22100409 Siaw Jia Yuin

Regular Expression:

identifiers: (a+$)($+_+a+.+1+E)*   [other] for delimiter

comment: [other] —(anything)*

assignment operator: [other] = [other]

keywords:
program+begin+int+if+print_line+end+else_if+end+while
[other]

number literal: 1*[other]

string literal: [other] " a* " [other]

punctuation - comma: [other] , [other]

statement terminator: [other];

less than operator: [other] < [other]

left parenthesis: [other] ([other]
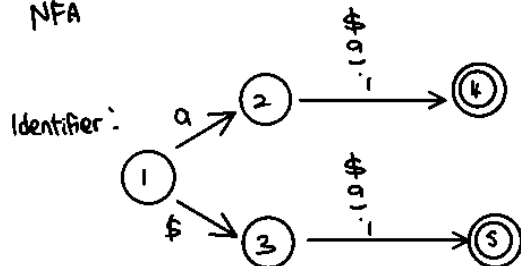
right parenthesis: [other]) [other]

multiplication operator: [other] * [other]
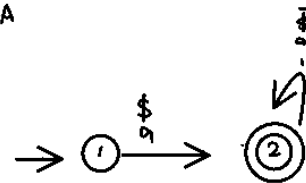
plus operator: [other] + [other]
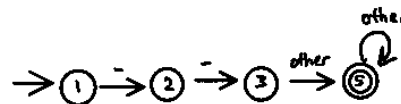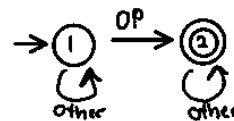
NFA and DFA
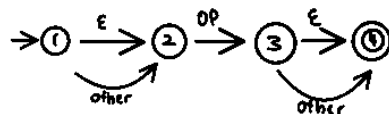
a : letters
1 : numbers

NFA

Identifier:

DFA

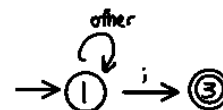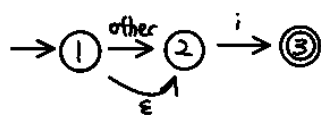comment :

left/right parenthesis, = + * < operators, comma ⇒ OP
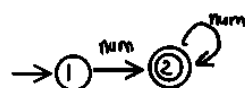

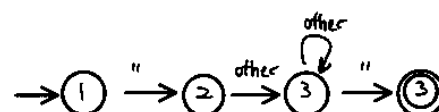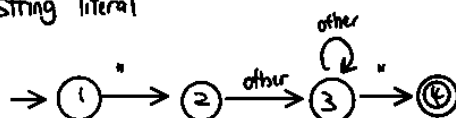
statement terminator

number literal :
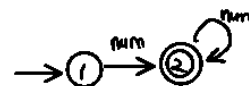
string literal

Output results:

input1.txt

```
program                 keyword
Exercise                identifier
begin                   keyword
-- comment1                         comment
int                     keyword
Time.10.24              identifier
=                       assignment operator
2206                    number literal
,                       punctuation - comma
Hour.In.Day            identifier
=                       assignment operator
0                       number literal
,                       punctuation - comma
$Minute.In.Hour        identifier
;                       statement terminator
if                      keyword
(                       left parenthesis
Time.10.24              identifier
<                       less than operator
10                      number literal
;                       statement terminator
begin                   keyword
print_line                          keyword
(                       left parenthesis
                 identifier
morning.               identifier
Hour.In.Day            identifier
=                       assignment operator
24                      number literal
;                       statement terminator
end                     keyword
else_if                 keyword
(                       left parenthesis
Time.10.24              identifier
<                       less than operator
20                      number literal
;                       statement terminator
begin                   keyword
$Minute.In.Hour        identifier
=                       assignment operator
Hour.In.Day            identifier
*                       multiplication operator
60                      number literal
;                       statement terminator
print_line                          keyword
(                       left parenthesis
                 identifier
day.             identifier
end                     keyword
else                    keyword
begin                   keyword
print_line                          keyword
(                       left parenthesis
                 identifier
evening.               identifier
end                     keyword
end                     keyword
```

Input1error.txt

```
program                 keyword
Exercise                identifier
3begin                  Illegal ID starting with digit
-- interpreted as an illegal identifier             comment
int                     keyword
Time.10.24              identifier
=                       assignment operator
2206                    number literal
,                       punctuation - comma
Hour.In.Day            identifier
=                       assignment operator
0                       number literal
,                       punctuation - comma
$Minute.In.Hour        identifier
;                       statement terminator
if                      keyword
(                       left parenthesis
Time.10.24              identifier
<                       less than operator
10                      number literal
;                       statement terminator
begin                   keyword
print_line                          keyword
(                       left parenthesis
                 identifier
morning.               identifier
Hour.In.Day            identifier
=                       assignment operator
24                      number literal
;                       statement terminator
end                     keyword
else_if                 keyword
(                       left parenthesis
Time.10.24              identifier
<                       less than operator
20                      number literal
;                       statement terminator
begin                   keyword
$Minute.In.Hour        identifier
=                       assignment operator
Hour.In.Day            identifier
*                       multiplication operator
60                      number literal
;                       statement terminator
print_line                          keyword
(                       left parenthesis
                 identifier
day.             identifier
end                     keyword
else$            identifier
-- interpreted as an identifier                 comment
begin                   keyword
print_line                          keyword
(                       left parenthesis
                 identifier
evening.               identifier
end                     keyword
end                     keyword
```

Input2.txt

```
program                 keyword
Exercise2               identifier
begin                   keyword
int                     keyword
i.Value.$               identifier
=                       assignment operator
0                       number literal
,                       punctuation — comma
$Minute.In.Hour         identifier
=                       assignment operator
1                       number literal
;                       statement terminator
-- comment first line                      comment
while                   keyword
(                       left parenthesis
i.Value.$               identifier
<                       less than operator
5                       number literal
;                       statement terminator
begin                   keyword
print_line                      keyword
(                       left parenthesis
i.Value.$               identifier
)                       right parenthesis
;                       statement terminator
i.Value.$               identifier
=                       assignment operator
i.Value.$               identifier
+                       plus operator
1                       number literal
;                       statement terminator
$Minute.In.Hour         identifier
=                       assignment operator
$Minute.In.Hour         identifier
+                       plus operator
10                      number literal
;                       statement terminator
end                     keyword
end                     keyword
```

Input2error.txt

```
<terminated> SmallLexer [Java Application] /Library/Java/JavaVirtualMachines/jdk-17.jdk/Contents/Home
program                 keyword
Exercise                identifier
3begin                  Illegal ID starting with digit
-- interpreted as an illegal identifier              comment
int                     keyword
@Time.10.24                     Illegal ID starting with wrong character
=                       assignment operator
2206                    number literal
,                       punctuation — comma
%Hour.In.Day                    Illegal ID starting with wrong character
=                       assignment operator
0                       number literal
,                       punctuation — comma
$Minute.In.Hour         identifier
;                       statement terminator
if                      keyword
(                       left parenthesis
Time.10.24              identifier
<                       less than operator
10                      number literal
;                       statement terminator
begin                   keyword
print_line                      keyword
(                       left parenthesis
                identifier
morning.                identifier
%Hour.In.Day                    Illegal ID starting with wrong character
=                       assignment operator
24                      number literal
;                       statement terminator
end                     keyword
^else_if                        Illegal ID starting with wrong character
(                       left parenthesis
@Time.10.24                     Illegal ID
<                       less than operator
20                      number literal
;                       statement terminator
*begin                  Illegal ID starting with wrong character
$Minute.In.Hour         identifier
=                       assignment operator
%Hour.In.Day                    Illegal ID starting with wrong character
*                       multiplication operator
60                      number literal
;                       statement terminator
print_line                      keyword
(                       left parenthesis
                identifier
day.            identifier
end                     keyword
else$           identifier
-- interpreted as an identifier              comment
begin                   keyword
print_line                      keyword
(                       left parenthesis
                identifier
evening.                identifier
end                     keyword
end                     keyword
```

Input3.txt

```
program                 keyword
Exercise2               identifier
begin                   keyword
int                     keyword
i.Value.$               identifier
=                       assignment operator
0,                      Illegal ID
$Minute.In.Hour         identifier
=                       assignment operator
1;                      Illegal ID
-- comment first line                   comment
while                   keyword
(                       left parenthesis
i.Value.$               identifier
<                       less than operator
5)                      Illegal ID
begin                   keyword
print_line                  keyword
(                       left parenthesis
i.Value.$               identifier
)                       right parenthesis
            identifier
i.Value.$               identifier
=                       assignment operator
i.Value.$               identifier
+                       plus operator
1                       number literal
;                       statement terminator
$Minute.In.Hour         identifier
=                       assignment operator
$Minute.In.Hour         identifier
+                       plus operator
10;                     Illegal ID
end                     keyword
end                     keyword
```

User's Manual:

To run small lexer program, just include the filename of the input text file as argument to the java program.