Kotoba

Regular Expressions

|  |  |
| --- | --- |
| Tokens | Regular Expressions |
| **RESERVED WORDS:** | |
| KOTOBA | “kotoba” |
| BEGIN | “begin” |
| END | “end” |
| READ | “kread” |
| WRITE | “kprint” |
| DEC | “declare” |
| BOOL | “bool” |
| NUMBER | “number” |
| WORD | “word” |
| SENTENCE | “sentence” |
| IF | “if” |
| ELSE | “else” |
| DO | “do” |
| WHILE | “while” |
| FUNC | “function” |
| RETURN | “return” |
| VOID | “void” |
| LENGTH | “length” |
| FREQUENCY | “frequency” |
| SEARCH | “search” |
| EXISTS | “exists” |
| MEAN | “mean” |
| MEDIAN | “median” |
| MODE | “mode” |
| WORDCOUNT | “wordCount” |
| TOKENIZE | “tokenize” |
| REMOVE | “remove” |
| **MORE TOKENS:** | |
| AND | \& |
| OR | \| |
| ID | [a-zA-Z][a-zA-Z0-9]\* |
| BOOLCTE | (“true”|“false”) |
| NUMBERCTE | [\+|\-]?[0-9]+(\.[0-9]+) |
| WORDCTE | \”[a-zA-Z0-9]+\” |
| SENTENCECTE | \"(.\*?)\" |
| RELOP | (< | > | “==” | “!=”) |
| PLUS | \+ |
| MINUS | \- |
| MULT | \\* |
| DIV | \/ |
| NOT | \! |
| ENDSTMT | \; |
| COMA | \, |
| DOT | \. |
| OPENCURL | \{ |
| CLOSECURL | \} |
| OPENPAREN | \( |
| CLOSEPAREN | \) |
| OPENBRAC | \[ |
| CLOSEBRAC | \] |
| EQUAL | \= |

Grammar

**START** → KOTOBA ID ENDSTMT **DECLARE STARTAUX** BEGIN **BLOCK** END

| KOTOBA ID ENDSTMT **STARTAUX** BEGIN **BLOCK** END

**STARTAUX** → **FUNCTION STARTAUX**

| empty

**BLOCK** → OPENCURL **BLOCKAUX** CLOSECURL

**BLOCKAUX** → **ACTION BLOCKAUX**

| empty

**ACTION** → **INPUT**

| **OUTPUT**

| **STATEMENT**

**INPUT**  → KREAD OPENPAREN ID CLOSEPAREN ENDSTMT

**OUTPUT** → KPRINT OPENPAREN **OUTPUTAUX**  CLOSEPAREN ENDSTMT

**OUTPUTAUX** → **EXPRESSION**

| **EXPRESSION** COMA **OUTPUTAUX**

**DECLARE** → DEC **DECAUX**

**DECAUX** → **TYPE** ID **DECLAREAUX**

| **TYPE** ID OPENBRAC **CTE** CLOSEBRAC **DECLAREAUX**

**DECLAREAUX** → ENDSTMT

| COMA **DECAUX**

**ASSIGN** →ID EQUAL **ASSIGNAUX**

| ID OPENBRAC **CTE** CLOSEBRAC EQUAL **ASSIGNAUX**

**ASSIGNAUX** → **EXP** ENDSTMT

| OPENCURL **ASSIAUX** CLOSECURL ENDSTMT

**ASSIAUX** → **EXP**

| **EXP** COMA **ASSIAUX**

**CTE** → ID

| BOOLCTE

| NUMBERCTE

| WORDCTE

| SENTENCECTE

**TYPE** → BOOL

| NUMBER

| WORD

| SENTENCE

**STATEMENT** → **ASSIGN**

| **EXPRESSION**

| **CONDITION**

| **CYCLE**

| **SPECIALFUNC**

**EXPRESSION** → **RELOPEXPRESSION**

| **RELOPEXPRESSION** AND **EXPRESSION**

| **RELOPEXPRESSION** OR **EXPRESSION**

**RELOPEXPRESSION**  → **EXP**

| **EXP** RELOP **EXP**

| NOT **EXP**

**EXP** →  **TERM**

| **TERM** PLUS **EXP**

| **TERM** MINUS **EXP**

**TERM** → **FACTOR**

| **FACTOR** MULT **TERM**

| **FACTOR** DIV **TERM**

**FACTOR** → OPENPAREN **EXPRESSION** CLOSEPAREN

| PLUS **CTE**

| MINUS **CTE**

| **CTE**

**CONDITION** → IF OPENPAREN **EXPRESSION** CLOSEPAREN **BLOCK**

| IF OPENPAREN **EXPRESSION** CLOSEPAREN **BLOCK** ELSE **BLOCK**

**CYCLE** → WHILE OPENPAREN **EXPRESSION** CLOSEPAREN **BLOCK**

| DO **BLOCK** WHILE OPENPAREN **EXPRESSION** CLOSEPAREN ENDSTMT

**FUNCTION** → FUNC **FUNCAUX** ID OPENPAREN **PARAMETERAUX** CLOSEPAREN OPENCURL **DECLARE** **BLOCKAUX RETURNAUX** ENDSTMT CLOSECURL

| FUNC **FUNCAUX** ID OPENPAREN **PARAMETERAUX** CLOSEPAREN OPENCURL **BLOCKAUX RETURNAUX** ENDSTMT CLOSECURL

**FUNCAUX** → **TYPE**

| VOID

**PARAMETERAUX** → **TYPE** ID

| **TYPE** ID COMA **PARAMETERAUX**

| empty

**RETURNAUX** → RETURN ID

| empty

**SPECIALFUNCTION** → ID DOT **SPECIAL** OPENPAREN **SPAUX** CLOSEPAREN ENDSTMT

**SPAUX** → cte

| empty

**SPECIAL** → LENGTH

| FREQUENCY

| SEARCH

| EXISTS

| MEAN

| MEDIAN

| MODE

| WORDCOUNT

| TOKENIZE

| REMOVE