

UNIT 2. Lexical Analysis

The main task of the lexical analysis is to read the source program from left-to-right and group the symbols into sequences of characters with a collective meaning, which are called tokens. The lexical analyzer or scanner must:

- identify errors in the source program;
- strip out from the source program comments and white space characters (tab, newline, space);
- associate a line number from the source program with a given error message.

The main objectives of Unit 2 are:

- To introduce the role of the lexical analyzer into the architecture of a language processor.
- To learn different techniques for the scanner implementation
- To review Regular Expressions and Finite Automata.
- To know how to implement a scanner by means of the recognition of tokens using Regular Expressions and Deterministic Finite Automata.