

Ejercicios sobre esquemas de traducción

Compiladores e intérpretes

Curso 2005/2006

Bloque 1: Desplazamiento de datos

Ejercicio 1

Dificultad: 1

Considera el esquema de traducción incompleto

$$\begin{aligned}\langle S \rangle &\rightarrow \langle A \rangle \textbf{ fin } \{ \text{Imprimir } \langle S \rangle.vlit \} \\ \langle A \rangle &\rightarrow \textbf{ abre } \langle A \rangle \textbf{ cierra} \\ \langle A \rangle &\rightarrow \textbf{ lit}\end{aligned}$$

y que el analizador léxico utilizado dota a los componentes de la categoría **lit** de un atributo *val* donde se almacena el valor del literal. Además, observa que toda secuencia de categorías léxicas aceptada por la gramática LL(1) implícita en el esquema incluye un único literal **lit**.

Se trata de completar el esquema de traducción con nuevas acciones semánticas para conseguir que lo que muestre la acción *Imprimir* $\langle S \rangle.vlit$ sea el valor del literal presente en la secuencia de componentes analizada. No necesitas utilizar ningún objeto global para ello.

El problema se resuelve fácilmente utilizando atributos sintetizados:

$$\begin{aligned}\langle S \rangle &\rightarrow \langle A \rangle \{ \langle S \rangle.vlit := \langle A \rangle.vlit \} \textbf{ fin } \{ \text{Imprimir } \langle S \rangle.vlit \} \\ \langle A \rangle &\rightarrow \textbf{ abre } \langle A \rangle_1 \textbf{ cierra} \{ \langle A \rangle.vlit := \langle A \rangle_1.vlit \} \\ \langle A \rangle &\rightarrow \textbf{ lit} \{ \langle A \rangle.vlit := \textbf{ lit.val} \}\end{aligned}$$

Ejercicio 2

Dificultad: 1

Considera la gramática

$$\begin{aligned}\langle Lis \rangle &\rightarrow \langle In \rangle \textbf{ elem } \langle Más \rangle \langle Fin \rangle \\ \langle Más \rangle &\rightarrow \textbf{ sep elem } \langle Más \rangle \mid \lambda \\ \langle In \rangle &\rightarrow \textbf{ apar } \mid \textbf{ acor} \\ \langle Fin \rangle &\rightarrow \textbf{ cpar } \mid \textbf{ ccor}\end{aligned}$$

donde las categorías léxicas **apar** y **cpar** corresponden a paréntesis (de apertura y cierre, respectivamente) y **acor** y **ccor**, análogamente, corresponden a corchetes. Observa que esta gramática permite utilizar un delimitador de cierre de tipo distinto al utilizado en la apertura (abrir un corchete y cerrar un paréntesis, por ejemplo), una circunstancia que se desea controlar semánticamente.

Se trata, por lo tanto, de añadir acciones semánticas a la gramática anterior de forma que el esquema de traducción resultante pueda sintetizar, mientras lleva a cabo un análisis LL(1) de su entrada, un atributo lógico $\langle Lis \rangle.ok$ que indique si se ha utilizado el mismo tipo de delimitador en apertura y cierre (valor Cierto) o no ha sido así (valor Falso). Además, ten en cuenta las siguientes restricciones:

- No puedes utilizar ningún objeto global.
- Todos los atributos deben ser de tipo lógico.
- Sólo se permiten acciones semánticas en el extremo derecho de la parte derecha de cada producción incontextual.

Y debes documentar de forma clara y concisa el significado de cada atributo. Por ejemplo:

- Atributo $\langle Lis \rangle.ok$: valor Cierto sólo si en la lista se utiliza el mismo tipo de delimitador para la apertura y el cierre.

Necesitamos dos atributos más, por ejemplo:

- Atributo $\langle Ini \rangle.par$: valor Cierto sólo si se utiliza una apertura de paréntesis.
- Atributo $\langle Fin \rangle.par$: valor Cierto sólo si se utiliza un cierre de paréntesis.

Y con ellos se puede construir el siguiente esquema de traducción:

$$\begin{aligned} \langle Lis \rangle &\rightarrow \langle Ini \rangle \mathbf{elem} \langle Más \rangle \langle Fin \rangle \{ \langle Lis \rangle.ok := (\langle Ini \rangle.par = \langle Fin \rangle.par) \} \\ \langle Más \rangle &\rightarrow \mathbf{sep elem} \langle Más \rangle \mid \lambda \\ \langle Ini \rangle &\rightarrow \mathbf{apar} \{ \langle Ini \rangle.par := Cierto \} \\ \langle Ini \rangle &\rightarrow \mathbf{acor} \{ \langle Ini \rangle.par := Falso \} \\ \langle Fin \rangle &\rightarrow \mathbf{cpar} \{ \langle Fin \rangle.par := Cierto \} \\ \langle Fin \rangle &\rightarrow \mathbf{ccor} \{ \langle Fin \rangle.par := Falso \} \end{aligned}$$

Obsérvese que, en la acción semántica de la primera producción, $\langle Ini \rangle.par$ y $\langle Fin \rangle.par$ sólo serán iguales si ambos atributos han tomado el valor Cierto, lo que corresponde a apertura y cierre de paréntesis, o si ambos han tomado el valor Falso, indicando con ello que se ha utilizado un par de corchetes.

Ejercicio 3

Dificultad: 2

En este caso, se parte de una gramática LL(1) que genera un lenguaje similar al del ejercicio anterior:

$$\begin{aligned} \langle Lis \rangle &\rightarrow \langle Ini \rangle \mathbf{elem} \langle Más \rangle \\ \langle Más \rangle &\rightarrow \mathbf{sep elem} \langle Más \rangle \mid \langle Fin \rangle \\ \langle Ini \rangle &\rightarrow \mathbf{apar} \mid \mathbf{acor} \mid \mathbf{alla} \\ \langle Fin \rangle &\rightarrow \mathbf{cpar} \mid \mathbf{ccor} \mid \mathbf{clla} \end{aligned}$$

El nuevo lenguaje permite la utilización de llaves como delimitadores para apertura y cierre (componentes **alla** y **clla**) y la estructura de la gramática es ligeramente distinta.

Se trata ahora de añadir acciones semánticas a la gramática de forma que se llame a una función especial `trata_error` si se ha utilizado un delimitador de cierre de tipo distinto al de apertura y, además, la llamada debe producirse desde acciones de la producción $\langle Más \rangle \rightarrow \langle Fin \rangle$, lo que impedirá resolver el ejercicio utilizando únicamente atributos sintetizados. Las restricciones adicionales que has de observar son éstas:

- No puedes utilizar ningún objeto global.
- Todos los atributos deben ser de tipo entero.

Finalmente, no olvides documentar todos los atributos utilizados.

Es necesario disponer de atributos en $\langle Ini \rangle$ y $\langle Más \rangle$ donde codificar mediante un valor entero el tipo de delimitador de apertura utilizado, por ejemplo: 1, paréntesis; 2, corchete; 3, llave. La información codificada se sintetizará en $\langle Ini \rangle$ y se irá heredando a través de los diferentes nodos $\langle Más \rangle$ del árbol de análisis. Utilizando la misma codificación, el tipo del delimitador de cierre puede sintetizarse en $\langle Fin \rangle$. Finalmente, si se utiliza el nombre *del* para todos los atributos anteriores, se puede llegar al siguiente esquema de traducción:

$$\begin{aligned} \langle Lis \rangle &\rightarrow \langle Ini \rangle \mathbf{elem} \{ \langle Más \rangle.del := \langle Ini \rangle.del \} \langle Más \rangle \\ \langle Más \rangle &\rightarrow \mathbf{sep elem} \{ \langle Más \rangle_1.del := \langle Más \rangle.del \} \langle Más \rangle_1 \\ \langle Más \rangle &\rightarrow \langle Fin \rangle \{ SI (\langle Fin \rangle.del \neq \langle Más \rangle.del): trata_error() FIN \} \\ \langle Ini \rangle &\rightarrow \mathbf{apar} \{ \langle Ini \rangle.del := 1 \} \\ \langle Ini \rangle &\rightarrow \mathbf{acor} \{ \langle Ini \rangle.del := 2 \} \\ \langle Ini \rangle &\rightarrow \mathbf{alla} \{ \langle Ini \rangle.del := 3 \} \\ \langle Fin \rangle &\rightarrow \mathbf{cpar} \{ \langle Fin \rangle.del := 1 \} \\ \langle Fin \rangle &\rightarrow \mathbf{ccor} \{ \langle Fin \rangle.del := 2 \} \\ \langle Fin \rangle &\rightarrow \mathbf{clla} \{ \langle Fin \rangle.del := 3 \} \end{aligned}$$

Bloque 2: Cómo contar

A partir de este bloque, y mientras no se indique lo contrario, se supondrá que todos los ejercicios han de resolverse cumpliendo las siguientes condiciones:

- El esquema de traducción resultante debe ser aplicable mientras se lleva a cabo un análisis LL(1) de la entrada.
- No puedes utilizar ningún objeto global.
- Los únicos tipos permitidos para los atributos de los no terminales son entero y lógico.
- Debes documentar todos los atributos que utilices.

Por lo tanto, no se insistirá en tales condiciones en los enunciados que siguen.

Ejercicio 4

Dificultad: 2

Debes añadir acciones semánticas a cada una de las siguientes gramáticas para que en el correspondiente atributo $\langle S \rangle.na$ se calcule el número de aes presentes en la entrada:

A. Una gramática de cadenas de aes:

$$\begin{aligned} \langle S \rangle &\rightarrow \langle X \rangle \\ \langle X \rangle &\rightarrow \mathbf{a} \langle X \rangle \mid \lambda \end{aligned}$$

Aparte de $\langle S \rangle.na$, se necesita un atributo entero asociado a cada no terminal $\langle X \rangle$, por ejemplo un $\langle X \rangle.na$ sintetizado donde se cuente el número de aes derivadas a partir de ese no terminal (es decir, que queden por debajo de él en el correspondiente árbol de análisis). El esquema de traducción resultante sería entonces éste:

$$\begin{aligned} \langle S \rangle &\rightarrow \langle X \rangle \{ \langle S \rangle.na := \langle X \rangle.na \} \\ \langle X \rangle &\rightarrow \mathbf{a} \langle X \rangle_1 \{ \langle X \rangle.na := \langle X \rangle_1.na + 1 \} \\ \langle X \rangle &\rightarrow \lambda \{ \langle X \rangle.na := 0 \} \end{aligned}$$

B. Una gramática de cadenas de aes, bes y ces:

$$\begin{aligned} \langle S \rangle &\rightarrow \mathbf{a} \langle X \rangle \mathbf{b} \\ \langle X \rangle &\rightarrow \mathbf{a} \langle X \rangle \langle X \rangle \mid \mathbf{b} \mathbf{a} \langle X \rangle \mathbf{a} \mid \mathbf{c} \langle Y \rangle \\ \langle Y \rangle &\rightarrow \mathbf{c} \langle Y \rangle \mid \mathbf{a} \end{aligned}$$

Pueden utilizarse exactamente los mismos atributos que en el esquema anterior (no es necesario tener un atributo $\langle Y \rangle.na$ porque de cualquier $\langle Y \rangle$ acabará derivándose una única a):

$$\begin{aligned} \langle S \rangle &\rightarrow \mathbf{a} \langle X \rangle \mathbf{b} \{ \langle S \rangle.na := \langle X \rangle.na + 1 \} \\ \langle X \rangle &\rightarrow \mathbf{a} \langle X \rangle_1 \langle X \rangle_2 \{ \langle X \rangle.na := \langle X \rangle_1.na + \langle X \rangle_2.na + 1 \} \\ \langle X \rangle &\rightarrow \mathbf{b} \mathbf{a} \langle X \rangle_1 \mathbf{a} \{ \langle X \rangle.na := \langle X \rangle_1.na + 2 \} \\ \langle X \rangle &\rightarrow \mathbf{c} \langle Y \rangle \{ \langle X \rangle.na := 1 \} \\ \langle Y \rangle &\rightarrow \mathbf{c} \langle Y \rangle \mid \mathbf{a} \end{aligned}$$

C. Una gramática donde cada una de las categorías léxicas asociadas a sus tres terminales, **may** (una letra mayúscula cualquiera), **min** (una minúscula) y **dig** (un dígito), cuenta con un atributo *lex* donde se guarda el correspondiente lexema:

$$\begin{aligned} \langle S \rangle &\rightarrow \mathbf{may} \langle X \rangle \mid \langle X \rangle \mid \mathbf{dig} \langle Y \rangle \\ \langle X \rangle &\rightarrow \mathbf{min} \langle X \rangle \mid \lambda \\ \langle Y \rangle &\rightarrow \mathbf{dig} \langle Y \rangle \mid \lambda \end{aligned}$$

De nuevo basta con utilizar atributos *na* en $\langle S \rangle$ y $\langle X \rangle$:

$$\begin{aligned} \langle S \rangle &\rightarrow \mathbf{may} \langle X \rangle \{ \text{SI } (\mathbf{may.lex} = \mathbf{A}): \langle S \rangle.na := \langle X \rangle.na + 1 \text{ SI_NO: } \langle S \rangle.na := \langle X \rangle.na \text{ FIN } \} \\ \langle S \rangle &\rightarrow \langle X \rangle \{ \langle S \rangle.na := \langle X \rangle.na \} \\ \langle S \rangle &\rightarrow \mathbf{dig} \langle Y \rangle \{ \langle S \rangle.na := 0 \} \\ \langle X \rangle &\rightarrow \mathbf{min} \langle X \rangle_1 \{ \text{SI } (\mathbf{min.lex} = \mathbf{a}): \langle X \rangle.na := \langle X \rangle_1.na + 1 \text{ SI_NO: } \langle X \rangle.na := \langle X \rangle_1.na \text{ FIN } \} \\ \langle X \rangle &\rightarrow \lambda \{ \langle X \rangle.na := 0 \} \\ \langle Y \rangle &\rightarrow \mathbf{dig} \langle Y \rangle \mid \lambda \end{aligned}$$

Ejercicio 5

Dificultad: 2

Considera la gramática

$$\begin{aligned}\langle S \rangle &\rightarrow \mathbf{apar} \langle A \rangle \mathbf{cpar} \\ \langle A \rangle &\rightarrow \mathbf{apar} \langle A \rangle \langle A \rangle \mathbf{cpar} \\ \langle A \rangle &\rightarrow \mathbf{x}\end{aligned}$$

donde las categorías léxicas **apar** y **cpar** corresponden, respectivamente, a paréntesis de apertura y cierre.

El *nivel de profundidad* de un punto de la entrada puede definirse entonces como el número de paréntesis que, habiendo sido abiertos con anterioridad a ese punto, aún no han sido cerrados.

Se trata de que, a partir de la gramática anterior, construyas dos esquemas de traducción distintos:

- A. En el primero, debes añadir acciones semánticas, exclusivamente en el extremo derecho de la parte derecha de cada producción, para que se calcule en un atributo entero $\langle S \rangle.mnp$ el máximo nivel de profundidad alcanzado en la entrada. Lógicamente, todos los atributos utilizados tendrán que ser sintetizados.

Aparte de $\langle S \rangle.mnp$, se necesita un atributo entero asociado a cada no terminal $\langle A \rangle$, por ejemplo un $\langle A \rangle.mnr$ sintetizado donde se calcule el máximo nivel relativo alcanzado dentro de la subcadena generada por ese no terminal, independientemente de a qué nivel se encuentre esa subcadena en el contexto de la entrada completa. Un posible esquema de traducción sería entonces éste:

$$\begin{aligned}\langle S \rangle &\rightarrow \mathbf{apar} \langle A \rangle \mathbf{cpar} \{ \langle S \rangle.mnp := \langle A \rangle.mnr + 1 \} \\ \langle A \rangle &\rightarrow \mathbf{apar} \langle A \rangle_1 \langle A \rangle_2 \mathbf{cpar} \{ \langle A \rangle.mnr := \max(\langle A \rangle_1.mnr, \langle A \rangle_2.mnr) + 1 \} \\ \langle A \rangle &\rightarrow \mathbf{x} \{ \langle A \rangle.mnr := 0 \}\end{aligned}$$

- B. En este caso, debes añadir las acciones necesarias para que, cada vez que se analice un componente **x**, se muestre por pantalla el nivel de profundidad en el que se ha encontrado. Necesitarás para ello algún atributo heredado.

Con un atributo entero $\langle A \rangle.np$ donde calcular el nivel al que aparece el correspondiente no terminal en el contexto de la entrada, podría construirse el siguiente esquema de traducción:

$$\begin{aligned}\langle S \rangle &\rightarrow \mathbf{apar} \{ \langle A \rangle.np := 1 \} \langle A \rangle \mathbf{cpar} \\ \langle A \rangle &\rightarrow \mathbf{apar} \{ \langle A \rangle_1.np := \langle A \rangle.np + 1 \} \langle A \rangle_1 \{ \langle A \rangle_2.np := \langle A \rangle.np + 1 \} \langle A \rangle_2 \mathbf{cpar} \\ \langle A \rangle &\rightarrow \mathbf{x} \{ \text{Imprimir } \langle A \rangle.np \}\end{aligned}$$

Ejercicio 6

Dificultad: 4

Considera ahora la gramática, también con paréntesis **apar** y **cpar**,

$$\begin{aligned}\langle S \rangle &\rightarrow \langle A \rangle \\ \langle A \rangle &\rightarrow \mathbf{apar} \mathbf{apar} \langle A \rangle \mathbf{cpar} \langle A \rangle \mathbf{cpar} \mid \mathbf{x} \langle A \rangle \mid \mathbf{y} \mathbf{apar} \mathbf{x} \langle A \rangle \mathbf{cpar} \mid \lambda\end{aligned}$$

y añádele las acciones semánticas necesarias para sintetizar en el símbolo inicial $\langle S \rangle$ los dos atributos enteros siguientes:

- El máximo nivel de profundidad alcanzado en la entrada, $\langle S \rangle.mnp$.
- El número de componentes de la categoría **x** que aparecen a profundidad 3, $\langle S \rangle.nxt$.

Como primera aproximación a la solución del problema, se presenta un posible esquema de traducción para la síntesis del atributo $\langle S \rangle.nxt$:

$$\begin{aligned}\langle S \rangle &\rightarrow \{ \langle A \rangle.np := 0 \} \langle A \rangle \{ \langle S \rangle.nxt := \langle A \rangle.nxt \} \\ \langle A \rangle &\rightarrow \mathbf{apar} \mathbf{apar} \{ \langle A \rangle_1.np := \langle A \rangle.np + 2 \} \langle A \rangle_1 \mathbf{cpar} \{ \langle A \rangle_2.np := \langle A \rangle.np + 1 \} \langle A \rangle_2 \mathbf{cpar} \{ \langle A \rangle.nxt := \langle A \rangle_1.nxt + \langle A \rangle_2.nxt \} \\ \langle A \rangle &\rightarrow \mathbf{x} \{ \langle A \rangle_1.np := \langle A \rangle.np \} \langle A \rangle_1 \{ \langle A \rangle.nxt := \langle A \rangle_1.nxt + \delta(\langle A \rangle.np, 3) \} \\ \langle A \rangle &\rightarrow \mathbf{y} \mathbf{apar} \mathbf{x} \{ \langle A \rangle_1.np := \langle A \rangle.np + 1 \} \langle A \rangle_1 \mathbf{cpar} \{ \langle A \rangle.nxt := \langle A \rangle_1.nxt + \delta(\langle A \rangle.np, 2) \} \\ \langle A \rangle &\rightarrow \lambda \{ \langle A \rangle.nxt := 0 \}\end{aligned}$$

En él se ha hecho uso de los siguientes atributos adicionales:

- Atributo $\langle A \rangle.np$: nivel al que aparece el correspondiente no terminal en el contexto de la entrada.

- Atributo $\langle A \rangle.nxt$: cuántos componentes x a profundidad 3 dentro del contexto de la entrada aparecen en la subcadena generada por ese no terminal $\langle A \rangle$.

Además, se ha utilizado la notación $\delta(m, n)$, donde la función devuelve 1 si coinciden los valores de m y n y, en caso contrario, devuelve 0.

Una posible solución al ejercicio consistiría en añadir al anterior esquema de traducción las acciones semánticas del siguiente, encaminadas a calcular $\langle S \rangle.mnp$:

$$\begin{aligned}
 \langle S \rangle &\rightarrow \langle A \rangle \{ \langle S \rangle.mnp := \langle A \rangle.mnr \} \\
 \langle A \rangle &\rightarrow \mathbf{apar\ apar} \langle A \rangle_1 \mathbf{cpar} \langle A \rangle_2 \mathbf{cpar} \{ \langle A \rangle.mnr := \max(\langle A \rangle_1.mnr + 2, \langle A \rangle_2.mnr + 1) \} \\
 \langle A \rangle &\rightarrow \mathbf{x} \langle A \rangle_1 \{ \langle A \rangle.mnr := \langle A \rangle_1.mnr \} \\
 \langle A \rangle &\rightarrow \mathbf{y\ apar\ x} \langle A \rangle_1 \mathbf{cpar} \{ \langle A \rangle.mnr := \langle A \rangle_1.mnr + 1 \} \\
 \langle A \rangle &\rightarrow \lambda \{ \langle A \rangle.mnr := 0 \}
 \end{aligned}$$

En este esquema se ha utilizado el atributo $\langle A \rangle.mnr$, máximo nivel relativo alcanzado dentro de la subcadena generada por ese no terminal $\langle A \rangle$.

Alternativamente, el ejercicio puede resolverse añadiendo al primer esquema de traducción acciones que aprovechen el valor del atributo $\langle A \rangle.np$ heredado por la regla $\langle A \rangle \rightarrow \lambda$ para, a partir de este punto, ir sintetizando atributos $\langle A \rangle.mnp$ donde, en cada uno de tales atributos, se calcule el máximo nivel absoluto alcanzado dentro de la subcadena generada por el no terminal correspondiente. El siguiente esquema de traducción, incompleto, muestra únicamente las acciones que habría que añadir al primero para obtener una posible solución a este ejercicio:

$$\begin{aligned}
 \langle S \rangle &\rightarrow \langle A \rangle \{ \langle S \rangle.mnp := \langle A \rangle.mnp \} \\
 \langle A \rangle &\rightarrow \mathbf{apar\ apar} \langle A \rangle_1 \mathbf{cpar} \langle A \rangle_2 \mathbf{cpar} \{ \langle A \rangle.mnp := \max(\langle A \rangle_1.mnp, \langle A \rangle_2.mnp) \} \\
 \langle A \rangle &\rightarrow \mathbf{x} \langle A \rangle_1 \{ \langle A \rangle.mnp := \langle A \rangle_1.mnp \} \\
 \langle A \rangle &\rightarrow \mathbf{y\ apar\ x} \langle A \rangle_1 \mathbf{cpar} \{ \langle A \rangle.mnp := \langle A \rangle_1.mnp \} \\
 \langle A \rangle &\rightarrow \lambda \{ \langle A \rangle.mnp := \langle A \rangle.np \}
 \end{aligned}$$

Bloque 3: Pues ya veremos...