

Re: Check for Common character sequence ( I will pay)?

## Re: Check for Common character sequence ( I will pay)?

---

*Source:*

<http://www.tech-archive.net/Archive/DotNet/microsoft.public.dotnet.framework/2006-05/msg00215.html>

---

- *From:* "Greg Young" <[DruckDruckGoose@xxxxxxxxxxxx](mailto:DruckDruckGoose@xxxxxxxxxxxx)>
  - *Date:* Mon, 8 May 2006 18:12:20 -0400
- 

Yes FoundString is the return type .. you listed the class definition of FoundString ...

Yes you are returning an array of FoundString objects.

Cheers,

Greg Young  
MVP - C#

"C" <[C@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](mailto:C@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)> wrote in message  
<news:2A3D6E79-3DDF-467A-A475-6AA11A0A7E99@xxxxxxxxxxxxxxxxxxxx>

OK. Thanks kevin.

One question for yiu.

What is FoundString[] countSubstrings(string[] rgszValues);

Do I need to return an array?

countSubstrings(string[] rgszValues); is the method.

Is FoundString[] the return type?

"Kevin Spencer" wrote:

Dude, programming is all problem-solving. Syntax is just language. If you want to be a programmer, you have to learn problem-solving. You've been given a problem. Solve it. Doing your homework for you simply nullifies the benefit that the homework is intended to convey. Homework is not assigned for the purpose of testing your ability to turn in homework on time. It is assigned to help you learn.

Here's a few clues:

Re: Check for Common character sequence ( I will pay)?

## Re: Check for Common character sequence ( I will pay)?

Big things are made up of lots of little things. – A large problem is made up of smaller problems. A program is made up of lots of small instructions. Each instruction is a solution to a small problem. Break the large problem up into smaller ones, then break each of the smaller ones into the smallest possible problems you can.

In this case, you're starting with an array of strings, which is made up of one string per element.

You need to identify character sequences of 3 or more characters that appear in more than one string.

This means that you have to identify sequences 1 character at a time, working with 3 characters to start, and seeing how much farther it goes with each one.

So, obviously, the first step is to iterate through the first string, one character at a time, and test each 3-character sequence that results.

Again, obviously, if the 3-character sequence doesn't match, neither will a

4-character sequence starting from the same point match. So, if the 3-character sequence isn't found in the test string, you move on.

If the 3-character sequence *\*does\** match, you need to (1) make a note of that 3-character sequence, and (2) expand it by one character and test again, against the same groups that matched the 3-character sequence.

You would continue doing this until a given 3+-character sequence is not found to match any of the formerly-matching sequences.

After that, you would return to 3 characters, move forward 1 character, and continue.

Okay, that's it in a nutshell. Software does one thing very well: It repeats

itself. So, the rest of it is applying this solution to the entire set of strings. Of course, you will need to account for and prevent finding the same match by comparing a string to a string it has previously been compared

to. But remember that once you've iterated through the first string, and compared it with all the others, you don't need to include it when you test the others.

Now, for extra credit, study how I analyzed the problem and learn how to analyze programming problems for yourself. If you plan to become a programmer, this is what you will be doing, day in and day out. And there will be nobody around to do it for you!

Re: Check for Common character sequence ( I will pay)?

--  
HTH,

Kevin Spencer  
Microsoft MVP  
Professional Numbskull

Hard work is a medication for which  
there is no placebo.

--  
HTH,

Kevin Spencer  
Microsoft MVP  
Professional Numbskull

Hard work is a medication for which  
there is no placebo.

"C" <C@xxxxxxxxxxxxxxxxxxxxxxxxxxxx> wrote in message  
[news:DA610AF0-E692-4739-8B2F-E5365D5FC74F@xxxxxxxxxxxxxxxxxxxx](mailto:news:DA610AF0-E692-4739-8B2F-E5365D5FC74F@xxxxxxxxxxxxxxxxxxxx)

Yeah. Its a Project I have to do in C#. And I really don't have  
a clue.

If you can do it I will pay you throuh Paypal.

Need it done tonight though.

Interested?

"Greg Young" wrote:

Homework?

"C" <C@xxxxxxxxxxxxxxxxxxxxxxxxxxxx>  
wrote in message  
[news:43D5A724-542B-49C0-8B22-72CEE887F86F@xxxxxxxxxxxxxxxxxxxx](mailto:news:43D5A724-542B-49C0-8B22-72CEE887F86F@xxxxxxxxxxxxxxxxxxxx)

Hi,

Anyone know how I can do  
this?

I will pay anyone via PayPal  
who can do this for me. Let

Re: Check for Common character sequence ( I will pay)?

me know?

Write a function in C# that takes in an array of ASCII strings and finds all common character sequences. A common character sequence is defined as 3 or more adjacent characters that appear in more than one string in the array. Use the following function signature for your response:

```
//C#  
  
class FoundString {  
  
    public int cOccurrences;  
    //count of times this  
    substring  
  
    //occurred in the value array  
    public string szSubstring;  
    //substring that was found  
  
}  
  
FoundString[]  
countSubstrings(string[]  
rgszValues);
```

Ex. input of ("will-01",  
"a\_williams",  
"will\_01\_iam") would  
return

Re: Check for Common character sequence ( I will pay)?

"wil" (cOccurrences = 3)

"will" (cOccurrences = 3)

"ill" (cOccurrences = 3)

"iam" (cOccurrences = 2)

Provide coded unit tests cases that exercise your solution.

For extra credit:

1a. Order the results in descending order based on the number of input values each sequence matched.

1b. Prune the results by removing all results that are a substring of another result that matched the same number of input values. In the above example "wil" and "ill" would both be discarded because they are both a substring of "will" and all 3 values match 3 input strings

Re: Check for Common character sequence ( I will pay)?