Системы счисления и двоичное представление информации

Информация и ее кодирование

Различные подходы к определению понятия «информация». Виды информационных процессов. Информационный аспект в деятельности человека

Информация (лат. *informatio* — разъяснение, изложение, набор сведений) — базовое понятие в информатике, которому нельзя дать строгого определения, а можно только пояснить:

- информация — это новые факты, новые знания;
- информация — это сведения об объектах и явлениях окружающей среды, которые повышают уровень осведомленности человека;
- информация — это сведения об объектах и явлениях окружающей среды, которые уменьшают степень неопределенности знаний об этих объектах или явлениях при принятии определенных решений.

Понятие «информация» является общенаучным, т. е. используется в различных науках: физике, биологии, кибернетике, информатике и др. При этом в каждой науке данное понятие связано с различными системами понятий. Так, в физике информация рассматривается как антиэнтропия (мера упорядоченности и сложности системы). В биологии понятие «информация» связывается с целесообразным поведением живых организмов, а также с исследованиями механизмов наследственности. В кибернетике понятие «информация» связано с процессами управления в сложных системах.

Основными социально значимыми свойствами информации являются:

- полезность;
- доступность (понятность);
- актуальность;
- полнота;
- достоверность;
- адекватность.

В человеческом обществе непрерывно протекают информационные процессы: люди воспринимают информацию из окружающего мира с помощью органов чувств, осмысливают ее и принимают определенные решения, которые, воплощаясь в реальные действия, воздействуют на окружающий мир.

Информационный процесс — это процесс сбора (приема), передачи (обмена), хранения, обработки (преобразования) информации.

Сбор информации — это процесс поиска и отбора необходимых сообщений из разных источников (работа со специальной литературой, справочниками;

проведение экспериментов; наблюдения; опрос, анкетирование; поиск в информационно-справочных сетях и системах и т. д.).

Передача информации — это процесс перемещения сообщений от источника к приемнику по каналу передачи. Информация передается в форме сигналов — звуковых, световых, ультразвуковых, электрических, текстовых, графических и др. Каналами передачи могут быть воздушное пространство, электрические и оптоволоконные кабели, отдельные люди, нервные клетки человека и т. д.

Хранение информации — это процесс фиксирования сообщений на материальном носителе. Сейчас для хранения информации используются бумага, деревянные, тканевые, металлические и другие поверхности, кино- и фотопленки, магнитные ленты, магнитные и лазерные диски, флэш-карты и др.

Обработка информации — это процесс получения новых сообщений из имеющихся. Обработка информации является одним из основных способов увеличения ее количества. В результате обработки из сообщения одного вида можно получить сообщения других видов.

Защита информации — это процесс создания условий, которые не допускают случайной потери, повреждения, изменения информации или несанкционированного доступа к ней. Способами защиты информации являются создание ее резервных копий, хранение в защищенном помещении, предоставление пользователям соответствующих прав доступа к информации, шифрование сообщений и др.

Язык как способ представления и передачи информации

Для того чтобы сохранить информацию и передать ее, с давних времен использовались знаки.

В зависимости от *способа восприятия* знаки делятся на:

- зрительные (буквы и цифры, математические знаки, музыкальные ноты, дорожные знаки и др.);
- слуховые (устная речь, звонки, сирены, гудки и др.);
- осязательные (азбука Брайля для слепых, жесты-касания и др.);
- обонятельные;
- вкусовые.

Для долговременного хранения знаки записывают на носители информации.

Для передачи информации используются знаки в виде *сигналов* (световые сигналы светофора, звуковой сигнал школьного звонка и т. д.).

По способу связи между формой и значением знаки делятся на:

- **иконические** — их форма похожа на отображаемый объект (например, значок папки «Мой компьютер» на «Рабочем столе» компьютера);
- **символы** — связь между их формой и значением устанавливается по общепринятому соглашению (например, буквы, математические символы $\int, \leq, \subseteq, \infty$; символы химических элементов).

Для представления информации используются знаковые системы, которые называются **языками**. Основу любого языка составляет **алфавит** — набор символов, из которых формируется сообщение, и набор правил выполнения операций над символами.

Языки делятся на:

- **естественные** (разговорные) — русский, английский, немецкий и др.;
- **формальные** — встречающиеся в специальных областях человеческой деятельности (например, язык алгебры, языки программирования, электрических схем и др.)

Системы счисления также можно рассматривать как формальные языки. Так, десятичная система счисления — это язык, алфавит которого состоит из десяти цифр 0..9, двоичная система счисления — язык, алфавит которого состоит из двух цифр — 0 и 1.

Методы измерения количества информации: вероятностный и алфавитный

Единицей измерения количества информации является *бит*. **1 бит** — это количество информации, содержащейся в сообщении, которое вдвое уменьшает неопределенность знаний о чем-либо.

Связь между количеством возможных событий N и количеством информации I определяется **формулой Хартли**:

$$N = 2^I.$$

Например, пусть шарик находится в одной из четырех коробок. Таким образом, имеется четыре равновероятных события ($N = 4$). Тогда по формуле Хартли $4 = 2^I$. Отсюда $I = 2$. То есть сообщение о том, в какой именно коробке находится шарик, содержит 2 бита информации.

Алфавитный подход

При алфавитном подходе к определению количества информации отвлекаются от содержания (смысла) информации и рассматривают ее как последовательность знаков определенной знаковой системы. Набор символов языка (алфавит) можно рассматривать как различные возможные события. Тогда, если считать, что появление символов в сообщении равновероятно, по формуле Хартли можно рассчитать, какое количество информации несет каждый символ:

$$I = \log_2 N.$$

Например, в русском языке 32 буквы (буква ё обычно не используется), т. е. количество событий будет равно 32. Тогда информационный объем одного символа будет равен:

$$I = \log_2 32 = 5 \text{ битов.}$$

Если N не является целой степенью 2, то число $\log_2 N$ не является целым числом, и для I надо выполнять округление в большую сторону. При решении задач в таком случае I можно найти как $\log_2 N'$, где N' — ближайшая к N степень двойки — такая, что $N' > N$.

Например, в английском языке 26 букв. Информационный объем одного символа можно найти так:

$$N = 26; N' = 32; I = \log_2 N' = \log_2(2^5) = 5 \text{ битов.}$$

Если количество символов алфавита равно N , а количество символов в записи сообщения равно M , то информационный объем данного сообщения вычисляется по формуле:

$$I = M \cdot \log_2 N.$$

Примеры решения задач

Пример 1. Световое табло состоит из лампочек, каждая из которых может находиться в одном из двух состояний («включено» или «выключено»). Какое наименьшее количество лампочек должно находиться на табло, чтобы с его помощью можно было передать 50 различных сигналов?

Решение. С помощью n лампочек, каждая из которых может находиться в одном из двух состояний, можно закодировать 2^n сигналов. $2^5 < 50 < 2^6$, поэтому пяти лампочек недостаточно, а шести хватит.

Ответ: 6.

Пример 2. Метеорологическая станция ведет наблюдения за влажностью воздуха. Результатом одного измерения является целое число от 0 до 100, которое записывается при помощи минимально возможного количества битов. Станция сделала 80 измерений. Определите информационный объем результатов наблюдений.

Решение. В данном случае алфавитом является множество целых чисел от 0 до 100. Всего таких значений 101. Поэтому информационный объем результатов одного измерения $I = \log_2 101$. Это значение не будет целочисленным. Заменим число 101 ближайшей к нему степенью двойки, большей 101. Это число $128 = 2^7$. Принимаем для одного измерения $I = \log_2 128 = 7$ битов. Для 80 измерений общий информационный объем равен:

$$80 \cdot 7 = 560 \text{ битов} = 70 \text{ байтов.}$$

Ответ: 70 байтов.

Вероятностный подход

Вероятностный подход к измерению количества информации применяют, когда возможные события имеют различные вероятности реализации. В этом случае количество информации определяют **по формуле Шеннона**:

$$I = -$$

$$\sum_{i=1}^N p_i \log_2 p_i,$$

где I — количество информации;

N — количество возможных событий;

p_i — вероятность i -го события.

Например, пусть при бросании несимметричной четырехгранной пирамидки вероятности отдельных событий будут равны:

$$p_1 =$$

$$\frac{1}{2}$$

$$, p_2 =$$

$$\frac{1}{4}$$

4

,p3=

1

8

,p4=

1

8

.

Тогда количество информации, которое будет получено после реализации одного из них, можно вычислить по формуле Шеннона:

$I = -$

1

2

$\cdot \log_2$

1

2

+

1

4

$\cdot \log_2$

1

4

+

1

8

$\cdot \log_2$

1

8

+

1
8

·log₂

1
8

)=

14
8

битов = 1.75 бита.

Единицы измерения количества информации

Наименьшей единицей информации является *бит* (англ. *binary digit (bit)* — двоичная единица информации).

Бит — это количество информации, необходимое для однозначного определения одного из двух равновероятных событий. Например, один бит информации получает человек, когда он узнает, опаздывает с прибытием нужный ему поезд или нет, был ночью мороз или нет, присутствует на лекции студент Иванов или нет и т. д.

В информатике принято рассматривать последовательности длиной 8 битов. Такая последовательность называется **байтом**.

Производные единицы измерения количества информации:

1 байт = 8 битов

1 килобайт (Кб) = 1024 байта = 2¹⁰ байтов

1 мегабайт (Мб) = 1024 килобайта = 2^{20} байтов

1 гигабайт (Гб) = 1024 мегабайта = 2^{30} байтов

1 терабайт (Тб) = 1024 гигабайта = 2^{40} байтов

Процесс передачи информации. Виды и свойства источников и приемников информации. Сигнал, кодирование и декодирование, причины искажения информации при передаче

Информация передается в виде сообщений от некоторого *источника* информации к ее *приемнику* посредством *канала связи* между ними.

В качестве источника информации может выступать живое существо или техническое устройство. Источник посылает передаваемое сообщение, которое кодируется в передаваемый *сигнал*.

Сигнал — это материально-энергетическая форма представления информации. Другими словами, **сигнал** — это переносчик информации, один или несколько параметров которого, изменяясь, отображают сообщение. Сигналы могут быть *аналоговыми* (непрерывными) или *дискретными* (импульсными).

Сигнал посылается по каналу связи. В результате в приемнике появляется принимаемый сигнал, который декодируется и становится принимаемым сообщением.

Передача информации по каналам связи часто сопровождается воздействием помех, вызывающих искажение и потерю информации.

Примеры решения задач

Пример 1. Для кодирования букв А, З, Р, О используются двухразрядные двоичные числа 00, 01, 10, 11 соответственно. Этим способом закодировали слово РОЗА и результат записали шестнадцатеричным кодом. Указать полученное число.

Решение. Запишем последовательность кодов для каждого символа слова РОЗА: 10 11 01 00. Если рассматривать полученную последовательность как двоичное число, то в шестнадцатеричном коде оно будет равно: $1011\ 0100_2 = B4_{16}$.

Ответ: $B4_{16}$.

Скорость передачи информации и пропускная способность канала связи

Прием/передача информации может происходить с разной скоростью. Количество информации, передаваемое за единицу времени, есть **скорость передачи информации**, или **скорость информационного потока**.

Скорость выражается в битах в секунду (бит/с) и кратных им Кбит/с и Мбит/с, а также в байтах в секунду (байт/с) и кратных им Кбайт/с и Мбайт/с.

Максимальная скорость передачи информации по каналу связи называется **пропускной способностью канала**.

Примеры решения задач

Пример 1. Скорость передачи данных через ADSL-соединение равна 256000 бит/с. Передача файла через данное соединение заняла 3 мин. Определите размер файла в килобайтах.

Решение. Размер файла можно вычислить, если умножить скорость передачи информации на время передачи. Выразим время в секундах: 3 мин = $3 \cdot 60 = 180$ с. Выразим скорость в килобайтах в секунду: 256000 бит/с = $256000 : 8 : 1024$ Кбайт/с. При вычислении размера файла для упрощения расчетов выделим степени двойки:

Размер файла = $(256000 : 8 : 1024) \cdot (3 \cdot 60) = (2^8 \cdot 10^3 : 2^3 : 2^{10}) \cdot (3 \cdot 15 \cdot 2^2) = (2^8 \cdot 125 \cdot 2^3 : 2^3 : 2^{10}) \cdot (3 \cdot 15 \cdot 2^2) = 125 \cdot 45 = 5625$ Кбайт.

Ответ: 5625 Кбайт.

Представление числовой информации. Сложение и умножение в разных системах счисления

Представление числовой информации с помощью систем счисления

Для представления информации в компьютере используется двоичный код, алфавит которого состоит из двух цифр — 0 и 1. Каждая цифра машинного двоичного кода несет количество информации, равное одному биту.

Система счисления — это система записи чисел с помощью определенного набора цифр.

Система счисления называется **позиционной**, если одна и та же цифра имеет различное значение, которое определяется ее местом в числе.

Позиционной является десятичная система счисления. Например, в числе 999 цифра «9» в зависимости от позиции означает 9, 90, 900.

Римская система счисления является **непозиционной**. Например, значение цифры X в числе XXI остается неизменным при вариации ее положения в числе.

Позиция цифры в числе называется **разрядом**. Разряд числа возрастает справа налево, от младших разрядов к старшим.

Количество различных цифр, употребляемых в позиционной системе счисления, называется ее **основанием**.

Развернутая форма числа — это запись, которая представляет собой сумму произведений цифр числа на значение позиций.

Например: $8527 = 8 \cdot 10^3 + 5 \cdot 10^2 + 2 \cdot 10^1 + 7 \cdot 10^0$.

Развернутая форма записи чисел произвольной системы счисления имеет вид

$-m$
Σ
$i=n-1$

$a_i q^i$,

где X — число;

a — цифры численной записи, соответствующие разрядам;

i — индекс;

m — количество разрядов числа дробной части;

n — количество разрядов числа целой части;

q — основание системы счисления.

Например, запишем развернутую форму десятичного числа 327.46:

$n=3, m=2, q=10$.

$X=$

-2
Σ
$i=2$

$a_i q^i = a_2 \cdot 10^2 + a_1 \cdot 10^1 + a_0 \cdot 10^0 + a_{-1} \cdot 10^{-1} + a_{-2} \cdot 10^{-2} = 3 \cdot 10^2 + 2 \cdot 10^1 + 7 \cdot 10^0 + 4 \cdot 10^{-1} + 6 \cdot 10^{-2}$

Если основание используемой системы счисления больше десяти, то для цифр вводят условное обозначение со скобкой сверху или буквенное обозначение: В — двоичная система, О — восьмеричная, Н — шестнадцатеричная.

Например, если в двенадцатеричной системе счисления $10 = A$, а $11 = B$, то число $7A,5B_{12}$ можно расписать так:

$$7A,5B_{12} = B \cdot 12^{-2} + 5 \cdot 2^{-1} + A \cdot 12^0 + 7 \cdot 12^1.$$

В шестнадцатеричной системе счисления 16 цифр, обозначаемых 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F, что соответствует следующим числам десятичной системы счисления: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15. Примеры чисел: $17D, ECH$; $F12AH$.

Перевод чисел в позиционных системах счисления

Перевод чисел из произвольной системы счисления в десятичную

Для перевода числа из любой позиционной системы счисления в десятичную необходимо использовать развернутую форму числа, заменяя, если это необходимо, буквенные обозначения соответствующими цифрами. Например:

$$1101_2 = 1 \cdot 2^3 + 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 13_{10};$$

$$17D, ECH = 12 \cdot 16^{-2} + 14 \cdot 16^{-1} + 13 \cdot 16^0 + 7 \cdot 16^1 + 1 \cdot 16^2 = 381,921875.$$

Перевод чисел из десятичной системы счисления в заданную

Для преобразования целого числа десятичной системы счисления в число любой другой системы счисления последовательно выполняют деление нацело на основание системы счисления, пока не получат нуль. Числа, которые возникают как остаток от деления на основание системы, представляют собой последовательную запись разрядов числа в выбранной системе счисления от младшего разряда к старшему. Поэтому для записи самого числа остатки от деления записывают в обратном порядке.

Например, переведем десятичное число 475 в двоичную систему счисления. Для этого будем последовательно выполнять деление нацело на основание новой системы счисления, т. е. на 2:

$$\begin{array}{r}
 475 \overline{) 2} \\
 \underline{1 \ 237} \\
 1 \ 118 \\
 \underline{0 \ 59} \\
 1 \ 29 \\
 \underline{1 \ 14} \\
 0 \ 7 \\
 1 \ 3 \\
 \underline{1 \ 1} \\
 1 \ 0
 \end{array}$$

Читая остатки от деления снизу вверх, получим 111011011.

Проверка:

$$1 \cdot 2^8 + 1 \cdot 2^7 + 1 \cdot 2^6 + 0 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 1 + 2 + 8 + 16 + 64 + 128 + 256 = 475_{10}.$$

Для преобразования десятичных дробей в число любой системы счисления последовательно выполняют умножение на основание системы счисления, пока дробная часть произведения не будет равна нулю. Полученные целые части являются разрядами числа в новой системе, и их необходимо представлять цифрами этой новой системы счисления. Целые части в дальнейшем отбрасываются.

Например, переведем десятичную дробь $0,375_{10}$ в двоичную систему счисления:

	целая часть числа
0,375	→ 0
<hr style="width: 100%; border: 0.5px solid black;"/>	
2	
0,750	→ 0
<hr style="width: 100%; border: 0.5px solid black;"/>	
2	
1,50	→ 1
<hr style="width: 100%; border: 0.5px solid black;"/>	
0,50	
<hr style="width: 100%; border: 0.5px solid black;"/>	
2	→ 1
<hr style="width: 100%; border: 0.5px solid black;"/>	
1,00	

Полученный результат — $0,011_2$.

Не каждое число может быть точно выражено в новой системе счисления, поэтому иногда вычисляют только требуемое количество разрядов дробной части.

Перевод чисел из двоичной системы счисления в восьмеричную и шестнадцатеричную и обратно

Для записи восьмеричных чисел используются восемь цифр, т. е. в каждом разряде числа возможны 8 вариантов записи. Каждый разряд восьмеричного числа содержит 3 бита информации ($8 = 2^3$; $I = 3$).

Таким образом, чтобы из восьмеричной системы счисления перевести число в двоичный код, необходимо каждую цифру этого числа представить триадой двоичных символов. Лишние нули в старших разрядах отбрасываются.

Например:

$$1234,777_8 = 001\ 010\ 011\ 100,111\ 111\ 111_2 = 1\ 010\ 011\ 100,111\ 111\ 111_2;$$

$$1234567_8 = 001\ 010\ 011\ 100\ 101\ 110\ 111_2 = 1\ 010\ 011\ 100\ 101\ 110\ 111_2.$$

При переводе двоичного числа в восьмеричную систему счисления нужно каждую триаду двоичных цифр заменить восьмеричной цифрой. При этом, если необходимо, число выравнивается путем дописывания нулей перед целой частью или после дробной.

Например:

$$1100111_2 = 001\ 100\ 111_2 = 147_8;$$

$$11,1001_2 = 011,100\ 100_2 = 3,44_8;$$

$$110,0111_2 = 110,011\ 100_2 = 6,34_8.$$

Для записи шестнадцатеричных чисел используются шестнадцать цифр, т. е. для каждого разряда числа возможны 16 вариантов записи. Каждый разряд шестнадцатеричного числа содержит 4 бита информации ($16 = 2^4$; $I = 4$).

Таким образом, для перевода двоичного числа в шестнадцатеричное его нужно разбить на группы по четыре цифры и преобразовать каждую группу в шестнадцатеричную цифру.

Например:

$$1100111_2 = 0110\ 0111_2 = 67_{16};$$

$$11,1001_2 = 0011,1001_2 = 3,9_{16};$$

$$110,0111001_2 = 0110,0111\ 0010_2 = 65,72_{16}.$$

Для перевода шестнадцатеричного числа в двоичный код необходимо каждую цифру этого числа представить четверкой двоичных цифр.

Например:

$$1234,AB77_{16} = 0001\ 0010\ 0011\ 0100,1010\ 1011\ 0111\ 0111_2 = 1\ 0010\ 0011\ 0100,1010\ 1011\ 0111\ 0111_2;$$

$$CE4567_{16} = 1100\ 1110\ 0100\ 0101\ 0110\ 0111_2.$$

При переводе числа из одной произвольной системы счисления в другую нужно выполнить промежуточное преобразование в десятичное число. При переходе из восьмеричного счисления в шестнадцатеричное и обратно используется вспомогательный двоичный код числа.

Например, переведем троичное число 211_3 в семеричную систему счисления. Для этого сначала преобразуем число 211_3 в десятичное, записав его развернутую форму:

$$211_3 = 2 \cdot 3^2 + 1 \cdot 3^1 + 1 \cdot 3^0 = 18 + 3 + 1 = 22_{10}.$$

Затем переведем десятичное число 22_{10} в семеричную систему счисления делением нацело на основание новой системы счисления, т. е. на 7:

$$\begin{array}{r} 22 \overline{) 7} \\ \underline{1 \quad 3 \quad 7} \\ \quad 3 \quad 0 \end{array}$$
$$22_{10} = 31_7.$$

Итак, $211_3 = 31_7$.

Примеры решения задач

Пример 1. В системе счисления с некоторым основанием число 12 записывается в виде 110. Указать это основание.

Решение. Обозначим искомое основание n . По правилу записи чисел в позиционных системах счисления $12_{10} = 110_n = 0 \cdot n^0 + 1 \cdot n^1 + 1 \cdot n^2$. Составим уравнение: $n^2 + n = 12$. Найдем натуральный корень уравнения (отрицательный корень не подходит, т. к. основание системы счисления, по определению, натуральное число большее единицы): $n = 3$. Проверим полученный ответ: $110_3 = 0 \cdot 3^0 + 1 \cdot 3^1 + 1 \cdot 3^2 = 0 + 3 + 9 = 12$.

Ответ: 3.

Пример 2. Указать через запятую в порядке возрастания все основания систем счисления, в которых запись числа 22 оканчивается на 4.

Решение. Последняя цифра в записи числа представляет собой остаток от деления числа на основание системы счисления. $22 - 4 = 18$. Найдем делители числа 18. Это числа 2, 3, 6, 9, 18. Числа 2 и 3 не подходят, т. к. в системах счисления с основаниями 2 и 3 нет цифры 4. Значит, искомыми основаниями являются числа 6, 9 и 18. Проверим полученный результат, записав число 22 в указанных системах счисления: $22_{10} = 34_6 = 24_9 = 14_{18}$.

Ответ: 6, 9, 18.

Пример 3. Указать через запятую в порядке возрастания все числа, не превосходящие 25, запись которых в двоичной системе счисления оканчивается на 101. Ответ записать в десятичной системе счисления.

Решение. Для удобства воспользуемся восьмеричной системой счисления. $101_2 = 5_8$. Тогда число x можно представить как $x = 5 \cdot 8^0 + a_1 \cdot 8^1 + a_2 \cdot 8^2 + a_3 \cdot 8^3 + \dots$, где $a_1, a_2, a_3, \dots$ — цифры восьмеричной системы. Искомые числа не должны превосходить 25, поэтому разложение нужно ограничить двумя первыми слагаемыми ($8^2 > 25$), т. е. такие числа должны иметь представление $x = 5 + a_1 \cdot 8$. Поскольку $x \leq 25$, допустимыми значениями a_1 будут 0, 1, 2. Подставив эти значения в выражение для x , получим искомые числа:

$$a_1 = 0; x = 5 + 0 \cdot 8 = 5;$$

$$a_1 = 1; x = 5 + 1 \cdot 8 = 13;$$

$$a_1 = 2; x = 5 + 2 \cdot 8 = 21;$$

Выполним проверку:

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

$$13_{10} = 1101_2;$$

$$21_{10} = 10101_2.$$

Ответ: 5, 13, 21.

Арифметические операции в позиционных системах счисления

Правила выполнения арифметических действий над двоичными числами задаются таблицами сложения, вычитания и умножения.

Сложение	Вычитание	Умножение
$0 + 0 = 0$	$0 - 0 = 0$	$0 \cdot 0 = 0$
$0 + 1 = 1$	$1 - 0 = 1$	$0 \cdot 1 = 0$
$1 + 0 = 1$	$1 - 1 = 0$	$1 \cdot 0 = 0$
$1 + 1 = 10$	$10 - 1 = 1$	$1 \cdot 1 = 1$

Правило выполнения операции сложения одинаково для всех систем счисления: если сумма складываемых цифр больше или равна основанию системы счисления, то единица переносится в следующий слева разряд. При вычитании, если необходимо, делают заем.

Пример выполнения *сложения*: сложим двоичные числа 111 и 101, 10101 и 1111:

$$\begin{array}{r} + \quad 111 \\ + \quad 101 \\ \hline 1100 \end{array} \qquad \begin{array}{r} + \quad 10101 \\ + \quad 1111 \\ \hline 100100 \end{array}$$

Пример выполнения *вычитания*: вычтем двоичные числа 10001 – 101 и 11011 – 1101:

$$\begin{array}{r} - \quad 10001 \\ - \quad 101 \\ \hline 1100 \end{array} \qquad \begin{array}{r} - \quad 11011 \\ - \quad 1101 \\ \hline 1110 \end{array}$$

Пример выполнения *умножения*: умножим двоичные числа 110 и 11, 111 и 101:

$$\begin{array}{r} \times \quad 110 \\ \quad 11 \\ \hline 110 \\ \quad 110 \\ \hline 10010 \end{array}$$

$$\begin{array}{r} \times \quad 111 \\ \quad 101 \\ \hline 111 \\ \quad 000 \\ \quad 111 \\ \hline 100011 \end{array}$$

Аналогично выполняются арифметические действия в восьмеричной, шестнадцатеричной и других системах счисления. При этом необходимо учитывать, что величина переноса в следующий разряд при сложении и заем из старшего разряда при вычитании определяется величиной основания системы счисления.

Например, выполним сложение восьмеричных чисел 36_8 и 15_8 , а также вычитание шестнадцатеричных чисел $9C_{16}$ и 67_{16} :

$$\begin{array}{r} + \quad 36 \\ \quad 15 \\ \hline 53 \end{array}$$

$$\begin{array}{r} - \quad 9C \\ \quad 67 \\ \hline 35 \end{array}$$

При выполнении арифметических операций над числами, представленными в разных системах счисления, нужно предварительно перевести их в одну и ту же систему.

Представление чисел в компьютере

Формат с фиксированной запятой

В памяти компьютера целые числа хранятся в формате с *фиксированной запятой*: каждому разряду ячейки памяти соответствует один и тот же разряд числа, «запятая» находится вне разрядной сетки.

Для хранения целых неотрицательных чисел отводится 8 битов памяти. Минимальное число соответствует восьми нулям, хранящимся в восьми битах ячейки памяти, и равно 0. Максимальное число соответствует восьми единицам и равно

$$1 \cdot 2^7 + 1 \cdot 2^6 + 1 \cdot 2^5 + 1 \cdot 2^4 + 1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 255_{10}.$$

Таким образом, диапазон изменения целых неотрицательных чисел — от 0 до 255.

Для n -разрядного представления диапазон будет составлять от 0 до $2^n - 1$.

Для хранения целых чисел со знаком отводится 2 байта памяти (16 битов). Старший разряд отводится под знак числа: если число положительное, то в знаковый разряд записывается 0, если число отрицательное — 1. Такое представление чисел в компьютере называется **прямым кодом**.

Для представления отрицательных чисел используется **дополнительный код**. Он позволяет заменить арифметическую операцию вычитания операцией сложения, что существенно упрощает работу процессора и увеличивает его

быстродействие. Дополнительный код отрицательного числа A , хранящегося в n ячейках, равен $2^n - |A|$.

Алгоритм получения дополнительного кода отрицательного числа:

1. Записать прямой код числа в n двоичных разрядах.
2. Получить **обратный код числа**. (Обратный код образуется из прямого кода заменой нулей единицами, а единиц — нулями, кроме цифр знакового разряда. Для положительных чисел обратный код совпадает с прямым. Используется как промежуточное звено для получения дополнительного кода.)
3. Прибавить единицу к полученному обратному коду.

Например, получим дополнительный код числа -2014_{10} для шестнадцатиразрядного представления:

Прямой код	Двоичный код числа 2014_{10} со знаковым разрядом	1000011111011110
Обратный код	Инвертирование (исключая знаковый разряд)	1111100000100001
	Прибавление единицы	1111100000100001 0000000000000001 +
Дополнительный код		1111100000100010

При алгебраическом сложении двоичных чисел с использованием дополнительного кода положительные слагаемые представляют в прямом коде, а отрицательные — в дополнительном коде. Затем суммируют эти коды, включая знаковые разряды, которые при этом рассматриваются как старшие разряды. При переносе из знакового разряда единицу переноса отбрасывают. В результате получают алгебраическую сумму в прямом коде, если эта сумма положительная, и в дополнительном — если сумма отрицательная.

Например:

1) Найдем разность $13_{10} - 12_{10}$ для восьмибитного представления. Представим заданные числа в двоичной системе счисления:

$$13_{10} = 1101_2 \text{ и } 12_{10} = 1100_2.$$

Запишем прямой, обратный и дополнительный коды для числа -12_{10} и прямой код для числа 13_{10} в восьми битах:

	13_{10}	-12_{10}
Прямой код	00001101	10001100
Обратный код	—	11110011
Дополнительный код	—	11110100

Вычитание заменим сложением (для удобства контроля за знаковым разрядом условно отделим его знаком «_»):

$$\begin{array}{r}
 0_0001101 \\
 + \\
 1_1110100 \\
 \hline
 10_0000001
 \end{array}$$

Так как произошел перенос из знакового разряда, первую единицу отбрасываем, и в результате получаем 00000001.

2) Найдем разность $8_{10} - 13_{10}$ для восьмибитного представления.

Запишем прямой, обратный и дополнительный коды для числа -13_{10} и прямой код для числа 8_{10} в восьми битах:

	8_{10}	-13_{10}
Прямой код	00001000	10001101

Обратный код	—	11110010
Дополнительный код	—	11110011

Вычитание заменим сложением:

$$\begin{array}{r}
 0_0001000 \\
 + 1_1110011 \\
 \hline
 1_111011
 \end{array}$$

В знаковом разряде стоит единица, а значит, результат получен в дополнительном коде. Перейдем от дополнительного кода к обратному, вычтя единицу:

$$1111011 - 00000001 = 1111010.$$

Перейдем от обратного кода к прямому, инвертируя все цифры, за исключением знакового (старшего) разряда: 10000101. Это десятичное число -5_{10} .

Так как при n -разрядном представлении отрицательного числа A в дополнительном коде старший разряд выделяется для хранения знака числа, минимальное отрицательное число равно: $A = -2^{n-1}$, а максимальное: $|A| = 2^{n-1}$ или $A = -2^{n-1} - 1$.

Определим диапазон чисел, которые могут храниться в оперативной памяти в формате *длинных целых чисел со знаком* (для хранения таких чисел отводится 32 бита памяти). Минимальное отрицательное число равно

$$A = -2^{31} = -2147483648_{10}.$$

Максимальное положительное число равно

$$A = 2^{31} - 1 = 2147483647_{10}.$$

Достоинствами формата с фиксированной запятой являются простота и наглядность представления чисел, простота алгоритмов реализации арифметических операций. Недостатком является небольшой диапазон представимых чисел, недостаточный для решения большинства прикладных задач.

Формат с плавающей запятой

Вещественные числа хранятся и обрабатываются в компьютере в формате *с плавающей запятой*, использующем экспоненциальную форму записи чисел.

Число в экспоненциальном формате представляется в таком виде:

$$A = m \cdot q^n,$$

где m — мантисса числа (правильная отличная от нуля дробь);

q — основание системы счисления;

n — порядок числа.

Например, десятичное число 2674,381 в экспоненциальной форме запишется так:

$$2674,381 = 0,2674381 \cdot 10^4.$$

Число в формате с плавающей запятой может занимать в памяти 4 байта (*обычная точность*) или 8 байтов (*двойная точность*). При записи числа выделяются разряды для хранения знака мантиссы, знака порядка, порядка и мантиссы. Две последние величины определяют диапазон изменения чисел и их точность.

Определим диапазон (порядок) и точность (мантиссу) для формата чисел обычной точности, т. е. четырехбайтных. Из 32 битов 8 выделяется для хранения порядка и его знака и 24 — для хранения мантиссы и ее знака.

Найдем максимальное значение порядка числа. Из 8 разрядов старший разряд используется для хранения знака порядка, остальные 7 — для записи величины порядка. Значит, максимальное значение равно $1111111_2 = 127_{10}$. Так как числа представляются в двоичной системе счисления, то

$$q^n = 2^{127} \approx 1.7 \cdot 10^{38}.$$

Аналогично, максимальное значение мантиссы равно

$$m = 2^{23} - 1 \approx 2^{23} = 2(10 \cdot 2.3) \approx 10002.3 = 10(3 \cdot 2.3) \approx 107.$$

Таким образом, диапазон чисел обычной точности составляет $\pm 1.7 \cdot 10^{38}$.

Кодирование текстовой информации. Кодировка ASCII. Основные используемые кодировки кириллицы

Соответствие между набором символов и набором числовых значений называется **кодировкой символа**. При вводе в компьютер текстовой информации происходит ее двоичное кодирование. Код символа хранится в оперативной памяти компьютера. В процессе вывода символа на экран производится обратная операция — **декодирование**, т. е. преобразование кода символа в его изображение.

Присвоенный каждому символу конкретный числовой код фиксируется в кодовых таблицах. Одному и тому же символу в разных кодовых таблицах могут соответствовать разные числовые коды. Необходимые перекодировки текста обычно выполняют специальные программы-конверторы, встроенные в большинство приложений.

Как правило, для хранения кода символа используется один байт (восемь битов), поэтому коды символов могут принимать значение от 0 до 255. Такие кодировки называют *однобайтными*. Они позволяют использовать 256

символов ($N = 2^1 = 2^8 = 256$). Таблица однобайтных кодов символов называется **ASCII (American Standard Code for Information Interchange — Американский стандартный код для обмена информацией)**. Первая часть таблицы ASCII-кодов (от 0 до 127) одинакова для всех IBM-PC совместимых компьютеров и содержит:

- коды управляющих символов;
- коды цифр, арифметических операций, знаков препинания;
- некоторые специальные символы;
- коды больших и маленьких латинских букв.

Вторая часть таблицы (коды от 128 до 255) бывает различной в различных компьютерах. Она содержит коды букв национального алфавита, коды некоторых математических символов, коды символов псевдографики. Для русских букв в настоящее время используется пять различных кодовых таблиц: **КОИ-8, CP1251, CP866, Mac, ISO**.

В последнее время широкое распространение получил новый международный стандарт **Unicode**. В нем отводится по два байта (16 битов) для кодирования каждого символа, поэтому с его помощью можно закодировать 65536 различных символов ($N = 2^{16} = 65536$). Коды символов могут принимать значение от 0 до 65535.

Примеры решения задач

Пример. С помощью кодировки Unicode закодирована следующая фраза:

Я хочу поступить в университет!

Оценить информационный объем этой фразы.

Решение. В данной фразе содержится 31 символ (включая пробелы и знак препинания). Поскольку в кодировке Unicode каждому символу отводится 2 байта памяти, для всей фразы понадобится $31 \cdot 2 = 62$ байта или $31 \cdot 2 \cdot 8 = 496$ битов.

Ответ: 32 байта или 496 битов.