EDIDIONG IME-ESSIEN

17/SCI01/041

CSC312

**Question**

Differentiate between lexemes and tokens. Give examples to support your discussion.

**Answers**

A lexeme is a sequence of alphanumeric characters in a token. The term is used in both the study of language and in the lexical analysis of computer program compilation. In the context of computer programming, lexemes are part of the input stream from which tokens are identified. An invalid or illegal token produces an error. A lexeme is one of the building blocks of language. A lexeme is a sequence of characters in the source program that matches the pattern for a token and is identified by the lexical analyzer as an instance of that token. Examples are: while y, (,) etc

A token is a pair consisting of a token name and an optional attribute value. The token name is an abstract symbol representing a kind of lexical unit, e.g., a particular keyword, or sequence of input characters denoting an identifier. The token names are the input symbols that the parser processes. Examples are : IF, void, ID , number, return etc. The token is a syntactic category that forms a class of lexemes that means which class the lexeme belong is it a keyword or identifier or anything else. One of the major tasks of the lexical analyzer is to create a pair of lexemes and tokens, that is to collect all the characters.