

Кодирование информации

Цель работы: ознакомиться с основами кодирования информации; освоить кодировки **ASCII**, **Windows-1251**, **UTF-8**, **UTF-16**, ознакомиться с их различиями.

1. Используйте при выполнении лабораторной работы материал лекции 2 [Л2].
2. Отчет по лабораторной работе оформить в виде документа в MS Word.
3. Дайте определение:
минимальной единицы хранения информации;
минимальной адресуемой единицы хранения информации.
4. Какие значения может принимать 1 бит?
5. Сколько битов содержится в 1 байте?
6. Составьте таблицу перевода десятичных чисел от 0 до 20 в двоичную и шестнадцатеричную системы счисления, представив их как однобайтовые целочисленные значения.
7. Исследуйте таблицы кодировок **US-ASCII** и **Windows-1251** (Л2).
8. Создайте проект консольного C++ приложения с именем **Lab03**.
9. Запишите в строках комментариев три строки:
 - 1) вашу фамилию, имя, год рождения на английском языке.
Например: IvanovAlex1998
 - 2) фамилию, имя, отчество, год рождения на русском языке.
Например: ИвановАлексейДмитриевич1998
 - 3) фамилию (на русском), год рождения, имя (на английском).
Например: Иванов1998Alex
10. Используя таблицы кодировок из Л2, запишите в комментариях шестнадцатеричное представление этих 3х строк (п.9) в кодировке **Windows-1251**.
11. Запишите в комментариях шестнадцатеричное представление этих трех строк в кодировках **UTF-8** и **UTF-16** (используя таблицы кодировок и алгоритм перевода символов юникода в кодировку UTF-8 из Л2).

12. Пример исходного кода **lab03.cpp**:

```
1  #include <iostream>
2  // IvanovAlex1998
3  // представление строки в Windows-1251
4  // в UTF-8
5  // в UTF-16
6
7  // ИвановАлексей1998
8  // ...
9
10 // Иванов1998Alex
11 // ...
12
13
14 int main()
15 {
16     int number = 0x12345678;
17     char hello[] = "Hello, ";
18     char lfie[] = "IvanovAlex1998";
19     char rfie[] = "ИвановАлексей1998";
20     char lr[] = "Иванов1998Alex";
21
22     wchar_t Lfie[] = L"IvanovAlex1998";
23     wchar_t Rfie[] = L"ИвановАлексей1998";
24     wchar_t LR[] = L"Иванов1998Alex";
25
26     std::cout << hello << lfie << std::endl;
27     return 0;
28 }
```

13. Запустите приложение и с помощью отладчика интегрированной среды MS VS и просмотрите области памяти, соответствующие шести строковым константам с вашими данными.
14. Убедитесь, что записанное в комментариях шестнадцатеричное представление строк в кодировках **Windows-1251** и **Unicode-16**, выполненное по таблицам кодировок самостоятельно, идентично полученному с помощью отладчика.
15. Ознакомьтесь с представлением в памяти компьютера целочисленной переменной **number**.
16. В интегрированной среде MS VS ознакомьтесь с представлением в памяти компьютера файла с исходным кодом программы **lab03**, открыв его с помощью двоичного редактора. Определите, в какой кодировке представлен файл (*.cpp) транслятора C++.
17. Определите разницу значений кодов в Windows-1251 для первых пяти букв вашей фамилии в прописном и строчном написании.

18. Опишите словесно алгоритм перевода прописной буквы в строчную для символов в кодировке **Windows-1251**.

Ответьте на следующие вопросы:

- что такое **таблица кодировки**?
- расшифруйте аббревиатуру **ASCII**;
- поясните структуру кодировки **Windows-1251**;
- что такое **UNICODE**?
- поясните структуру **UNICODE**;
- что такое **UTF-8** и **UTF-16**?
- определите разницу значений кодов **следующих** символов в **UTF-16**:
Ё и **ё**, **Ş** и **ş**, **Ł** и **ł**, **Б** и **б**, **Г** и **г**, **Э** и **э**
Каким способом можно получить из кода буквы нижнего регистра букву верхнего регистра?