

Source:

1 | **The Role of the Lexical Analyzer**

1.1 | **Lexical vs Syntactic analysis**

1. Syntax and raw text are different and can be treated separately
2. it may be more efficient
3. better portability

1.2 | **terms: tokens, patterns, lexemes**

- #definition token: is a name and a value, where the name like a keyword or an identifier and the value is a section of the source text?
- #definition pattern: basically a regex of what string structures are allowed
- #definition lexeme: part of the source text that is matched by a pattern as an instance of a token

1.3 | **common token breakdown**

1. keywords (usually one per keyword)
2. operators (sometimes in operator classes)
3. identifiers
4. constants (sometimes one per type)
5. punctuation (usually one per each, including parens, comma, and semicolon)

**