

Segunda Práctica

Teoría de Autómatas y Lenguajes Formales I
Escuela Politécnica Superior, Universidad Autónoma de Madrid
ENTREGA: semana del 16 al 20 de mayo

Curso 2004-2005

1 Parte obligatoria

El objetivo de esta segunda práctica es aprovechar la primera práctica para construir una pequeña interfaz de mandatos para un programa de visualización de imágenes, en español simplificado.

En esta interfaz, toda instrucción comenzará con un verbo y necesitará un objeto directo. Ese objeto directo puede ser un sintagma nominal simple o compuesto. Los nombres pueden ir o no modificados por adjetivos. Finalmente, varias oraciones pueden ir unidas por una conjunción.

Se trata de implementar un programa que, a partir de una gramática independiente del contexto, reconozca la estructura sintáctica de la instrucción del usuario, de acuerdo a la gramática que os proporcionamos en este enunciado.

Se admitirán las siguientes peticiones:

- Peticiones de visualización de imágenes. Por ejemplo,

`mostrar el gato verde`

- Petición de visualización relativas a la imagen actual. Por ejemplo,

`avanzar cinco fotos
retroceder 2 fotos`

- Peticiones de modificar el aumento de la imagen. Por ejemplo,

`ampliar la imagen
reducir la foto`

- Varias peticiones unidas por una conjunción. Por ejemplo,

`mostrar el hombre y el perro y avanzar una imagen`

En este caso, el programa de visualización localizaría una imagen cuyo nombre coincidiera con la solicitada, por ejemplo, `el hombre triste` y `el perro gris.gif`. Mostraría esta imagen al usuario, esperaría a que éste pulsara la barra espaciadora, y pasaría a la siguiente acción, para mostrar la imagen siguiente dentro del directorio.

- Finalmente, el siguiente mandato se incluye también para finalizar la ejecución del programa.

`terminar el programa`

A continuación se muestra la gramática del lenguaje de órdenes:

```
S ::= verbo NP
S ::= S conj S
NP ::= det nom AP
      | det nom
NP ::= num nom AP
      | num nom
NP ::= NP conj NP
AP ::= adj
AP ::= AP conj AP
```

Figure 1: Gramática del lenguaje de órdenes, sin concordancia. Los símbolos no terminales serían `S` (oración, *sentence*), `NP` (sintagma nominal, *noun phrase*), y `AP` (sintagma adjetival, *adjectival phrase*).

Para construir el analizador, será necesario, en primer lugar, transformarla en una gramática equivalente en forma normal de Chomsky. Una vez hecho este cambio, se utilizará un algoritmo llamado CKY o CYK, cuyo pseudocódigo se detalla a continuación:

Para analizar una frase s con N palabras,

1. Inicializar una matriz m con N filas y N columnas. Cada celdilla de la matriz contendrá un array unidimensional, inicialmente vacío.
2. Para i variando de 0 a $N - 1$,
 - (a) Rellenar la casilla $m[i][i]$ con un par $(c, s[i])$, donde c es la categoría sintáctica (verbo, conj, det, nom o adj) de la palabra $s[i]$.
3. Para j variando de 1 a $N - 1$,
 - (a) Para i variando de $j - 1$ a 0,
 - i. Para k variando de i a $j - 1$,
 - A. Si $(A_1, t_1) \in m[i][k]$, $(A_2, t_2) \in m[k + 1][j]$, y hay una regla en la gramática de la forma
$$A_0 ::= A_1 A_2$$
entonces añadir (A_0, t) a la casilla $m[i][j]$, donde t es un árbol que tiene A_0 como raíz, t_1 como hijo izquierdo, y t_2 como hijo derecho.
4. Si (S, t) pertenece a la casilla $m[0][N - 1]$, la oración se considerará bien formada, y t será un posible árbol de derivación sintáctica de la misma.

Si la oración no estaba bien formada, se notificará al usuario con un mensaje de error.

En caso de que la oración haya sido considerada correcta, se mostrará en pantalla el árbol de derivación sintáctica obtenido.

2 Parte optativa

La parte optativa consiste en incorporar esta interfaz a un sistema de manipulación de imágenes, para que efectúe realmente las acciones indicadas por el usuario, del siguiente modo:

1. Si es un mandato compuesto, por ejemplo **avanzar una foto y terminar el programa**, se separará en subárboles correspondientes a mandatos separados, para ejecutar uno detrás de otro por separado. Entre mandato y mandato se esperará a que el usuario pulse la barra espaciadora.
2. Para cada uno de los mandatos, se tomará el verbo, para saber qué tipo de acción hay que realizar:
 - (a) Si es **terminar**, finalizará la ejecución del programa.
 - (b) Si es **avanzar**, extraerá del complemento directo el número n de imágenes que ha de avanzar, y procederá a mostrar la imagen que se encuentre n puestos más adelante en el listado de archivos del directorio de fotos.

- (c) Si es **retroceder**, extraerá del complemento directo del verbo el número de imágenes n que ha de retroceder, y procederá a mostrar la imagen solicitada.
- (d) Si es **ampliar** mostrará la foto ampliada, aumentando el zoom.
- (e) Si es **reducir**, mostrará la foto reducida.
- (f) Si es **mostrar**, deberá tomar el complemento directo del verbo, para saber qué contenido desea ver el usuario. Éste va a ser el mandato más complicado de realizar, dado que será necesario examinar los nombres de todas las fotos del directorio (recorriendo éste en orden, y comenzando en la imagen que está mostrada actualmente), localizar la primera cuyo nombre pueda coincidir con la consulta solicitada, y mostrarla.

Consideraremos que el nombre de una foto coincide con el nombre solicitado si se cumple la siguiente condición:

- Para todos los sintagmas nominales de la consulta, existe un sintagma nominal en el nombre de la foto que coincide con él.

Igualmente, consideraremos que un sintagma nominal de una consulta coincide con un sintagma nominal de una foto siempre y cuando:

- El nombre en ambos sintagmas nominales coincide.
- Los adjetivos del sintagma nominal de la consulta son un subconjunto de los adjetivos del sintagma nominal del nombre del archivo.

Por ejemplo, si tenemos la consulta **Mostrar el cazador furtivo y el perro gris**, supongamos que los archivos del directorio son los siguientes (en el orden indicado):

- i. **el cazador y el perro gris.jpg**: Este archivo no sería seleccionado, pues el cazador tendría que tener el adjetivo “furtivo”.
- ii. **el cazador furtivo.jpg**: Este archivo no sería seleccionado, pues debería aparecer también el perro gris.
- iii. **el perro blanco y gris y el cazador furtivo.jpg**: Este archivo sí sería seleccionado, pues tiene dos sintagmas nominales, **el perro blanco y gris**, que tiene el mismo nombre que **el perro gris** de la consulta, y más adjetivos, y **el cazador furtivo**, que tiene el mismo nombre que **el cazador furtivo** de la consulta, y el mismo adjetivo.

El programa los analizaría uno por uno, con el analizador sintáctico, y retornaría el tercero, pues es el primer archivo del directorio cuyo nombre ha cumplido todas las condiciones.

Como otro ejemplo, si el usuario escribe la consulta **mostrar el gato**, entonces todas las imágenes que contengan el nombre **gato** podrían ser seleccionadas.

3 Material complementario

Para la realización de esta práctica se proporcionará el archivo **practica3.tar.gz**, que contiene:

- `imagenes/*`, un directorio con imágenes para probar el programa, y `imagenes.lst`, un archivo con las descripciones de todas ellas.
- `lenguajes`, que contiene el archivo de entrada con palabras, y los lenguajes a los que pertenece cada una de ellas.
- Dos archivos, `display.c` y `display.h`, que contienen el código para visualizar una imagen. Sólo serán necesarios si se realiza la parte optativa de la práctica. Si se utilizan, para compilar, se deberán utilizar las opciones de compilación que necesita ImageMagick. Tenéis un ejemplo sencillo de estas opciones en el archivo `compile.sh`
- Un ejecutable como ejemplo del programa completo. Para ponerlo en marcha, se ha de escribir:

```
./albumDeFotos lenguajes imagenes.lst
```

Para salir, escribir `terminar` el programa.

Estos archivos estarán disponibles en la página web de prácticas de la asignatura, junto con el enunciado de la práctica.

4 Otras prácticas opcionales

Se pueden realizar las siguientes prácticas opcionales:

- Extender tanto el lenguaje de mandatos como las funcionalidades del programa de acuerdo a las opciones de manipulación de imágenes de la biblioteca utilizada, por ejemplo, añadiendo la facultad de modificar brillo o contraste, guardar la imagen modificada, o visualizar el histograma de colores.
- Añadir al analizador sintáctico concordancia de género y número entre determinantes, nombres y adjetivos, dentro de los sintagmas nominales, para que una frase como la siguiente

`mostrar los perro`

sea considerada incorrecta, pues el determinante no concuerda en número con el nombre.

- Añadir al programa algún procedimiento para traducir cada palabra a su forma canónica, de modo que si tenemos una imagen descrita como `los perros`, y el usuario escribe el mandato `mostrar el perro`, también sea seleccionada, por corresponder a la misma palabra, pero en plural.