Curso: Compiladores Tema 6. Traducción dirigida por Sintaxis (Parte II)

En la clase pasada vimos cómo definir traducciones sencillas incluyendo símbolos de acción en las gramáticas libres de contexto para obtener una gramática de traducción. Sin embargo en la mayoría de las gramáticas de traducción no es suficiente incluir sólo los símbolos de acción para que defina traducciones más complejas o que defina las tareas del Analizador Semántico.

Ahora veremos cómo definir, en la gramática de traducción, una de las tareas del Analizador Semántico.

Análisis Semántico a través del Análisis Sintáctico

Una de las tareas del analizador semántico en lenguajes tipificados, es revisar que las variables en un programa estén declaradas (indicando su tipo), antes de ser usadas. Para realizar esta tarea, el analizador semántico debe actualizar la tabla de símbolos para colocar el tipo de variable en el campo correspondiente. Por ejemplo, en el lenguaje C, si tenemos la siguiente declaración: *float x,y;* la tabla de símbolos después del análisis semántico de esa sentencia declarativa, deberá verse de la siguiente forma:

Posición	nombre	tipo
0	X	float
1	у	float

Una vez que esté actualizada la tabla de símbolos, el analizador semántico puede revisar, cada que se encuentre con las variables x y y, si las operaciones que se realizan sobre ellas, van de acuerdo al tipo de dato declarado, o inclusive si fueron declaradas.

Ejercicio 6.3.1

Incluir los símbolos de acción a la siguiente gramática que define las sentencias declarativas de un cierto lenguaje, para que defina, además, la tarea de incluir el tipo de variables en la tabla de símbolos.

Donde los átomos son:

1: D \rightarrow TaV 4: T \rightarrow c a es un identificador 2: T \rightarrow i 5: V \rightarrow , aV i es int 3: T \rightarrow f 6: V \rightarrow ; f es float c es char

Una primera solución sería incluir el símbolo de acción {AT} que Asigna el Tipo a la variable **a** al momento de declarase:

1: D \rightarrow Ta{AT}V 4: T \rightarrow c 2: T \rightarrow i 5: V \rightarrow , a{AT}V 3: T \rightarrow f 6: V \rightarrow ;

Ejercicio 6.3.1 (continuación)

Antes de construir su tabla de Parser, hay que calcular el conjunto de selección de cada producción:

1: D \rightarrow Ta{AT}V c.s.(1)={i f c} 4: T \rightarrow c c.s.(4)={c} 2: T \rightarrow i c.s.(2)={i} 5: V \rightarrow , a{AT}V c.s.(5)={,} 3: T \rightarrow f c.s.(3)={f} 6: V \rightarrow ; c.s.(6)={;}

Entonces, la tabla de Parser es:

	а	i	f	С	,	;	4
D		Re	Reemp(Ta{AT}V) ^r retén				
Т		Pop avanza	Pop avanza	Pop avanza			
V					Reemp(a{AT}V) ^r avanza	Pop avanza	
а	Pop avanza						
{AT}			Po	op, AsignaTip	o(), retén		
Λ							Acepta

Ejercicio 6.3.1 (continuación):

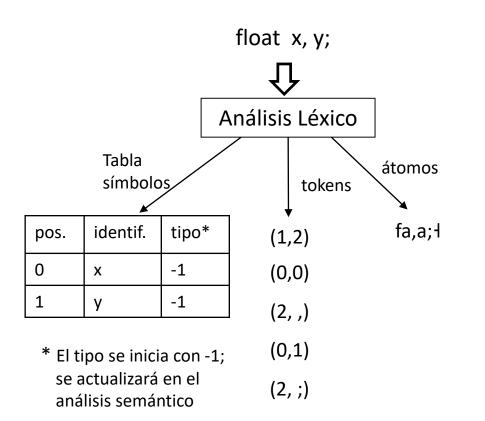
Ahora veamos si al revisar sintácticamente la sentencia fa,a;-l, actualiza la tabla de símbolos:

Cadena	Pila	acción
fa,a;+	DΛ	
fa,a;+	Ta{AT}V∧	
a,a; 1	a{AT}V∧	
,a;ŧ	{AT}V∧	Asigna Tipo

Ya no podemos continuar porque ¿a qué variable de la tabla de símbolos se le asigna qué tipo? ¿Dónde obtenemos esta información?

Ejercicio 6.3.1 (continuación):

Para contestar estas preguntas, revisemos el Análisis Léxico del proceso de compilación. Consideremos sólo la siguiente sentencia declarativa en el programa fuente:



Catálogos de clases y palabras reservadas

Clase	Componente léxico
0	identificador
1	Palabras reservadas
2	Símbolos especiales

Palabra reservada	Valor
char	0
int	1
float	2

Podemos observar que las tres salidas que genera el Analizador Léxico tienen información que requerimos para actualizar la tabla de símbolos en la etapa del análisis semántico.

Vimos en los temas de análisis sintáctico tanto descendente como ascendente, que sólo se ocupaba como entrada la cadena de átomos. Esta cadena sólo nos da información para la revisión sintáctica. Para las tareas de análisis semántico y traducciones más complejas, se requiere de información que está en las otras salidas de Analizador Léxico, o sea, en los tokens y en la tabla de símbolos.

Cuando es un token de un identificador, el componente *valor* indica la posición de dicho identificador dentro de la tabla de símbolos.

Cuando se refiere a un token de una palabra reservada, el componente *valor* nos indica la posición de dicha palabra en la tabla correspondiente.

Entonces ahora la entrada al Analizador Semántico, incluido en el Analizador Sintáctico, serán los átomos, la información de los tokens y de la tabla de símbolos.

Para el caso que se presenta en el Ejercicio 6.3.1, el token del identificador dará la posición dentro de la tabla de símbolos y el token de la palabra reservada float, nos indicará de qué palabra reservada se trata, por medio de su posición, y se colocará en la columna *tipo* de la tabla de símbolos, ya sea la posición o la palabra reservada como tal.

Para incluir dichos datos faltantes en la gramática de traducción, se usarán otros elementos denominados *atributos*.

Consideraciones al incluir los *atributos* en una gramática de traducción.

- Los atributos pueden estar asociados a elementos terminales, no-terminales y símbolos de acción de la gramática. Se colocarán como subíndices.
- No todas las producciones tendrán atributos, sólo aquellas que requieran información adicional al elemento de la gramática para realizar las actividades semánticas o de traducción.
- Cuando existan atributos en una producción, habrá reglas de asignación que indicarán cómo los atributos adquieren su valor, que puede ser de otros atributos, o bien de una función externa. Estas reglas son en cada producción.
- Los atributos de una producción son independientes a los que se tengan en otra producción, aunque tengan el mismo nombre.

Ejercicio 6.3.2

A la gramática del Ejercicio 6.3.1, incluir los atributos para que el analizador semántico actualice el tipo de identificador en la tabla de símbolos.

1: D \rightarrow Ta{AT}V	$4: T \rightarrow c$
$2: T \rightarrow i$	5: $V \rightarrow$, a{AT}V
$3: T \rightarrow f$	$6: V \rightarrow ;$

Vamos a incluir los siguientes atributos:

 $\mathbf{p} \rightarrow$ indica la posición de un identificador en la tabla de símbolos.

 $\mathbf{t} \rightarrow$ indica el tipo de dato de la lista de identificadores.

p1 y **p2** → atributos auxiliares para que el símbolo de acción pueda saber la posición del identificador en la tabla de símbolos, al que le pondrá su tipo.

t1 y **t2** → atributos auxiliares para que el símbolo de acción pueda poner el tipo del identificador en la tabla de símbolos.

$$\begin{array}{c} \textit{Ejercicio 6.3.2 (continuación)} \\ 1: D \rightarrow T_t \ a_p \ \{\text{AT}\}_{p1,t1}, \ V_{t2} \end{array} \begin{array}{c} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{array} \begin{array}{c} \text{Estas son las reglas de} \\ \text{asignación para la} \\ \text{producción 1} \end{array}$$

$$2: T_{t1} \rightarrow i_t \qquad \qquad \left\{ \begin{array}{l} \text{Esta es la regla de asignación} \\ \text{para la producción 2} \end{array} \right.$$

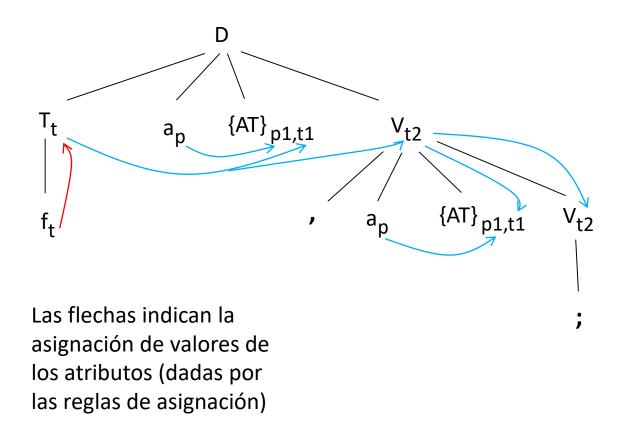
$$3{:}\ T_{t1} \to f_t \qquad \qquad \left\{\begin{array}{ll} \text{Esta es la regla de asignación} \\ \text{para la producción 3} \end{array}\right.$$

$$4 \colon T_{t1} \to c_t \qquad \qquad \left\{ \begin{array}{l} \text{Esta es la regla de asignación} \\ \text{para la producción 4} \end{array} \right.$$

$$5: V_t \rightarrow \text{, a}_p \text{ {AT}}_{p1,t1} V_{t2} \qquad \begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases} \qquad \text{Estas son las reglas de asignación para la producción 5}$$

6:
$$V_t \rightarrow$$
; Esta producción no tiene reglas de asignación porque, aunque tenga el atributo t, éste no da o recibe su valor de otros atributos de la misma producción. Más adelante se entenderá este caso.

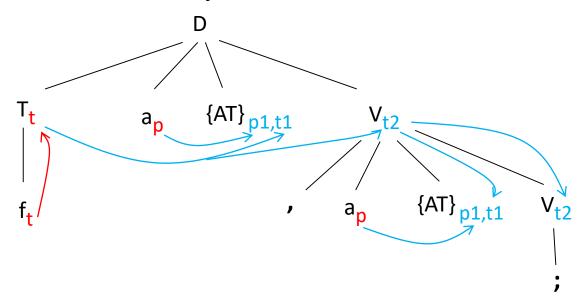
Ahora obtengamos la sentencia fa,a; construyendo el árbol sintáctico correspondiente.



Las gramáticas de traducción con atributos, maneja dos clases de atributos:

- Atributos heredados. Los valores que reciben, vienen en dirección de la raíz a las hojas o bien de nodos hermanos.
- Atributos sintetizados. Los valores que reciben, vienen en dirección de las hojas a la raíz, o son atributos asociados a elementos terminales.

Entonces identifiquemos en la gramática del Ejercicio 6.3.2, tanto en el árbol sintáctico de la sentencia dada como en sus producciones, los atributos heredados en azul y los atributos sintetizados en rojo.



M.C. Laura Sandoval Montaño

Ejercicio 6.3.2

1: D
$$\rightarrow$$
 T_t a_p {AT}_{p1,t1}, V_{t2}
$$\begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases}$$
 Estas son las reglas de asignación para la producción 1

2:
$$T_{t1} \rightarrow i_t$$
 $\{ t1 \leftarrow t \}$

$$\{$$
 t1 \leftarrow t

Esta es la regla de asignación para la producción 2

$$3: T_{t1} \rightarrow f_t$$

$$\left\{ t1 \leftarrow t \right\}$$

Esta es la regla de asignación para la producción 3

4:
$$T_{t1} \rightarrow c_t$$

Esta es la regla de asignación para la producción 4

5:
$$V_t \rightarrow$$
, $a_p \{AT\}_{p1,t1} V_{t2}$

$$\begin{cases}
p1 \leftarrow p \\
t1 \leftarrow t \\
t2 \leftarrow t
\end{cases}$$

$$\begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases}$$

Estas son las reglas de asignación para la producción 5

6:
$$V_{t} \rightarrow ;$$

Recordar:

- Atributos heredados en azul.
- Atributos sintetizados en rojo.

Para implementar el analizador semántico para una gramática de traducción con atributos deben cumplirse las siguientes condiciones:

- 1. Cada elemento de la gramática, puede tener asociado un conjunto finito de atributos.
- Los atributos de un no-terminal o de un símbolo de acción, pueden ser heredados o sintetizados. Los elementos terminales sólo pueden tener atributos sintetizados.
- 3. Los atributos heredados del símbolo inicial de la gramática y los atributos sintetizados de los terminales, se especifican con un valor inicial, dado de una función externa.
- 4. Dada una producción, el valor del atributo heredado de un elemento del lado derecho, puede ser una constante, el valor de algún atributo heredado del lado izquierdo, o el valor de algún atributo sintetizado de un elemento que aparezca a la izquierda del elemento al que pertenece el atributo heredado dado.
- 5. Dada una producción, el valor de un atributo sintetizado del no-terminal del lado izquierdo puede ser una constante, el valor de algún atributo heredado del lado izquierdo, o el valor de algún atributo sintetizado del lado derecho.
- 6. Para los símbolos de acción, el valor de su atributo sintetizado, se calcula en función de atributos heredados del mismo símbolo de acción.

Para entender más dichas condiciones, revisemos si se cumplen para algunas producciones de la gramática del Ejercicio 6.3.2.

$$1: D \to T_t \ a_p \ \{AT\}_{p1,t1}, \ V_{t2} \qquad \begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases} \qquad \text{Estas son las reglas de asignación para la producción 1}$$

t es un atributo sintetizado del no-terminal del lado derecho \mathbf{T} ; cumple con la condición 2. Como es sintetizado y observando el árbol sintáctico del Ejercicio 6.3.2, su valor lo obtiene del atributo de su hoja, que es la producción $3: T_{t2} \rightarrow f_t$ $t2 \leftarrow t$ p es un atributo sintetizado del elemento terminal \mathbf{a} ; cumple la condición 2 ya que los terminales sólo pueden tener atributos sintetizados. Su valor lo obtendrá de una función que lea el token correspondiente y del componente *valor*, obtenga la posición del identificador dentro de la tabla de símbolos.

p1,t1 son atributos heredados del símbolo de acción {AT}, que se encuentra del lado derecho de la producción. El valor de p1 lo obtiene del atributo sintetizado del terminal a que aparece a la izquierda del símbolo de acción; al igual que t1 cuyo valor viene del atributo de T que también está a la izquierda del símbolo de acción; por lo que cumplen las condiciones 2 y 4.

t2 es atributo heredado del no-terminal del lado derecho V y su valor lo obtiene del atributo sintetizado de T que está a la izquierda de V. Cumple con las condiciones 2 y 4.

M.C. Laura Sandoval Montaño

Veamos otra producción:

$$3: T_{\underbrace{t1}} \to f_{\underbrace{t}} \qquad \qquad \begin{cases} \text{ t1} \leftarrow t \\ \text{ para la producción 3} \end{cases}$$

t es un atributo sintetizado del elemento terminal **a** (por la condición 2). Su valor lo obtendrá de una función que lea el token correspondiente y de su componente *valor*, obtenga la posición de la palabra reservada float en la tabla correspondiente.

t1 es un atributo sintetizado del no-terminal del lado izquierdo **T**; el valor lo obtiene del atributo sintetizado del elemento terminal **f** que está en el lado derecho de la producción. Cumple con las condiciones 2 y 5.

Si revisamos el resto de producciones de la gramática, podemos concluir que sí es una gramática de traducción con atributos a la que se le puede construir su tabla de parser LL para análisis sintáctico y semántico.

Construcción de la tabla de parser LL para revisión semántica a través del análisis sintáctico.

Como introdujimos a los atributos en la gramática de traducción, éstos deberán estar involucrados en las operaciones de la tabla de parser.

Cuando el no-terminal, terminal o símbolo de acción que entra en la pila, tiene atributos, cada atributo ocupará una localidad independiente en la pila, en la cual irá teniendo el valor que vaya adquiriendo el atributo.

De tal manera que deberá ocurrir lo siguiente cuando un elemento de la gramática esté en el tope de la pila, antes de un *Pop*:

- a) Las localidades correspondientes a sus atributos heredados, contendrán su valor.
- b) Las localidades correspondientes a sus atributos sintetizados, contendrán un apuntador a una lista ligada de localidades en las que deberá guardar su valor.

En virtud de que son varias las operaciones realizadas en las celdas, se enumerarán en la tabla de parser y se desarrollarán aparte.

Ejercicio 6.3.2.

Construye la tabla Parser LL de la gramática de traducción con atributos del Ejercicio 6.3.2.

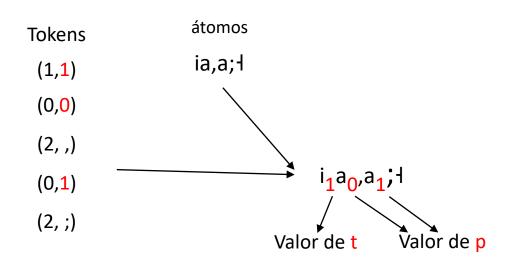
	а	i	f	С	,	;	4
D			Transición #1	Ĺ			
Т		Transición #2	Transición #3	Transición #4			
V					Transición #5	Transición #6	
а	Transición #7						
{AT}				Transició #8	n		
Λ							Acepta

Vamos a ir desarrollando y/o aplicando las operaciones de las transiciones #1, #2, #7, #8, #5, #7, #8 y #6, que son las que va reconociendo el Parser LL para la cadena ia,a; †; así veremos cómo se realiza la asignación de tipo en la transición #8.

Ejercicio 6.3.3 (continuación)

Recordemos que en el análisis semántico de la cadena ia,a;·l, se requiere el conocimiento del campo *valor* del token correspondiente a la palabra reservada i y de cada identificador a.

Por lo anterior, se van a ir leyendo cada átomo y su token correspondiente a la par, para todos los terminales de la cadena de entrada.



Catálogos de clases y palabras reservadas

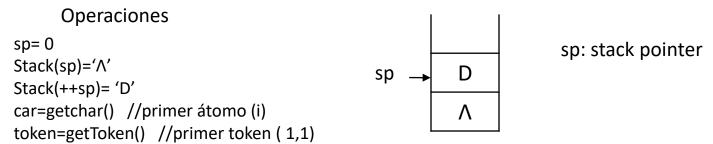
Clase	Componente léxico
0	identificador
1	Palabras reservadas
2	Símbolos especiales

Palabra reservada	Valor
char	0
int	1
float	2

Ejercicio 6.2.3 (continuación)

Desarrollemos las operaciones para el reconocimiento sintáctico y análisis semántico de la cadena ia,a; l

Situación inicial de la pila:



Como el símbolo inicial no tiene atributos, sólo entra dicho símbolo. Si tuviera un atributo sintetizado, ocuparía una localidad con valor de apuntador nulo (null); si tuviera un atributo heredado, ocuparía una localidad con su valor inicial (condición 3).

Para las operaciones de las transiciones, cada elemento de la gramática que tenga atributos y entre a la pila, sus atributos sintetizados apuntarán a la lista de localidades de los atributos a los que darán su valor; de los atributos heredados, se regirán por la condición 4, que puede ser una constante, o si reciben valor de atributos sintetizados, su valor será un apuntador al siguiente atributo de dicha lista y si es el último, su valor será nulo.

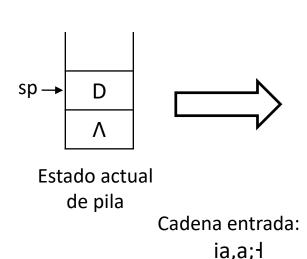
Ejercicio 6.3.3 (continuación)

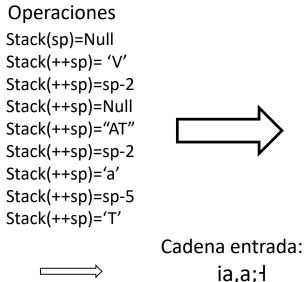
Transición #1. En el tope del stack está D y en la cadena de átomos se apunta a i

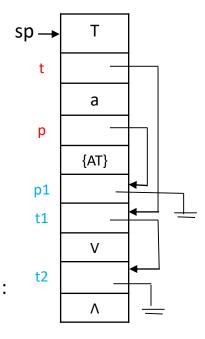
Corresponde a la producción:

1: D
$$\rightarrow$$
 T_t a_p {AT}_{p1,t1}, V_{t2}
$$\begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases}$$
 Equivale a: $t2 \leftarrow t1 \leftarrow t$

Y sus operaciones son: Reemp $(T_t a_p \{AT\}_{p1,t1}, V_{t2})^r$, retén. Hay que incluir las reglas de asignación de los atributos de la siguiente forma:





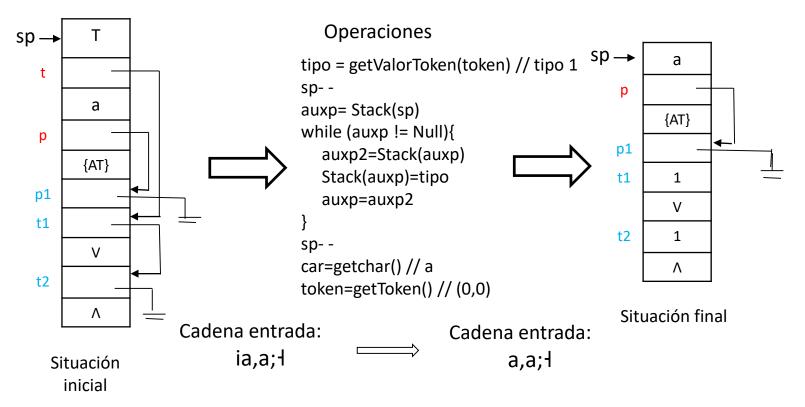


Situación final

M.C. Laura Sandoval Montaño

Ejercicio 6.3.3 (continuación)

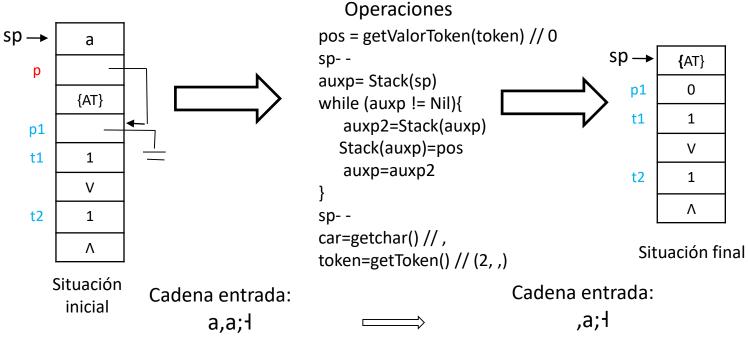
- Transición #2. Pertenece a la producción 2: $T_{t1} \rightarrow i_t$ $t1 \leftarrow t$
- Si no tuviera atributos, sólo se haría pop y avanza, pero como T tiene un atributo, que indica el tipo de dato; entonces, antes de hacer el pop, hay que obtener el tipo de dato del token correspondiente al átomo i y se colocará en las localidades de la lista apuntada por t



M.C. Laura Sandoval Montaño

Ejercicio 6.3.3 (continuación)

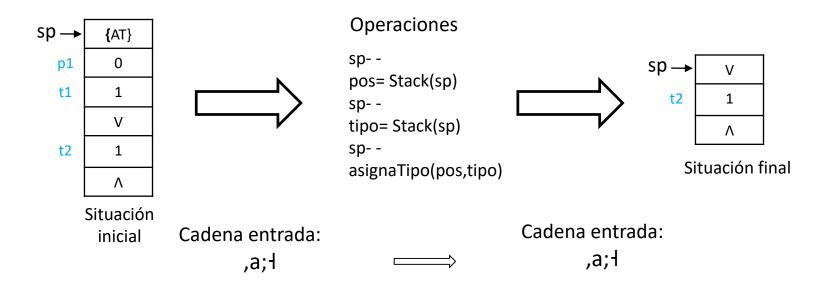
- Transiciones #3 y #4 son análogas a la transición #2, sólo que en la cadena de átomos se apunta a f o a c respectivamente.
- Transición #7. Corresponde a la celda TP[a,a]. Como el terminal a tiene un atributo, que indica la posición de la variable en la tabla de símbolos, antes de hacer el pop, hay que obtener el componente valor del token correspondiente al átomo a y se colocará en las localidades de la lista apuntada por p.



M.C. Laura Sandoval Montaño

Ejercicio 6.3.3 (continuación)

Transición #8. Corresponde cuando en el tope del stack está el símbolo de acción {AT}, sin importar lo que se esté apuntando en la cadena de entrada. Las operaciones, a groso modo, son Pop, Asigna Tipo y retén. Como el símbolo de acción está en el tope de la pila, sus atributos ya tienen valores, por lo que en la posición p1 de la tabla de símbolos, se pondrá el tipo t1.



Ejercicio 6.3.3 (continuación)

• Transición #5. En el tope del stack está V y en la cadena de átomos se apunta a ",

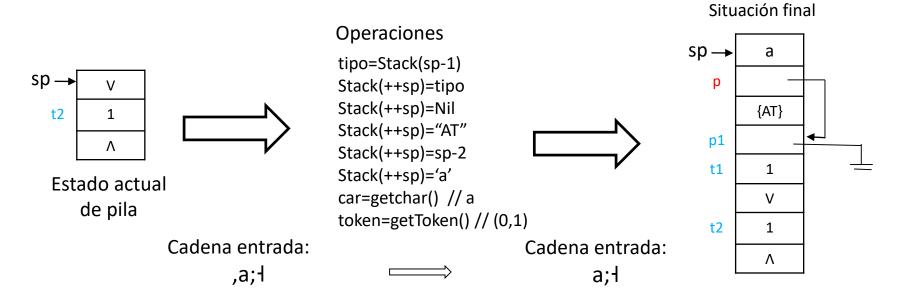
Corresponde a la producción:

5:
$$V_t \rightarrow$$
, $a_p \{AT\}_{p1,t1} V_{t2}$

$$\begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases}$$
Equivale a: $t2 \leftarrow t1 \leftarrow t$

Y sus operaciones son: Reemp($a_p \{AT\}_{p1,t1}, V_{t2}$)^r, avanza.

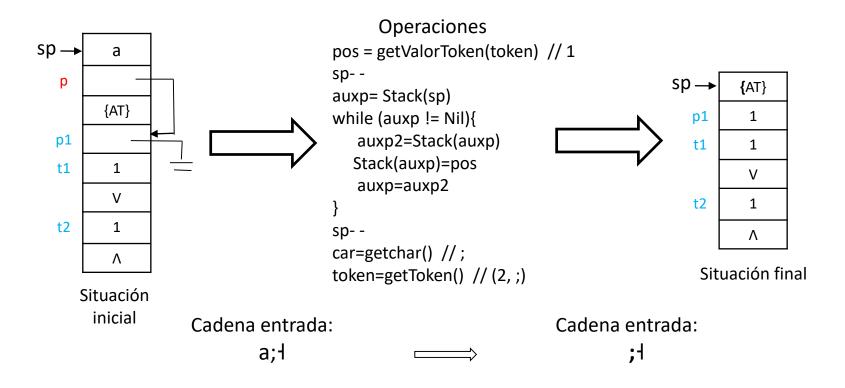
Hay que incluir las reglas de asignación de los atributos de la siguiente forma:



M.C. Laura Sandoval Montaño

Ejercicio 6.3.3 (continuación)

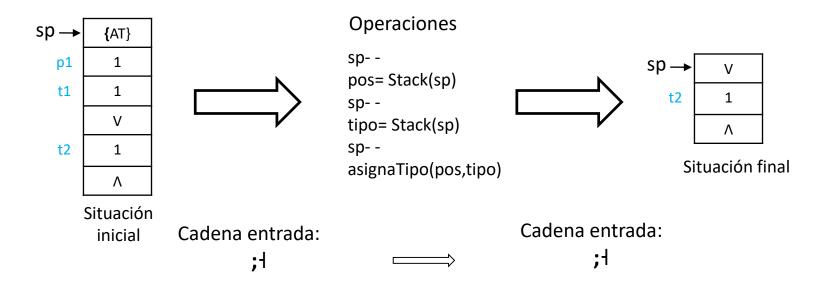
Transición #7. Nuevamente la aplicamos. Corresponde a la celda TP[a,a]. Como el terminal a tiene un atributo, que indica la posición de la variable en la tabla de símbolos, antes de hacer el pop, hay que obtener el componente valor del token del átomo a y se colocará en las localidades de la lista apuntada por p



M.C. Laura Sandoval Montaño

Ejercicio 6.3.3 (continuación)

 Transición #8. Nuevamente la aplicamos. Corresponde cuando en el tope del stack está el símbolo de acción {AT}, sin importar qué se esté apuntando en la cadena de entrada. Las operaciones, a groso modo, son Pop, Asigna Tipo y retén. Como el símbolo de acción está en el tope de la pila, sus atributos ya tienen valores, por lo que en la posición p1 de la tabla de símbolos, se pondrá el tipo t1.

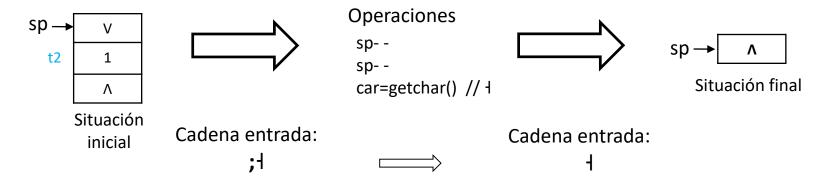


Ejercicio 6.3.3 (continuación)

Transición #6. En el tope del stack está V y en la cadena de átomos se apunta a;

Corresponde a la producción: 6: $V_t \rightarrow$;

Sus operaciones son: Pop y avanza ; como V tiene un atributo (cuyo valor ya no se va a usar), se tiene que hacer otro pop



Finalmente, como se apunta a indicador de stack vacío y al fin de cadena, se
 Acepta. En las dos transiciones #8 aplicadas, se hizo la asignación del tipo 1 (int) a
 las dos variables en la tabla de símbolos.

pos.	identif.	tipo
0	х	1
1	У	1

M.C. Laura Sandoval Montaño

Actividad 6.3.1

De la gramática del Ejercicio 6.3.2, revisa si se cumplen las condiciones de los atributos requeridas para la construcción del Parser LL, para cada uno de los atributos de la siguiente producción:

5:
$$V_t \rightarrow$$
, $a_p \{AT\}_{p1,t1} V_{t2} = \begin{cases} p1 \leftarrow p \\ t1 \leftarrow t \\ t2 \leftarrow t \end{cases}$

Actividad 6.3.2

La siguiente gramática define las sentencias declarativas del lenguaje del analizador

sintáctico que entregaron:

D' → DD'
$D' \rightarrow \xi$
$D \rightarrow V L$:
$L \rightarrow aGC$
$C \rightarrow L$
$C \rightarrow \xi$
$V \rightarrow b$
$V \rightarrow c$
$V \rightarrow f$
$V \rightarrow n$
$V \rightarrow g$
$G \rightarrow [e]$
$G \rightarrow \xi$

6. Traducción Dirigida por Sintaxis6.1 Gramáticas de traducción.

Actividad 6.3.2 (continuación)

- a) Incluye los símbolos de acción y los atributos (con sus reglas de asignación por producción) para que el analizador semántico realice la actualización de la columna tipo en las variables que se declaren.
- b) Elabora la tabla de Parser y desarrolla las transiciones. Para esto te puedes apoyar revisando una cadena de átomos. Utiliza la tabla de clases y de palabras reservadas definidas para la elaboración del analizador sintáctico ya entregado.