**ANÁLISIS LÉXICO**

El análisis léxico es aquella primera fase de un compilador donde se encarga de evaluar los caracteres ingresados en el compilador y con el objetivo de formar los componentes léxicos, del cual para eso ocupamos definir un lenguaje, que componentes léxicos se pueden formar a través del lenguaje y los posibles errores que pueden generarse por la generación de dichos componentes léxicos, por ende cada uno de los elementos mencionados con anterioridad se definen a continuación.

**LENGUAJE DEL COMPILADOR**

El lenguaje de un compilador son aquellos caracteres o símbolos que acepta como válidos en el mismo, por ende los caracteres que son válidos para este compilador se muestran en los siguientes grupos mostrados en la siguiente tabla:

|  |  |
| --- | --- |
| **GRUPO** | **CARÁCTER** |
| **Símbolo** | [ ,.=()<>#{}+-:&\_] |
| **Comilla** | [ ‘ ] |
| **Letra** | [a-zA-Z] |
| **Dígito** | [0-9] |

**COMPONENTES LÉXICOS DEL COMPILADOR**

Los componentes léxicos son aquellas cadenas de caracteres que concuerdan con un patrón, dicho patrón describe la forma en cómo se forma dicho componente léxico, por lo tanto los componentes léxicos de este compilador estarán formados por el lenguaje del mismo o otros componentes léxicos, y los componentes léxicos se muestran en la siguiente tabla:

|  |  |
| --- | --- |
| **COMPONENTE LÉXICO** | **PATRÓN** |
| **NUMERO\_ENTERO** | ({Digito})({Digito})\* |
| **NUMERO\_REAL** | ({Digito})+.({Digito})({Digito})\* |
| **CONDICIONAL** | (true | false) |
| **IDENTIFICADOR\_CADENA** | ({Comilla})({Letra} | {Dígito})({Letra} | {Dígito} | {Símbolo})\*({Comilla}) |
| **IDENTIFICADOR** | ({Letra})({Letra} | {Digito})\* |
| **VALORES** | (NUMERO\_ENTERO | NUMERO\_REAL | IDENTIFICADOR\_CADENA | CONDICIONAL) |
| **PUERTO** | (port\_A0 | port\_A1 | port\_A2 | port\_A3 | port\_A4 | port\_A5 | port\_B0 | port\_B1 | port\_B2 | port\_B3 | port\_B4 | port\_B5 | port\_B6 | port\_B7 | port\_C2 | port\_C3 | port\_C4 | port\_C5 | port\_C6 | port\_C7 | port\_D0 | port\_D1 | port\_D2 | port\_D3 | port\_D4 | port\_D5 | port\_D6 | port\_D7) |
| **TIPO\_PUERTO** | (proximity | temperature | LED | LCD | button | motor) |
| **TIPO\_DATO** | (int | string | decimal | boolean) |
| **MOTOR** | (move | restart | start) |
| **METODO\_SENSOR** | (distance | state | time | degree) |
| **ESTRUCTURA\_DE\_CONTROL** | (function | begin | loop | if | else) |
| **PALABRA\_RESERVADA** | (ligther | delay | operation | call | return | print | console | turn\_off) |
| **DECLARACION** | (var | const) |
| **OPERADOR\_LOGICO** | (and | or | not) |
| **OPERADOR\_RELACIONAL** | (> | < | >= | <= | ==) |
| **OPERADOR\_ARITMETICO** | (+ | - | \* | /) |
| **OPERADOR\_ASIGNACION** | (=) |
| **SIGNO\_AGRUPACION** | ( { | } | ( | )) |
| **SIGNO\_PUNTUACION** | (, | .) |
| **FIN\_DE\_LINEA** | (!) |

**ERRORES LÉXICOS DEL COMPILADOR**

Los errores léxicos en el compilador son aquellas cadenas de caracteres incorrectas que no terminan de generar un componente léxico o bien es un carácter no valido en el lenguaje, del cual contamos con los siguientes caracteres que no son válidos para el lenguaje de este compilador:

CaracterNoValido = [$¬@%?¡;¿"|"]

Los errores léxicos que genera este compilador son los siguientes:

|  |  |
| --- | --- |
| **NOMBRE DEL ERROR** | **DESCRIPCIÓN** |
| **Error\_léxico\_X** | Error desconocido, verifique como está escrito el código |
| **Error\_léxico\_0** | El carácter o caracteres no es válido en el lenguaje |
| **Error\_léxico\_1** | Un carácter o caracteres del identificador cadena no pertenece a un símbolo válido |
| **Error léxico\_2** | No hay ningún elemento entre ‘ ’ |
| **Error\_léxico\_3** | La comilla o comillas no está asociada a un identificador cadena |
| **Error\_léxico\_4** | El carácter o caracteres guion bajo no está asociado a un puerto, o a la palabra reservada turn\_off |
| **Error\_léxico\_5** | El número real contiene un carácter o caracteres no validos |
| **Error\_léxico\_6** | El número real inicia con un punto o varios puntos |
| **Error\_léxico\_7** | El número real inicia con un punto o varios puntos y cuenta con un carácter o caracteres no validos |
| **Error\_léxico\_8** | El número real inicia con un punto o varios puntos y tiene más de uno |
| **Error\_léxico\_9** | El número real inicia con un punto o varios puntos, tiene más de uno y cuenta con un carácter o caracteres no validos |
| **Error\_léxico\_10** | El número real tiene más de un punto |
| **Error\_léxico\_11** | El número real tiene más de un punto y cuenta con un carácter o caracteres no validos |
| **Error\_léxico\_12** | El número real no continua después del punto |
| **Error\_léxico\_13** | El numero entero contiene un carácter o caracteres no validos |