверсия разрядов) + 1  1 1110100 (-12доп) | Преобразование из дополнит. в прямой код  1 1110100 (-12*доп*)  0 0001011 (инверсия разрядов)  + 1  0 0001100 (12) |

Теперь, когда мы определили основные формулы соответствия код-значение, рассмотрим алгоритмы выполнения арифметических действия сложения и вычитания с этими кодами

### Арифметические действия над числами со знаком

В большинстве компьютеров операция вычитания не используется. Вместо нее производится сложение обратных или дополнительных кодов уменьшаемого и вычитаемого. Это позволяет существенно упростить конструкцию АЛУ.

Удобно пользоваться ***модифицированными кодами.*** У них знак разряда занимает 2 бита: 00 – число положительное, 11 – число отрицательное, 01 и 10 – запрещенные комбинации знака, они возникают при переполнении разрядной сетки в результате выполнения операции сложения (вычитания)

Преимущества работы с модифицированными дополнительными или обратными кодами:

* + 1. операции сложения (вычитания) выполняются точно так же как с прямыми кодами, при этом знак и значение результата получается автоматически в том же коде;
    2. легко определить факт переполнения по запрещенным комбинациям знака.

**2.3.1. Сложение (вычитание) обратных кодов**

Правила выполнения операции сложения (вычитания) ***в обратном коде***:

1) выполняется сложение (вычитание) двоичных разрядов чисел, в том числе и знаковых, переносы (заем в) из старшего знакового разряда игнорируются;

2) если возникает перенос в знаковый разряд, то к результату операции нужно добавить 1;

3) если возникает заем из знакового разряда, то из результата операции вычитается 1.

Результат операции получается в обратном коде. Если знаковые разряды результата: 00 – ответ положительный, 11 – отрицательный, 01 или 10 – переполнение.

Пример: Даны два числа A и B. Найти сумму A+B и разность A-B в обратном коде

Перенос (заем) в знаковый разряд обозначим символом < (>).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № п/п | Данные | Модифицированный обратный код | Сумма | Разность |
| 1 | A=18,  B=-45 | 00 0010010 (18)  11 1010010 (-45обр) | 00 0010010 (18)  +11 1010010 (-45об)  11 1100100 (-27об)  Переобраз.в прямой код:  **11 0011011 (-27прям)**  Ответ: -27 | 00 0010010 (18)  -11 1010010 (-45об)  00**>**1000000  - 1  **00 0111111 (63)**  Ответ: 63 |
| 2 | A=-18  B=-45 | 11 1101101 (-18обр)  11 1010010 (-45обр) | 11 1101101 (-18обр)  +11 1010010 (-45обр)  11**<**0111111  + 1  **11 1000000 (-63обр)**  Переобраз.в прямой код:  11 0111111 (-63прям)  Ответ: -63 | 11 1101101 (-18обр)  -11 1010010 (-45обр)  **00 0011011 (27)**  Ответ: 27 |
| 3 | A=-118  B=-45 | 11 0001001 (-118обр)  11 1010010 (-45обр) | 11 0001001 (-118обр)  +11 1010010 (-45обр)  **10** 1011011 **(переп-е)**  Ответ: переполнение | 11 0001001 (-118обр)  -11**>**1010010 (-45обр)  11 0110111  - 1  **11 0110110 (-73обр)**  Переобраз.в прямой код:  11 1001001 (-73прям)  Ответ: -73 |
| 4 | A=45  B=118 | 00 0101101 (45)  00 1110110 (118) | 00 0101101 (45)  +00 1110110 (118)  **01** 0100011**(переп-е)**  Ответ: переполнение | 00 0101101 (45)  -00 1110110 (118)  11**>**0110111  - 1  **11 0110110 (-73обр)**  Переобраз.в прямой код:  11 1001001 (-73прям)  Ответ: -73 |

Если результат получен со знаком минус (с "11"), то для проверки вычислений результат преобразуем в прямой код!!!

**2.3.2. Сложение (вычитание) дополнительных кодов**

Правило выполнения операции сложения (вычитания) в дополнительном коде:

1) выполняется сложение (вычитание) двоичных разрядов чисел, в том числе и знаковых, переносы (заем в) из старшего знакового разряда игнорируются.

Результат операции получается в дополнительном коде. Если знаковые разряды результата: 00 – ответ положительный, 11 – отрицательный, 01 или 10 – переполнение.

Пример: Даны два числа A и B. Найти сумму A+B и разность A-B в дополнительном коде

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| № п/п | Данные | Модифицированный обратный код | Сумма  A+B | Разность  A-B |
| 1 | A=18,  B=-45 | 00 0010010 (18)  11 1010011 (-45доп) | 00 0010010 (18)  +11 1010011 (-45доп)  **11 1100101 (-27доп)**  Переобраз.в прямой код:  11 1100101 (-27доп)  11 0011010 (инверс)  + 1  11 0011011 (27прям)  Ответ: -27 | 00 0010010 (18)  -11 1010011 (-45доп)  **00 0111111 (63)**  Ответ: 63 |
| 2 | A=-18  B=-45 | 11 1101110 (-18доп)  11 1010011 (-45доп) | 11 1101110 (-18доп)  +11 1010011 (-45доп)  **11 1000001 (-63доп)**  Переобраз.в прямой код:  11 1000001 (-63доп)  11 0111110 (инверсия)  + 1  11 0111111 (-63прям)  Ответ: -63 | 11 1101110 (-18доп)  -11 1010011 (-45доп)  **00 0011011 (27)**  Ответ: 27 |
| 3 | A=-118  B=-45 | 11 0001010 (-118доп)  11 1010011 (-45доп) | 11 0001010 (-118доп)  +11 1010011 (-45доп)  **10** 1011101 (**переп-е**)  Ответ: переполнение | 11 0001010 (-118доп)  -11 1010011 (-45доп)  **11 0110111 (-73доп)**  Переобраз.в прямой код:  11 0110111 (-73доп)  11 1001000 (инверсия)  + 1  11 1001001 (-73прям)  Ответ: -73 |
| 4 | A=45  B=118 | 00 0101101 (45)  00 1110110 (118) | 00 0101101 (45)  +00 1110110 (118)  **01** 0100011(**переп-е**)  Ответ: переполнение | 00 0101101 (45)  -00 1110110 (118)  **11 0110111 (-73доп)**  Переобраз.в прямой код:  11 0110111 (-73доп)  11 1001000 (инверсия)  + 1  11 1001001 (-73прям)  Ответ: -73 |

Если результат получен со знаком минус (с "11"), то для проверки вычислений результат преобразуем в прямой код!!!

1. Прямой, обратный и дополнительный коды

Преобразовать десятичные числа обратный и дополнительный двоичные коды Выполнить алгебраическое сложение и вычитание в двоичной системе счисления для обратного и дополнительного кодов. Результат представить в прямом коде. Разрядная сетка 8 бит. Указать на переполнение разрядной сетки, если оно есть. Убедиться в правильности полученного результата с учетом диапазона кодов для данного способа кодирования и числа разрядов сетки.

Вариант выбирается по вашему номеру в списке группы

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Вариант | Числа | Вариант | Числа | Вариант | Числа | Вариант | Числа |
| **1** | 9, -2 | **10** | -101, 43 | **19** | -23, -64 | **28** | -100, 126 |
| **2** | -9, 2 | **11** | -127, -1 | **20** | -7, 127 | **29** | -4, 65 |
| **3** | -5, -125 | **12** | -73, 45 | **21** | -12, 105 | **30** | -80, 44 |
| **4** | 5, -125 | **13** | 120, -120 | **22** | -73, 45 | **31** | -22, 100 |
| **5** | 20, -21 | **14** | 120, -122 | **23** | 50, -25 | **32** | -3,95 |
| **6** | -75, -12 | **15** | 16, -32 | **24** | 25, -50 | **33** | 65, 66 |
| **7** | 15, -125 | **16** | 100, -100 | **25** | -75, 44 | **34** | -22, 10 |
| **8** | 120, -21 | **17** | 40, -122 | **26** | 50, -35 | **35** | -64,-95 |
| **9** | -65, -12 | **18** | 19, -32 | **27** | -75, -50 | **36** | 62, 66 |

**Пример выполнения**

|  |  |  |  |
| --- | --- | --- | --- |
| В-24 | Прямой код | Обратный код | Дополнительный |
| 65 | 0 1000001 | 00 1000001 | 00 1000001 |
| 66 | 0 1000010 | 00 1000010 | 00 1000010 |
| (65)+(66) |  | 00 1000001 (65)  +00 1000010 (66)  01 0000011  + 1 (корр)  **01** 0000100 (**переп**) | 00 1000001  +00 1000010  **01** 0000011 (**переп**) |
| (65)-(66) =  - (66-65) |  | 00 1000001  - 00 1000010  11 1111111  - 1 (корр)  **11 1111110** (-1)  Проверка:  **1111110** (-1)  0000001 (1) | 00 1000001  - 00 1000010  **11 1111111** (-1)  Проверка  1111111 (-1доп)  0000000 (инвер)  + 1  0000001 (1) |

## Тема 3. Числа с плавающей запятой.

Любое вещественное число может быть представлено в виде

*X* = *mqp****,***

где ***m*** — мантисса числа, ***р*** — порядок, ***q*** — основание системы счисления.

В настоящее время используется стандарт IEEE-754 кодирования вещественных чисел.

# Формальное представление чисел в стандарте IEEE 754 для любого формата точности

Операции над числами в формате с плавающей запятой имеют существенные отличия от аналогичных операций целочисленной арифметики, поэтому их обычно реализуют с помощью самостоятельного операционного устройства. Как и целочисленное ОПУ, операционное устройство для чисел в формате с плавающей запятой.

Рассмотрим основные положения записи чисел в стандарте IEEE 754.

Чтобы записать число в стандарте IEEE 754 или восстановить его, необходимо знать три параметра:

* S- бит знака (31-й бит)
* E- смещенная экспонента (30-23 биты)
* M - остаток от мантиссы (22-0 биты)

Это целые числа, которые записанные в числе IEEE 754 в двоичном виде.

Приведём формулу для получения десятичного числа из числа IEEE754 одинарной точности:

, **(1)**

где *A* – десятичное число.

Основное применение в технике и программировании получили форматы 32 и 64 бита. Например, в VB используют типы данных single (32 бита) и double (64 бита). В Си аналогично используют float (32 бита) и double (64 бит).

Чтобы представить число в формате IEEE 754 необходимо привести его к двоичному нормализованному виду.

Описание преобразования в 32 битный формат IEEE 754:

1. Число может быть положительным или отрицательным .  
   Поэтому отводится 1 бит для обозначения знака числа:   
   0 ­­- положительное  
   1 - отрицательное  
   Этот самый старший бит в 32 битной последовательности.
2. Далее пойдут биты экспоненты, для этого выделяют 1 байт (8 бит).  
   Экспонента может быть, как и число, со знаком “+” или “-“.  
   Для определения знака экспоненты, чтобы не вводить ещё один бит знака, добавляют смещение к экспоненте в половину байта +127(0111 1111). То есть, если наша экспонента = +7 (+111 в двоичной), то смещенная экспонента = 7+127=134. А если бы, наша экспонента была -7 , то смещенная экспонета=127-7 =120. Смещенную экспоненту записывают в отведенные 8 бит. При этом, когда нам будет нужно получить экспоненту двоичного числа, мы просто отнимем 127 от этого байта.
3. Оставшиеся 23 бита отводят для мантиссы.

Но, у нормализованной двоичной мантиссы первый бит всегда равен 1, так как число лежит в диапазоне 1<=M<2.

Нет смысла записывать единицу в отведенные 23 бита, поэтому в отведенные 23 бита записывают остаток от мантиссы.

(sign bit)

Бит знака

W – бит (bit)

S

E

M

1

b – бит (bit)

n – бит (bit)

(offset exponent)

Смещенная экспонента

(Mantissa)

мантисса

**Рис. 2.**Представление числа в формате IEEE 754

где:

* S - бит знака, если S=0 - положительное число; S=1 - отрицательное число
* E - смещенная экспонента двоичного числа;  
  exp2 = E - (2(b-1) - 1) - экспонента двоичного нормализованного числа с плавающей точкой   
  (2(b-1) -1) - заданное смещение экспоненты (в 32-битном IEEE754 оно равно +127 см.выше)
* M - остаток мантиссы двоичного нормализованного числа с плавающей точкой

**Представление нормализованного числа в формате IEEE 754**

Формула вычисления нормализованных чисел с плавающей точкой, из чисел представленных в стандарте IEEE754:

**(2)**

Используя формулу (1) вычислим формулы для нахождения десятичных чисел из форматов одинарной (32 бита) и двойной (64 бита) точности IEEE 754:

0

22

30

31

1

8

23

32

0

51

32

63

1

11

52

64

**Рис.2**. Формат числа одинарной (32бита-float) и двойной точности (64 бита-double)

### Представление денормализованного числа и других чисел в формате IEEE 754

Если применить формулу (2) для вычисления минимального и максимального числа одинарной точности представленного в IEEE754, то получим следующие результаты:

* 00 00 00 00hex= 5,87747175411144e-39 (минимальное положительное число)
* 80 00 00 00hex=-5,87747175411144e-39 (минимальное отрицательное число)
* 7f ff ff ffhex= 6,80564693277058e+38 (максимальное положительное число)
* ff ff ff ffhex=-6,80564693277058e+38 (максимальное отрицательное число)

Отсюда видно, что невозможно представить число нуль или бесконечность в заданном формате.

#### Поэтому формула (2) не применяется в следующих случаях:

1. Число IEEE754=00 00 00 00hex считается числом +0

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 00hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | |

Число IEEE754=80 00 00 00hex считается числом -0

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 00hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | |

1. Число IEEE754=7F 80 00 00hex считается числом +∞

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 7Fhex | | | | | | | | 80hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | |

Число IEEE754=FF 80 00 00hex считается числом -∞

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| FFhex | | | | | | | | 80hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | |

1. Числа IEEE754=FF (1xxx)X XX XXhex не считается числами (NAN), кроме случая п.2  
   Числа IEEE754=7F (1xxx)X XX XXhex не считается числами (NAN), кроме случая п.2  
   Число представленное в битах с 0...22 могут быть любым числом кроме 0 (т.е.+∞ и -∞ ).

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
|  | | | | | | | |  | | | | | | | |  | | | | | | | |  | | | | | | | |

1. Числа IEEE754=(x000) (0000) (0xxx)X XX XXhex считаются денормализованными числами, за исключением чисел п.1(то есть -0 и +0)

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
|  | | | | | | | |  | | | | | | | |  | | | | | | | |  | | | | | | | |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x | x |
|  | | | | | | | |  | | | | | | | |  | | | | | | | |  | | | | | | | |

### Формула расчета денормализованных чисел:

**(3)**

**Пояснения к исключительным числам:**

1. **- ∞/ +∞** числа, которые больше границ диапазона представления чисел считаются бесконечными.
2. **Не числа NAN(NoaNumbers)**к ним относятся символы, или результаты недопустимых операций.
3. **Денормализованные числа**. Это числа, мантиссы которых лежат в диапазоне 0.1 <= M <1.  
   Денормализованные числа находятся ближе к нулю, чем нормализованные. Денормализованные числа как бы разбивают минимальный разряд нормализованного числа на некоторое подмножество. Сделано так потому, что в технической практике чаще встречаются величины близкие к нулю.

# Диапазон чисел с плавающей запятой

### Вычисление границ диапазона для чисел одинарной точности IEEE 754

Зная формат чисел с одинарной точностью стандарта IEEE 754 можно посчитать границы диапазона представления действительных чисел в этом формате. Для этого подставим значения максимальных и минимальных абсолютных чисел IEEE 754 в формулы (2) и (3).

Минимальное нормализованное число по (абсолютное):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 80 или 00 | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | | 00hex | | | | | | | |

00 80 00 00 = 2-126∙(1+0/223)= 2-126 ≈ 1,17549435∙e-38  
80 80 00 00 = -2-126∙(1+0/223)=-2-126 ≈ -1,17549435∙e-38  
Максимальное денормализованое число (абсолютное):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 80 или 00 | | | | | | | | 7F | | | | | | | | FF | | | | | | | | FF | | | | | | | |

00 7F FF FF = 2-126∙(1-2-23) ≈ 1,17549421∙e-38  
80 7F FF FF = -2-126∙(1-2-23) ≈ -1,17549421∙e-38  
Отсюда видно что минимальное нормализированное число граничит с максимальным денормализированным.

Минимальное денормализованное число (абсолютное):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| x | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| 80 или 00 | | | | | | | | 00 | | | | | | | | 00 | | | | | | | | 01 | | | | | | | |

00 00 00 01 = 2-126∙ 2-23= 2-149 ≈ 1,40129846∙e-45  
80 00 00 01 = -2-126∙ 2-23= 2-149 ≈ -1,40129846∙e-45  
Это число граничит с нулем.

Максимальное нормализированное число (абсолютное):

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 31 | |  |  |  |  |  |  | 22 | |  |  |  |  |  |  | 15 | |  |  |  |  |  |  | 7 |  |  |  |  |  |  | 0 |
| x | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| FFили 7F | | | | | | | | 7F | | | | | | | | FF | | | | | | | | FF | | | | | | | |

7F 7F FF FF = 2127∙(2-2-23) ≈ 3,40282347∙e+38  
FF 7F FF FF = -2127∙(2-2-23) ≈ -3,40282347∙e+38  
Это число граничит с бесконечностью.

### Полный диапазон чисел одинарной точности (32 бит) по стандарту IEEE754

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Отрицательные числа | | | | | | | | | | | | | | Положительные числа | | | | | | | | | | | | | |
| FFFFFFFF | FF800001 | FF800001 | FF800000 | FF7FFFFF | FF800001 | FF800001 | FF800001 | FF800001 | 80800000 | 807FFFFF | FF800001 | 80000001 | 80000000 | 00000000 | 00000001 | FF800001 | 007FFFFF | 00800000 | FF800001 | FF800001 | FF800001 | FF800001 | 7F7FFFFF | 7F800000 | 7F900001 | FF800001 | 7FFFFFFF |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  | —∞ | —3,40282347e+38 |  |  |  |  | —1,17549435e—38 | —1,17549421e—38 |  | —1,40129846e—45 | —0 | +0 | 1,40129846e—45 |  | 1,17549421e—38 | 1,17549435e—38 |  |  |  |  | 3,40282347e+38 | +∞ |  |  |  |
|  | | — НУЛЕВЫЕ ЧИСЛА | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — ДЕНОРМАЛИЗОВАННЫЕ ЧИСЛА (РАССЧИТЫВАЮТСЯ ПО ФОРМУЛЕ №3) | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — НОРМАЛИЗОВАННЫЕ ЧИСЛА (РАССЧИТЫВАЮТСЯ ПО ФОРМУЛЕ №2) | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — БЕСКОНЕЧНОСТЬ | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — НЕ ЧИСЛА (NANs) | | | | | | | | | | | | | | | | | | | | | | | | | |

**Рис.4 .**Диапазон чисел формата одинарной точности (32 бита) представленных по стандарту

IEEE 754

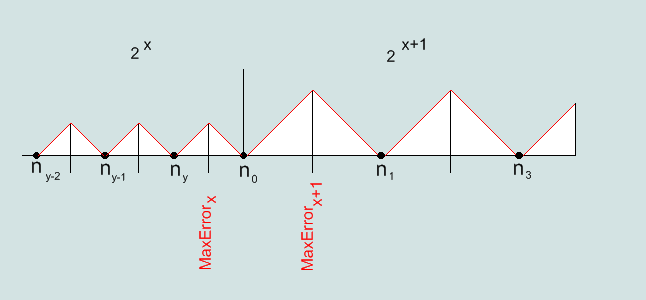
### Полный диапазон чисел двойной точности (64 бит) по стандарту IEEE754

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Отрицательные числа | | | | | | | | | | | | | | Положительные числа | | | | | | | | | | | | | |
| FFFFFFFF FFFFFFFF | FF800001 | FFF00000 00000001 | FFF00000 00000000 | FFFEFFFF FFFFFFFF | FF800001 | FF800001 | FF800001 | FF800001 | 80100000 00000000 | 800FFFFF FFFFFFFF | FF800001 | 80000000 00000001 | 80000000 00000000 | 00000000 00000000 | 00000000 00000001 | FF800001 | 000FFFFF FFFFFFFF | 00100000 00000000 | FF800001 | FF800001 | FF800001 | FF800001 | 7FEFFFFF FFFFFFFF | 7FF00000 00000000 | 7FF00000 00000001 | FF800001 | 7FFFFFFF FFFFFFFF |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| NAN | | | —∞ | —1,797693134862315708e+308 |  |  |  |  | —2,225073858507201383e—308 | —2,225073858507200889e —308 |  | —4,940656458412465441e—324 | —0 | +0 | 4,940656458412465441e—324 |  | 2,225073858507200889e —308 | 2,225073858507201383e—308 |  |  |  |  | 1,797693134862315708e+308 | +∞ | NAN | | |
|  | | — НУЛЕВЫЕ ЧИСЛА | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — ДЕНОРМАЛИЗОВАННЫЕ ЧИСЛА (РАССЧИТЫВАЮТСЯ ПО ФОРМУЛЕ №3) | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — НОРМАЛИЗОВАННЫЕ ЧИСЛА (РАССЧИТЫВАЮТСЯ ПО ФОРМУЛЕ №2) | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — БЕСКОНЕЧНОСТЬ | | | | | | | | | | | | | | | | | | | | | | | | | |
|  | | — НЕ ЧИСЛА (NANs) | | | | | | | | | | | | | | | | | | | | | | | | | |

**Рис.5 .**Диапазон чисел формата двойной точности (64 бита) представленных по стандарту IEEE 754

# Точность представления вещественных чисел в формате IEEE754.

Числа представленные в формате IEEE754 представляют конечное множество, на которое отображается бесконечное множество вещественных чисел. Поэтому исходное число может быть представлено в формате IEEE754 с ошибкой.



**Рис.5** Функция ошибки точности представления числа в IEEE754

Абсолютная максимальная ошибка для числа в формате IEEE754 равна в пределе половине шага чисел. Шаг чисел удваивается с увеличением экспоненты двоичного числа на единицу. То есть, чем дальше от нуля, тем шире шаг чисел в формате IEEE754 по числовой оси.  
Шаг денормализованных чисел равен **2(E-149)** (Single) и **2(E-1074)**(Double). Соответственно предел макс. абсолютной ошибки будет равен 1/2 шага числа:

Δабс=**2(E-150)** (Single) и **2(E-1075)** (Double).

Относительная ошибка в % будет равна:

**δотн=(2(E-150)/X)\*100%**(Single) и **(2(E-1075)/X)\*100%** (Double).  
Шаг нормализованных чисел равен **2(E-150)** (Single) и **2(E-1075)**(Double). Соответственно предел макс. абсолютной ошибки будет равен 1/2 шага числа: **2(E-151)** (Single) и **2(E-1076)** (Double).  
Относительная ошибка в % будет равна: **(2(E-151)/F)\*100%**(Single) и **(2(E-1076)/F)\*100%** (Double).

Максимальная относительная ошибка для денормализованного числа (single/double):

Максимальная относительная ошибка нормализованного числа (single):

Максимальная относительная ошибка нормализованного числа (double):

Расчеты показывают, что основная масса чисел в формате IEEE754 имеет стабильную небольшую относительную погрешность: Максимально возможная относительная погрешность для числа **Single** составляет 2-23\*100% =11,920928955078125e—6 % , что соответсвует точности в **6-7 значащих десятичных цифр**.  
Максимально возможная относительная погрешность для числа **Double** составляет 2-52\*100% =2,2204460492503130808472633361816e— 14 %. Для формата Double имеем **14-15 значащих десятичных цифр.** На границе перехода от нормализованных к денормализованным числам относительная погрешность может возрастать до 50%.

1. Числа с плавающей запятой

*Представить десятичное число X в формате с плавающей запятой стандарта IEEE 754, четырехбайтный формат (float); определить абсолютную |Δабс| и относительную погрешность δmax числа (можно пользоваться калькулятором). Использовать формулы для нормализованных чисел*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Вари-ант | X10 | Вари-  ант | X10 | Вари-  ант | X10 |
| 1 | 143,125 | 11 | 0,00056 |  | 0,00056 |
| 2 | -167,625 | 12 | -5,377 |  | -5,377 |
| 3 | 15,0625 | 13 | -12,18 |  | -12,18 |
| 4 | -80,5625 | 14 | 44,82 |  | 44,82 |
| 5 | 11,38 | 15 | 16,9 |  | 16,9 |
| 6 | 7,15 | 16 | -10,4 |  | -10,4 |
| 7 | -0,021 | 17 | 5643 |  | 5643 |
| 8 | 167235 | 18 | -8,67 |  | -8,67 |
| 9 | -163,3 | 19 | -10,2 |  | -10,2 |
| 10 | -0,0003 | 20 | 0,02 |  | 0,02 |

**Контрольные вопросы**

1. В какой системе счисления представлена информация при ее обработке и хранении на компьютере?
2. Как выглядит прямой код?
3. Что из себя представляет дополнительный код и зачем он нужен?
4. Какие алгоритмы преобразования числа в дополнительный код вы знаете?
5. Какие ошибки могут возникнуть при сложении двух чисел:

а) целых чисел;

б) вещественных чисел?

1. Что описывает стандарт IEEE-754?
2. В каком формате хранятся отрицательные целые числа?
3. Как выгляди нормализованный экспоненциальный вид записи числе по стандарту IEEE-754?
4. Как выглядит денормализованный экспоненциальный вид записи числа?
5. Что такое мантисса, как этот термин обозначается по-английски?
6. Что такое показатель степени числа, как этот термин обозначается по-английски?
7. Что обозначает индекс 16 в записи 110А16?
8. Как записать число -∞ в памяти ЭВМ?
9. Почему в формате с плавающей запятой существуют два способа записи нуля?
10. Какие диапазоны чисел рассчитываются по формуле:
11. Какие диапазоны чисел рассчитываются по формуле:
12. Что есть «не числа»
13. Почему вместо показателя степени хранится смещенный показатель степени?

**Библиографический список**

1. Грошев, А. С. Информатика : учебник / А. С. Грошев, П. В. Закляков. — 3-е изд. — Москва : ДМК Пресс, 2015. — Глава 1, С. 8-22. — URL: https://e.lanbook.com/book/69958 (дата обращения: 12.08.2023). — Режим доступа: для авториз. пользователей.