CMPT 379 - Lexical Analysis Practice

(1) a. Let $\Sigma = \{0, 1\}$. How many elements in the set Σ?

b. Provide Σ^3.

c. Explain what Σ^* represents.

d. Give regexp for all strings in Σ^* equal to decimal number 6.

e. Give regexp for all strings in Σ^* that are powers of two.

f. Give regexp for all strings in Σ^* that are even numbers.

g. Give regexp for all strings in Σ^* that are Binary Coded Decimal (BCD) numbers (include the empty string). A BCD number is a decimal number where each decimal digit is encoded using a 4-bit representation of its binary value. For example, the BCD number of 2509 is 0010010100001001

(2) You are given the following ordered list of token definitions:

`TOKEN_A`	cda^*
`TOKEN_B`	c^*a^*c
`TOKEN_C`	c^*b

Provide the tokenized output for the following input strings using the greedy longest match lexical analysis method. Provide the list of tokens and the lexeme values.

a. *cdaaab*

b. *cdccc*

c. *ccc*

d. *cdccd*