**Homework: Strings and Text Processing**

**Problem 1. Strings in C**

* Describe the strings in C#.
* What is typical for the string data type?
* Describe the most important methods of the String class.

**Problem 2. Reverse string**

* Write a program that reads a string, reverses it and prints the result at the console.

*Example:*

| **input** | **output** |
| --- | --- |
| sample | elpmas |

**Problem 3. Correct brackets**

* Write a program to check if in a given expression the brackets are put correctly.

*Example of correct expression:* ((a+b)/5-d). *Example of incorrect expression:* )(a+b)).

**Problem 4. Sub-string in text**

* Write a program that finds how many times a sub-string is contained in a given text (perform case insensitive search).

*Example:*

The target sub-string is in

The text is as follows: We are liv**in**g **in** an yellow submar**in**e. We don't have anyth**in**g else. **in**side the submar**in**e is very tight. So we are dr**in**k**in**g all the day. We will move out of it **in** 5 days.

The result is: 9

**Problem 5. Parse tags**

* You are given a text. Write a program that changes the text in all regions surrounded by the tags <upcase> and</upcase> to upper-case.
* The tags cannot be nested.

*Example:* We are living in a yellow submarine. We don't have anything else.

*The expected result:* We are living in a YELLOW SUBMARINE. We don't have ANYTHING else.

**Problem 6. String length**

* Write a program that reads from the console a string of maximum 20 characters. If the length of the string is less than 20, the rest of the characters should be filled with \*.
* Print the result string into the console.

**Problem 7. Encode/decode**

* Write a program that encodes and decodes a string using given encryption key (cipher).
* The key consists of a sequence of characters.
* The encoding/decoding is done by performing XOR (exclusive or) operation over the first letter of the string with the first of the key, the second – with the second, etc. When the last key character is reached, the next is the first.

**Problem 8. Extract sentences**

* Write a program that extracts from a given text all sentences containing given **word**.

*Example:*

*The word is:* **in**

*The text is:* We are living **in** a yellow submarine. We don't have anything else. Inside the submarine is very tight. So we are drinking all the day. We will move out of it **in** 5 days.

*The expected result is:* We are living in a yellow submarine. We will move out of it in 5 days.

*Consider that the sentences are separated by . and the words – by****non-letter symbols****.*

**Problem 9. Forbidden words**

* We are given a string containing a list of forbidden words and a text containing some of these words.
* Write a program that replaces the forbidden words with asterisks.

*Example text:* Microsoft announced its next generation PHP compiler today. It is based on .NET Framework 4.0 and is implemented as a dynamic language in CLR.

*Forbidden words:* PHP, CLR, Microsoft

*The expected result:* \*\*\*\*\*\*\*\*\* announced its next generation \*\*\* compiler today. It is based on .NET Framework 4.0 and is implemented as a dynamic language in \*\*\*.

**Problem 10. Unicode characters**

* Write a program that converts a string to a sequence of C# Unicode character literals.
* Use format strings.

*Example:*

| **input** | **output** |
| --- | --- |
| Hi! | \u0048\u0069\u0021 |

**Problem 11. Format number**

* Write a program that reads a number and prints it as a decimal number, hexadecimal number, percentage and in scientific notation.
* Format the output aligned right in 15 symbols.

**Problem 12. Parse URL**

* Write a program that parses an URL address given in the format: [protocol]://[server]/[resource] and extracts from it the [protocol], [server] and [resource] elements.

*Example:*

| **URL** | **Information** |
| --- | --- |
| http://telerikacademy.com/Courses/Courses/Details/212 | [protocol] = http  [server] = telerikacademy.com  [resource] = /Courses/Courses/Details/212 |

**Problem 13. Reverse sentence**

* Write a program that reverses the words in given sentence.

*Example:*

| **input** | **output** |
| --- | --- |
| C# is not C++, not PHP and not Delphi! | Delphi not and PHP, not C++ not is C#! |

**Problem 14. Word dictionary**

* A dictionary is stored as a sequence of text lines containing words and their explanations.
* Write a program that enters a word and translates it by using the dictionary.

*Sample dictionary:*

| **input** | **output** |
| --- | --- |
| .NET | platform for applications from Microsoft |
| CLR | managed execution environment for .NET |
| namespace | hierarchical organization of classes |

**Problem 15. Replace tags**

* Write a program that replaces in a HTML document given as string all the tags <a href="…">…</a> with corresponding tags [URL=…]…/URL].

*Example:*

| **input** | **output** |
| --- | --- |
| <p>Please visit <a href="http://academy.telerik. com">our site</a> to choose a training course. Also visit <a href="www.devbg.org">our forum</a> to discuss the courses.</p> | <p>Please visit [URL=http://academy.telerik. com]our site[/URL] to choose a training course. Also visit [URL=www.devbg.org]our forum[/URL] to discuss the courses.</p> |

**Problem 16. Date difference**

* Write a program that reads two dates in the format: day.month.year and calculates the number of days between them.

*Example:*

Enter the first date: 27.02.2006

Enter the second date: 3.03.2006

Distance: 4 days

**Problem 17. Date in Bulgarian**

* Write a program that reads a date and time given in the format: day.month.year hour:minute:second and prints the date and time after 6 hours and 30 minutes (in the same format) along with the day of week in Bulgarian.

**Problem 18. Extract e-mails**

* Write a program for extracting all email addresses from given text.
* All sub-strings that match the format @… should be recognized as emails.

**Problem 19. Dates from text in Canada**

* Write a program that extracts from a given text all dates that match the format DD.MM.YYYY.
* Display them in the standard date format for Canada.

**Problem 20. Palindromes**

* Write a program that extracts from a given text all palindromes, e.g. ABBA, lamal, exe.

**Problem 21. Letters count**

* Write a program that reads a string from the console and prints all different letters in the string along with information how many times each letter is found.

**Problem 22. Words count**

* Write a program that reads a string from the console and lists all different words in the string along with information how many times each word is found.

**Problem 23. Series of letters**

* Write a program that reads a string from the console and replaces all series of consecutive identical letters with a single one.

*Example:*

| **input** | **output** |
| --- | --- |
| aaaaabbbbbcdddeeeedssaa | abcdedsa |

**Problem 24. Order words**

* Write a program that reads a list of words, separated by spaces and prints the list in an alphabetical order.

**Problem 25. Extract text from HTML**

* Write a program that extracts from given HTML file its title (if available), and its body text without the HTML tags.

*Example input:*

<html>

<head><title>News</title></head>

<body><p><a href="http://academy.telerik.com">Telerik

Academy</a>aims to provide free real-world practical

training for young people who want to turn into

skilful .NET software engineers.</p></body>

</html>

*Output:*

Title: News

Text: Telerik Academy aims to provide free real-world practical training for young people who want to turn into skilful .NET software engineers.