

Tabla de símbolos

```
# /src/lexer/lexer.py
import re

identifier = r'^[A-Za-z]+[0-9]*$'
decimal = r'^decimal$'
numero = r'^numero$'
palabra = r'^palabra$'

str_re = r'^\"[a-zA-Z0-9\\s]*\"$'
int_re = r'^[0-9]+$'
float_re = r'^[0-9]+\.[0-9]+$'

ope_re = r'^\+|\-|\*|\/|\%|\=|\=$'

regex_patterns = {
    identifier: 'identifíer',
    decimal: 'decimal',
    numero: 'numero',
    palabra: 'palabra',
    str_re: 'palabra',
    int_re: 'numero',
    float_re: 'decimal',
    ope_re: "
}

def match_token(token: str, prev_token: str):
    for pattern, data_type in regex_patterns.items():
        if re.match(pattern, token):
            return data_type
    return "

def tokenize(source_code: list[list[str]]):
    tokens = {}
    prev_token = None

    for line in source_code:
        for word in line:
            data_type = match_token(word, prev_token)
            tokens[word] = data_type
            prev_token = data_type

    return tokens

if __name__ == "__main__":
```

```
source_code = [  
    ["numero", "_A1", "=", "_A2", "+", "_A3", ";"],  
    ["palabra", "_Bueno1", ";"],  
    ["palabra", "_Malo1", ";"],  
    ["decimal", "_NumeroDecimal1", ";"],  
    ["palabra", "_oracion1", "=", "_Bueno1", "+", "_Malo1", ";"],  
    ["_A2", "=", "_A1", "-", "_A3", ";"],  
    ["decimal", "_A4", "=", "_A3", "/", "_A3", ";"],  
    ["palabra", "_oracion2", "=", "_Malo1", "+", "_Bueno1", ";"],  
]
```

```
tokens = tokenize(source_code)  
print(tokens)
```