**TP compilateurs 1 :**

**Prise en main de PLY**

Programme mylex1.py

#file: mylex1.py

#usage: python mylex1.py prog1.txt

import ply.lex as lex

tokens = (

'NUMBER',

'PLUS',

'MINUS'

)

t\_PLUS = r'\+'

t\_MINUS= r'\-'

def t\_NUMBER(t):

r'\d+'

t.value = int(t.value)

return t

def **t\_newline**(t):

r'\n+'

t.lexer.lineno += len(t.value)

**t\_ignore** = ' \t'

def **t\_error**(t):

print ("Illegal character '%s'" % t.value[0])

t.lexer.skip(1)

lex.lex()

if \_\_name\_\_ == "\_\_main\_\_":

import sys

prog = open(sys.argv[1]).read()

lex.input(prog)

while 1:

tok = lex.token()

if not tok:

print ("NO TOKEN")

break

else:

print ("line %d: %s(%s)" % (tok.lineno, tok.type, tok.value))

file : Prog1.txt

10 + 5

+ 5 667

a (\* 10 )

**python mylex1.py prog1.txt**

line 1: NUMBER(10)

line 1: PLUS(+)

line 1: NUMBER(5)

line 2: PLUS(+)

line 2: NUMBER(5)

line 2: NUMBER(667)

Illegal character 'a'

Illegal character '('

Illegal character '\*'

line 3: NUMBER(10)

Illegal character ')'

NO TOKEN

File : Prog2.txt

10.3 + 5.44554

+ 5 667.02092

a \* 10 - (17 / 15)

**>python mylex1.py prog2.txt**

line 1: NUMBER(10)

Illegal character '.'

line 1: NUMBER(3)

line 1: PLUS(+)

line 1: NUMBER(5)

Illegal character '.'

line 1: NUMBER(44554)

line 2: PLUS(+)

line 2: NUMBER(5)

line 2: NUMBER(667)

Illegal character '.'

line 2: NUMBER(2092)

Illegal character 'a'

Illegal character '\*'

line 3: NUMBER(10)

line 3: MINUS(-)

Illegal character '('

line 3: NUMBER(17)

Illegal character '/'

line 3: NUMBER(15)

Illegal character ')'

NO TOKEN

Programme mylex2.py

#file: mylex2.py

#usage: python mylex2 prog2.py

import ply.lex as lex

tokens = (

'NUMBER',

'ADD\_OP',

'MUL\_OP'

)

**literals = '()'**

**t\_MUL\_OP** = r'/|\\*' #\* reservé dans ER -> \\*

def **t\_ADD\_OP**(t):

r'\+|-'

return t

def t\_NUMBER(t):

r'\d+(\.\d+)?'

**try**:

t.value = float(t.value)

**except ValueError**:

print ("Line %d: Problem while parsing %s!" % (t.lineno,t.value))

t.value=0

return t

def t\_newline(t):

r'\n+'

t.lexer.lineno += len(t.value)

t\_ignore = ' \t'

def t\_error(t):

print ("Illegal character '%s'" % t.value[0])

t.lexer.skip(1)

lex.lex()

if \_\_name\_\_ == "\_\_main\_\_":

import sys

prog = open(sys.argv[1]).read()

lex.input(prog)

while 1:

tok = lex.token()

if not tok: break

print ("line %d: %s(%s)" % (tok.lineno, tok.type, tok.value))

file : Prog1.txt

10 + 5

+ 5 667

a (\* 10 )

**>python mylex2.py prog1.txt**

line 1: NUMBER(10.0)

line 1: ADD\_OP(+)

line 1: NUMBER(5.0)

line 2: ADD\_OP(+)

line 2: NUMBER(5.0)

line 2: NUMBER(667.0)

Illegal character 'a'

line 3: **((()**

line 3: MUL\_OP(\*)

line 3: **NUMBER(10.0)**

line 3: **)())**

NO TOKEN

File : Prog2.txt

10.3 + 5.44554

+ 5 667.02092

a \* 10 - (17 / 15)

**>python mylex2.py prog2.txt**

line 1: NUMBER(10.3)

line 1: ADD\_OP(+)

line 1: NUMBER(5.44554)

line 2: ADD\_OP(+)

line 2: NUMBER(5.0)

line 2: NUMBER(667.02092)

Illegal character 'a'

line 3: MUL\_OP(\*)

line 3: NUMBER(10.0)

line 3: ADD\_OP(-)

line 3: ((()

line 3: NUMBER(17.0)

line 3: MUL\_OP(/)

line 3: NUMBER(15.0)

line 3: )())

NO TOKEN