Using

Regular Expressions

in

Borland Delphi

Renato Mancuso - BUG UK Meeting

What Regular Expressions are

A Regular Expression is a string that describes a target text by defining the features that the target text must posses

ex: the target text must start with a lower case letter, followed by 1 to 3 digits, and it must be terminated by a dot:

^ [a- z]\d{1,3}.*\.$

Common Uses for Regular Expressions

- Validate data
- Pull pieces of text out of larger blocks
- Substitute new text for old text

Syntax

Characters and Metacharacters

- character shorthands: \a \n \r \t
- octal escapes: \012
- hex and Unicode escapes: \x0A
- control characters: \cH (backspace)

Character Classes and Class-like Construct

- normal classes: [...], [^ ...]
- almost any character (dot): .
- class shorthands: \d, \s, \w, \S, \D, \W
- POSIX character classes: [[:alpha:]], [[:upper:]]

Anchors and Zero-width Assertions

- start of line/string: ^ , \A
- end of line/string: $, \Z
- start of match: \G
- word boundary: \b, \B, \<, \>
- lookahead: (?=...), (?!...)
- lookbehind: (?<=...), (?<!...)

Comments and Mode Modifiers

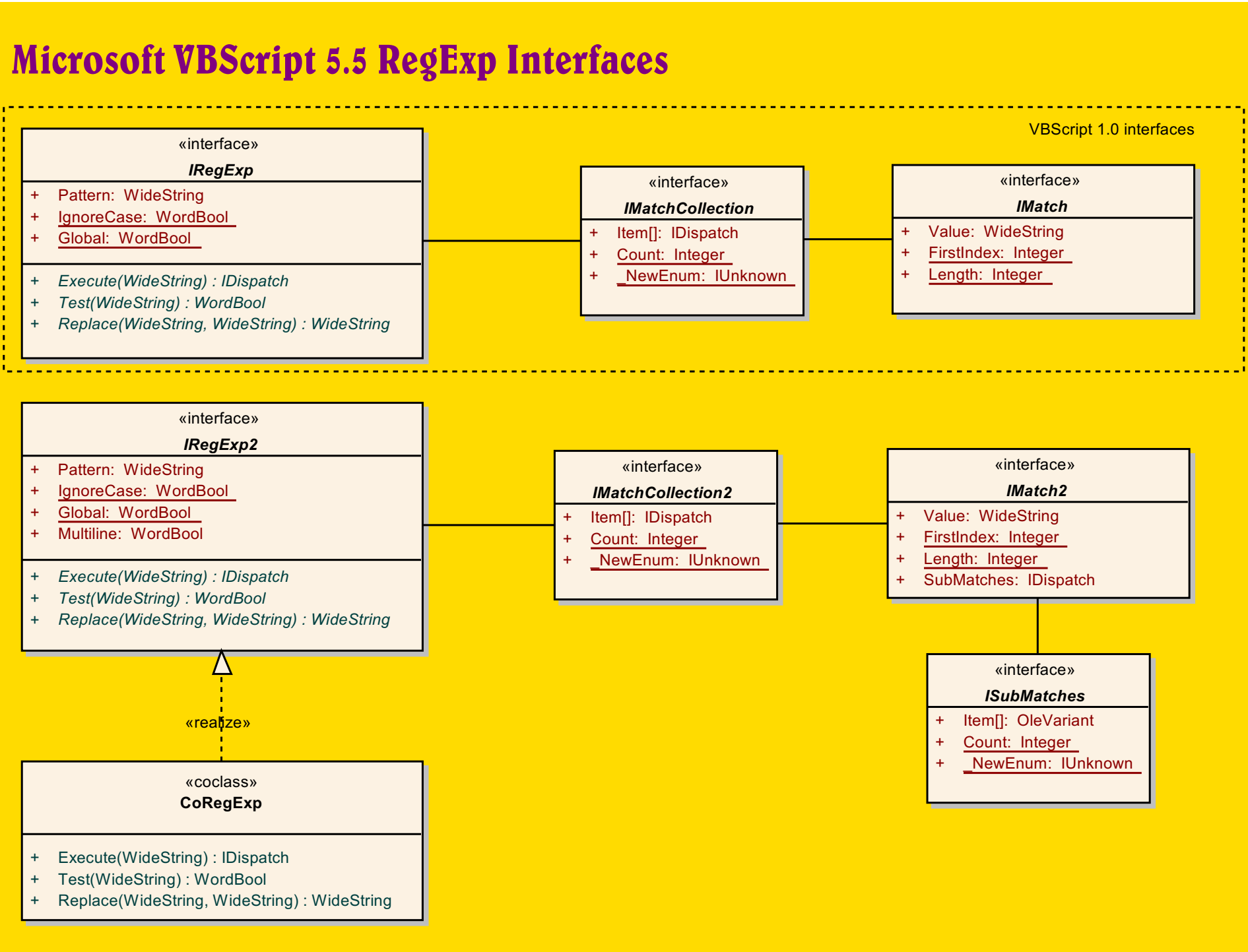
- multi-line mode: m
- single-line mode: s (DOTALL)
- case insensitive mode: i
- free spacing mode: x
- inline mode modifiers: (?x), (?-x)
- comments: (?#...), # (free spacing mode)
- literal text spans: \Q...\E

Grouping, Capturing, Conditionals e Control

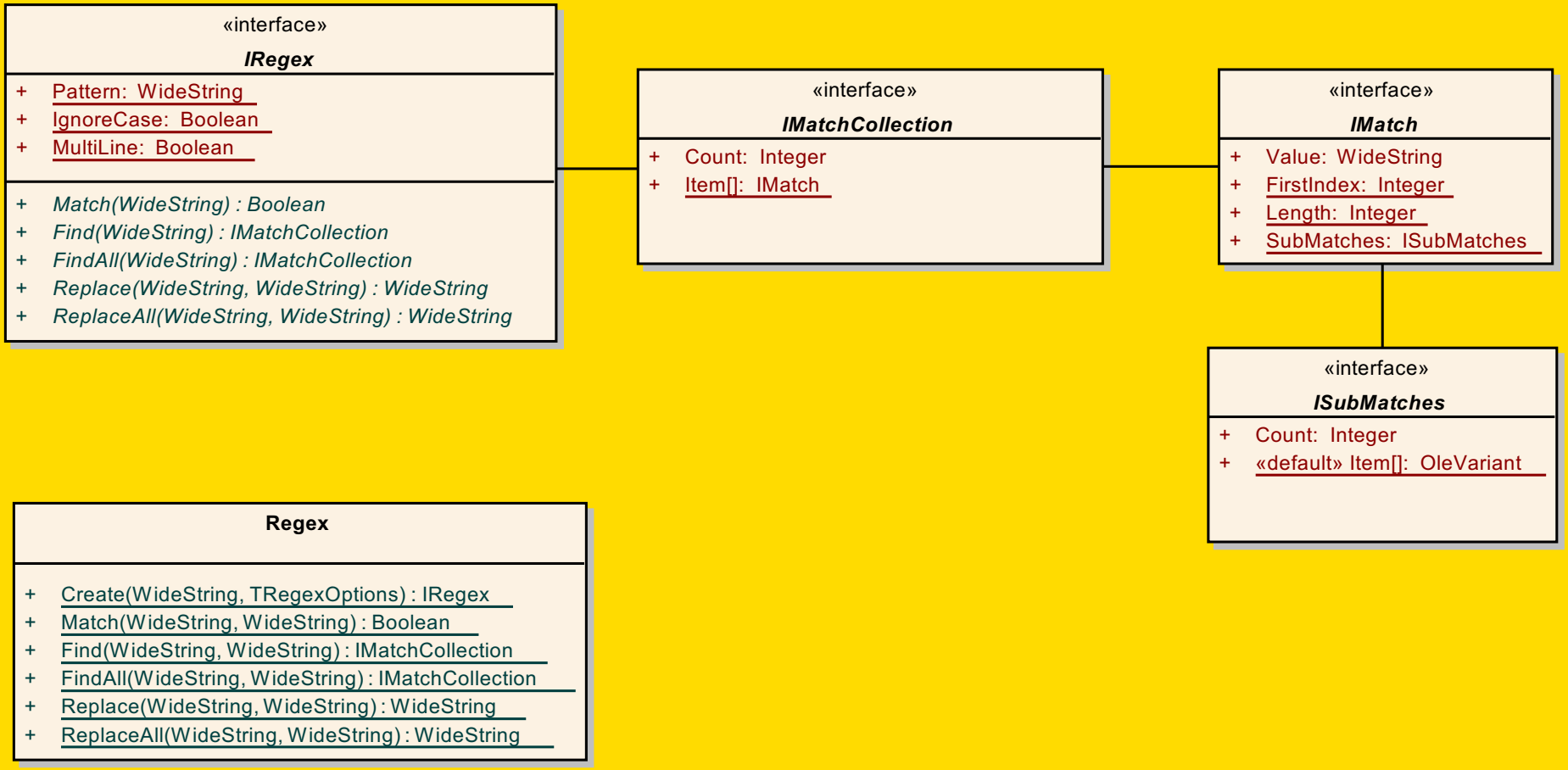
- capturing and grouping parentheses: (...)
- back references: \1 \2
- grouping only parentheses: (?:...)
- named captures: (?<name>...) [.NET], (?P<name>...) [PCRE]
- atomic grouping: (?>...)
- alternation: ...|...
- conditional: (?if then|else)
- greedy quantifiers: *, +, ?, {n,m}
- lazy quantifiers: *?, +?, ??, {n,m}?
- possessive quantifiers: *+, ++, ?+, {n,m}+

Using the VBScript 5.5 RegExp in Delphi

Microsoft VBScript 5.5 RegExp Interfaces

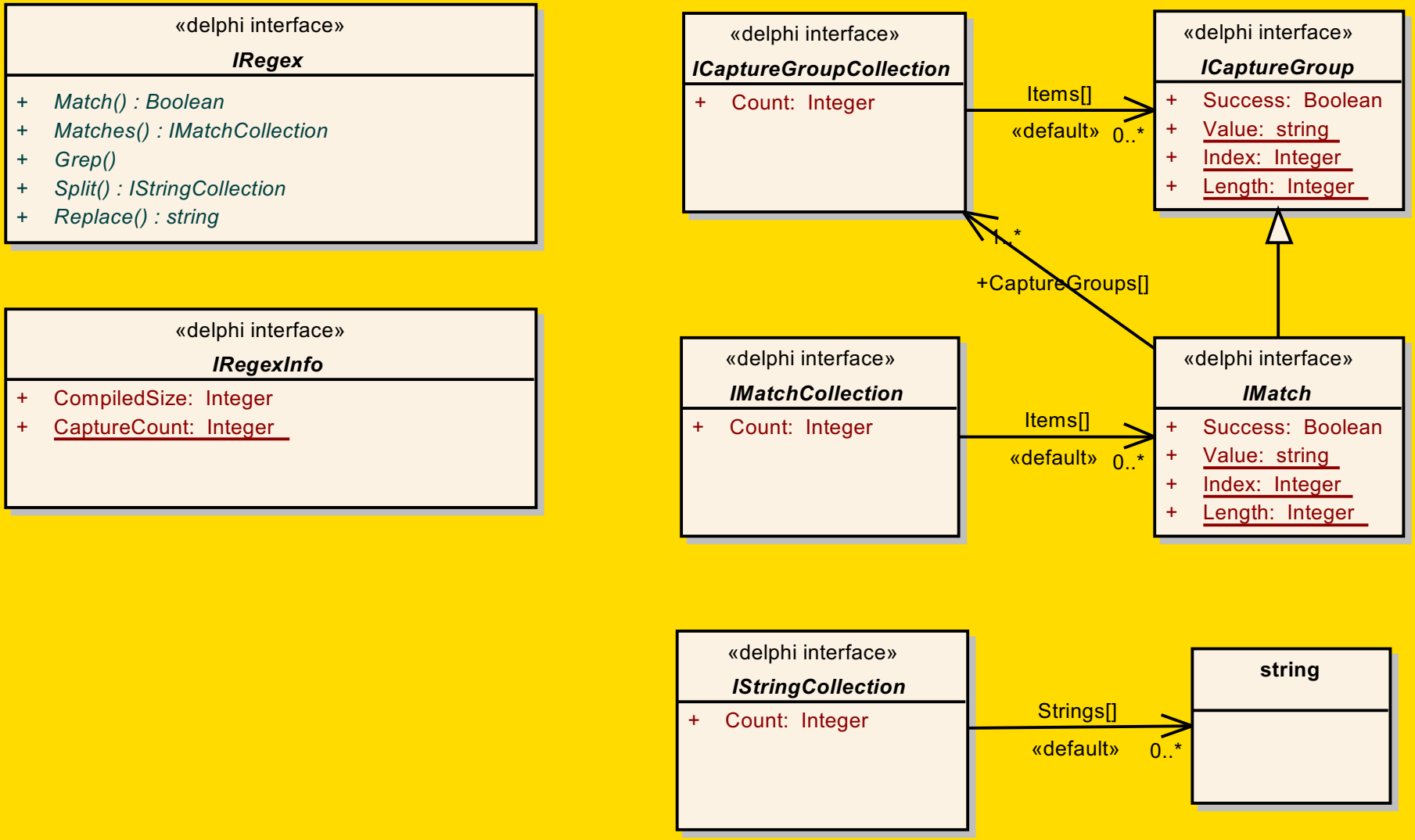


Wrapping the VBScript 5.5 RegExp Interfaces



Using Pcre in Delphi

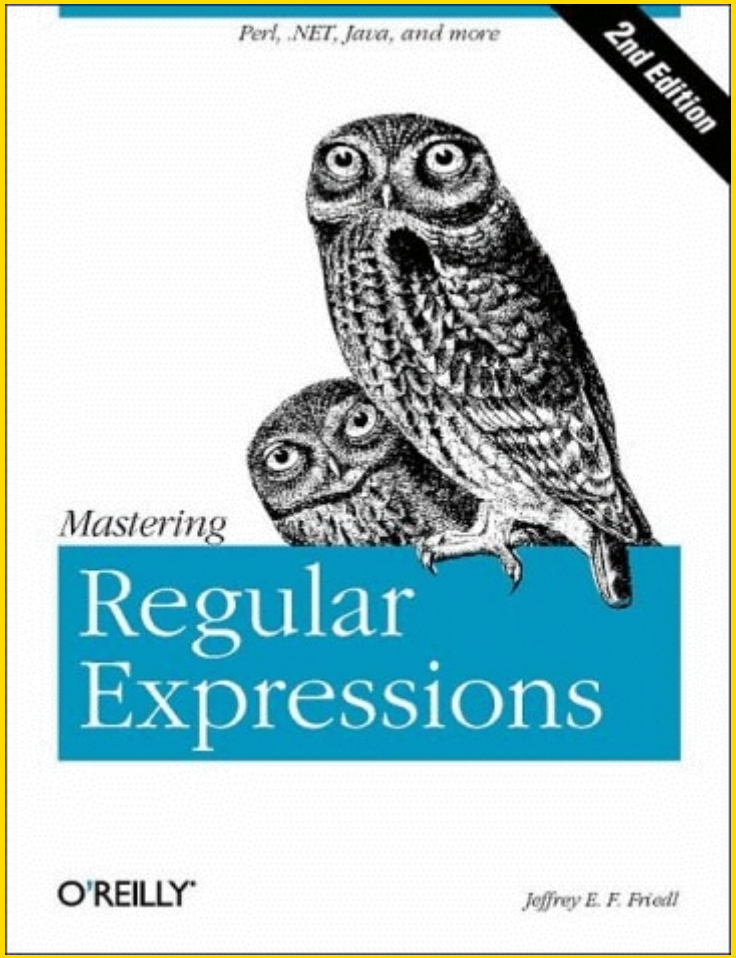
PCRE 4.4 Delphi wrapper

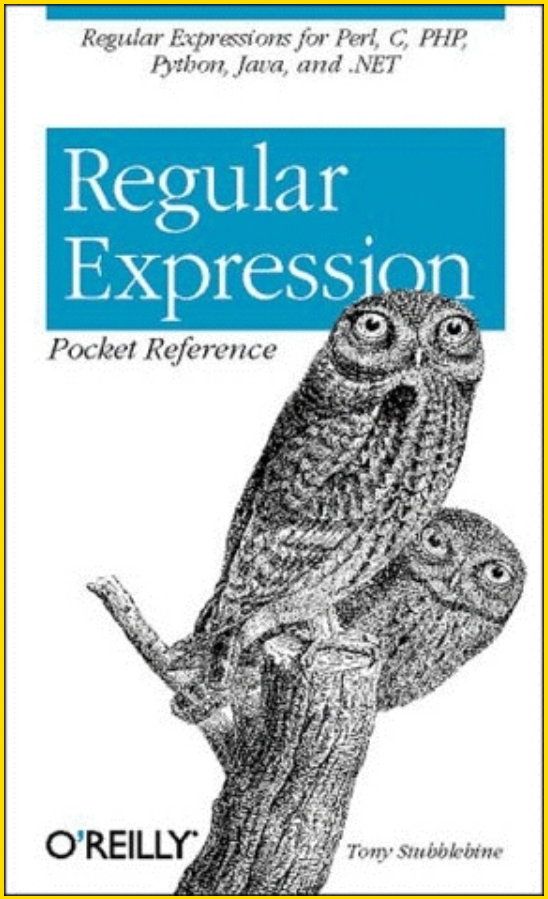


References

Books

- Jeffrey E. F. Friedl - Mastering Regular Expressions (2nd edition) - O'Reilly
- Tony Stubblebine - Regular Expression Pocket Reference - O'Reilly





Web

- http://www.pcre.org - PCRE
- http://msdn.microsoft.com/library - VBScript RegExp docs & .NET Regex docs
- http://www.boost.org/libs/regex/doc/index.html - boost.regex documentation [C++](John Maddock)
- http://www.renatomancuso.com - Delphi wrappers for PCRE and VBScript RegExp

mancuso@renatomancuso.com