

Введение в компьютерные науки

ЛЕКТОР К.Т.Н. МОХОВ В.А.

ГЛАВА 1. ХРАНЕНИЕ ДАННЫХ

Раздел 1: Хранение данных

1-2

- ▶ 1.1 Хранение битов
- ▶ 1.2 Основная память
- ▶ 1.3 Массовая память
- ▶ 1.4 Представление информации в виде комбинации двоичных разрядов
- ▶ 1.5 Двоичная система счисления
- ▶ 1.6 Представление целых чисел
- ▶ 1.7 Представление дробных значений
- ▶ 1.8 Универсальные методы сжатия данных
- ▶ 1.9 Ошибки при передаче информации

Биты и битовые шаблоны

1-3

- ▶ **Бит:** Двоичная цифра (0 или 1)
- ▶ Битовые шаблоны используются для представления информации.
 - ▶ Числа
 - ▶ Текстовые символы
 - ▶ Изображения
 - ▶ Звук
 - ▶ и другие...

Булевы (логические) операции

1-4

- ▶ **Булева операция:** Операция, которая манипулирует значениями *истина/ложь*
- ▶ Конкретные булевы операции:
 - ▶ AND
 - ▶ OR
 - ▶ XOR (исключающее ИЛИ)
 - ▶ NOT

Рисунок 1.1 Булевы операции AND, OR и XOR

1-5

а) Операция AND

$$\begin{array}{r} 0 \\ \text{AND } 0 \\ \hline 0 \end{array}$$

$$\begin{array}{r} 0 \\ \text{AND } 1 \\ \hline 0 \end{array}$$

$$\begin{array}{r} 1 \\ \text{AND } 0 \\ \hline 0 \end{array}$$

$$\begin{array}{r} 1 \\ \text{AND } 1 \\ \hline 1 \end{array}$$

б) Операция OR

$$\begin{array}{r} 0 \\ \text{OR } 0 \\ \hline 0 \end{array}$$

$$\begin{array}{r} 0 \\ \text{OR } 1 \\ \hline 1 \end{array}$$

$$\begin{array}{r} 1 \\ \text{OR } 0 \\ \hline 1 \end{array}$$

$$\begin{array}{r} 1 \\ \text{OR } 1 \\ \hline 1 \end{array}$$

в) Операция XOR

$$\begin{array}{r} 0 \\ \text{XOR } 0 \\ \hline 0 \end{array}$$

$$\begin{array}{r} 0 \\ \text{XOR } 1 \\ \hline 1 \end{array}$$

$$\begin{array}{r} 1 \\ \text{XOR } 0 \\ \hline 1 \end{array}$$

$$\begin{array}{r} 1 \\ \text{XOR } 1 \\ \hline 0 \end{array}$$

Вентили

1-6

- ▶ **Вентиль:** устройство для выполнения булевой операции
 - ▶ Чаще реализуются в виде (малых) электронных схем
 - ▶ Представляют собой строительные блоки, из которых конструируются компьютеры
 - ▶ СБИС (сверхбольшие интегральные схемы)

Рисунок 1.2 Схематическое представление вентилей AND, OR, XOR и NOT и таблицы ВХОДНЫХ И ВЫХОДНЫХ ДАННЫХ

1-7

AND



Входы	Выход
0 0	0
0 1	0
1 0	0
1 1	1

OR



Входы	Выход
0 0	0
0 1	1
1 0	1
1 1	1

XOR



Входы	Выход
0 0	0
0 1	1
1 0	1
1 1	0

NOT



Входы	Выход
0	1
1	0

Триггеры

1-8

- ▶ **Триггер:** Схема, построенная из вентилей, которая может хранить один бит.
 - ▶ Один вход используется для установки хранимого значения в 1
 - ▶ Один вход используется для установки хранимого значения в 0
 - ▶ Пока оба входа равны 0, сохраняется последнее записанное значение

Рисунок 1.3 Схема простого триггера

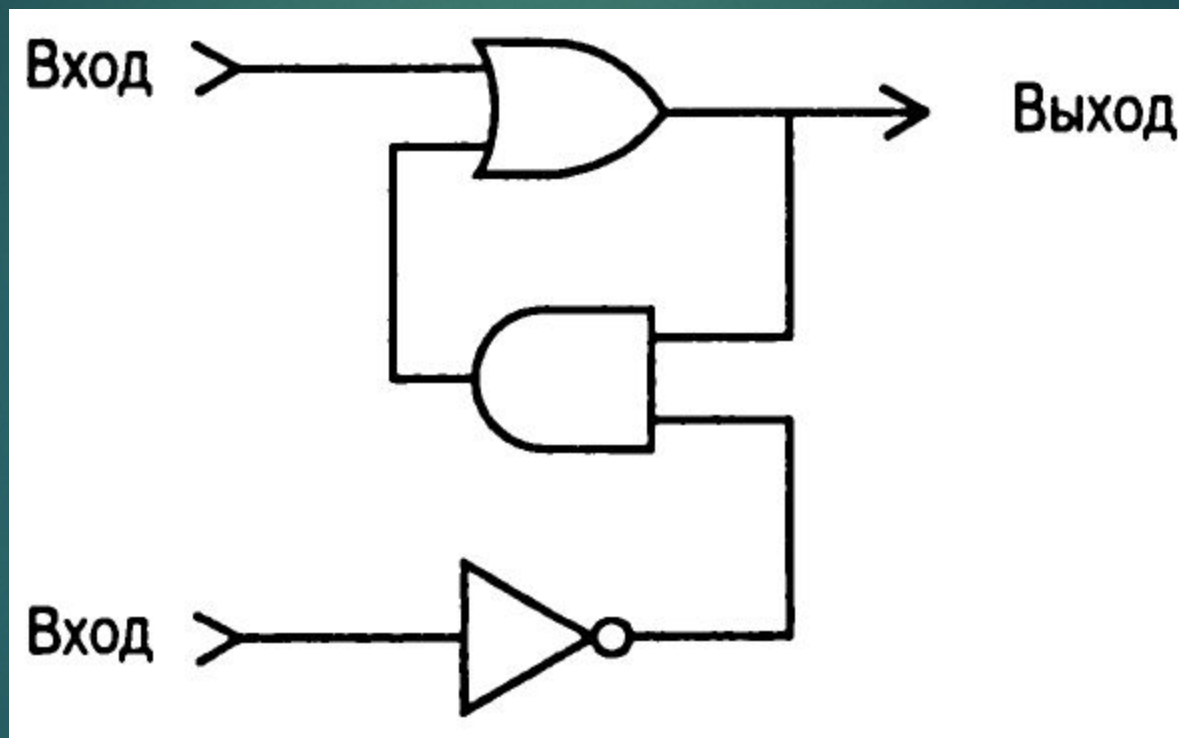
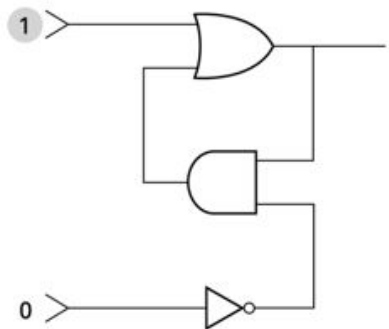


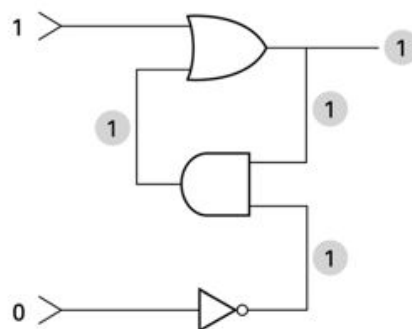
Рисунок 1.4 Установка выходного значения триггера равным 1

1-10

а) Единица поступает на верхний вход



б) Это вызывает появление единицы на выходе вентиля OR, что, в свою очередь, вызывает появление единицы на выходе вентиля AND



в) Наличие единицы на выходе вентиля AND удерживает вентиль OR от изменения его состояния и после снятия единичного сигнала с верхнего входа

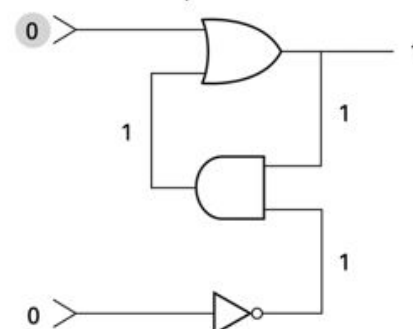
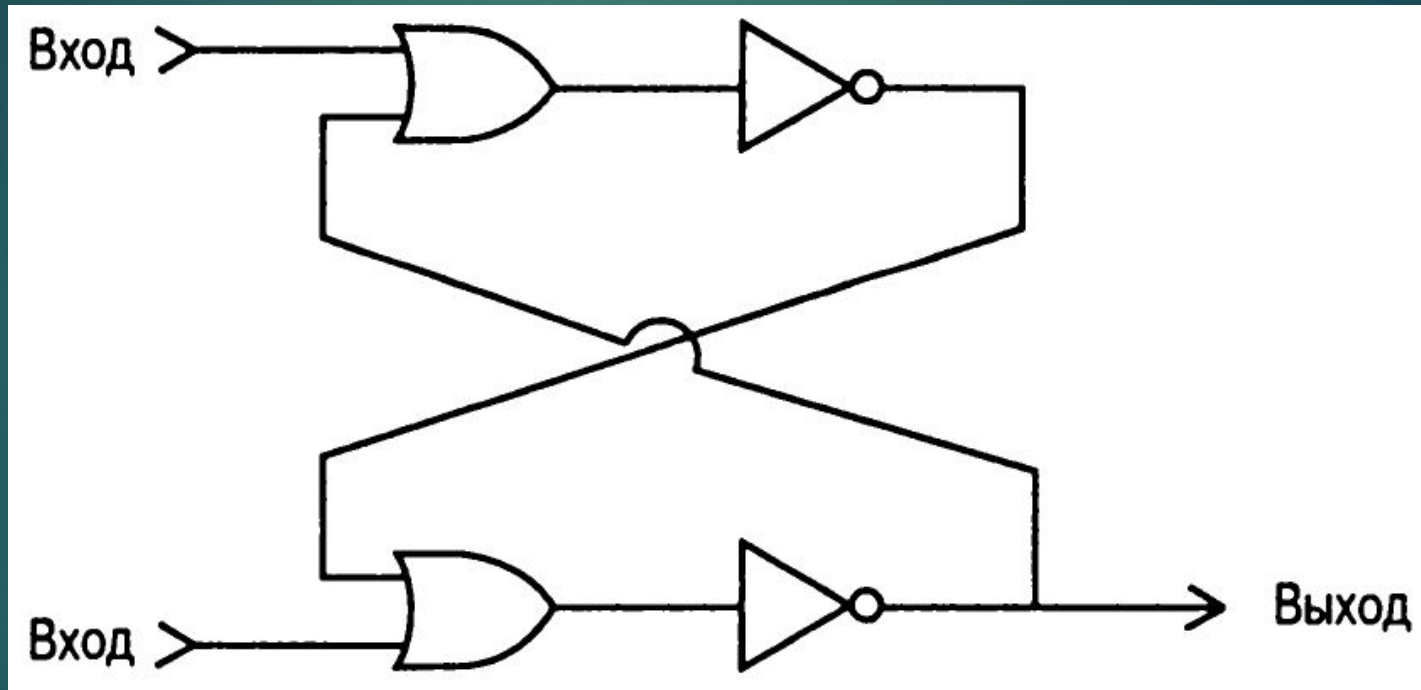


Рисунок 1.5 Ещё один вариант конструкции триггера

1-11



Шестнадцатеричная система счисления

1-12

- ▶ **Шестнадцатеричная нотация:** Сокращенная форма представления длинных комбинаций двоичных разрядов
 - ▶ Разбиваем комбинацию двоичных разрядов на группы по 4 бита в каждой
 - ▶ Представляем каждую группу одним символом
- ▶ Пример: **10100011** превращается в **A3**

Рисунок 1.6 Шестнадцатеричное кодирование

1-13

Комбинация битов	Шестнадцатеричное представление
0000	0
0001	1
0010	2
0011	3
0100	4
0101	5
0110	6
0111	7
1000	8
1001	9
1010	A
1011	B
1100	C
1101	D
1110	E
1111	F

Ячейки основной памяти

1-14

- ▶ **Ячейка:** Элемент основной памяти (как правило 8 бит в составе одного **байта**)
 - ▶ **Наиболее значимый бит:** крайний слева (старший) в битовой цепочке ячейки памяти
 - ▶ **Наименее значимый бит:** крайний справа (младший) в битовой цепочке ячейки памяти

Рисунок 1.7 Организация ячейки памяти размером 1 байт

1-15



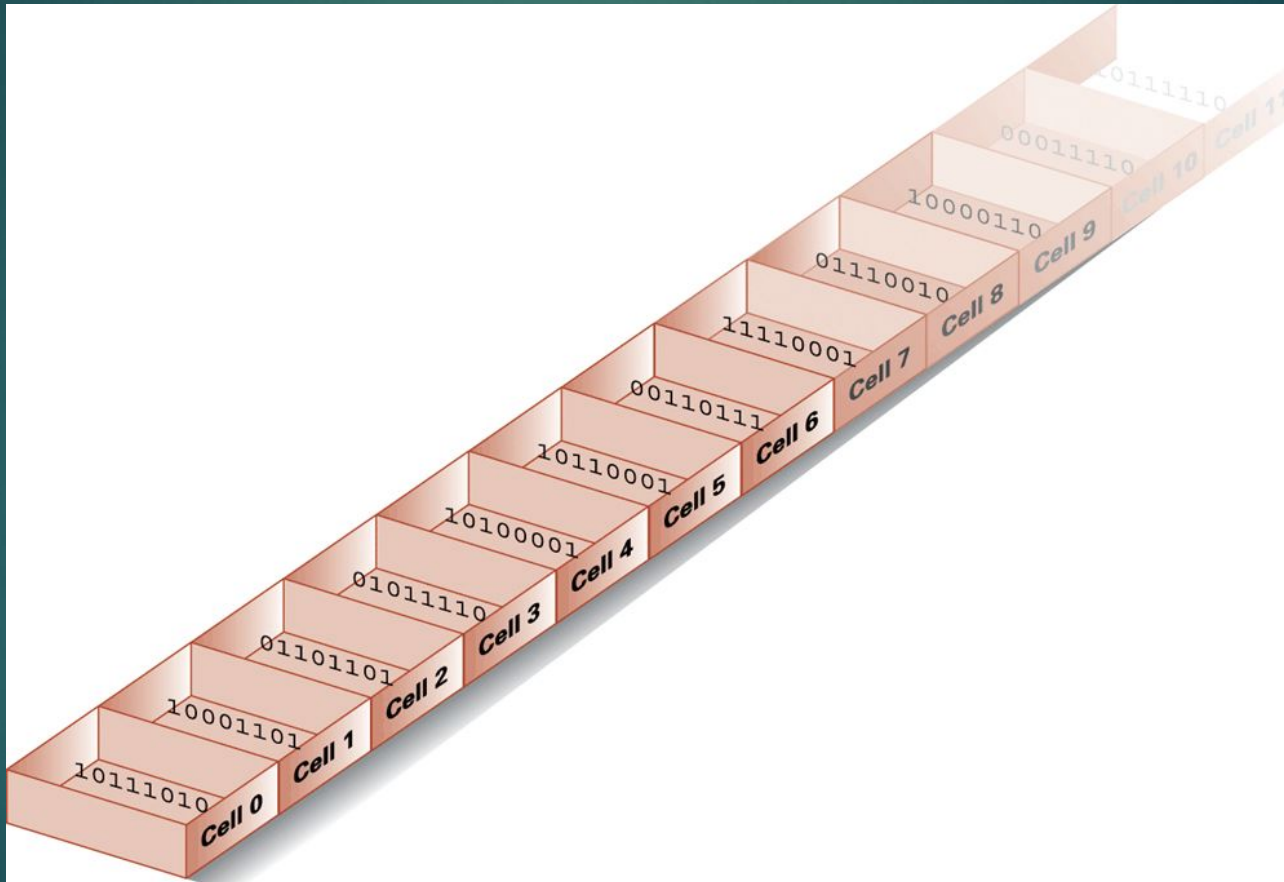
Адресация основной памяти

1-16

- ▶ **Адрес:** «Имя», которое уникальным образом идентифицирует ячейку в основной памяти компьютера
 - ▶ Имена являются числами (номерами)
 - ▶ Нумерация ячеек выполняется последовательно, начиная с нуля
 - ▶ Нумерация упорядочивает ячейки памяти

Рисунок 1.8 Образное представление ячеек памяти, упорядоченных по адресам

1-17



Терминология

1-18

- ▶ **Random Access Memory (RAM):** Память с произвольной выборкой (отдельные ячейки памяти могут быть доступны в любом порядке)
- ▶ **Dynamic Memory (DRAM):** RAM, созданная с использованием технологии динамической (энергозависимой) памяти

Измерение объёма ПАМЯТИ

1-19

- ▶ **Килобайт:** 2^{10} байт = 1024 байт
 - ▶ Пример: 3 Кб = 3 раза по 1024 байт
- ▶ **Мегабайт:** 2^{20} байт = 1,048,576 байт
 - ▶ Пример: 3 Мб = 3 раза по 1,048,576 байт
- ▶ **Гигабайт:** 2^{30} байт = 1,073,741,824 байт
 - ▶ Пример: 3 Гб = 3 раза по 1,073,741,824 байт

Массовая память

1-20

- ▶ Постоянно подключённые (on-line) и автономные (off-line) устройства
- ▶ Обычно больше основной памяти
- ▶ Обычно надёжнее, чем основная память
- ▶ Обычно медленнее основной памяти

Устройства массового хранения

0-21

- ▶ Магнитные
 - ▶ Диски
 - ▶ Ленты
- ▶ Optical Systems
 - ▶ CD
 - ▶ DVD
- ▶ Флэш-технологии
 - ▶ Флэш-устройства
 - ▶ Безопасные цифровые (Secure Digital, SD) карты памяти

Рисунок 1.9 Запоминающее устройство на магнитном диске

1-22



Рисунок 1.10 Запоминающее устройство на магнитной ленте

1-23

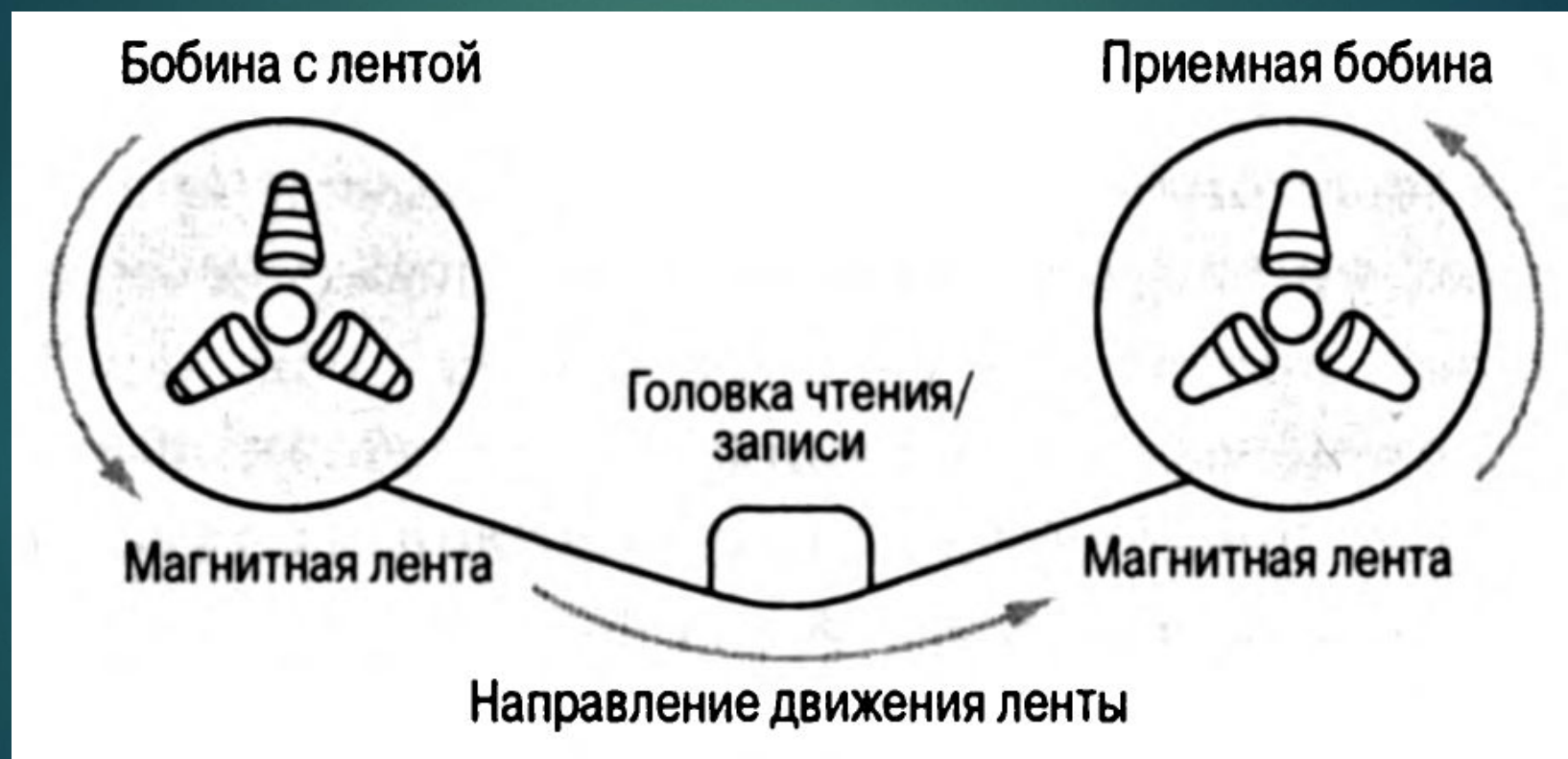
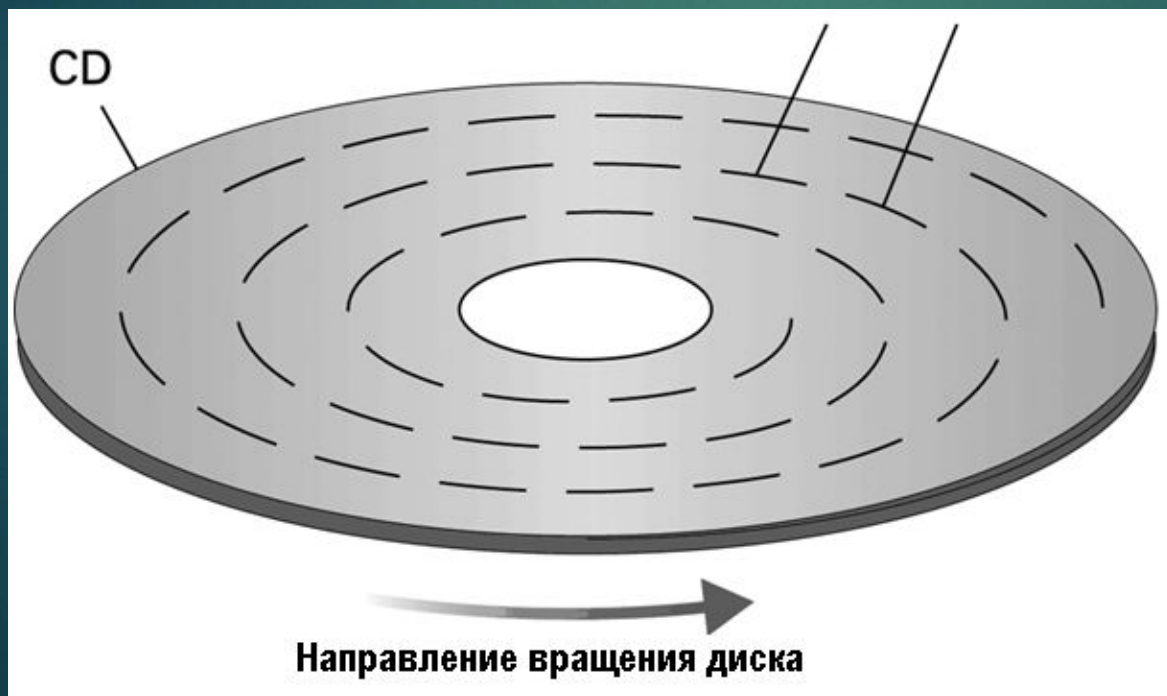


Рисунок 1.11 Запоминающее устройство на компакт диске

0-24

Данные записываются на единственную дорожку, включающую собственные сектора, причём длина оборота спиральной дорожки увеличивается по направлению от внутренней части диска к внешней



Файлы

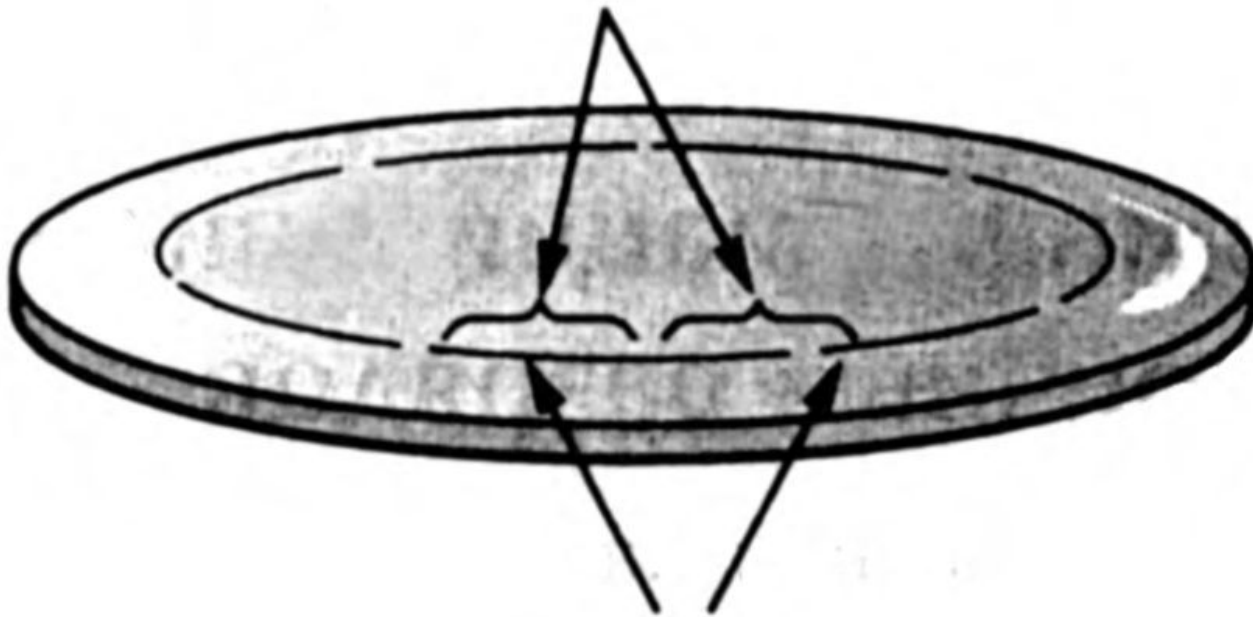
1-25

- ▶ **Файл:** Блок данных хранящийся в массовой памяти
 - ▶ **Поля** и **ключевые поля**
- ▶ Физическая и логическая записи
- ▶ **Буфер:** Область памяти, используемая для временного хранения данных (как правило, этапа в передаче данных)

Рисунок 1.12 Логическая и физическая записи на диске

1-26

Размер логических записей соответствует истинным размерам элементов данных



Размер физических записей соответствует размеру сектора

Представление текста

1-27

- ▶ **Каждому символу (букве алфавита, знаку пунктуации и т.д.) присваивается уникальная битовая комбинация двоичных разрядов.**
 - ▶ ASCII: Используется 7-битовая модель представления символов английского алфавита
 - ▶ ISO разработала ряд из 8-битных расширений ASCII, каждое из которых предназначено для представления символов крупной языковой группы
 - ▶ Unicode: Использует 16-битную модель для представления основных символов, используемых в языках всего мира

Рисунок 1.13 Представление «Hello.» в кодах ASCII

1-28

01001000

H

01100101

e

01101100

l

01101100

l

01101111

o

00101110

.

Представление числовых значений

1-29

- ▶ **Двоичное представление:** Способ выражения числовых величин с помощью только двух цифр (0 и 1)
- ▶ Имеет ограничения на компьютерное представление числовых величин
 - ▶ **Переполнение:** происходит, когда значение слишком велико для представления
 - ▶ **Усечение (потеря значений):** происходит, когда величина не может быть представлена точно

Представление изображений

1-30

- ▶ Растровые методы
 - ▶ Pixel: наименьший элемент изображения
 - ▶ RGB: трёхкомпонентный пиксель
 - ▶ Яркость и цветность
- ▶ Векторные методы
 - ▶ Масштабируемость
 - ▶ TrueType (Microsoft и Apple Computer) и PostScript (Adobe Systems)

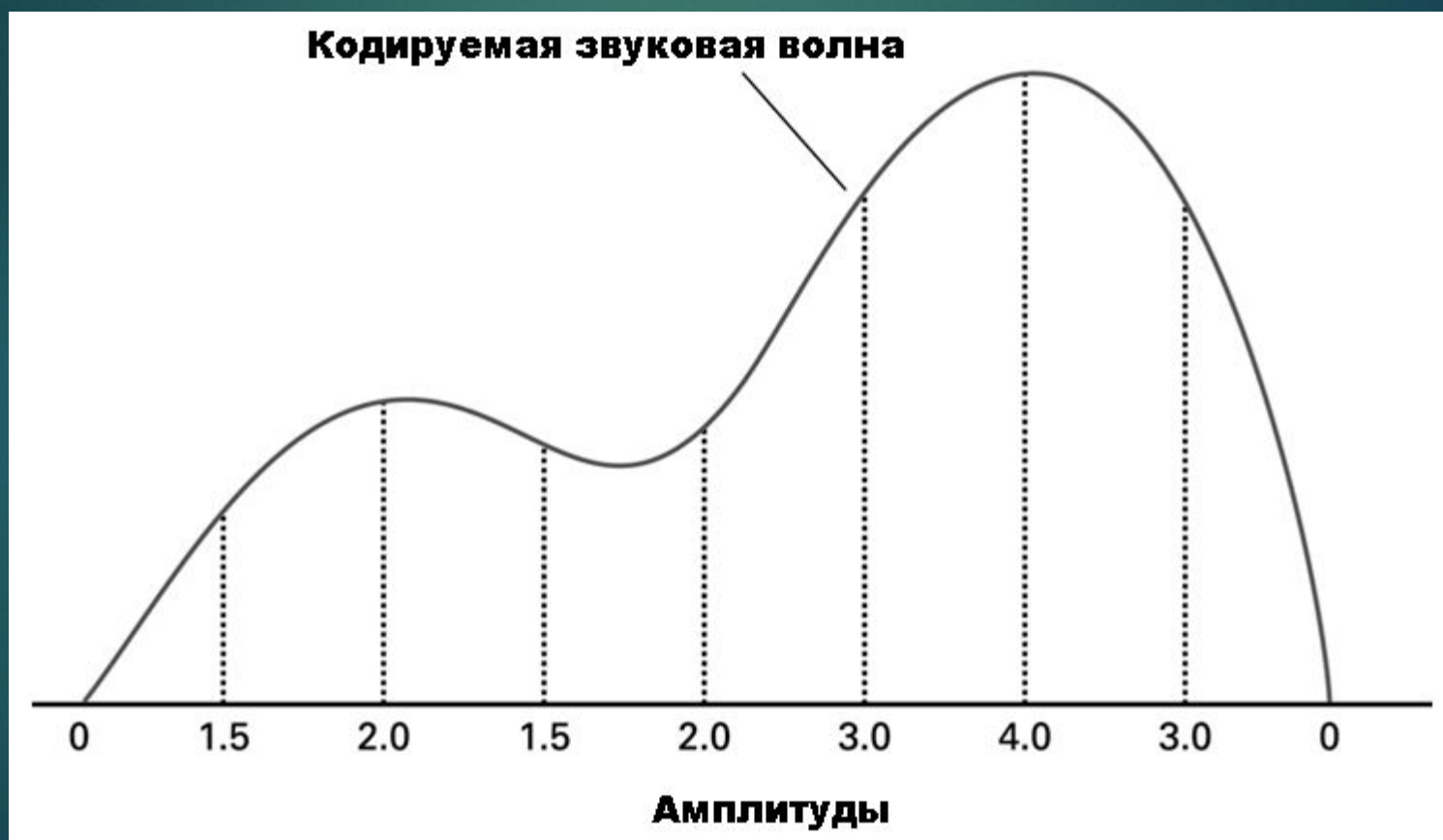
Представление звука

1-31

- ▶ Выборочные методы
 - ▶ Используются для высококачественной записи
 - ▶ Запись реального аудио
- ▶ MIDI (Musical Instrument Digital Interface) – цифровой интерфейс музыкальных инструментов
 - ▶ Используется в музыкальных синтезаторах
 - ▶ Записи музыкального сопровождения

Рисунок 1.14 Звуковая волна,
представлена последовательностью
0, 1.5, 2.0, 1.5, 2.0, 3.0, 4.0, 3.0, 0

1-32



Двоичная система счисления

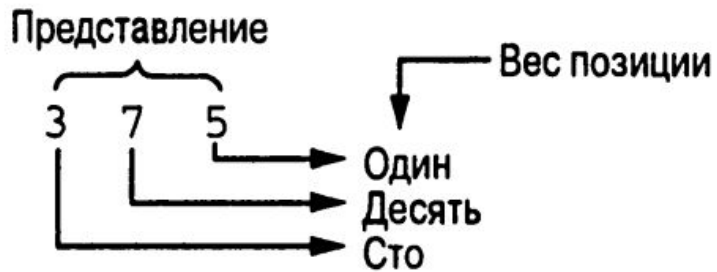
1-33

Традиционная десятичная система использует в качестве основания 10.

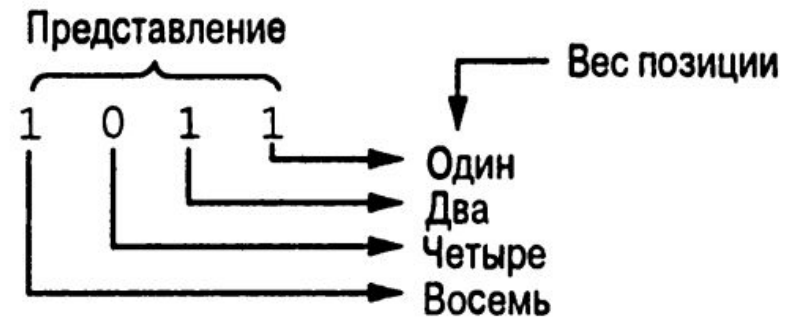
Двоичная система использует в качестве основания 2.

Рисунок 1.15 Основы десятичной и двоичной систем

1-34



а) Десятичная система счисления



б) Двоичная система счисления

Рисунок 1.16

Преобразование двоичного числа 100101

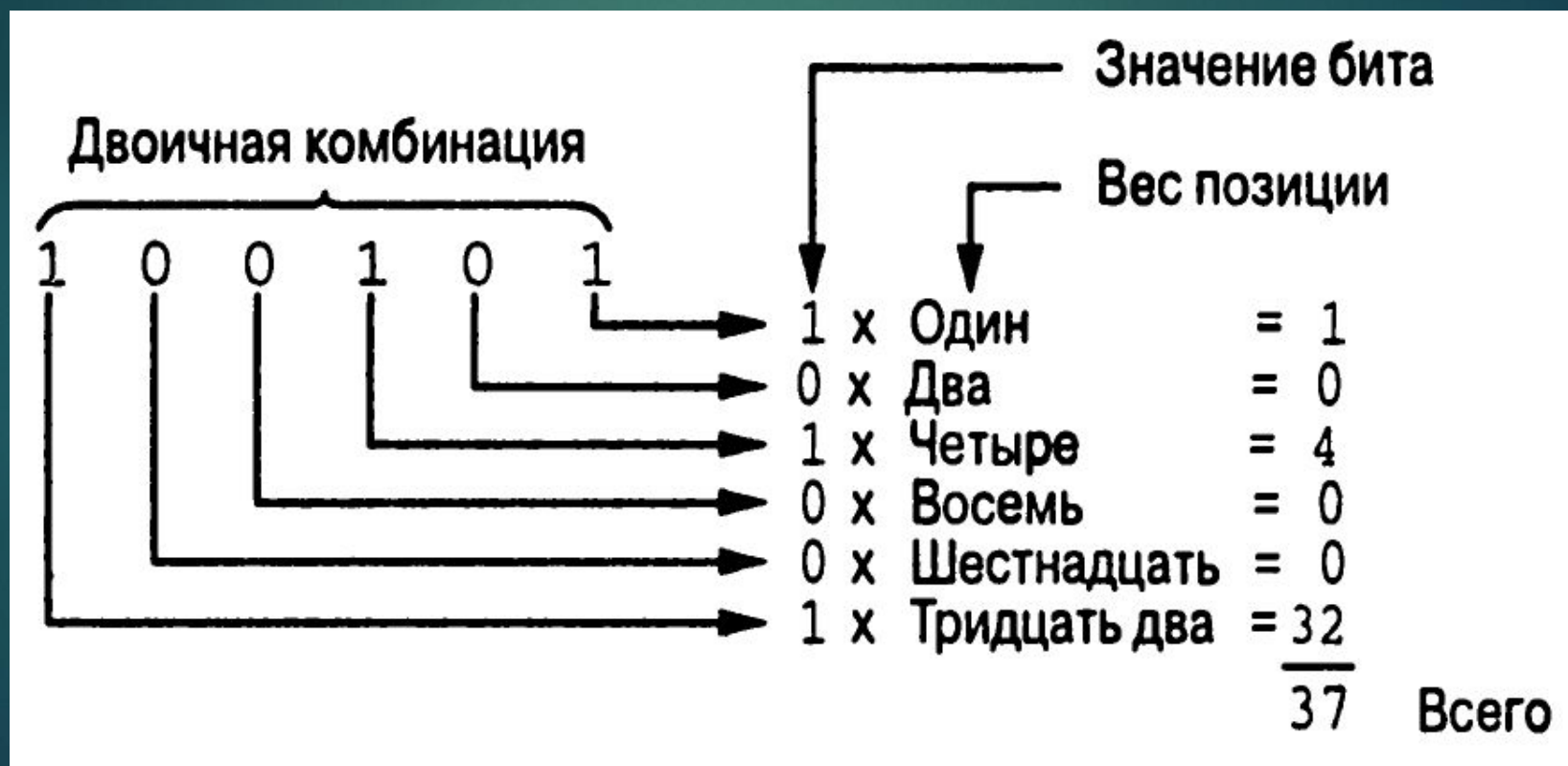


Рисунок 1.17 Алгоритм вычисления двоичного представления для произвольного положительного числа

Этап 1. Поделите число на два и запишите остаток.

Этап 2. Пока не будет получено частное, равное нулю, выполняйте операцию деления предыдущего частного на два и записывайте полученный при каждом делении остаток.

Этап 3. Как только будет получено частное, равное нулю, двоичное представление исходного числа можно будет записать как последовательность всех полученных остатков от деления, последовательно расположенных в направлении справа налево.

Рисунок 1.18 применение алгоритма с рисунка 1.15 для представления двоичного числа 13

1-37

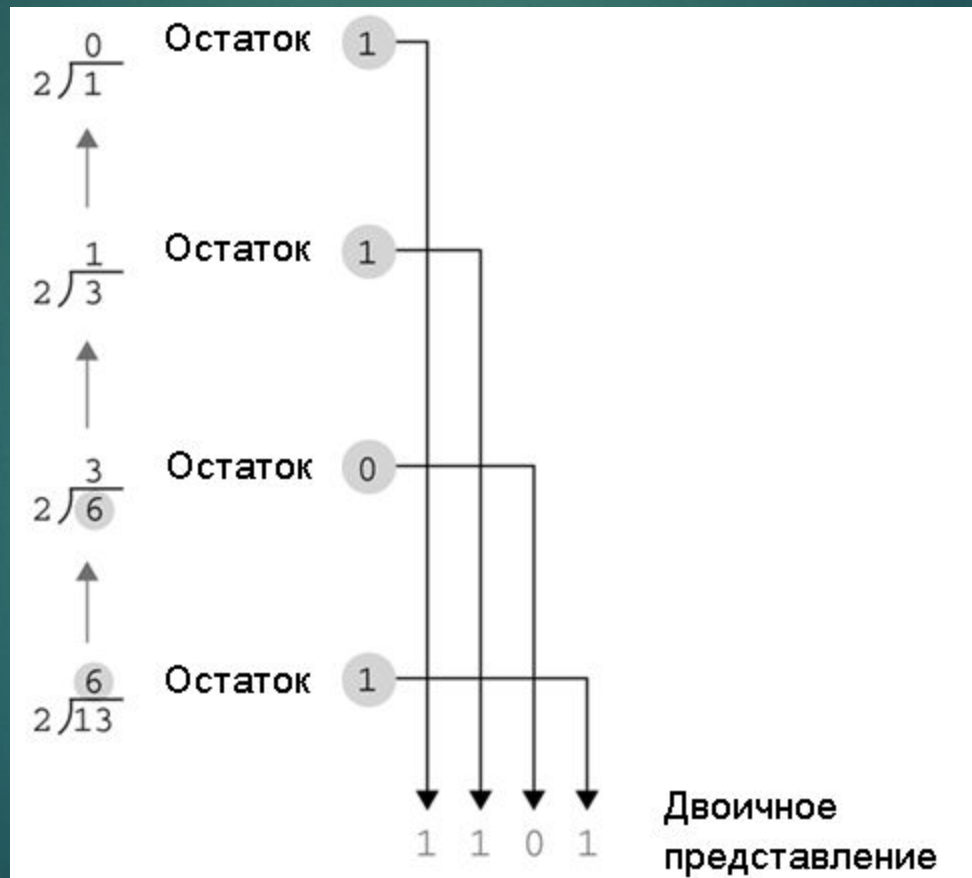


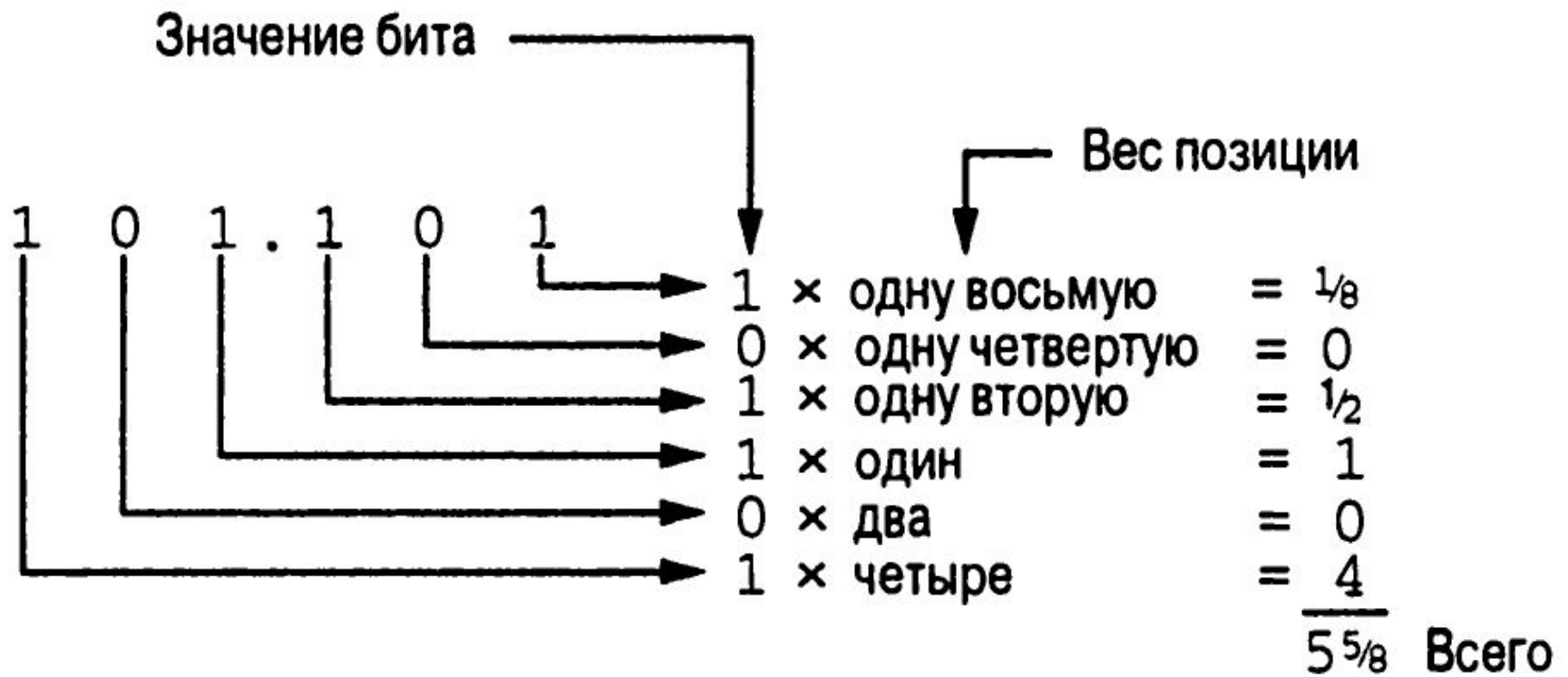
Рисунок 1.19 Правила ДВОИЧНОГО СЛОЖЕНИЯ

1-38

0	1	0	1
+ 0	+ 0	+ 1	+ 1
0	1	1	10

Рисунок 1.20 Преобразование двоичного числа 101.101

1-39



Представление целых чисел

1-40

- ▶ **Двоичный дополнительный код:** Наиболее распространенная система представления целых чисел
- ▶ **Код с избытком:** Ещё один способ представления целых значений
- ▶ Оба могут страдать от ошибок переполнения.

Рисунок 1.21 Схемы кодирования в двоичном дополнительном коде

1-41

а) Трехразрядный дополнительный код

<u>Двоичное представление</u>	<u>Десятичное представление</u>
011	3
010	2
001	1
000	0
111	-1
110	-2
101	-3
100	-4

б) Четырехразрядный дополнительный код

<u>Двоичное представление</u>	<u>Десятичное представление</u>
0111	7
0110	6
0101	5
0100	4
0011	3
0010	2
0001	1
0000	0
1111	-1
1110	-2
1101	-3
1100	-4
1011	-5
1010	-6
1001	-7
1000	-8

Рисунок 1.22 Представление числа -6 в четырёхразрядном дополнительном коде

1-42

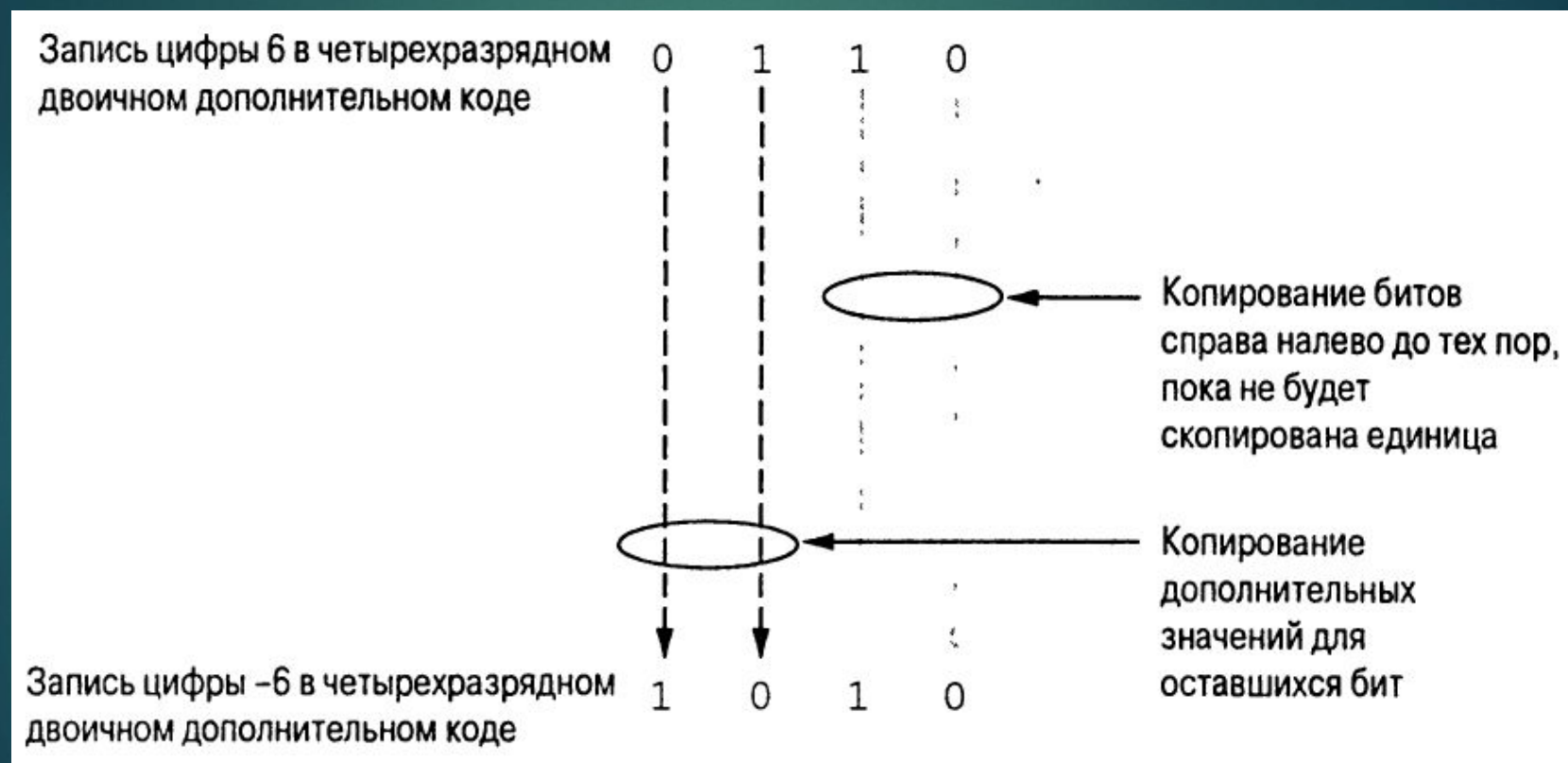


Рисунок 1.23 Сложение чисел в ДВОИЧНОМ ДОПОЛНИТЕЛЬНОМ КОДЕ

1-43

3
+ 2
——
→

0011
+ 0010
——
0101 → 5

-3
+ -2
——
→

1101
+ 1110
——
1011 → -5

7
+ -5
——
→

0111
+ 1011
——
0010 → 2

Рисунок 1.24 Значения двоичных 4-разрядных битовых комбинаций в коде с избытком

1-44

Bit pattern	Value represented
1111	7
1110	6
1101	5
1100	4
1011	3
1010	2
1001	1
1000	0
0111	-1
0110	-2
0101	-3
0100	-4
0011	-5
0010	-6
0001	-7
0000	-8

Рисунок 1.25 Значения двоичных 3-разрядных битовых комбинаций в коде с избытком

1-45

Bit pattern	Value represented
111	3
110	2
101	1
100	0
011	-1
010	-2
001	-3
000	-4

Представление дробных значений

1-46

- ▶ **Представление с плавающей точкой:** Состоит из знакового бита, поля порядка, и поля мантиссы.
- ▶ Данное представление имеет
 - ▶ Нормализованную форму
 - ▶ Ошибки усечения

Рисунок 1.26 Двоичное представление числа в форме с плавающей точкой

1-47

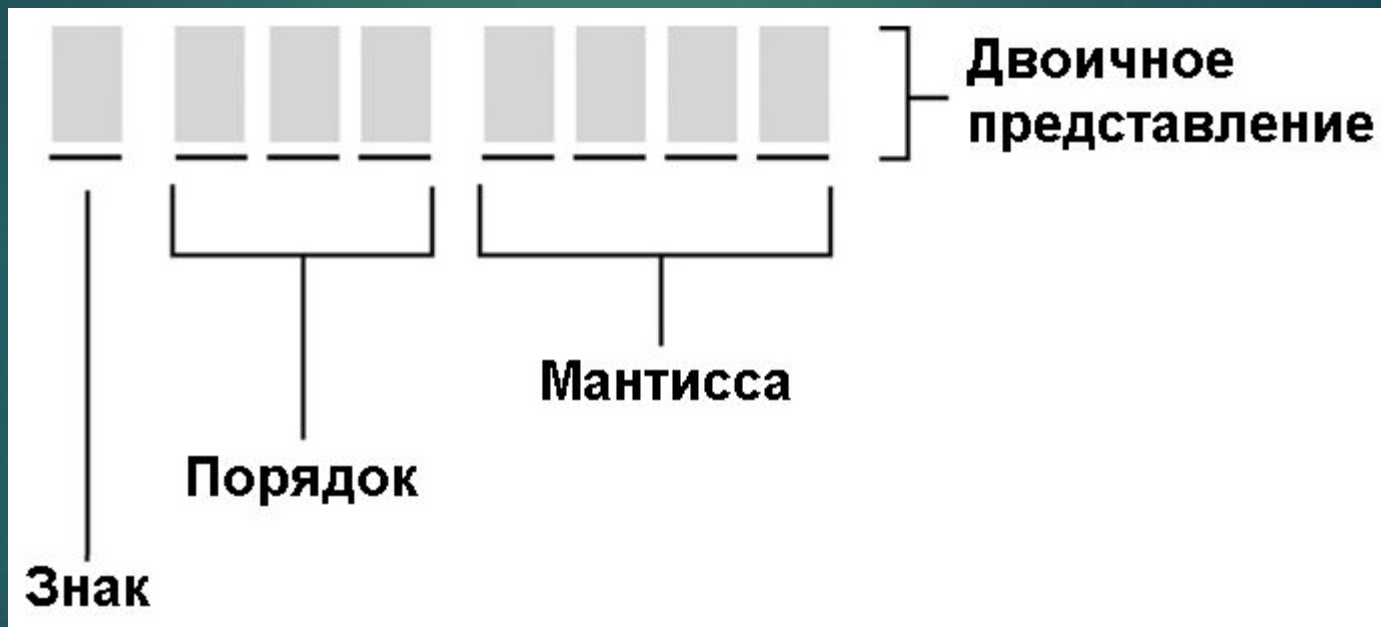
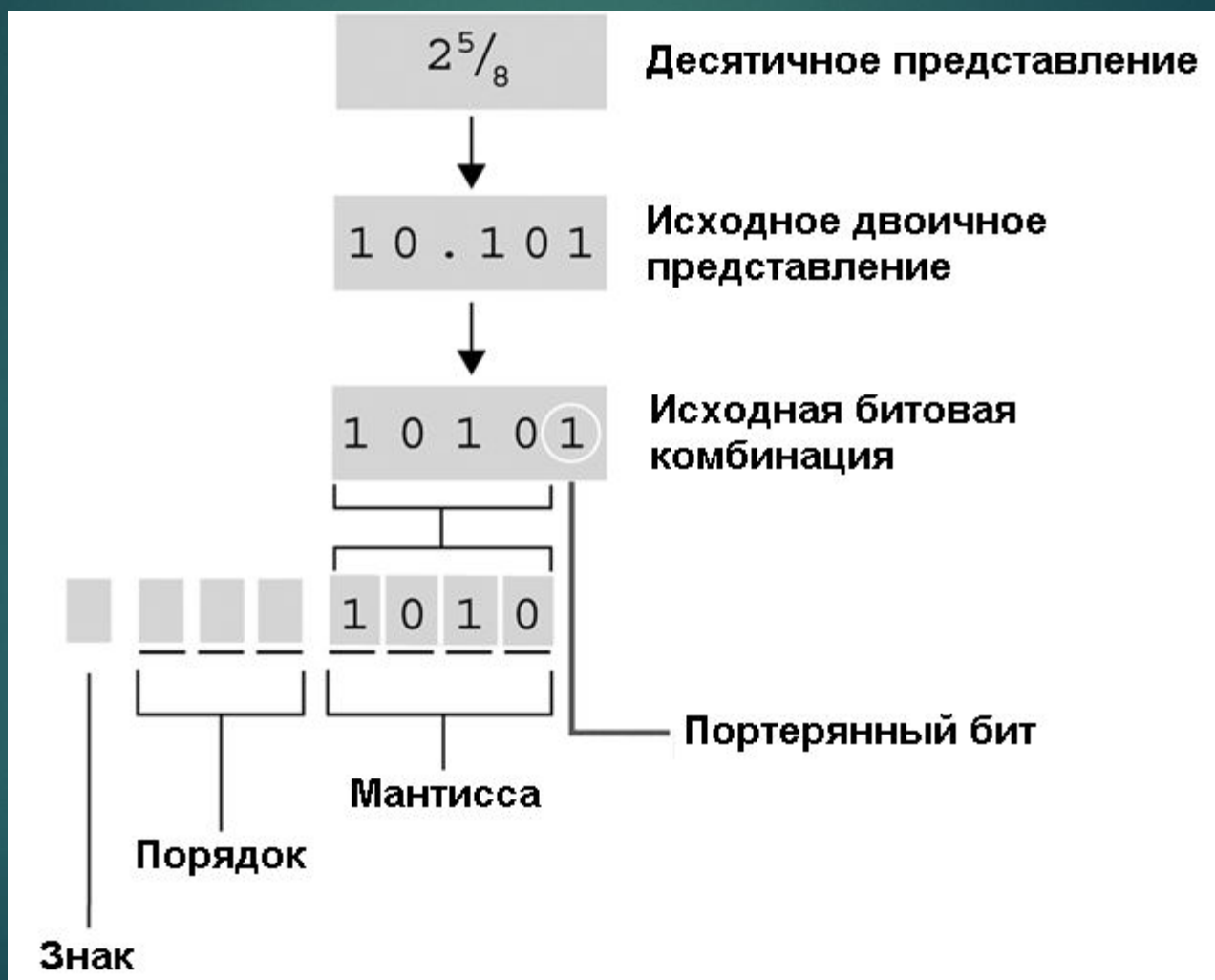


Рисунок 1.27 Схема кодирования числа $2\frac{5}{8}$

1-48



Сжатие данных

1-49

- ▶ С потерями и без потерь
- ▶ Метод кодирования длин серий
- ▶ Относительное кодирование
- ▶ Частотно-зависимое кодирование (Коды Хоффмана)
- ▶ Кодирование с применением адаптивного словаря

Сжатие изображений

1-50

- ▶ GIF: Удобен для мультфильмов
- ▶ JPEG: Хорош для фотографий
- ▶ TIFF: Удобен для архивирования изображений

Сжатие звука и видео

1-51

- ▶ MPEG
 - ▶ Трансляции телевидения высокой четкости
 - ▶ Видео-конференции
- ▶ MP3
 - ▶ Эффект маскировки
 - ▶ Частотная маскировка

Ошибки при передаче информации

1-52

- ▶ Биты чётности (и нечётности)
- ▶ Контрольные байты
- ▶ Коды с исправлением ошибок

Рисунок 1.28 Представление символов А и F в коде ASCII с использованием контрольного бита

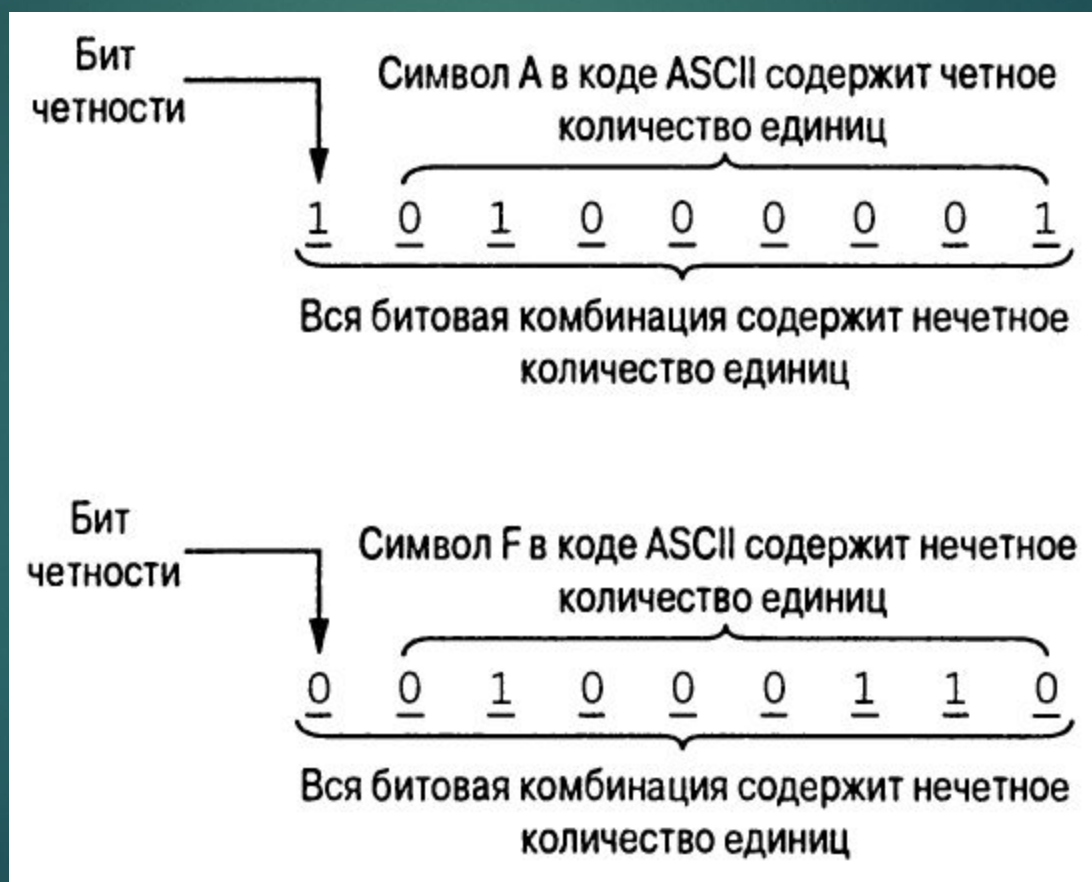


Рисунок 1.29 Код с исправлением ошибок

1-54

Symbol	Code
A	000000
B	001111
C	010011
D	011100
E	100110
F	101001
G	110101
H	111010

Рисунок 1.30 Декодирование битовой комбинации 010100 с помощью кода на рисунке 1.29 на основе дистанции Хэмминга

Character	Code	Pattern received	Distance between received pattern and code	
A	0 0 0 0 0 0	0 1 0 1 0 0	2	
B	0 0 1 1 1 1	0 1 0 1 0 0	4	
C	0 1 0 0 1 1	0 1 0 1 0 0	3	
D	0 1 1 1 0 0	0 1 0 1 0 0	1	Smallest distance
E	1 0 0 1 1 0	0 1 0 1 0 0	3	
F	1 0 1 0 0 1	0 1 0 1 0 0	5	
G	1 1 0 1 0 1	0 1 0 1 0 0	2	
H	1 1 1 0 1 0	0 1 0 1 0 0	4	