

## Section 9 Handout

## Regular Expression Reference Sheet

The follow table provides an overview of the common pattern options you may use for regex:

Regex quick reference

[abc]	A single character of: a, b, or c	.	Any single character	(...)	Capture everything enclosed
[^abc]	Any single character except: a, b, or c	\s	Any whitespace character	(a b)	a or b
[a-z]	Any single character in the range a-z	\S	Any non-whitespace character	a?	Zero or one of a
[a-zA-Z]	Any single character in the range a-z or A-Z	\d	Any digit	a*	Zero or more of a
^	Start of line	\D	Any non-digit	a+	One or more of a
\$	End of line	\w	Any word character (letter, number, underscore)	a{3}	Exactly 3 of a
\A	Start of string	\W	Any non-word character	a{3,}	3 or more of a
\Z	End of string	\b	Any word boundary	a{3,6}	Between 3 and 6 of a

options:    i case insensitive    m make dot match newlines    x ignore whitespace in regex    o perform #(...) substitutions only once

On Wednesday's lecture, we will introduce regular expression functions you may use in PHP. We will use these in Thursday's lab. Keep this sheet handy!

## PHP Regex Functions

Function	Description
preg_match(regex, str)	Returns whether str matches regex
preg_replace(regex, repl, str)	Returns a new string with all substrings of str that match regex replaced by repl
preg_split(regex, str)	Returns an array of strings from given str split apart using given regex as delimiter

## Regex Section Problems

Write regex for each of the following:

- UW Student ID numbers (non-negative 7-digit numbers).
- DNA (strings containing only A, C, G, T, ignoring letter-casing).
- camelCased strings containing only letters and at least one capitalized letter (e.g. "superCuteBabyCuttlefish" or "toString" but not "tostring" or "bookData23").
- A 16-digit credit card number, with optional dashes for every 4 numbers (e.g. 1234-1234-1234-1234 or 1234123412341234).

- (e) A version of "Google" with an odd-number of o's.
- (f) A https URL ending with "edu", "com", "org", or "gov" (e.g. "https://www.cs.washington.edu").
- (g) A real number such as 3.14 or -42.8775.
- (h) Dollar amounts of at least \$100.00.
- (i) Strings that contain at least two consecutive vowels (e.g. "foo", "queue", "QUERY", but not "bar", "panini", "Rhythm").
- (j) Strings that contain only consonants (e.g. "sky", "PHP", "SQL", "rhythm" but not "panini", "vowel", "foo").
- (k) Strings that contain "cat" or "dog" but not both (ignoring letter-casing).
- (l) Digimon names: strings starting with a capital letter and ending in "mon", ignoring casing (e.g. "Agumon", "Gabumon", "MON", but not "Agu" or "Montana").
- (m) Binary strings with at least one 1 and at most two 0's.
- (n) A capitalized first name, single (upper-case) middle initial, and last name with spaces between each and with the first and last name both having lengths fewer than 10 letters.