

### Actividad Evaluada de Tema 3

Un tokenizador es un proceso que permite dividir la entrada (texto) en unidades o tokens. Estas unidades pueden ser por ejemplo: signos de puntuación, palabras, enteros, reales, oraciones, párrafos, etc. Se quiere que programe un tokenizador usando el generador de analizadores léxicos Lex de linux que reciba como entrada un texto y genere una salida texto indicando en cada línea la especificación de cada token (palabras, verbos en infinitivo, signos de puntuación, reales, enteros, enteros con porcentajes, reales con porcentajes). Por ejemplo, dado el siguiente texto como entrada:

```
La oferta mundial de cacao puede estar entre 40 y 50%,  
llegando a niveles de 5% de participación mundial (120.000  
tm/año).
```

Genere como salida la siguiente:

```
La --> palabra  
oferta --> palabra  
mundial --> palabra  
de --> palabra  
cacao --> palabra  
puede --> palabra  
estar --> verbo infinitivo  
entre --> palabra  
40 --> entero  
y --> palabra  
50,5% --> real porcentaje  
, --> signo de puntuacion  
llegando --> palabra  
a --> palabra  
niveles --> palabra  
de --> palabra  
5% --> entero porcentaje  
de --> palabra  
participación --> palabra  
mundial --> palabra  
( --> signo de puntuación  
120.000 --> entero  
tm/año --> palabra  
) --> signo de puntuación  
. --> signo de puntuación
```

```
Número de palabras = 16  
Número de verbos infinitos = 1  
Número de enteros = 2  
Número de real con porcentaje = 1  
Número de entero con porcentaje = 1
```

Número de signos de puntuación = 4  
Número total de tokens = 25

Para la entrega de esta tarea por favor envíe un archivo comprimido identificado con NombreApellido-tema3.\* (donde \* puede ser zip, rar o gz) que contenga:

- (1) el archivo fuente con las instrucciones Lex con extensión .l
- (2) el archivo compilado, archivos de entrada de prueba y los archivos de salida de cada archivo de entrada.