PYTHON ADVANCED PROGRAMMING

Basic Concepts

Chitra G M and P Rama Devi
Department of Computer Science Engineering

PYTHON ADVANCED PROGRAMMING

Basic Concepts

P Rama Devi
Department of Computer Science and Engineering

PYTHON ADVANCED PROGRAMMING

re—Regular Expressions

Regular expressions are text-matching patterns described with a formal syntax. The patterns are interpreted as a set of instructions, which are then executed with a string as input to produce a matching subset or modified version of the original

Expressions can include literal text matching, repetition, pattern composition, branching, and other sophisticated rules

Regular expressions are typically used in applications that involve a lot of text processing. For example, they are commonly used as search patterns in text-editing programs used by developers, including vi, emacs, and modern IDEs

PYTHON ADVANCED PROGRAMMING

Finding Patterns in Text

The most common use for re is to search for patterns in text. The search() function takes the pattern and text to scan, and returns a Match object when the pattern is found. If the pattern is not found, search() returns None .

```
import re
pattern = ’this’
text = ’Does this text match the pattern?’
match = re.search(pattern, text)
s = match.start()
e = match.end()
```

PYTHON ADVANCED PROGRAMMING

Compiling Expressions

re includes module-level functions for working with regular expressions as text strings, but it is more efficient to compile the expressions a program uses frequently. The compile() function converts an expression string into a RegexObject .

```
sentence='start with a beautiful smile end.'
pattern=re.compile(r'.')
matches=pattern.finditer(sentence)
for match in matches:
    print(match)
```

PYTHON ADVANCED PROGRAMMING

Special Characters

1	. - Any Character Except New Line
2	\d - Digit (0-9)
3	\D - Not a Digit (0-9)
4	\w - Word Character (a-z, A-Z, 0-9, _)
5	\W - Not a Word Character
6	\s - Whitespace (space, tab, newline)
7	\S - Not Whitespace (space, tab, newline)

PYTHON ADVANCED PROGRAMMING

Special Characters

1	\b - Word Boundary
2	\B - Not a Word Boundary
3	^ - Beginning of a String
4	$ - End of a String

PYTHON ADVANCED PROGRAMMING

Special Characters

1	[] - Matches Characters in brackets
2	[^] - Matches Characters NOT in brackets
3	| - Either Or
4	( ) - Group

PYTHON ADVANCED PROGRAMMING

Special Characters

	Quantifiers:
1	* - 0 or More
2	+ - 1 or More
3	? - 0 or One
4	{3} - Exact Number
5	{3,4} - Range of Numbers (Minimum, Maximum)

THANK YOU

Chitra G M and P Rama Devi
Department of CSE

pramadevi@pes.edu
chitragm@pes.edu