# Análisis Léxico

Examen octubre de 2014 Agencia de Espías

Juan Pedro Caraça-Valente Hernández septiembre 2020

Una agencia internacional de espías ha decidido introducir un nuevo sistema para que los espías informen de sus movimientos. Para ello, el espía deberá emitir su código clave y las coordenadas geográficas de su destino base, junto a un desplazamiento opcional respecto a dicha base. Si el espía está trabajando en equipo junto a otros espías, deberá enviar en su informe el código clave de todos los espías del equipo.

- El código clave de espía está formado por una secuencia de entre 3 y 5 dígitos y solo son válidos los códigos de espías almacenados en la tabla de espías de la agencia.
- Las coordenadas geográficas son un sistema de referencia que utiliza las dos coordenadas angulares, latitud y longitud, y sirve para determinar la posición exacta de un punto sobre la superficie terrestre:
  - La latitud mide el ángulo entre cualquier punto y el ecuador, de tal manera que será latitud Norte (N) si está al norte del ecuador, y latitud Sur (S) si está al sur. El valor de la latitud será un número real entre 0 y 90 grados (aunque se puede omitir la parte decimal), seguido inmediatamente de la letra 'N' o 'S'.
  - La longitud mide el ángulo a lo largo del Ecuador desde cualquier punto de la Tierra al meridiano de Greenwich. Se acepta que Greenwich en Londres es la longitud 0, por lo que longitud Este (E) se sitúa al este de Greenwich y longitud Oeste (W) al oeste de Greenwich. La longitud será un número real entre 0 y 180 grados (aunque se puede omitir la parte decimal), seguido inmediatamente de la letra 'E' o 'W'.
- Para el desplazamiento respecto al destino base, se utiliza un sistema de cuadrícula secreta centrada en el destino base y formado por una letra y dos dígitos.

La agencia quiere procesar al final del día todos los informes recibidos de los espías y que se han ido almacenando en un fichero, a razón de un informe en cada línea, teniendo en cuenta que los diferentes elementos del informe pueden venir o no separados por blancos o tabuladores. La agencia necesita los valores de las coordenadas para poder calcular la distancia básica entre los espías. Para todo ello, ha encargado a la ETSIInf la realización del diseño de un Analizador Léxico que reconozca todos los elementos de los informes, para lo cual hay que detallar la definición de los tokens, la Gramática Tipo 3, el Autómata Finito Determinista, acciones semánticas y errores.

Seguidamente, se muestra un ejemplo de un fichero con cinco informes correctos:

Y un ejemplo de un fichero en el que todos los elementos son incorrectos (asumiendo que el espía 'doble-cero tres' no pertenece a la agencia):

## 1) Definición de Tokens

Espía: <ESP, pos> lexema: 3 a 5 dígitos, atributo: posición en la tabla de agentes

<Norte, valor> lexema: dígitos formando número real entre 0 y 90 grados (aunque se

puede omitir la parte decimal), seguido inmediatamente de la letra 'N'

<Sur, valor>

<Oeste, valor> lexema: dígitos formando número real entre 0 y 180 grados (aunque se

puede omitir la parte decimal), seguido inmediatamente de la letra 'W'

<Este, valor>

<Desplazamiento, lexema> lexema: una letra seguida por2 dígitos

atributo: lexema de la cuadrícula

<cr, -> lexema: carácter salto de linea

Alternativa:

<Latitud, valor> Atributo valor positivo para Norte, negativo para Sur

<Longitud, valor> Atributo valor positivo para Este, negativo para Oeste

### 1) Tokens:

- <ESP, pos>: espía, siendo el atributo la posición en la tabla de agentes
- <N, valor>: latitud norte, siendo el atributo el valor del ángulo
- <S, valor>: latitud sur, siendo el atributo el valor del ángulo
- <0, valor>: longitud oeste, siendo el atributo el valor del ángulo
- <E, valor>: longitud este, siendo el atributo el valor del ángulo
- < Desplazamiento, lex>: posición de la cuadrícula, siendo el atributo el lexema de la cuadrícula
- <CR, ->: salto de línea

### 2) Gramática Regular:

 $A \rightarrow dB \mid IE \mid delA \mid cr$ 

 $B \rightarrow dB \mid .C \mid n \mid s \mid w \mid e \mid \lambda$ 

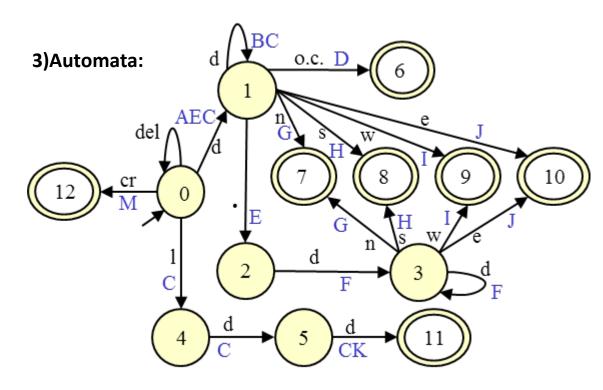
 $C \rightarrow dD$ 

 $D \rightarrow dD \mid n \mid s \mid w \mid e \mid \lambda$ 

 $E \rightarrow dF$ 

 $F \rightarrow d$ 

d::= dígito, l::= letra, del::= blanco/tabulador, cr::= salto de línea, n,s,w,e:== letras 'N', 'S', 'W', 'E'



#### 4) Acciones semánticas

• Error: todas las transiciones no indicadas, producirían una acción de Error.

•Leer: realiza la lectura del siguiente carácter del fichero de entrada. Se lee en todas las transiciones (excepto en

la transición al estado 6), aunque no se ha indicado en el AFD por simplificar.

Nombre	Definición	1
Α	pal:= φ; cont:= 1; núm:= d	d l
В	núm:= núm * 10 + d; cont:= cont + 1	del AEC
С	concat (pal)	
D	If (cont<3 or cont>5) Then Error ("longitud del código de espía incorrecta")  Else p:= BuscaT (pal)//busca en la tabla de espías If (p=NULL) Then Error ("espía", pal, " inexistente") Else GenToken (ESP, p)	$\begin{array}{c c} 12 & cr & 0 \\ \hline & 1 & c \\ \hline \end{array}$
E	dec:= 0	c] (
F	dec:= dec + 1; núm:= núm * 10 + d	d d
G	If (núm/10 <sup>dec</sup> > 90) Then Error ("latitud norte fuera de rango") Else GenToken (N, núm/10 <sup>dec</sup> )	4 C
Н	If (núm/10 <sup>dec</sup> > 90) Then Error ("latitud sur fuera de rango") Else GenToken (S, núm/10 <sup>dec</sup> )	
1	If (núm/10 <sup>dec</sup> > 180) Then Error ("longitud oeste fuera de rango") Else GenToken (O, núm/10 <sup>dec</sup> )	
J	If (núm/10 <sup>dec</sup> > 180) Then Error ("longitud este fuera de rango") Else GenToken (E, núm/10 <sup>dec</sup> )	
K	GenToken (Desplazamiento, pal)	
М	GenToken (CR, -)	

