



Wie erkennt man ein Lexem?

Die Menge aller gültigen Zeichen und Bezeichner einer Programmiersprache ist üblicherweise eine reguläre Sprache.

digit ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9
letter ::= A | B | C | ... | Z | a | b | ... | z
sign... ::= + | - | : | * | < | > | = | := | == | ! | && | ; | (|) | { | } | [|]
integer ::= digit digit*
identifizier ::= letter (letter | digit)*
if ::= if | IF
while ::= while | WHILE

Die gültigen Symbole bilden die reguläre Sprache
 $L(\text{sign+} \mid \dots \mid \text{sign}] \mid \text{integer} \mid \text{identifizier} \mid \text{if} \mid \text{while})$