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. A token is a pair consisting of a token name and an optional attribute value.

One example of a lexeme would be the word 'create'. When appearing alone, it conveys a single meaning. But when used as the stem, it can take on inflected forms such as 'created' and 'creating', which alter the grammatical function of the original stem. It can also take a on a derived form such as 'creator', which is a more distinct, though related, lexical unit. When the same lexeme appears in the form of a 'catena', or chain of words, such as the Biblical phrase, 'crown of creation' to describe mankind, one can still see its relationship with the original lexeme.

Example of tokens

There are five categories of tokens: 1) constants, 2) identifiers, 3) operators, 4) separators, and 5) reserved words. For example, the reserved words "new" and "function" are tokens of the JavaScript language. Operators, such as +, -, \*, and /, are also tokens of nearly all programming languages.