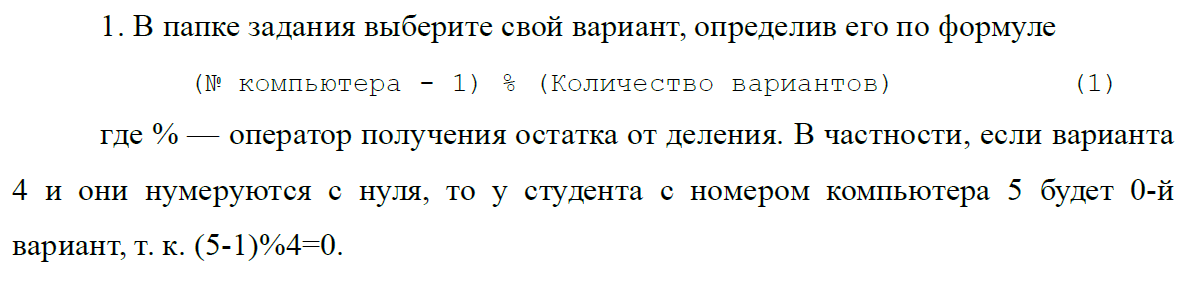
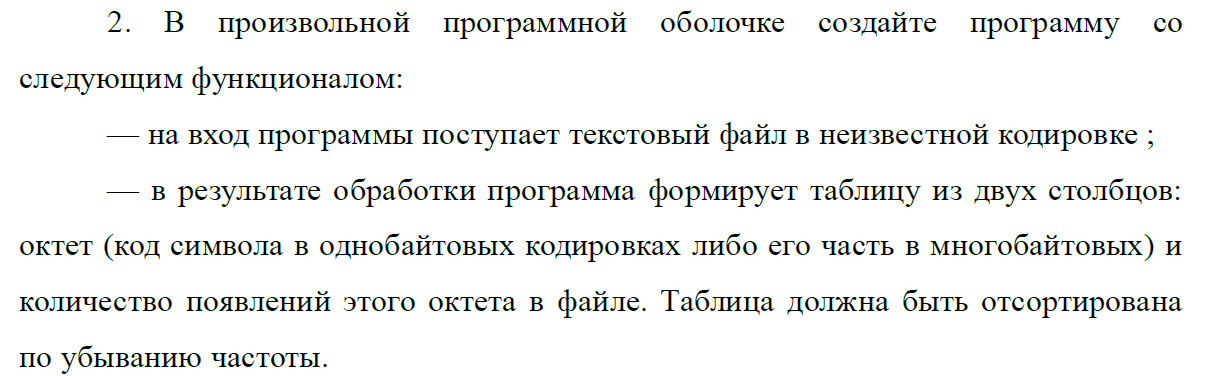
**Отчет по лабораторной работе № 3**

**Работа с кодовыми таблицами русского языка**

Выполнили: Москвитина Е.В., Якупова А.А

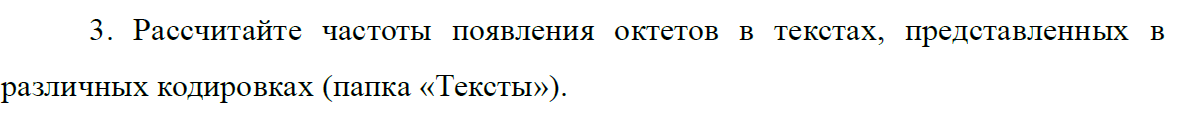


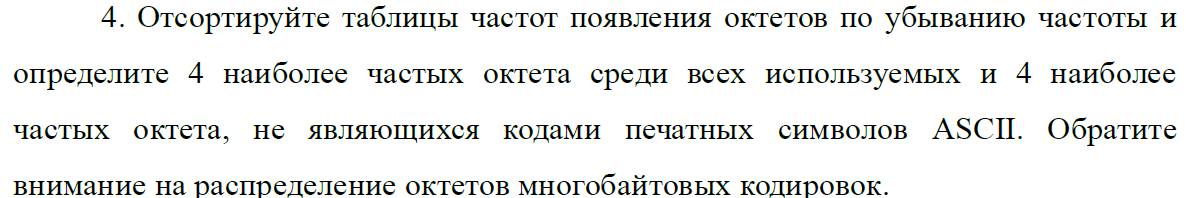
**Вариант (26-1) % 9 = 7**



#!/bin/bash

hexdump -ve '/1 "%02X\n"' "$1" | sort | uniq -c | sort -nr | awk '{ print $2 " " $1}'





**Windows CP1251**

4 наиболее частых октета: 20, EE, E5, E8

4 наиболее частых, не являющихся кодами печатных символов ASCII: EE, E5, E8, E0

**Dos CP866**

4 наиболее частых октета: 20, AE, A5, A8

4 наиболее частых, не являющихся кодами печатных символов ASCII: AE, A5, A8, A0

**MacCyrillic**

4 наиболее частых октета: 20, EE, E5, E8

4 наиболее частых, не являющихся кодами печатных символов ASCII: EE, E5, E8, E0

**ISO8859**

4 наиболее частых октета: 20, DE, D5, D8

4 наиболее частых, не являющихся кодами печатных символов ASCII: DE, D5, D8, D0

**KOI8**

4 наиболее частых октета: 20, CF, C5, C9

4 наиболее частых, не являющихся кодами печатных символов ASCII: CF, C5, C9, C1

**UTF8**

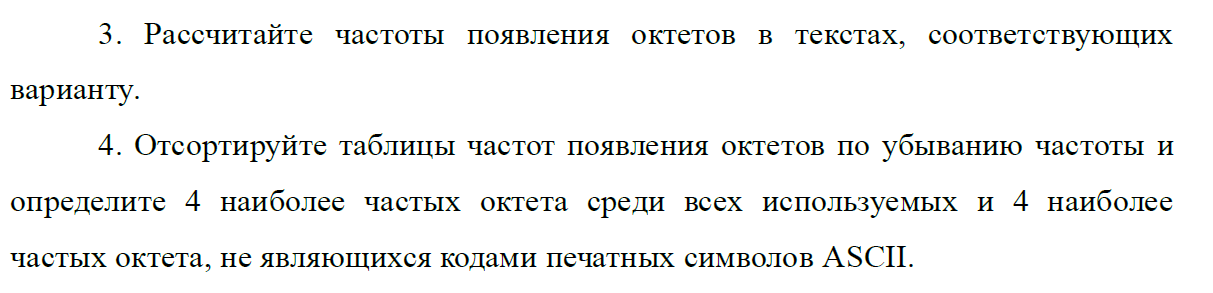
4 наиболее частых октета: D0, D1, 20, BE

4 наиболее частых, не являющихся кодами печатных символов ASCII: D0, D1, BE, B5

**UTF16**

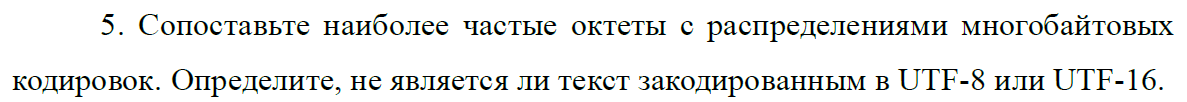
4 наиболее частых октета: 04, 00, 20, 3E

4 наиболее частых, не являющихся кодами печатных символов ASCII:

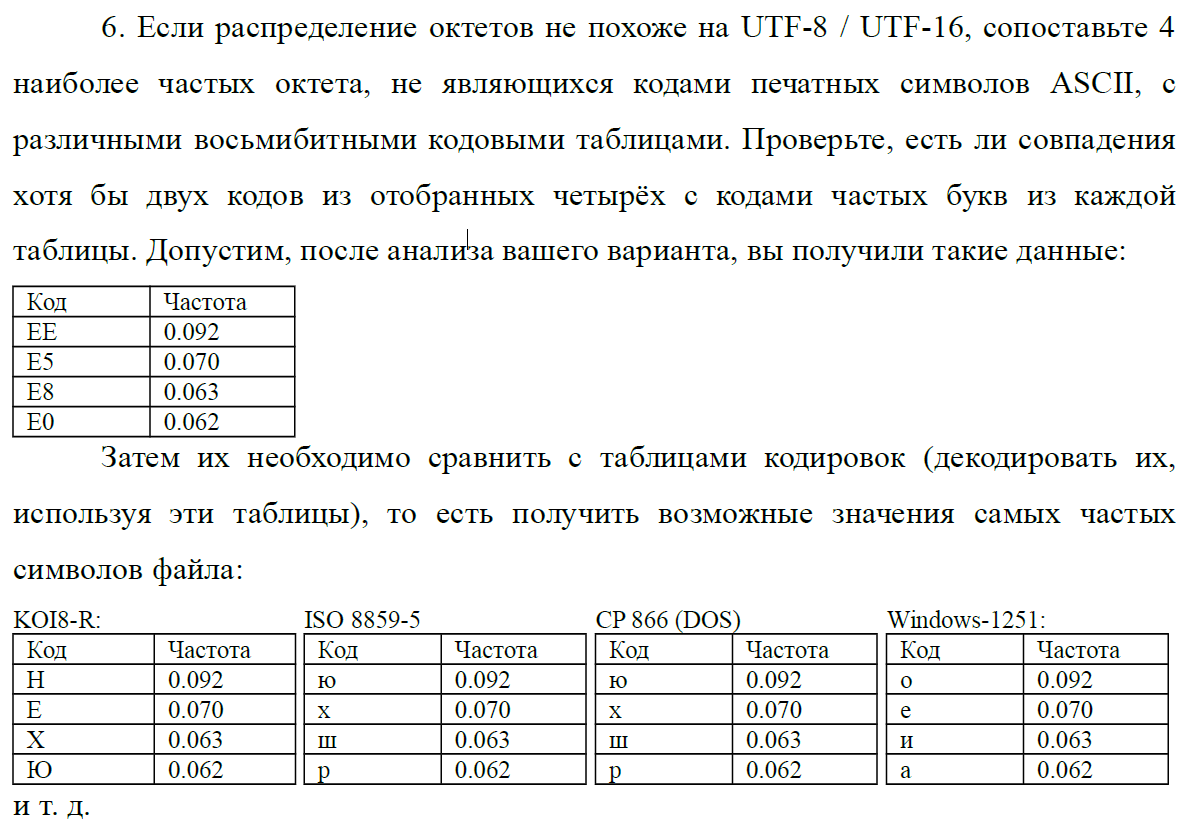


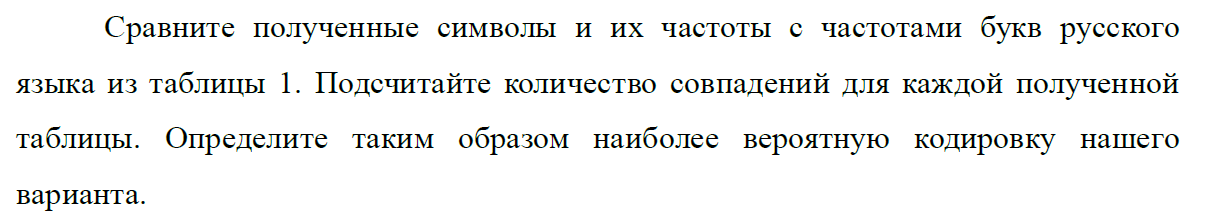
4 наиболее частых октета: 20, EE, E0, E5

4 наиболее частых, не являющихся кодами печатных символов ASCII: E0, EE, E5, E8



Очевидно, что текст варианта закодирован не в UTF-16 и не в UTF-8, тк отсутствуют октеты 04, 00 (для UTF-16) и D0, D1 (для UTF-8)

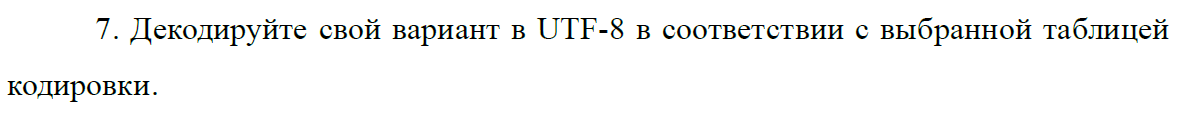




Наиболее вероятные кодировки – Windows CP1251 или Mac Cyrillic

Но в нашем случае в тексте проблемы только с заглавными буквами (при открытии текста в кодировке Windows CP 1521). Ведь таблицы кодов для Windows и Mac отличаются именно в прописных буквах кириллицы.

**Таким образом заключаем что кодировка текста – Mac Cyrillic**



iconv -f MACCYRILLIC -t UTF-8 7.txt > 7\_mac.txt

dos2unix -n 7\_mac.txt 7\_mac\_u.txt