CMPT 379: Principles of Compiler Design

Lexical Analysis Worksheet

Given the following token definitions as regular expressions:

TOKEN_A c d a *
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.

1. cdaaab

2. cdccc

3. ccc

4. cdccd