Regular Expressions

Peter Levinsky IT, Roskilde

24.02.2022

Intro Regular Expressions

Standard textual syntax for representing **patterns** for matching text.
Can be used for (Purpose):

- **Validate**
  Validate text to ensure that it matches a predefined pattern (such as an email address)

- **Find**
  Find specific character patterns

- **Extract**
  Extract substrings from a text

- **Replace etc**
  Replace, edit or delete text substrings

Regular Expressions - Patterns

Examples of pattern:

Pattern	Matching
hello	contains {hello}
gr(a|e)y	contains {gray, grey}
b[aeiou]bble	contains {babble, bebble, bibble, bobble, bubble}
rege(x(es)?|xps?)	contains {regex, regexes, regexp, regexps}
go*gle	contains {ggle, gogle, google, gooogle, goooogle, ...}
go+gle	contains {gogle, google, gooogle, goooogle, ...}

+ => 1-* | => OR   {n} => n times   ex: (0|1){8} => byte 01101010
? => 0-1   () => Cluster   {n, } => at least n
* => 0-*   [] => Choice between   {n,m} => Between n and m times

Regular Expressions – More Patterns

Special groups of characters

\d   => digital cipher   ~= [0-9]
\D   => Not digital cipher   ~= ^\d or [^0-9]
\w   => word-character
\W   => non word

More Special Patterns

^   => begins with
$ => end with
^[abc]{1,2}$
          => exact both begin-with and end-with   ~= a,b,c,aa,ab,ac,ba,bb,bc,ca,cb,cc

Regular Expressions – More Patterns

Some links to overview of Patterns

- **https://cs.lmu.edu/~ray/notes/regex/**
- **https://www.rexegg.com/regex-quickstart.html**
- **https://networking.ringofsaturn.com/Web/regex.php** **(Posix a standard)**

C# Related regular expressions Link:
- **https://docs.microsoft.com/en-us/dotnet/standard/base-types/regular-expression-language-quick-reference**

- **https://www.c-sharpcorner.com/article/c-sharp-regex-examples/**

Exercise regex#1

Regular Expressions – How to use in C#

```
String pattern = "… regular expression …";
Regex regex = new Regex(pattern);
String someThingToValidate = "…someText…";


// the validation
Regex.IsMatch(someTringToValidate) true |
false
```

Regular Expressions – How to use in C# - part2

Splitting:

```
string Text = "1 One, 2 Two, 3 Three is good.";
string[] digits = Regex.Split(Text, @"\d\s*");
```

One,
Two,
Three is good.

Regular Expressions – How to use in C# - part3

Replacing:

string Text = "1 One, 2 Two, 3 Three is good.";
String str = Regex.Replace(Text, @"[^0-9a-zA-Z]+", "")

1One2Two3Threeisgood

Demo

Exercise regex#2