

## Especificación léxica

El análisis léxico consiste en dividir una cadena de caracteres en objetos llamados «tokens», que son las unidades mínimas de un lenguaje de programación con significado propio. En el fichero `ejercicio1` hay la implementación de un pequeño analizador léxico.

Ejecutando el código, el programa imprime por pantalla:

```
('ID', 'abc')
('SPACE', ' ')
('NUMBER', '123')
('SPACE', ' ')
('ID', 'cde')
('SPACE', ' ')
('NUMBER', '456')
```

- 1) *En muchos lenguajes de programación, los espacios y saltos de línea son ignorados. Modifique el fichero para que se ignoren estos tokens.*

Si se cambia la variable «text» de la siguiente manera

```
text = 'abc 123 ! cde 456'
```

El programa imprime el siguiente error:

```
('ID', 'abc')
('NUMBER', '123')
Traceback (most recent call last):
  File "ex1.py", line 29, in <module>
    for tok in tokenize(text):
  File "ex1.py", line 25, in tokenize
    raise SyntaxError('Bad char %r' % text[index])
SyntaxError: Bad char '!'
```

- 2) *Notesé que, tal como está programado, el analizador lanza una excepción y acaba su ejecución cuando aparece un carácter no previsto. Modifique el fichero para que imprima un mensaje cada vez que encuentra un carácter desconocido y continúe la ejecución.*