Regular Expressions in Python and Perl

I Introduction

((Coming soon))

2 Basic Symbols

| Wat do? | How Perl do? | How Python do? |
|------------------------|----------------------------|----------------------------|
| Custom character class | [,] | [] |
| Negated custom | [^] | [^] |
| character class | | |
| Ranges | [a-z] (with '-' escaped if | [a-z] (with '-' escaped if |
| | it comes last) | it comes last) |
| Alternation ("or") | | |

3 Zero-width assertions

| Wat do? | How Perl do? | How Python do? |
|--------------------|--------------|----------------|
| Word boundary | \b | \b |
| Anywhere but word | \B | \B |
| boundary | | |
| Beginning of | ^ / \A | ^ / \A |
| line/string | | |
| End of line/string | \$ / \Z | \$ / \Z |

4 Captures and Groups

| Wat do? | How Perl do? | How Python do? |
|---------------------|--|--------------------------|
| Capturing group | () or (? <name>)</name> | () or (?P <name>)</name> |
| Non-capturing group | (?:) | (?:) |
| Backreference to a | \1, \g1 | \1 |
| specific group | | |
| Named backreference | $\ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $ | (?P=name) |

5 Character Classes

| Wat do? | How Perl do? | How Python do? |
|-------------------------|---------------------------------------|-------------------------|
| Any character (except | | |
| newline) | | |
| Match a non-"word" | \W | \W |
| character | · | · |
| Match a "word" | \w or [[:word:]] | \w |
| character | | |
| Case | [[:upper:]] or [[:lower:]] | N/A |
| Whitespace (not | N/A | N/A |
| including newlines) | | |
| Whitespace (including | $\setminus s \text{ or } [[:space:]]$ | \s |
| newline) | | |
| Match a non-whitespace | \S | \S |
| character | | |
| Match a digit character | \d or [[:digit:]] | \d |
| Match a non-digit | \D | $\backslash \mathrm{D}$ |
| character | | |
| Any hexadecimal digit | [[:xdigit:]] | N/A |
| Any octal digit | N/A | N/A |
| Any graphical character | [[:punct:]] | N/A |
| excluding "word" | | |
| characters | | |
| Any alphabetical | [[:alpha:]] | N/A |
| character | | |
| Any alphanumerical | [[:alnum:]] | N/A |
| character | FF 11 | |
| ASCII character | [[:ascii:]] | N/A |

6 Lookarounds

((Coming soon))

7 Multiplicity

((Coming soon))

8 Examples

((Coming soon))

9 What type of regex does my text editor use?

((Coming soon))

10 Which type of regex does \$LINUX_UTIL use?

((Coming soon))

11 Conclusion

((Coming soon))