

# WUOLAH



Iridescent

[www.wuolah.com/student/Iridescent](http://www.wuolah.com/student/Iridescent)



7510

## ExamenPDL2019.pdf

*Examen Resuelto 23-10-2019(1Parcial)*



3º Procesadores de Lenguajes



Grado en Ingeniería Informática



Escuela Técnica Superior de Ingenieros Informáticos  
Universidad Politécnica de Madrid

### -50€ OFF

¡Viaje sorpresa en camper  
con tus amigos!

Descubre tu destino 2 días antes

Código: WAYNABOXSTUDENT



# PROCESADORES DE LENGUAJES

## ANÁLISIS LÉXICO Y TABLA DE SÍMBOLOS

23 de octubre de 2019

**Observaciones:** 1. Las calificaciones se publicarán hacia el 6 de noviembre.  
2. La revisión será hacia el 8 de noviembre.  
3. En la web se avisará de las fechas exactas.  
4. La duración de este examen es de 40 minutos.

Un supermercado en línea quiere tener cada día un registro de todos los productos comprados por sus clientes para controlar las existencias de sus productos. Para ello, la aplicación de la tienda proporciona al final del día un fichero que contiene, en este orden, la siguiente información (cada elemento puede ir separado por delimitadores):

- Nombre completo del comprador: está formado por letras, aunque también existen nombres compuestos (con un único guion en su interior para separar ambos nombres). Ej.: Juan Moreno-Romero. (Si un comprador adquiere varios productos, solamente aparecerá su nombre una vez delante de todos sus productos.)
- Código EAN (*European Article Number*) del producto: el código de barras de cada artículo se corresponde con una secuencia de 13 dígitos, donde los 3 primeros dígitos identifican el país del producto (solo hay 675 prefijos válidos y cada prefijo con su país se encuentran accesibles en una tabla). Ej.: 8412042502367 (el prefijo 841 corresponde a España), 7702993035023 (el prefijo 770 corresponde a Colombia), 1997531598520 (199 es uno de los prefijos que no corresponde a ningún país, por lo que es un EAN erróneo).
- Descripción del producto: viene en una o varias cadenas encerradas entre comillas simples ('). Ej.: 'Agua Mineral Natural' '500 ml.' 'Pontevedra (España)'. (Una cadena no puede tener más de 128 caracteres.)
- Precio del producto: es un número que puede tener una coma seguida de uno o dos decimales. Ej.: 11,99 103 2,5. (El supermercado no tiene productos que cuesten más de 5000€.)



Ejemplo de fichero de compras correcto:

```
Juan Luis Moreno-Romero y Alonso 8412042502367 'agua
mineral -natural-' '500ml.' 'Pontevedra (España)' 1,99
7702993035023'12 Barritas de coco con chocolate, 12 un.'3,5 Peter
Navarro0692771022044 '' 'American block 120 hojas, 6*4.3" 4º' 10
```

Ejemplo de fichero de compras con elementos incorrectos:

```
Juan-Luis-Pedro Moreno- Fdez. 12345678 'abcdefghijklmnopqrstuvwxy
abcdefghijklmnopqrstuvwxy abcdefghijklmnopqrstuvwxy
abcdefghijklmnopqrstuvwxy abcdefghijklmnopqrstuvwxy' ,99
1997515389521 '(Madrid) 8,
```

Se pide realizar el Diseño del **Analizador Léxico**, en todas sus fases: *Tokens*, Gramática Regular, Autómata Finito, Acciones Semánticas y Errores.





## Aprende desde casa, como si estuvieras en el aula. Cursos on-line, trato personalizado a distancia.

Contacto personalizado, material actualizado, videos explicativos, sesiones de dudas y tutorías.

Especializados en estudios de ingeniería informática. Computación, Software, Videojuegos. Dobles grados en ADE y Matemáticas.

① Se pide realizar el Diseño del analizador léxico, en todos sus fases: Tokens, Gramática Regular, Automata Finito, Acciones semánticas y errores.

② Tokens:

- <EAN, lexema>: Código del producto, siendo el atributo el lexema del EAN.
- <DESC, lexema>: Cadena con la descripción, siendo el atributo el lexema con el contenido de la cadena.
- <NOM, lexema>: Representa un nombre o un apellido, siendo el atributo el lexema completo
- <Precio, valor>: Representa el precio del producto, siendo el atributo el valor del número.

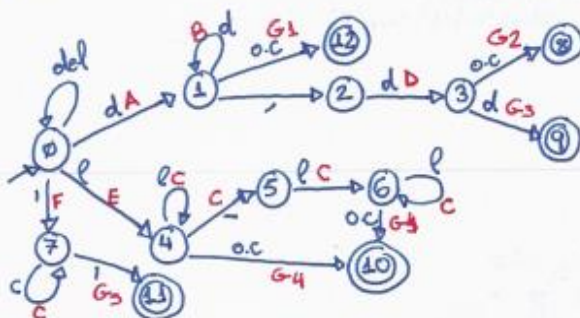
③ Gramática Regular:

$A \rightarrow d B | E | 'H' | del A$   
 $B \rightarrow d B | C | \lambda$   
 $C \rightarrow d B$   
 $D \rightarrow d | A$   
 $E \rightarrow E | F | \lambda$   
 $F \rightarrow E G$   
 $G \rightarrow E G | A$   
 $H \rightarrow C H | \lambda$

Donde:

- $d ::=$  dígito
- $E ::=$  letra
- $C ::=$  un carácter excepto comilla
- $del ::=$  blanco, tabulador, salto de línea

④ AFD:



④ Acciones semánticas y errores:

- Leer: Realiza la lectura del siguiente carácter del fichero de entrada. Se lee en todas sus transiciones (excepto en las etiquetadas con o.c)
- Error: Cualquier transición no prevista en el AFD sería un error.

- A - Concat; cont := 1; valor := d
- B - Concat; cont := cont + 1; valor := valor \* 10 + d  
if (cont = 3) pais := Busca(lex)
- C - Concat: concatena el carácter leído en la variable lex
- D - valor := valor + d/10
- E - cont := 0; Concat
- F - cont := 0
- G1 - if (cont = 13 and pais != NULL) then GenToken (EAN, lexema)  
else if (valor > 5000) then Error ("Precio o EAN incorrecto")  
else GenToken (Precio, valor)
- G2 - if (valor > 5000) then Error ("Precio incorrecto")  
else GenToken (Precio, valor)
- G3 - valor := valor + d/100  
if (valor > 5000) then Error ("Precio incorrecto")  
else GenToken (Precio, valor)
- G4 - GenToken (NOM, lexema)
- G5 - if (cont > 128) then Error ("Cadena demasiado larga")  
else GenToken (DESC, lexema)