

CMPT 379: Principles of Compiler Design

Lexical Analysis Worksheet (Answers)

Given the following token definitions as regular expressions:

TOKEN.A `cd*`
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`

TOKEN.A `cdaaa`
TOKEN.C `b`

2. `cdccc`

TOKEN.A `cd`
TOKEN.B `ccc`

3. `ccc`

TOKEN.B `ccc`

4. `cdccd`

TOKEN.A `cd`
TOKEN.B `cc`
ERROR