

Lesson 08

Java String Libraries



Java String

❖ Phân loại String.

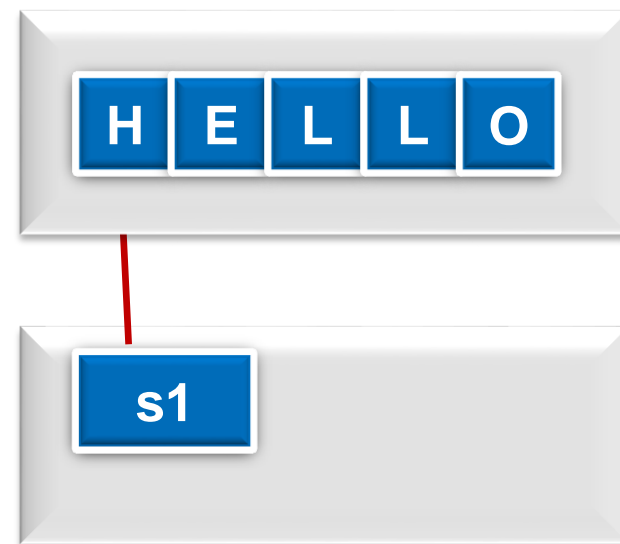
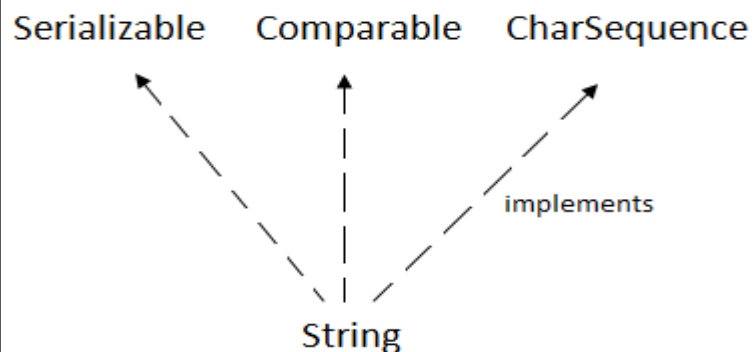
- Immutable
 - String literal
 - String object
- Mutable
 - StringBuilder
 - StringBuffer

String literal vs object

❖ Khai báo chuỗi

- String: là chuỗi, tập các ký tự
- Trong Java, String là lớp quản lý dữ liệu văn bản và là class với các phương thức đi kèm
- Khai báo: Có 2 cách

```
String s1 = new String("Hello");  
String s2 = "Hello";
```



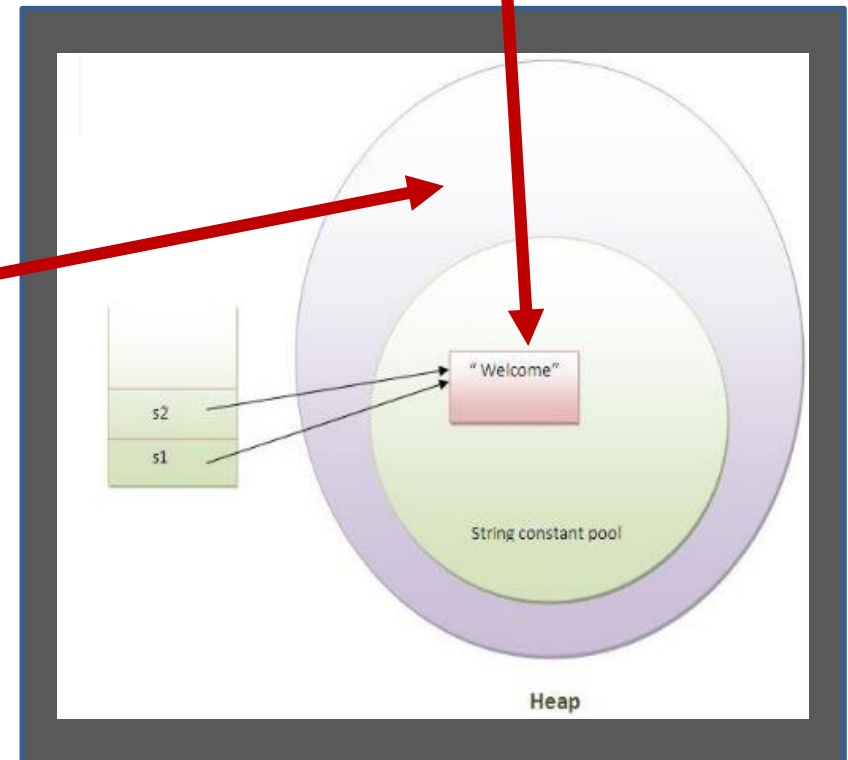
String literal vs object

String constant pool

❖ Khai báo chuỗi

```
String s1 = "Welcome";  
String s2 = "Welcome"; // will not create an instance  
String s3 = "Student"; // create a new instance
```

```
String o1 = new String("Welcome");  
String o2 = new String("Welcome");  
String o3 = new String("Student");  
// create a new instance in heap area
```

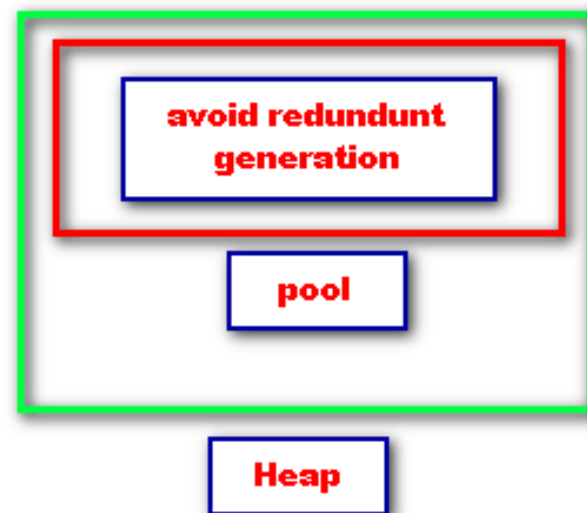


String literal vs object

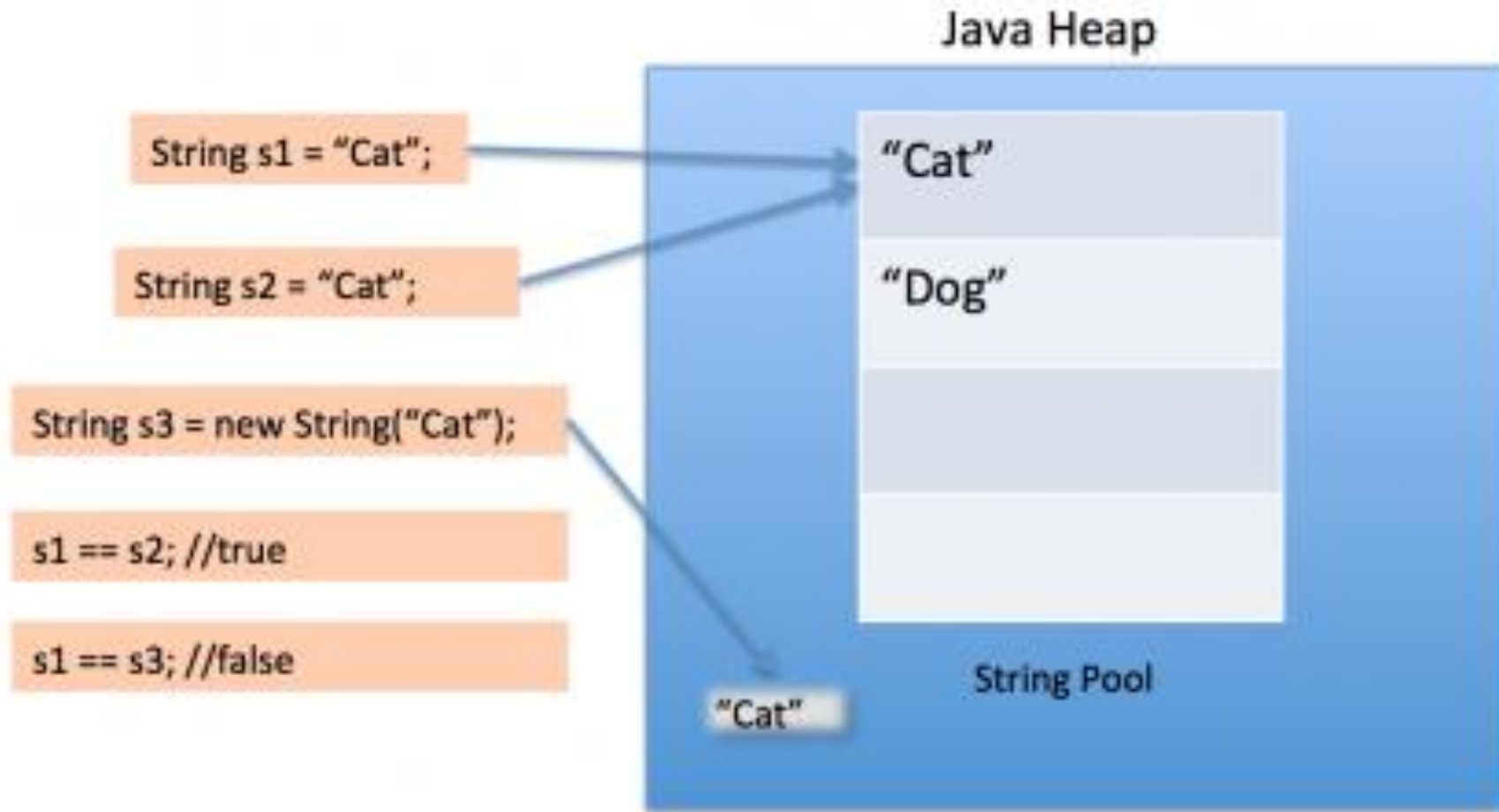
❖ Khai báo chuỗi

```
String s = "text";  
String x = new String("text");
```

```
/**  
 * Giống nhau:  
 *   +> Cả 2 đều là tham chiếu  
 *   +> Thuộc kiểu dữ liệu đối tượng là String  
 * Khác nhau:  
 *   +> new String("text"):  
 *       . Luôn luôn tạo [new] mới một tham chiếu, địa chỉ  
 *       . Không quan tâm giá trị đã tồn tại hay chưa  
 *       . Lưu giá trị tại Heap  
 *   +> s = "text"  
 *       . Tăng khả năng reuse các thể hiện [instance] từ vùng nhớ String constant pool  
 *       . Tiết kiệm bộ nhớ  
 * **/
```




What is a constant pool





String constant pool

- ❖ String Pool is a pool of Strings stored in Heap Memory. Because String is **immutable** in Java and it's implementation of String **interning concept**
 - ❖ String pool helps in **saving a lot of space** for Java Runtime although it **takes more time to create** the String.
- 
- ❖ When we use **double quotes – string literal** to create a String, it first looks for String with same value in the String pool, if found it just returns the reference else it creates a new String in the pool and then returns the reference.
 - ❖ However using **new operator – string object**, we force String class to create a new String object in heap space. We can use **intern()** method to put it into the pool or refer to other String objects from string pool having same value.



Object or Literal

```
203     ....public void test() {  
204         ....final String l1 = "cat";  
205         ....final String l2 = "dog";  
206  
207         ....final String o1 = new String("tiger");  
208         ....final String o2 = new String("mouse");  
...
```

'new String("mouse")' is redundant [less...](#) (Ctrl+F1)

Reports any attempt to instantiate a new **String** object by copying an existing string. Constructing new **String** objects in this way is rarely necessary, and may cause performance problems if done often enough.

```
212     ....}
```




So sánh chuỗi

❖ Tham chiếu (So sánh địa chỉ - Ít khi được sử dụng)

- Sử dụng toán tử ==

❖ Nội dung

- CompareTo
- Equals

So sánh chuỗi

❖ Toán tử ==

- Dùng để so sánh giá trị của
 - Kiểu dữ liệu nguyên thủy
 - Địa chỉ của đối tượng
- Kiểm tra 2 tham chiếu có cùng trỏ đến một ô nhớ hay không

❖ Phương thức equals

- Dùng để so sánh giá trị của 2 đối tượng
- Phương thức equals() chỉ áp dụng cho kiểu đối tượng, không áp dụng cho kiểu nguyên thủy.

❖ Phương thức compareTo()

- Trả về giá trị kiểu int
 - = 0
 - != 0

Ví dụ

1. Cùng giá trị → true

```
new String("test").equals("test")
```

2. Khác object, instance, reference → false

```
new String("test") == "test"
```

3. Khác reference → false

```
new String("test") == new String("test")
```

4. Cùng tham chiếu – reference → true

```
"test" == "test" : String constant pool
```

5. Test

```
Objects.equals("test", new String("test")) // --> true
```

```
Objects.equals(null, "test") // --> false
```



So sánh chuỗi

❖ Ex01

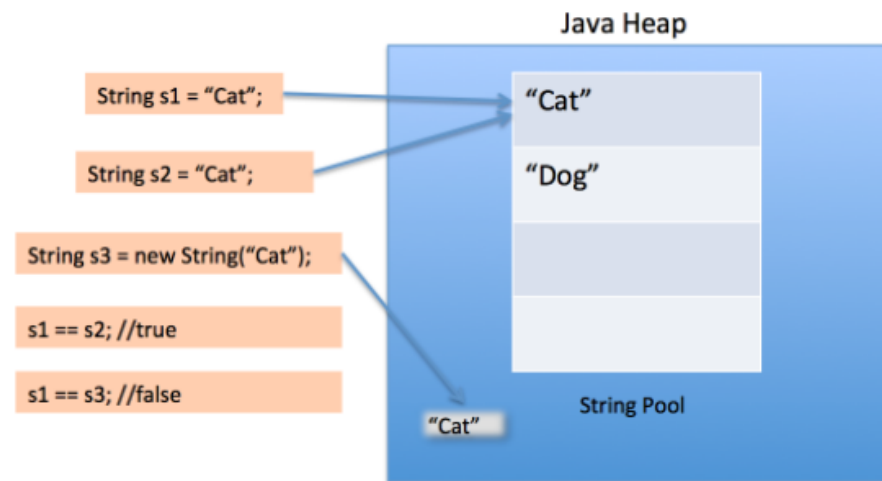
```
// Original String => String pool
String s1 = "bkit-java";
/**
    Vấn đề = và += : đang tạo 1 thể hiện mới
    có dạng new String ("literal") => tạo 1 địa chỉ mới
*/
String test = "bkit";
System.out.println("1: " + ((test+="-java")==s1)); // false
```

❖ Ex02

```
// Original String => String pool
String s1 = "bkit-java"; // POOL
/**
    Vấn đề = và += : đang tạo 1 thể hiện mới
    có dạng new String ("literal") => tạo 1 địa chỉ mới
*/
String test = "bkit";
test += "-java"; // <=> test = test + java <=> test = new String(test+java) // HEAP
// Có thể same hashCode same address memory nhưng khác vùng nhớ => FALSE
System.out.println("1: " + (test==s1)); // false
```

❖ Ex03

```
System.out.println("1: " + (test.intern()==s1.intern())); // true
// Đưa về String pool để so sánh
```





Libraries

❖ Các phương thức xử lý trên kiểu String

Vấn đề:

Cần có các giá trị để phục vụ cho việc hiển thị và tính toán

Ví dụ:

- Tính chiều dài của chuỗi s
- Nối chuỗi s1 vào chuỗi s
- Lấy một ký tự tại vị trí index(3) trong chuỗi s
- Duyệt từng phần tử trong chuỗi
- Tìm vị trí – chỉ số xuất hiện đầu tiên, cuối cùng của ký tự “a” trong chuỗi s

Giải quyết:

Sử dụng hàm chuỗi trong thư viện hàm của Java



Libraries

❖ Các phương thức xử lý trên kiểu String

Vấn đề:

Cần có các giá trị để phục vụ cho việc hiển thị và tính toán

Ví dụ:

- Kiểm tra kí tự, chuỗi bắt đầu, kết thúc trong chuỗi s có phải là s1 không.
- Thay thế chuỗi s1 bằng chuỗi s2 trong chuỗi s
- Loại bỏ các khoảng trắng thừa của chuỗi s3
- Tạo chuỗi con s4 từ chuỗi s bắt đầu từ vị trí số 2
- Xác định chuỗi s2 có tồn tại trong chuỗi s1 hay không

Giải quyết:

Sử dụng hàm chuỗi trong thư viện hàm của Java



String with Split method

- ❖ Có 2 instances của split() method trong java string
 - String split(String regex)
 - String split(String regex, int limit)
 - Regex: biểu thức chính quy được áp dụng trong string
 - **PatternSyntaxException** nếu pattern không tồn tại
 - Limit: giới hạn số lượng chuỗi sau khi split

```
String s1 = "xin chào các bạn";  
  
System.out.println("-----");  
for (String w : s1.split(" ", 3))  
    System.out.println(w);  
}
```

String parsing

❖ Parsing string with split

- Chia một string vào trong **tokens** dựa trên các **delimiters** đã cho
- Tokens: one piece of information, a "word"
- Delimiters: one (or more) characters used to separate tokens
- Example:

```
String phrase = "Java is one of the best languages";  
String delims = "[ ]+";  
String [] tokens = phrase.split(delims);  
  
/* "[delim_characters]+" */
```




Xử lý chuỗi split.

❖ Convert từ String thành Stream với Pattern split

```
String phrase = "Join Beta, Anne Hora, Steffen Baka, Join Son";  
String delims = ",";  
String [] tokens = Pattern.compile(delims).split(phrase);
```

Ques 01. List the name has first name is Join.

```
String result = Pattern.compile(", ")  
    .splitAsStream(phrase)  
    .filter(obj -> obj.contains("Join"))  
    .collect(Collectors.joining("\n"))
```

StringTokenizer

❖ Cho phép chương trình break a string vào trong tokens

```
StringTokenizer st = new StringTokenizer(charSq, delims);  
while (st.hasMoreTokens()) {  
    System.out.println(st.nextToken());  
}
```

```
String[] result = "this is a test".split("\\s");  
for (int x=0; x<result.length; x++){  
    System.out.println(result[x]);  
}
```



Libraries

B1: Đếm xem trong chuỗi " aBcD172#@! " sau đây có bao nhiêu kí tự in Hoa, Thường và Số.

B2: Viết chương trình in ra danh sách tên các bài nhạc mp3 có trong thư mục music

Input: String bh1 = "d:/music/shapeofyou.mp3"

String bh2 = "d:/music/allweknow.mp3"

String bh3 = "d:/music/onmyway.mov"

String [] paths = {bh1, bh2, bh3};

Output: String[] songs = {"shapeofyou", "onmyway"};

B3: Cho mảng danh sách các bài hát như sau

Song[] songs = {new Song(1, "Shape Of You"), new Song(2, "All We Know"), new Song(3, "On My Way")) }



String BB

- ❖ Khi làm việc với kiểu dữ liệu văn bản Java cung cấp 3 class:
 - String (Object & Literal):
 - StringBuffer
 - StringBuilder
 - **Immutable String:** cannot change the original string value in heap
 - **Mutable String:** could change the original string value in heap

StringBuilder	StringBuffer
Single-Thread, no synchronized, Unsafe	Multi-Thread, safe synchronized, thread safe
	Avoidable conflict threads
Single-Thread => faster	Faster than String
Mutable: chuỗi có thể thay đổi về length, content	



String vs StringBuffer, StringBuider

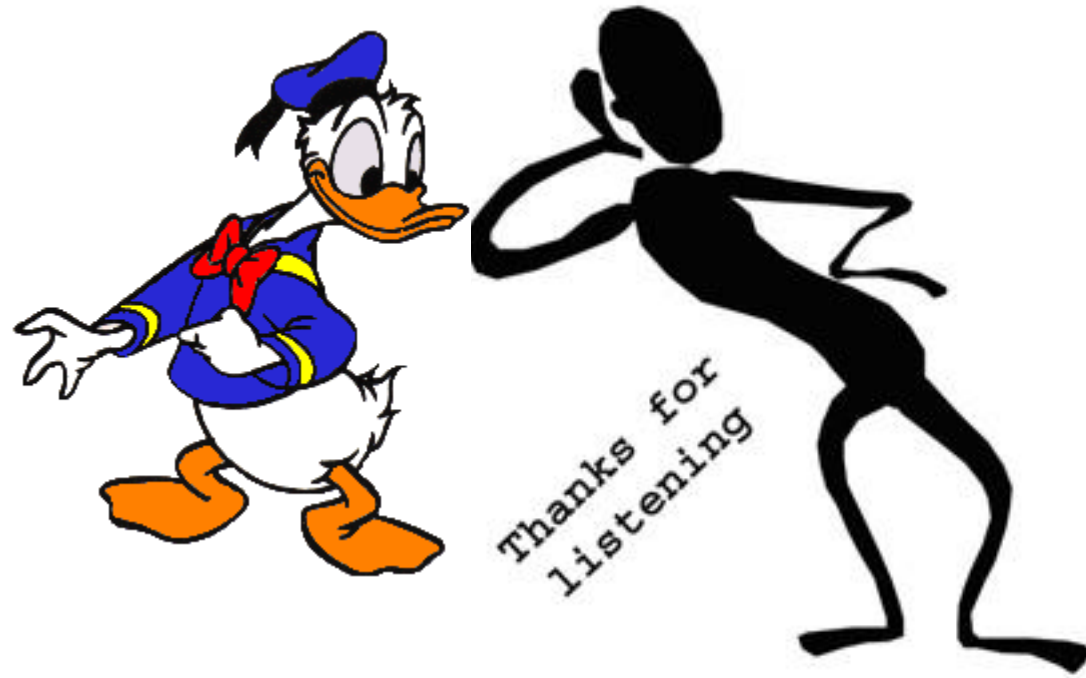
```
public class BkitAva04 {  
    // Java program to demonstrate difference between  
    // String, StringBuilder and StringBuffer  
  
    public static void concat1(String s1) {  
        System.out.println(System.identityHashCode(s1));  
        s1 = s1 + "ava04";  
        // s1 = new String(s1 + "ava04");  
        System.out.println(System.identityHashCode(s1));  
    }  
  
    public static void concat2(StringBuilder s2) {  
        s2.append("av04");  
    }  
  
    public static void concat3(StringBuffer s3) {  
        s3.append("ava04");  
    }  
  
    public static void main(String[] args) {  
        String s1 = "bkit";  
        concat1(s1); // s1 is not changed  
        System.out.println("String: " + s1);  
  
        StringBuilder s2 = new StringBuilder("bkit");  
        concat2(s2); // s2 is changed  
        System.out.println("StringBuilder: " + s2);  
  
        StringBuffer s3 = new StringBuffer("bkit");  
        concat3(s3); // s3 is changed  
        System.out.println("StringBuffer: " + s3);  
    }  
}
```



String vs StringBuffer, StringBuider

	String	StringBuffer	StringBuilder
Storage Area	Constant String Pool	Heap	Heap
Modifiable	No (immutable)	Yes (mutable)	Yest (Mutable)
Thread Safe	Yes	Yes	No
Performance	Fast	Very slow	Fast

- ‡ Khi nào nên sử dụng String, StringBuffer, StringBuilder.
- ❖ Các kiểu đối tượng **StringBuffer** và **StringBuilder** được sử dụng khi cần thao tác nhiều với các chuỗi ký tự. Không giống như **Strings**, các đối tượng kiểu **StringBuffer** và **StringBuilder** có thể thay đổi giá trị nhiều lần mà không tạo ra nhiều đối tượng không sử dụng.
- ❖ Kiểu **StringBuilder** được giới thiệu từ Java 5 và điểm khác biệt lớn nhất giữa **StringBuilder** và **StringBuffer** là các phương thức của **StringBuilder** không phải là *thread safe*.
- ❖ Nếu đặc tính *thread safe* là cần thiết thì **StringBuffer** là lựa chọn tốt nhất. Trong các trường hợp khác, chúng ta nên sử dụng **StringBuilder** để cho tốc độ xử lý nhanh hơn.



END