## Практическая работа № 13. Обработка строк в Java

**Цель:** закрепить знания в области обработки строк, научиться применять методы класса String и других классов для обработки строк.

#### Теоретические сведения

Для работы с текстовыми данными в Джава есть три класса: String, StringBuffer и StringBuilder.

#### Особенности использования строк в Джава

В Джава строки представляют собой неизменяемую последовательность символов Unicode. В отличие от представления в C / C ++, где строка является просто массивом типа char, любая Java, строка является объектом класса java.lang.

Однако Джава строка, представляет собой в отличие от других используемых классов особый класс, который обладает довольно специфичными характеристиками. Отличия класса строк от обычных классов:

• строка в Джава представляет из себя строку литералов (текст), помещенных в двойные кавычки, например:

"Hello, World! ". Вы можете присвоить последовательность строковых литералов непосредственно переменной типа String, вместо того чтобы вызывать конструктор для создания экземпляра класса String.

- Оператор '+' является перегруженным, для объектов типа String, и всегда используется, чтобы объединить две строки операндов. В данном контексте мы говорим об операции конкатенации или сложения строк. Хотя '+' не работает как оператор сложения для любых других объектов, кроме строк, например, таких как Point и Circle.
- Строка является неизменяемой, то есть, символьной константой. Это значит, что ее содержание не может быть изменено после ее (строки как объекта) создания. Например, метод toUpperCase () преобразования к верхнему регистру создает и возвращает новую строку вместо изменения содержания существующей строки.

Обратитесь к API JDK для того чтобы ознакомиться с полным списком возможностей класса String в java.lang.String. Наиболее часто используемые методы класса String приведены ниже.

```
boolean equals (String another)
    // НЕЛЬЗЯ использовать'==' или'!=' для сравнения
объектов String в Java
    boolean equalsIgnoreCase(String another)
    int compareTo(String another)// возвращает 0 если эта
строка совпадает с another;
    //<0 если лексикографически меньше another; or >0
    int compareToIgnoreCase(String another)
    boolean startsWith(String another)
    boolean startsWith(String another, int fromIndex) //
поиск начинается с fromIndex
    boolean endsWith(String another)
    // поиск & индексирование
    int indexOf(String search)
    int indexOf(String search, int fromIndex)
    int indexOf(int character)
    int indexOf(int character, int fromIndex) // поиск
вперед от fromIndex
    int lastIndexOf(String search)
    int lastIndexOf(String search, int fromIndex) //
поиск назад от fromIndex
    int lastIndexOf(int character)
    int lastIndexOf(int character, int fromIndex)
                                                   String
    // выделение char или части строки из
(подстрока)
    char charAt(int index)
                                       OД 0 то кирикоп \backslash \backslash
(длина строки-1)
    String substring(int fromIndex)
    String substring(int fromIndex, int endIndex)
                                                        //
exclude endIndex
    // создается новый String или char[] из исходного
(Strings не изменяются!)
    String toLowerCase()
    //преобразование к нижнему регистру
    String toUpperCase()
    //преобразование к верхнему регистру
    String trim()
    // создается новый String с помощью удаления пробелов
```

```
спереди и сзади
    String replace(char oldChar, char newChar)
    //создание нового
                        String
                                               oldChar
                                 CO
                                       старым
перемещается посредством буфера newChar
    String concat(String another)
    // то же самое какthisString + другое
    char[] toCharArray()
    // создается char[] из string
    void getChars(int srcBegin, int srcEnd, char[] dst,
int dstBegin)
    // копируется в массив назначения dst char[]
    /* статические методы для преобразования примитивов в
String*/
    static String ValueOf(type arg)
     // тип может быть примитивный или char[]
    // статические методы дают форматированный
используя //спецификаторы форматирования
    static
                       format(String formattingString,
              String
Object... args)
    // Tak we kak printf()
    // регулярные выражения (JDK 1.4)
    boolean matches (String regexe)
    String replaceAll(String regexe, String replacement)
    String replaceAll(String regexe, String replacement)
    String[] split(String regexe)
    // разделяет String используя regexe как разделитель,
    // возвращает массив String
    String[] split(String regexe, int count)
    // для подсчета количества раз только (count)
```

## Статический метод String.format()

Статический метод String.format() (введен в JDK 1.5) может быть использован для получения форматированного вывода, таким же образом как это делается в языке Си с использованием функции printf() и спецификаторов вывода для различных типов данных. Метод format() делает то же самое, что функция printf(). Например:

```
String.format("%.1f", 1.234);
// возвращает String "1.2"
```

Удобно использовать String.format(), если вам нужно получить простую отформатированную строку для некоторых целей (например, для использования в методе ToString()). Для сложных строк, нужно использовать StringBuffer/StringBuilder с Formatter. Если вам просто нужно отправить простую отформатированную строку на консоль, то просто восльзуйтесь методом System.out.printf(), например:

```
System.out.printf("%.1f", 1.234);
```

Пример использования методов lastIndexOf() и substring() в пользовательском классе Filename

```
Листинг программы 13.1 – Пример класс Filename
public class Filename {
private String fullPath;
private char pathSeparator, extensionSeparator;
public Filename(String str, char sep, char ext) {
fullPath = str;
pathSeparator = sep;
extensionSeparator = ext;
}
public String extension() {
int dot = fullPath.lastIndexOf(extensionSeparator);
return fullPath.substring(dot + 1);
// получение имени файла без расширения
public String filename() {
int dot = fullPath.lastIndexOf(extensionSeparator);
int sep = fullPath.lastIndexOf(pathSeparator);
return fullPath.substring(sep + 1, dot);
 }
public String path() {
int sep = fullPath.lastIndexOf(pathSeparator);
return fullPath.substring(0, sep);
Теперь рассмотрим программу Main, которая использует класс Filename:
Листинг программы 13.2 – Пример программы обрабатывающей строки
public class Main{
public static void main(String[] args) {
  final String FPATH = "/home/user/index.html";
  Filename myHomePage = new Filename(FPATH, '/', '.');
```

```
System.out.println("Extension = " +
myHomePage.extension());
System.out.println("Filename = " +
myHomePage.filename());
System.out.println("Path = " + myHomePage.path());
}
Peзультат работы программы:
Extension = html
Filename = index
Path = /home/user
```

#### Особенности класса String

Строки получили специальное значение в Java, потому что они часто используются в любой программе. Следовательно, эффективность работы с ними (с точки зрения вычислений и хранения) имеет решающее значение. Что нужно помнить про класс **String** 

- это immutable (неизменный) класс
- это final класс

Разработчики Джава все-таки решили сохранить примитивные типы в объектно-ориентированном языке вместо того, чтобы сделать вообще все в виде объектов. Нужно сказать, что сделано это в первую очередь, для того чтобы повысить производительность языка. Ведь примитивы хранятся в стеке вызовов, и следовательно, требуют меньше пространства для хранения, и ими легче управлять. С другой стороны, объекты хранятся в области памяти, которую используют программы, и которая называется "куча" (heap), а этот механизм требуют сложного управления памятью и потребляет гораздо больше места для хранения.

По соображениям производительности, класс String в Джаваразработан, так, чтобы быть чем-то промежуточным между примитивными типами данных и типами данных типа класс. Как уже было отмечено выше специальные характеристики типа String включают в себя:

- '+' оператор, который выполняет сложение примитивных типов данных (таких, как int и double), и перегружен, чтобы работать на объектах String. Операция '+' выполняет конкатенацию двух операндов типа String.
- Джава не поддерживает механизма перегрузки операций по разработке программного обеспечения. В языке, который поддерживает перегрузку операций, например С ++, вы можете превратить оператор '+' (с помощью перегрузки) в оператор для выполнения сложения или вообще

вычитания, например двух матриц, кстати это будет примером плохого кода. В Джава оператор '+' является единственным оператором, который внутренне перегружен, чтобы поддержать конкатенацию (сложение) строк в Java. Нужно принять к сведению, что '+' не работает на любых других произвольных объектах, помимо строк, например, таких как рассмотренные нами ранее классы Point или Circle.

Теперь, собственно, о строках. Существуют несколько способов создания строк. Строка String может быть получена одним из способов:

- непосредственно из присвоения строкового литерала ссылке типа String – таким же способом как примитивные типы данных;
- или с помощью оператора new и конструктора класса String, аналогично вызову конструктора любого другого класса. Тем не менее, этот способ не часто используется и использовать его не рекомендуется.

```
Для примера:
```

```
String str1 = "Java is Hot";

// неявный вызов конструктора через присваивание строкового литерала

String str2 = new String("I'm cool");

// явный вызов конструктора через new char[] array = { 'h', 'e', 'l', 'l', 'o', '.' };

//еще один способ создания строки
```

В первом случае str1 объявлена как ссылка типа String и инициализируется строкой "Java is Hot". Во второй строке, str2 объявлена как ссылка на строку и инициализируется с помощью вызова оператора new и конструктора, который инициализирует ее значением "I'm cool". Строковые литералы хранятся в общем пуле. Это облегчает совместное использование памяти для строк с тем же содержанием в целях сохранения памяти. Объекты строк, выделенные с помощью оператора new, хранятся в куче (heap), а там нет разделяемого хранилища для того же самого контента (содержания).

# Строковые литералы и объекты типа String

Как уже упоминалось, есть два способа создания строк: неявное создание путем присвоения строкового литерала переменной или явного создания объекта String, через вызов оператора new и вызов конструктора. У класса String есть еще одна особенность. Все строковые литералы, определенные в Джава коде, вроде "asdf", на этапе компиляции кэшируются и добавляются в так называемый пул строк. Например:

```
String s1 = "Hello"; // String литерал
```

```
String s2 = "Hello"; // String литерал
String s3 = s1; // одинаковые ссылки
String s4 = new String("Hello"); // String объект
String s5 = new String("Hello"); // String объект
```

Джава предоставляет программистам специальный механизм для хранения последовательностей символьных литералов (строк), так называемый общий пул строк. Если две последовательности литералов (строки) имеют одинаковое содержание, то они разделяют общее пространство для хранения внутри общего пула. Такой подход принят для того, чтобы сохранить место для хранения часто используемых строк. С другой стороны, объекты типа String (строки), созданные с помощью оператора new и конструктора хранятся в куче. Механизм работы со строками в Джава представлен на рис.13.1.

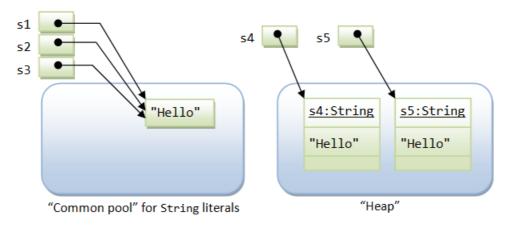


Рисунок 13.1 Механизм работы со строками в Джава

Каждый объект String в куче имеет свое собственное место для хранения, как и любой другой объект. Там нет обмена хранения в куче, даже если два объекта Строковые имеют, то же содержание.

А в куче нет разделяемого пространства для хранения двух объектов, даже если эти два объекта являются объектами типа String и имеют одинаковое содержание.

Вы можете использовать метод equals() класса String для сравнения содержимого двух строк. Вы можете использовать оператор сравнения на равенство '==', чтобы сравнить ссылки (или указатели) двух объектов. Изучите следующие программные примеры:

```
s1 == s1; // true, одинаковые ссылки
s1 == s2; // true, s1 and s1 разделяют общий пул
s1 == s3; /* true, s3 получает то же самое значение
что ссылка s1*/
s1.equals(s3); // true, одинаковое содержимое
s1 == s4; // false, различные ссылки
```

```
s1.equals(s4); // true, одинаковое содержимое s4 == s5; // false, различные ссылки в куче s4.equals(s5); // true, одинаковое содержимое
```

#### Важные замечания

- В приведенном выше примере, используется оператор отношения для того, чтобы проверить на равенство '==' ссылки двух объектов String. Это сделано, чтобы показать различия между строковыми последовательностями литералов, которые используют совместное пространство для хранения в общем пуле строк и объектов String, созданных в куче. Это логическая ошибка в использовании выражения ( str1 == str2) в программе, чтобы сравнить содержимое двух объектов типа String.
- Строка может быть создана непосредственно путем присваивания последовательности литералов (строки), которая разделяет общий пул строк. Не рекомендуется использовать оператор new для создания объектов String в куче.

Помните о том: что строки в Джава являются неизменяемыми!

С тех самых пор, когда в языке Джава появились возможности по использованию разделяемого пространства для хранения строк с одинаковым содержанием в виде строкового пула, String в Джава стали неизменяемыми. То есть, как только строка создается (как объект в памяти программы), ее содержание не может быть изменено никаким образом (по аналогии с Си – строковые литералы — это символьные константы). В противном случае, если этого не сделать, другие ссылки String разделяющие ту же самую ячейку памяти будут зависеть от изменений, которые могут быть непредсказуемыми и, следовательно, является нежелательными. Такой метод, как например, toUpperCase () казалось-бы может изменить содержимое объекта String. Хотя на самом деле, создается совершенно новый объект String и возвращается как раз он в точку вызова. Исходный объект-строка будет впоследствии удален сборщиком мусора (Garbage-collected), как только не окажется больше ссылок, которые ссылаются на него.

Вот поэтому-то объект типа String и считается неизменяемым объектом, вследствие этого, считается не эффективным использовать тип String, например, в том случае, если вам нужно часто модифицировать строку (так вы в таком случае будете создавать много новых объектов типа String, которые каждый раз будут занимать новые места для хранения). Например,

```
// неэффективный код
String str = "Hello";
for (int i = 1; i < 1000; ++i) {
```

```
str = str + i;
}
```

Выводы.

Если содержимое строки должно часто меняться в вашей программе, используйте классы StringBuffer или StringBuilder вместо класса String.

# Классы StringBuffer и StringBuilder

Классы StringBuffer и StringBuilder в Джаваиспользуются, когда возникает необходимость сделать много изменений в строке символов.В отличие от String, объекты типа StringBuffer и StringBuilder могут быть изменены снова и снова. В ДжаваStringBuilder был введен начиная с Джава5.

Как объяснялось выше, строки String являются неизменяемыми, поэтому строковые литералы с таким контентом хранятся в пуле строк. Изменение содержимого одной строки непосредственно может вызвать нежелательные побочные эффекты и может повлиять на другие строки, использующие ту же память.

JDK предоставляет два класса для поддержки возможностей по изменению строк: это классы StringBuffer и StringBuilder (входят в основной пакет java.lang ).

Объекты StringBuffer или StringBuilder так же, как и любые другие обычные объекты, которые хранятся в куче, а не совместно в общем пуле, и, следовательно, могут быть изменены, не вызывая нехороших побочных эффектов на другие объекты.

Класс StringBuilder как класс был введён в JDK 1.5. Это то же самое, как использование класса StringBuffer, за исключением того, что StringBuilder не синхронизирован по многопоточным операциям. Тем не менее, для программы в виде одного потока или нити управления, использование класса StringBuilder, без накладных расходов на синхронизацию, является более эффективным.

## Использование StringBuffer

Класс находится в пакете **java.lang**. Прочитайте спецификацию API JDK для использования java.lang.StringBuffer.

#### Методы класса:

```
// конструкторы
StringBuffer() // инициализация пустым
anStringBuffer
StringBuffer(int size) // определяет размер при
инициализации
```

```
StringBuffer(String s)
                                         //инициализируетя
содержимым
    // длина строки
    int length()
    // Методы для конструирования содержимого
    StringBuffer append(type arg)
                                      // тип может
                                                      быть
примитивным, char[], String, StringBuffer, и т.д.
    StringBuffer insert(int offset, arg)
    // Методы для манипуляции содержимым
    StringBuffer delete(int start, int end)
    StringBuffer deleteCharAt(int index)
    void setLength(int newSize)
    void setCharAt(int index, char newChar)
    StringBuffer replace (int start, int end, String s)
    StringBuffer reverse()
    // Методы для выделения целого/части содержимого
    char charAt(int index)
    String substring(int start)
    String substring(int start, int end)
    String toString()
    // Методы для поиска
    int indexOf(String searchKey)
    int indexOf(String searchKey, int fromIndex)
    int lastIndexOf(String searchKey)
```

int lastIndexOf(String searchKey, int fromIndex)

Обратите внимание, что объект класса StringBuffer является обычным объектом в прямом понимании этого слова. Вам нужно будет использовать конструктор для создания объектов типа класс StringBuffer (вместо назначения в строку буквальном). Кроме того, оператор '+' не применяется к объектам, в том числе и к объектам StringBuffer. Вы должны будете использовать такой метод, как append() или insert() чтобы манипулировать StringBuffer.

Чтобы создать строку из частей, более эффективно использовать класс StringBuffer (для многопоточных программ) или StringBuilder (для однопоточных), вместо конкатенации строк. Например

Листинг 13.3 – Пример создания строки с помощью StringBuilder и String

Компилятор JDK, по сути, использует оба класса как String, так и StringBuffer для обработки конкатенации через операцию сложения строк '+'.

```
Для примера, рассмотрим строчку кода:
```

```
String msg = "a" + "b" + "c";
```

Представленная выше строчка кода будет скомпилирован в следующую для повышения эффективности:

```
String msg = new
StringBuffer().append("a").append("b").append("c").toStri
ng();
```

В этом процессе создаются промежуточный объект StringBuffer и возвращаемый объект String.

# Использование класса StringBuilder

Класс находится в пакете **java.lang**. Прочитайте спецификацию API JDK для использования. java.lang.StringBuilder

Программа ниже демонстрирует разные способы инвертирования длинных строк. Сравниваются три способа работы со строками: как с объектами класса String, так и с помощью StringBuffer и StringBuilder с использованием метода reverse(). Для измерения времени выполнения различных участков кода в примере используется метод

public static native long nanoTime();

Этот метод возвращает текущее значение наиболее точное время системных часов (таймера), в наносекундах. Его можно использовать только для

измерения затраченного времени на выполнение операций, и он никак не связан с системным временем и текущим мировым временем.

Возвращаемое методом значение представлено в виде наносекунд, с момента фиксации, на любой произвольный момент времени. Например, нам необходимо определить, сколько времени занимает выполнение некоторого кода, для этого необходимо выполнить следующее (см Листинг 13.4):

```
Листинг 13.4—Пример программы
    /*
    long startTime = System.nanoTime();
    // ... the code being measured ...
    long estimatedTime = System.nanoTime() - startTime;
    * @return The current value of the system timer, in
nanoseconds.
    * @since 1.5
// Риверс длинной строки с помощью String и StringBuffer
    public class StringsBenchMark {
    public static void main(String[] args) {
    long beginTime, elapsedTime;
    // Build a long string
    String str = "";
    int size = 16536;
    char ch = 'a';
    // Эталонное время в наносекундах
    beginTime = System.nanoTime();
    for (int count = 0; count < size; ++count) {</pre>
      str += ch;
      ++ch;
      if (ch > 'z') {
        ch = 'a';
    elapsedTime = System.nanoTime() - beginTime;
    System.out.println("Elapsed
                                     Time
                                              is
                                                   * *
elapsedTime/1000 + " usec (Build String)");
    /* Риверс строки, строим другую строку за символом в
обратном порядке*/
    String strReverse = "";
```

```
beginTime = System.nanoTime();
    for (int pos = str.length() - 1; pos \geq 0; pos--) {
    strReverse += str.charAt(pos); // Конкатенация
    }
    elapsedTime = System.nanoTime() - beginTime;
    System.out.println("Elapsed
                               Time
elapsedTime/1000 + " usec (Using String to reverse)");
    // Риверс строки через пустой StringBuffer путем
добавления символов в обратном порядке
    beginTime = System.nanoTime();
    StringBuffer sBufferReverse = new StringBuffer(size);
    for (int pos = str.length() - 1; pos >= 0 ; pos--) {
    sBufferReverse.append(str.charAt(pos)); // append
    elapsedTime = System.nanoTime() - beginTime;
    System.out.println("Elapsed
                                 Time
                                          is
elapsedTime/1000
                        usec (Using
                                      StringBuffer to
reverse)");
    /* Reverse a String by creating a StringBuffer with
the given String and invoke its reverse()*/
    beginTime = System.nanoTime();
    StringBuffer
                    sBufferReverseMethod =
                                                   new
StringBuffer(str);
    method
    elapsedTime = System.nanoTime() - beginTime;
    System.out.println("Elapsed Time
elapsedTime/1000 + " usec (Using StringBuffer's reverse()
method)");
    /* Reverse a String via an empty StringBuilder by
appending characters in the reverse order*/
    beginTime = System.nanoTime();
    StringBuilder
                      sBuilderReverse
                                                   new
StringBuilder(size);
    for (int pos = str.length() - 1; pos >= 0 ; pos--) {
```

```
sBuilderReverse.append(str.charAt(pos));
    elapsedTime = System.nanoTime() - beginTime;
    System.out.println("Elapsed
                                    Time
                                             is
elapsedTime/1000
                  +
                         usec
                                (Using
                                        StringBuilder
                                                        to
reverse)");
    // Reverse a String by creating a StringBuilder with
the given String and invoke its rever its reverse()*/
    beginTime = System.nanoTime();
                      sBuilderReverseMethod
    StringBuffer
                                                       new
StringBuffer(str);
    sBuilderReverseMethod.reverse();
    elapsedTime = System.nanoTime() - beginTime;
    System.out.println("Elapsed
                                  Time
elapsedTime/1000 + " usec (Using StringBuidler's reverse()
method)");
       }
    }
```

### Задания на практическую работу № 13

#### Задание 1 на класс String (5%)

- 1. Напишите метод, который принимает в качестве параметра любую строку, например "I like Java!!!".
- 2. Распечатать последний символ строки. Используем метод String.charAt().
- 3. Проверить, заканчивается ли ваша строка подстрокой "!!!". Используем метод String.endsWith().
- 4. Проверить, начинается ли ваша строка подстрокой "I like". Используем метод String.startsWith().
- 5. Проверить, содержит ли ваша строка подстроку "Java". Используем метод String.contains().
  - 6. Найти позицию подстроки "Java" в строке "I like Java!!!".
  - 7. Заменить все символы "а" на "о".
  - 8. Преобразуйте строку к верхнему регистру.
  - 9. Преобразуйте строку к нижнему регистру.
  - 10. Вырезать строку Java с помощью метода String.substring().

Задание 2. (5%) Разработать класс Person, в котором имеется функция, возвращающая Фамилию И.О. Функция должна учитывать возможность отсутствия значений в полях Имя и Отчество. Программу оптимизировать с точки зрения быстродействия.

Задание 3. (15%) Доработать класс адреса, который из полученной строки формата "Страна[d] Регион[d] Город[d] Улица[d] Дом[d] Корпус[d] Квартира" ([d] — разделитель, например, «запятая») выбирает соответствующие части и записывает их в соответствующие поля класса Address. Учесть, что в начале и конце разобранной части адреса не должно быть пробелов. Все поля адреса строковые. Разработать проверочный класс не менее чем на четыре адресных строки. В программе предусмотреть две реализации этого метода:

- а) разделитель только запятая (использовать метод split()); Внимание, при разработке нужно учесть, что
  - б) разделитель любой из символов ',' '.' ';' (класс StringTokenizer<sup>8</sup>).

**Задание 4. (25%)** Реализуйте класс Shirt: Метод toString() выводит объяснение и значение полей построчно.

Дан также строковый массив: shirts[0] = "S001,Black Polo Shirt,Black,XL"; shirts[1] = "S002,Black Polo Shirt,Black,L"; shirts[2] = "S003,Blue Polo Shirt,Blue,XL"; shirts[3] = "S004,Blue Polo Shirt,Blue,M"; shirts[4] = "S005,Tan Polo Shirt,Tan,XL"; shirts[5] = "S006,Black T-Shirt,Black,XL"; shirts[6] = "S007,White T-Shirt,White,XL"; shirts[7] = "S008,White T-Shirt,White,L"; shirts[8] = "S009,Green T-Shirt,Green,S"; shirts[9] = "S010,Orange T-Shirt,Orange,S"; shirts[10] = "S011,Maroon Polo Shirt,Maroon,S";

Преобразуйте строковый массив в массив класса Shirt и выведите его на консоль.

**Задание 5. (25%)** Разработайте класс, который получает строковое представление телефонного номера в одном из двух возможных строковых форматов:

+<Код страны><Номер 10 цифр>, например "+79175655655" или

"+104289652211" или

8<Номер 10 цифр> для России, например "89175655655"

и преобразует полученную строку в формат:

+<Код страны><Три цифры>--<Четыре цифры>

Задание 6. (25%) В методе main считай с консоли имя файла, который содержит слова, разделенные пробелом. В методе getLine() используя

 $<sup>^{8}</sup>https://docs.oracle.com/javase/7/docs/api/java/util/StringTokenizer.html\#:\sim:text=The\%20string\%20tokenizer\%20class\%20allows,they\%20recognize\%20and\%20skip\%20comments.$ 

StringBuilder расставьте все слова в таком порядке, чтобы последняя буква данного слова совпадала с первой буквой следующего не учитывая регистр. Каждое слово должно участвовать 1 раз.