Lab Assignment Week 4 : Compiler Design

Lexical Analysis

Common Instruction for all questions

Write regular expression and equivalent code in any programming language, STRING based BUILT IN functions/classes for given job shall be used

Q 1. Write a function named is_integer() which inputs a string (may contain any ascii character) and returns true(1) if input string is a valid integer, otherwise false(0).

Example :

(i) I/P "23"	O/P 1	(ii) I/P "-23"	O/P 1
(iii) I/P "2a3"	O/P 0	(iv) I/P "2.3"	O/P 0

Q 2. Write a function named is_float() which inputs a string (may contain any ascii character) and returns true(1) if input string is a valid floating point number, otherwise false(0).

Example :

(i) I/P "2.3"	O/P 1	(ii) I/P "-2.3"	O/P 1
(iii) I/P "23"	O/P 0	(iv) I/P "-23"	O/P 0
(V) I/P "0.23"	O/P 1	(vi) I/P "-0.23"	O/P 1
(vii) I/P "0.2.3"	O/P 0	(viii) I/P "2..3"	O/P 0
(ix) I/P ".23"	O/P 1	(x) I/P "23."	O/P 1

Q3. Write a function named chk_ipaddress() which inputs a string (may contain any ascii character) and checks if the given string is a valid IP address or not, and output type of IP address, i.e. (IPv4/IPv6/Invalid-IP)

Note: Take care of IPv4 and IPv6, including collapsed/compressed format also

(i)	I/P	123.234.43.2	O/P	IPv4
(ii)	I/P	123.234.43.2.4.3	O/P	Invalid
(iii)	I/P	2002:4559:1FE2:0000:0000:0000:4559:1FE2	O/P	IPv6
(iv)	I/P	2002:4559:1FE2:0:0:0:4559:1FE2	O/P	IPv6
(v)	I/P	2002:4559:1FE2::4559:1FE2	O/P	IPv6
(vi)	I/P	2002:4559:1FE2::4559:1Fx2	O/P	Invalid
(vii)	I/P	123.284.43.2	O/P	Invalid

Q4. Write a function named is_email() which inputs a string (may contain any ascii character) and returns true(1) if input string is a valid email-id, otherwise false(0).

Note: Confirm rules from authenticated document and mention the same in comments.

Followings are test cases obtained by
http://blogs.msdn.com/b/testing123/archive/2009/02/05/email-address-test-cases.aspx

Valid Email address	Reason
email@domain.com	Valid email
firstname.lastname@domain.com	Email contains dot in the address field
email@subdomain.domain.com	Email contains dot with subdomain
firstname+lastname@domain.com	Plus sign is considered valid character
email@123.123.123.123	Domain is valid IP address
email@[123.123.123.123]	Square bracket around IP address is considered valid
"email"@domain.com	Quotes around email is considered valid
1234567890@domain.com	Digits in address are valid
email@domain-one.com	Dash in domain name is valid
_____@domain.com	Underscore in the address field is valid
email@domain.name	.name is valid Top Level Domain name
email@domain.co.jp	Dot in Top Level Domain name also considered valid (use co.jp as example here)
firstname-lastname@domain.com	Dash in address field is valid
Invalid Email address	**Reason**
plainaddress	Missing @ sign and domain
#@%^%#$@#$@#.com	Garbage
@domain.com	Missing username
Joe Smith <email@domain.com>	Encoded html within email is invalid
email.domain.com	Missing @
email@domain@domain.com	Two @ sign
.email@domain.com	Leading dot in address is not allowed
email.@domain.com	Trailing dot in address is not allowed
email..email@domain.com	Multiple dots
あいうえお@domain.com	Unicode char as address
email@domain.com (Joe Smith)	Text followed email is not allowed
email@domain	Missing top level domain (.com/.net/.org/etc)
email@-domain.com	Leading dash in front of domain is invalid
email@domain.web	.web is not a valid top level domain
email@111.222.333.44444	Invalid IP format
email@domain..com	Multiple dot in the domain portion is invalid