

به نام خدا

استاد: دکتر سعید پارسا
درس اصول طراحی کامپایلر

نام: فاطمه زهرا بخشنده
شماره دانشجویی: 98522157

گزارش تمرین کلاسی 1:

Exercises

- Modify the code given in the previous slides to:
 1. Calculate the length of the tokens.
 2. Accept special characters such as '?' and '*' in tokens.
 3. Write the type name such as "identifier" instead of the numeric code.
 4. Add comments to the lexical rules.

کد این تمرین در آدرس gen\test\ test_v1_main.py قرار دارد. گرامر آن در grammars\ test.g4 است. فایل های lexer و مابقی فایل های ساخته شده از گرامر نیز در gen\test\v1_python موجود است.

1. برای قسمت اول، علاوه بر خود توکن، length هر توکن را نیز چاپ می کنیم.
2. برای قسمت دوم، گرامر را طوری تغییر داده که در نام توکن های identifier، علاوه بر حروف الفبا، کاراکترهای خاص مانند & و @ و \$ و ... را نیز بپذیرد.
3. برای قسمت سوم علاوه بر چاپ کردن token.type، testLexer.ruleNames[token.type - 1] را نیز چاپ می کنیم که نام rule آن است، مثلا identifier، یا plus.
4. برای قسمت چهارم گرامر مربوط به کامنت خطی و بلاک کامنت را به فایل گرامر اضافه می کنیم تا کامنت هارا نیز تشخیص دهد.

پس از این تغییرات یک input_string به آن می دهیم تا نتیجه را مشاهده کنیم.

```
input_string = """y& = a? + b% * c# / er$@ - // line comment  
/*block  
comment*/"""
```

21 ای توکن در خروجی تشخیص داد که نتیجه مربوط به چند تا از آن هارا اینجا میگذارم:

token 0:

token: y&

token length: 2

token type number: 8

token type rule name: IDENTIFIER

token 1:

token:

token length: 1

token type number: 12

token type rule name: Whitespace

token 2:

token: =

token length: 1

token type number: 7

token type rule name: ASSIGN

.

.

.

token 16:

token: er\$@

token length: 4

token type number: 8

token type rule name: IDENTIFIER

.
. .
.

token 20:

token: // line comment

token length: 15

token type number: 14

token type rule name: LineComment

token 21:

token: /*block

comment*/

token length: 17

token type number: 13

token type rule name: Comment