Universidad de San Carlos de Guatemala

Facultad de Ingeniería

Carrera: Ingeniería en ciencias y sistemas Catedrático: Ing. Mario Bautista

Auxiliar: José Puac

Curso: Organización de lenguajes y compiladores 1

Sección "N"



GRAMATICA

José Abraham Solórzano Herrera 201800937

Guatemala 4 de Julio del 2021

INTRODUCCION

El software es un intérprete de alto nivel, es un lenguaje exclusivo de la Universidad de San Carlos de Guatemala, el cual se llama JPR, consiste en un editor, cuya finalidad es proporcionar ciertas funcionalidades, características, herramientas que serán de utilidad al usuario. La funcionalidad del editor será el ingreso del código fuente que será analizado, donde podrá aceptar archivos con extensión ".jpr" y mostrará la línea actual, está diseñado en el lenguaje de programación Python, por medio de una librería PLY.

OBJETIVOS

Objetivo General:

Aplicar los conocimientos sobre la fase de análisis léxico, sintáctico y semántico de un compilador para la realización de un intérprete completo, con las funcionalidades principales para que sea funcional.

Objetivo Especifico:

- 1. Reforzar los conocimientos de análisis léxico, sintáctico y semántico para la creación de un lenguaje de programación.
- 2. Aplicar los conceptos de compiladores para implementar el proceso de interpretación de código de alto nivel.
- 3. Aplicar los conceptos de compiladores para analizar un lenguaje de programación y producir las salidas esperadas.
- 4. Aplicar la teoría de compiladores para la creación de soluciones de software.

GRAMATICA

Figura No. 1: Reservadas.

Se inicia creado las reservadas, son las palabras ya definidas por el programa, las cuales serán parte de los tokens Figura No. 2.

```
reservadas =
     'print'
                    'RPRINT',
                    'RVAR',
    'var'
    'true'
                    'RTRUE',
'RFALSE',
    'false'
                    'RIF',
'RELSE',
'RWHILE',
    'else'
                     'RBREAK'
     'break'
                     'RNULL',
                    'RMAIN',
     'main'
                    'RFUNC',
    'func'
    'for'
                     'RFOR',
    'switch'
                     'RSWITCH',
                     'RCASE',
    'case
                     'RDEFAULT',
    'default'
    'return'
                     'RRETURN',
                     'RINT',
    'int'
     'double'
                     'RDOUBLE'
    'string'
                     'RSTRING',
                    'RCHAR',
                    'RBOOLEAN',
'RCONTINUE'
    'boolean'
     continue'
     'read'
                    'RREAD',
```

Figura No.2: Tokens

```
tokens = [
    'PUNTOCOMA',
    'COMA',
    'PARA',
    'PARA',
    'PARC',
    'LLAVEC',
    'IGUAL',
    'MAS',
    'MENOS',
    'POR',
    'DIV',
    'POT',
    'MODULO',
    'MENORQUE',
    'MENORQUE',
    'MENORIGUAL',
    'MAYORIGUAL',
    'MAYORIGUAL',
    'DIFERENCIA',
    'AND',
    'OR',
    'NOT',
    'DECIMAL',
    'ENTERO',
    'CADENA',
    'CHAR',
    'ID',
    'COMENTARIO_SIMPLE',
    'COMENTARIO_VARIAS_LINEAS',
    'INCREMENTO',
    'DOSPUNTOS',
] + List(reservadas.values())
```

Print: Tendrá la función de mostrar el contenido.

Var: La variable nos sirve para iniciar declarando una variable.

True: Es una variable primitiva Booleana.

False: Es una variable primitiva Booleana.

If: Es una sentencia de control.

Else: Es una sentencia de control.

While: Es una sentencia de control.

Break: Es una sentencia de trasferencia.

Null: Es un tipo de dato primitiva, que representa el valor nulo o vacío.

Main: Es el encargado de poder ejecutar todo el código generado dentro del lenguaje.

Func: Una función es una subrutina de código que se identifica con un nombre, un conjunto de parámetros y de instrucciones.

For: Es una sentencia cíclica.

Switch: Es una sentencia cíclicla.

Case: Es parte del Switch y es una sentencia cíclica.

Default: Es parte del Switch y es una sentencia cíclica.

Return: Es una sentencia de trasnferencia.

Int: Es un tipo de dato primitivo, que representa el valor entero.

Double: Es un tipo de dato primitivo, que representa el valor de decimal.

String: Es un tipo de dato primitivo, que representa el valor de una cadena.

Char: Es un tipo de dato primitivo, que representa el valor de un carácter.

Boolean: Es un tipo de dato primitivo, que representa el valor de un boolean.

Continue: Es una sentencia de trasferencia.

Read: Esta función nos permite obtener valores que queramos ingresar en

el momento de ejecución del código.

Figura No. 3: Asignación al valor tokens.

En la Figura No. 3 se asigna a cada respectivo token su valor.

```
# Caracteres ignorados
t_ignore = " \t"
```

Figura No. 4: Decimal.

```
def t_DECIMAL(t):
    r'\d+\.\d+'
    try:
        t.value = float(t.value)
    except ValueError:
        print("Float value too large %d", t.value)
        t.value = 0
    return t
```

En esta figura se encarga de reconocer si la expresión es un decimal y es parte del t_decimal.

Figura No. 5: Entero.

```
def t_ENTERO(t):
    r'\d+'
    try:
        t.value = int(t.value)
    except ValueError:
        print("Integer value too large %d", t.value)
        t.value = 0
    return t
```

En esta figura se encarga de reconocer si la expresión es un entero y es parte del t entero.

Figura No. 6: Id

```
def t_ID(t):
    r'[a-zA-Z][a-zA-Z_0-9]*'
    t.type = reservadas.get(t.value.lower(),'ID')
    return t
```

En esta figura se encarga de reconocer si la expresión es una variable y es parte del t_id.

Figura No.7:

```
def t_CADENA(t):
    #r'(\".*?\")'
    # t.value = t.value[1:-1] # remuevo las comillas
    # return t
    r'\"(\\"|.)*?\"'
    t.value = t.value[1:-1] # remover comillas
    t.value = t.value.replace('\\n', '\n')
    t.value = t.value.replace('\\\', '\\')
    return t
```

En esta figura se encarga de reconocer si la expresión es una cadena y es parte del t_cadena.

Figura No. 8:

En esta figura se encarga de reconocer si la expresión es un carácter y es parte del t char.

Figura No. 9:

```
def t_COMENTARIO_VARIAS_LINEAS(t):
    r'\#\*(.|\n)*?\*\#'
    t.lexer.lineno += t.value.count("\n")
```

En esta figura se encarga de reconocer si es un comentario multi línea.

Figura No. 10:

```
# Comentario simple // ...
def t_COMENTARIO_SIMPLE(t):
    r'\#.*\n'
    t.lexer.lineno += 1
```

En esta figura se encarga de reconocer si es un comentario de una línea.

Figura No. 11:

```
def t_newline(t):
    r'\n+'
    t.lexer.lineno += t.value.count("\n")
```

Este token se encarga de reconocer un salto de línea.

Figura No. 12:

```
def t_error(t):
    errores.append(Exception("Lexico","Error léxico." + t.value[0] , t.lexer.lineno, find_column(input, t)))
    t.lexer.skip(1)
# Compute column
```

Este token se encarga de reconocer si

Figura No. 13:

```
def find_column(inp, token):
    line_start = inp.rfind('\n', 0, token.lexpos) + 1
    return (token.lexpos - line_start) + 1
```

Este token se encarga de reconocer la fila.

Figura No. 14:

```
# Construyendo el analizador léxico
import Interprete.ply.lex as lex
lexer = lex.lex()
# Asociacion
precedence = (
    ('left','OR'),
    ('left','AND'),
    ('right','UNOT'),
    ('left','GUALIGUAL','DIFERENCIA','MENORQUE','MENORIGUAL','MAYORQUE','MAYORIGUAL'),
    ('left','MAS','MENOS'),
    ('left','DIV','POR','MODULO'),
    ('nonassoc', 'POT'),
    ('right','UMENOS'),
    )
}
```

En esta figura se asigna la precedencia de la gramática.

Figura No. 15:

Esta Figura es donde inicia la gramática, y representa la gramática ascendente recursiva por la izquierda, usando LALR.

Figura No. 16:

```
#
def p_instruccion(t):
    '''instruccion : imprimir_ fin_instruccion
    | declaracion_ins fin_instruccion
    | incre_decre_ins fin_instruccion
    | if_ins
    | while_ins
    | switch_ins
    | for_ins
    | main_ins
    | break_ins fin_instruccion
    | return_ins fin_instruccion
    | continue_ins fin_instruccion
    | funcion_ins
    | llamada_ins fin_instruccion
    | cOMENTARIO_VARIAS_LINEAS
    | COMENTARIO_SIMPLE
```

En esta figura representa todas las posibles instrucciones que puede contener la gramática.

Figura No. 17:

En esta figura inicia buscando si es una declaración o una asignación.

Figura No. 18:

En esta figura representa la declaración y la asignación.

Figura No. 19:

En esta gramática se crea la declaración o asignación para el for.

Figura No. 20:

```
# 
def p_imprimir(t):
    'imprimir_ : RPRINT PARA expression PARC'
    t[0] = Imprimir(t[3], t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte de la gramática, reconoce la palabra reservada print y su función principal es mostrar el resultado.

Figura No. 21:

En esta parte de la gramática reconoce las posibles alternativas de la sentencia if, ya sea un else if o un else.

Figura No. 22:

```
SENTENCIA SWITCH
   p_condicion_switch_case_list_default(t): # Aqui verifiac que la condicion venga [<CASES_LIST>] [<DEFAULT>]
'switch_ins : RSWITCH PARA expresion PARC LLAVEA case_switch_ins default_switch LLAVEC'
t[0] = Switch(t[3], t[6], t[7], t.lineno(1), find_column(input, t.slice[1]))
ef p_condicion_switch_case_list(t): # Aqui verifiac que la condicion venga [<CASES_LIST>]
    switch_ins : RSWITCH PARA expresion PARC LLAVEA case_switch_ins LLAVEC
   t[0] = Switch(t[3], t[6], None, t.lineno(1), find_column(input, t.slice[1]))
   p_condicion_switch_default(t): # Aqui verifiac que la condicion venga
    switch_ins : RSWITCH PARA expresion PARC LLAVEA default_switch LLAVEC
   t[0] = Switch(t[3], None, t[6], t.lineno(1), find_column(input, t.slice[1]))
   p_casos_switch_ins_caso_switch(t): # Aqui hace a que sea recursivo y puedan venir infinitos [<CASES_LIST>]
   'case_switch_ins : case_switch_ins case_switch'
   if t[2] != "":
t[1].append(t[2])
   t[0] = t[1]
   p_caso_switch_(t):
   'case_switch_ins : case_switch'
if t[1] == "":
        t[0] = []
        t[0] = [t[1]]
   p_condicion_switch_case(t):
   'case_switch : RCASE expresion DOSPUNTOS instrucciones'
t[0] = Case(t[2], t[4], t.lineno(1), find_column(input, t.slice[1]))
  p_condicion_default_switch(t):
    'default_switch : RDEFAULT DOSPUNTOS instrucciones'
    # t[0] = t[
   t[0] = Default(t[3], t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte se vuelve a utilizar la gramática ascendente por la izquierda, ya que en el switch pueden venir una infinidad de case.

Figura No. 23:

```
#
def p_sentencia_while(t) :
   'while_ins : RWHILE PARA expresion PARC LLAVEA instrucciones LLAVEC'
   t[0] = While(t[3], t[6], t.lineno(1), find_column(input, t.slice[1]))
```

en esta figura la gramática puede reconocer la sentencia cíclica del while.

Figura No. 24:

En esta parte de la gramática reconoce la sentencia de cíclica del for.

Figura No. 25:

```
p_funcion_2(t):
    funcion_ins
                    : RFUNC ID PARA PARC LLAVEA instrucciones LLAVEC'
   t[0] = Funcion(t[2], [], t[6], t.lineno(1), find_column(input, t.slice[1]))
def p_funcion_parametros(t) :
                  : RFUNC ID PARA parametros PARC LLAVEA instrucciones LLAVEC'
   t[0] = Funcion(t[2], t[4], t[7], t.lineno(1), find_column(input, t.slice[1]))
def p_parametros_parametros(t) :
                   : parametros COMA parametro'
    'parametros
   t[1].append(t[3])
   t[0] = t[1]
def p_parametros_parametro(t) :
    parametros
                  : parametro
   t[0] = [t[1]]
   p_parametro(t) :
    parametro
                  : tipo_funcion ID'
   t[0] = \{'tipoDato':t[1], 'identificador':t[2]\} # Se crea un diccionario tipoDato: tipo, identificador
```

En esta parte de la gramática puede reconocer una función, y hacer una declaración por medio de los parámetros, en este caso en los parámetros se vuelve a utilizar la gramatica ascendente por la izquierda.

Figura No. 26:

```
p_llamada_de_funcion(t) :
'llamada_ins : ID PARA PARC'
t[0] = Llamada(t[1], [], t.lineno(1), find_column(input, t.slice[1]))
p_llamada_de_fincion_parametros(t) :
'llamada_ins : ID PARA parametros_llamada PARC'
t[0] = Llamada(t[1], t[3], t.lineno(1), find_column(input, t.slice[1]))
p_parametros_llamadas_parametros_llamadas(t) :
 'parametros_llamada
                           : parametros_llamada COMA parametro_llamada'
t[1].append(t[3])
t[0] = t[1]
p_parametros_llamadas_parametro_llamada(t) :
 parametros_llamada
                         : parametro_llamada'
t[0] = [t[1]]
p_parametro_llamada(t) :
'parametro_llamada :
                          : expresion'
t[0] = t[1]
```

En esta parte de la gramática se hace el llamado de la función y se empieza a utilizar la gramática ascendente por la izquierda.

Figura No. 27:

```
#
def p_sentencia_break(t) :
   'break_ins : RBREAK'
   t[0] = Break(t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte reconoce la gramática la sentencia break.

Figura No. 28:

```
#
def p_return_instruccion(t):
    'return_ins : RRETURN expression'
    t[0] = Return(t[2], t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte reconoce la gramática la sentencia return.

Figura No. 29:

```
#
def p_continue_instruccion(t) :
    'continue_ins : RCONTINUE'
    t[0] = Continue(t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte la gramática reconoce la sentencia continúe.

Figura No. 30:

En esta figura representa todos los posibles datos, sin embargo, el primer tipo sirve para las declaraciones, en cambio el tipo función sirve para crear funciones donde van a resivir un tipo "tipo_funcion".

Figura No. 31:

```
# INCREMENTO O DECREMENTO

def p_incremento_decremento(t):

''' incre_decre_ins : ID INCREMENTO

if t[2] == '++':

t[0] = IncrementoDecremento(t[1], Operador_Aritmetico.INCREMENTO, t.lineno(1), find_column(input, t.slice[1]))

elif t[2] == '--':

t[0] = IncrementoDecremento(t[1], Operador_Aritmetico.DECREMENTO, t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte de la gramática reconoce los incrementos y decrementos que pueden llegar a venir.

Figura No. 32:

```
_expresion_binaria(t):
   expresion : expresion MAS expresion
         expresion POR expresion
         expresion POT expresion
         expresion MODULO expresion
         expresion MENORIGUAL expresion
         expresion MAYORQUE expresion
         expresion DIFERENCIA expresion
         expresion AND expresion
         expresion OR expresion
if t[2] == '+':
   t[0] = Aritmetica(Operador_Aritmetico.SUMA, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '-':
   t[0]
        = Aritmetica(Operador_Aritmetico.RESTA, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '*':
t[0] = Aritmetica(Operador_Aritmetico.POR, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '/':
   t[0] = Aritmetica(Operador_Aritmetico.DIV, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] ==
   t[0] = Aritmetica(Operador_Aritmetico.POTE, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '%':
    t[0] = Aritmetica(Operador_Aritmetico.MODU, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] ==
   t[0] = Relacional(Operador_Relacional.IGUALACION, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '<':
t[\emptyset] = Relacional(Operador_Relacional.MENORQUE, t[1],t[3], t.lineno(2), find_column(input, t.slice[2])) elif t[2] == '<=':
   t[0] = Relacional(Operador_Relacional.MENORIGUAL, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] ==
   t[0] = Relacional(Operador_Relacional.MAYORQUE, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '>=':
   t[0] = Relacional(Operador_Relacional.MAYORIGUAL, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
    t[0] = Relacional(Operador_Relacional.DIFERENCIA, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '&&':
    t[0] = Logica(Operador_Logico.AND, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
elif t[2] == '||':
   t[0] = Logica(Operador_Logico.OR, t[1],t[3], t.lineno(2), find_column(input, t.slice[2]))
```

En esta parte empieza a reconocer las expresiones, ya sea aritmética, ralacional, o lógica.

Figura No. 33:

En esta parte de la gramática reconoce las expresiones por la izquierda, como vienen siendo.

Figura No. 34:

En esta parte de la gramática hace las agrupaciones siendo recursivo.

Figura No. 35:

```
p_expresion_identificador(t):
'''expresion : ID'''
    t[0] = Identificador(t[1], t.lineno(1), find_column(input, t.slice[1]))
def p_expresion_entero(t):
    t[0] = Primitivos(Tipo.ENTERO,t[1], t.lineno(1), find_column(input, t.slice[1]))
def p_primitivo_decimal(t):
       expresion : DECIMA
    t[0] = Primitivos(Tipo.DECIMAL, t[1], t.lineno(1), find_column(input, t.slice[1]))
def p_primitivo_cadena(t):
    t[0] = Primitivos(Tipo.CADENA,str(t[1]).replace('\\n', '\n'), t.lineno(1), find_column(input, t.slice[1]))
def p_primitivo_char(t):
    t[0] = Primitivos(Tipo.CHAR,str(t[1]).replace('\\n', '\n'), t.lineno(1), find_column(input, t.slice[1]))
def p_primitivo_true(t):
    t[0] = Primitivos(Tipo.BOOLEANO, True, t.lineno(1), find_column(input, t.slice[1]))
def p_primitivo_false(t):
    t[0] = Primitivos(Tipo.BOOLEANO, False, t.lineno(1), find_column(input, t.slice[1]))
def p_primitivo_null(t):
    t[0] = Primitivos(Tipo.NULO, None, t.lineno(1), find_column(input, t.slice[1]))
def p_expresion_llamada(t):
       'expresion : llamada ins'''
    t[0] = t[1]
def p_expresion_casteo(t):
    t[0] = Casteo(t[2], t[4], t.lineno(1), find_column(input, t.slice[1]))
def p_expresion_read(t):
      'expresion : RREAD PARA PARC'''
    t[0] = Read(t.lineno(1), find_column(input, t.slice[1]))
```

En esta parte de la gramática guarda todas las expresiones que sean tipo primitivo, identificador, casteo y read.

Figura No. 36:N

```
def crearNativas(ast):
                               # CREACION Y DECLARACION DE LAS FUNCIONES NATIVAS
   nombre = "toupper'
   parametros = [{'tipoDato':Tipo.CADENA,'identificador':'toUpper##Param1'}]
   instrucciones = []
   toUpper = ToUpper(nombre, parametros, instrucciones, -1, -1)
   ast.addFuncion(toUpper)
   nombre = "tolower"
   parametros = [{'tipoDato':Tipo.CADENA,'identificador':'toLower##Param1'}]
   instrucciones = []
   toLower = ToLower(nombre, parametros, instrucciones, -1, -1)
                             # GUARDAR LA FUNCION EN "MEMORIA" (EN EL ARBOL)
   ast.addFuncion(toLower)
   nombre = "length"
   parametros = [{'tipoDato':Tipo.CADENA,'identificador':'length##Param1'}]
   instrucciones = []
   length = Length(nombre, parametros, instrucciones, -1, -1)
   ast.addFuncion(length)
   nombre = "truncate"
   parametros = [{'tipoDato':Tipo.ENTERO,'identificador':'truncate##Param1'}]
   instrucciones = []
   truncate = Truncate(nombre, parametros, instrucciones, -1, -1)
   ast.addFuncion(truncate)
                              # GUARDAR LA FUNCION EN "MEMORIA" (EN EL ARBOL)
   nombre = "round"
   parametros = [{'tipoDato':Tipo.ENTERO,'identificador':'round##Param1'}]
   instrucciones = []
   rround = Round(nombre, parametros, instrucciones, -1, -1)
ast.addFuncion(rround)  # GUARDAR LA FUNCION EN "MEMOR:
                             # GUARDAR LA FUNCION EN "MEMORIA" (EN EL ARBOL)
   nombre = "typeof"
   parametros = [{'tipoDato':Tipo.NULO,'identificador':'typeOf##Param1'}]
   ef interprete(entrada, consola):
```

En esta parte se crean las funciones nativas, que ya están definidas.