



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИНФОРМАТИКА И СИСТЕМЫ УПРАВЛЕНИЯ

КАФЕДРА КОМПЬЮТЕРНЫЕ СИСТЕМЫ И СЕТИ (ИУ6)

НАПРАВЛЕНИЕ ПОДГОТОВКИ 09.04.01 Информатика и вычислительная техника

МАГИСТЕРСКАЯ ПРОГРАММА 09.04.01/07 Интеллектуальные системы анализа,
обработки и интерпретации больших данных

О Т Ч Е Т

по лабораторной работе № 7

Название: Строки, регулярные выражения. Вариант 9

Дисциплина: Языки программирования для работы с большими
данными

Студент ИУ6-22М
(Группа)

К.Ю. Каташинский
(Подпись, дата) (И.О. Фамилия)

Преподаватель

П.В. Степанов
(Подпись, дата) (И.О. Фамилия)

Москва, 2024

Цель работы

Изучить работу со строками и регулярными выражениями на языке программирования Java.

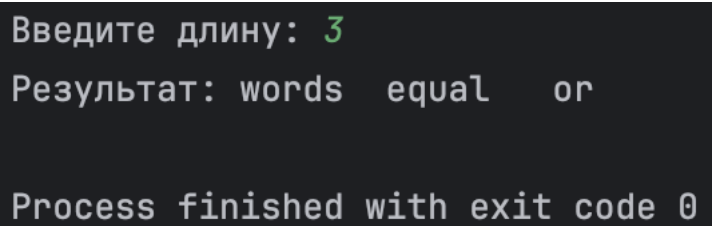
Ход работы

Задание 1. Из текста удалить все слова заданной длины, начинающиеся на согласную букву.

Код программы представлен в листинге 1, результат – на рисунке 1.

Листинг 1 – Код программы задания 1

```
package Lab7;
import java.util.Scanner;
public class Variant19 {
    public static void main(String[] args) {
        Scanner input = new Scanner(System.in);
        System.out.print("Введите длину: ");
        int len = input.nextInt();
        String text = "words are equal len two or its";
        System.out.println("Результат: " +
text.replaceAll(String.format("\\b[\\w^aeiouAEIOU]\\w{%d}\\b", len - 1), ""));
    }
}
```



```
Введите длину: 3
Результат: words equal or
Process finished with exit code 0
```

Рисунок 1 – Результат задания 1

Задание 2. Удалить из текста его часть, заключенную между двумя символами, которые вводятся (например, между скобками ‘(’ и ‘)’ или между звездочками ‘*’ и т.п.).

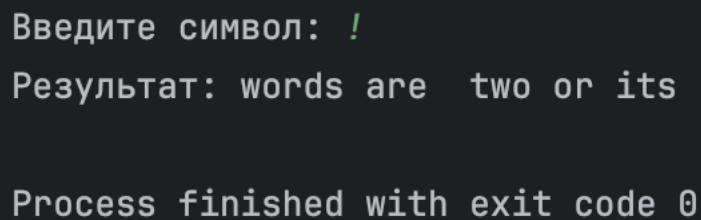
Код программы представлен в листинге 2, результат – на рисунке 2.

Листинг 2 – Код программы задания 2

```

package Lab7;
import java.util.Scanner;
public class Variant110 {
    public static void main(String[] args) {
        Scanner input = new Scanner(System.in);
        System.out.print("Введите символ: ");
        var symbol = input.next();
        String text = "words are !equal len! two or its";
        System.out.println("Результат: " +
text.replaceAll(String.format("%s.*%s", symbol, symbol),
""));
    }
}

```



```

Введите символ: !
Результат: words are two or its
Process finished with exit code 0

```

Рисунок 2 – Результат задания 2

Задание 3. В стихотворении найти одинаковые буквы, которые встречаются во всех словах.

Код программы представлен в листинге 3, результат – на рисунке 3.

Листинг 3 – Код программы задания 3

```

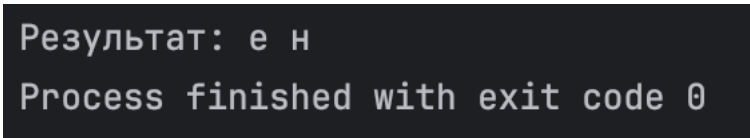
package Lab7;
public class Variant29 {
    public static void main(String[] args) {
        String[] words = "Очень интересное крупное
здание".split(" ");
        boolean contains = true;
        System.out.print("Результат: ");
        for (int i = 0; i < words[0].length(); ++i) {
            contains = true;

```

```

        for (int j = 1; j < words.length; ++j) {
            contains &=
words[j].contains(String.valueOf(words[0].charAt(i)));
        }
        if (contains)
System.out.print(words[0].charAt(i) + " ");
    }
}
}

```



```

Результат: е н
Process finished with exit code 0

```

Рисунок 3 – Результат задания 3

Задание 4. В тексте найти первую подстроку максимальной длины, не содержащую букв.

Код программы представлен в листинге 4, результат – на рисунке 4.

Листинг 4 – Код программы задания 4

```

package Lab7;

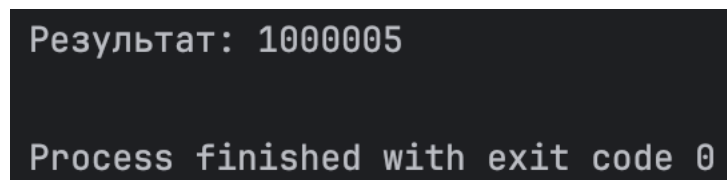
public class Variant210 {
    public static void main(String[] args) {
        int maxCount = 0;
        String text = "Some 2024 text with numbers
1000005!";
        for (int i = 0; i < text.length(); ++i)
        {
            if
(String.valueOf(text.charAt(i)).matches("\\d"))
            {
                StringBuilder value = new StringBuilder();
                while
(String.valueOf(text.charAt(i)).matches("\\d"))

```

```

        {
            value.append(String.valueOf(text.charAt(i)));
            ++i;
        }
        int number =
Integer.parseInt(value.toString());
        if (maxCount < number) maxCount = number;
    }
}
System.out.println("Результат: " + maxCount);
}
}

```



```

Результат: 1000005
Process finished with exit code 0

```

Рисунок 4 – Результат задания 4

Задание 5. Напечатать слова русского текста в алфавитном порядке по первой букве. Слова, начинающиеся с новой буквы, печатать с красной строки.

Код программы представлен в листинге 5, результат – на рисунке 5.

Листинг 5 – Код программы задания 5

```

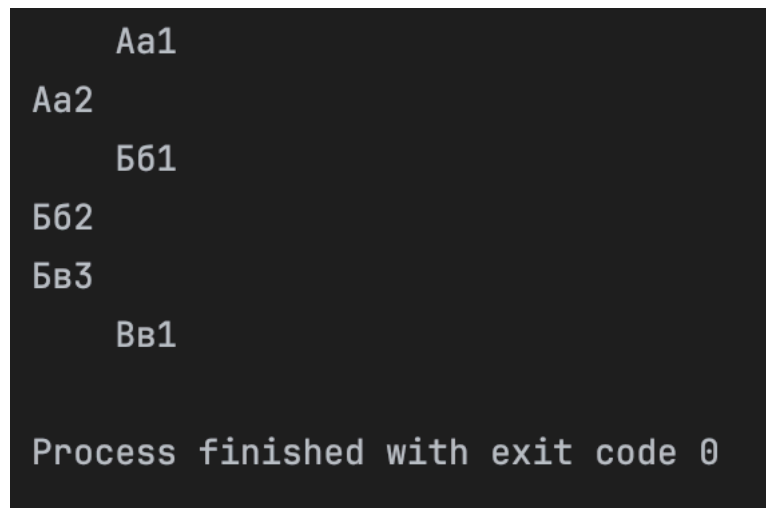
package Lab7;
import java.util.ArrayList;
import java.util.Collections;
import java.util.List;
public class Variant39 {
    public static void main(String[] args) {
        var text = "Aa1 Бб2 Вв1 Бв3 Бб1 Aa2";
        var list = new
ArrayList<String>(List.of(text.split(" ")));
        Collections.sort(list);

```

```

String letter = list.getFirst().substring(0, 1);
boolean isRedline = true;
for (String word:list) {
    if (!word.substring(0, 1).equals(letter)) {
        isRedline = true;
        letter = word.substring(0, 1);
    }
    if (isRedline) {
        System.out.println("\t"+word);
    } else {
        System.out.println(word);
    }
    isRedline = false;
}
}
}

```



```

    Aa1
Aa2
    Бб1
Бб2
    Бв3
Вв1

Process finished with exit code 0

```

Рисунок 5 – Результат задания 5

Задание 6. Рассортировать слова русского текста по возрастанию доли гласных букв. (отношение количества гласных к общему количеству букв в слове).

Код программы представлен в листинге 6, результат – на рисунке 6.
Листинг 6 – Код программы задания 6

```

package Lab7;

import java.util.*;
import java.util.function.Function;

public class Variant310 {

    public static void main(String[] args) {
        var text = "Aa1 Bб2 BБ1 ББ3 Бб1 Aa2";
        var list = new
ArrayList<String>(List.of(text.split(" ")));

        Function<String, Integer> getCount = new
Function<String, Integer>() {
            private final Set<Character> set = new
HashSet<>(Arrays.asList('a', 'e', 'y', 'ы', 'o', 'я'));

            @Override
            public Integer apply(String s) {
                var count = 0;
                for (int i = 0; i < s.length(); ++i) {
                    count +=
set.contains(s.toLowerCase().charAt(i)) ? 1 : 0;
                }
                return count;
            }
        };

        Comparator<String> comparator = (s1, s2) ->
Integer.compare(getCount.apply(s1), getCount.apply(s2));
        Collections.sort(list, comparator);
        System.out.print("Результат: ");
        for (var word : list) {
            System.out.print(word + " ");
        }
    }
}

```

```
Результат: Б62 Вв1 Бв3 Б61 Аа1 Аа2
Process finished with exit code 0
```

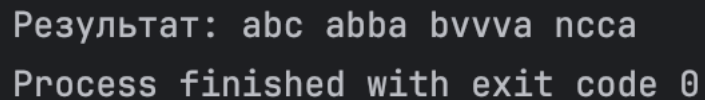
Рисунок 6 – Результат задания 6

Задание 7. Преобразовать каждое слово в тексте, удалив из него все последующие (предыдущие) вхождения первой (последней) буквы этого слова.

Код программы представлен в листинге 7, результат – на рисунке 7.

Листинг 7 – Код программы задания 7

```
package Lab7;
import java.util.ArrayList;
import java.util.List;
public class Variant49 {
    public static void main(String[] args) {
        String text = "abcc ababa bvvvaa nnccaa";
        var list = new
ArrayList<String>(List.of(text.split(" ")));
        var newList = new ArrayList<String>();
        for (var word : list) {
            var firstChar = word.charAt(0);
            var lastChar = word.charAt(word.length() - 1);
            var newWord =
word.replaceAll(String.valueOf(firstChar),
"").replaceAll(String.valueOf(lastChar), "");
            newList.add(firstChar + newWord + lastChar);
        }
        System.out.print("Результат: ");
        for (var word : newList) {
            System.out.print(word + " ");
        }
    }
}
```

```
Результат: abc abba bvvva нсса
Process finished with exit code 0
```

Рисунок 7 – Результат задания 7

Задание 8. Исключить из текста подстроку максимальной длины, начинающуюся и заканчивающуюся одним и тем же символом.

Код программы представлен в листинге 8, результат – на рисунке 8.

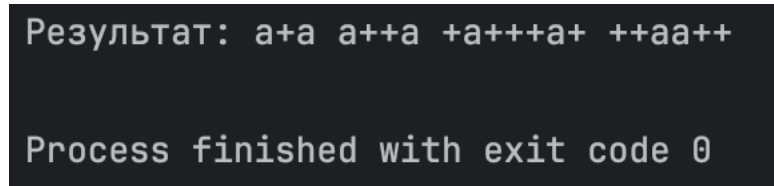
Листинг 8 – Код программы задания 8

```
package Lab7;

import java.util.regex.Pattern;
import java.util.regex.Matcher;

public class Variant410
{
    public static void main(String[] args)
    {
        String text = "a+a a++a +a+++a+ ++a++++a++";
        String regexPattern = "(?<=a)\\S+(?=a)";
        int maxLength = 0, startIndex = -1, endIndex = -1;
        Matcher matcher =
Pattern.compile(regexPattern).matcher(text);
        while (matcher.find()) {
            var length = matcher.end() - matcher.start();
            if (length > maxLength) {
                startIndex = matcher.start();
                endIndex = matcher.end();
            }
        }
        if (startIndex > 0) {
            text = text.substring(0, startIndex) +
text.substring(endIndex);
        }
    }
}
```

```
        System.out.println("Результат: " + text);  
    }  
}
```



```
Результат: a+a a++a +a+++a+ ++aa++  
  
Process finished with exit code 0
```

Рисунок 8 – Результат задания 8

Вывод

В ходе выполнения лабораторной работы была изучена работа со строками и регулярными выражениями на языке программирования Java.