CS480
Translators

Introduction to Lexical Analysis
Chap. 3
Odds and Ends

• Assignment #2 is posted
• Demo your Assignment #1 next week

- not on Monday

5 pts of Assign 1

next Monday
Quiz #2

Question 1

• What is the language of the following CFG:
  \[ S \rightarrow bSbb \mid A \]
  \[ A \rightarrow aA \mid \epsilon \]
• Provide the parse tree for bbbaaabbba.

Question 2

• Provide the abstract syntax tree for the following:
  \[-1 + 2 \times 3.0^4.7 / 6\]
• What is the post-order traversal of the tree.
• Explain how you would implement this in gforth.
• How would the tree look if you embed the gforth?
Lexical Analysis

• What is its purpose?
• What is the difference in a token vs. lexeme?
• Example:
  – The Brown Fox
  – if (i !=32) then j := 12
• Are spaces important?

OSU Oregon State University
The Role of the Scanner...

Figure 3.1: Interactions between the lexical analyzer and the parser
while (i > 0)
  i = i - 2;

Mini-Translator

Source Language → Lexical Analyzer → Token Stream → Error Messages

Tokens:
- T_WHILE
- T_LPAREN
- T_IDENTIFIER
- T_LESSTHAN
- T_INTCONSTANT
- T_RPAREN
- T_IDENTIFIER
- T_EQUALS
- T_MINUS
- T_INTCONSTANT
- T_SEMICOLON
What’s new in this grammar?

\[
\begin{align*}
expr & \rightarrow expr + term \quad \{ \text{print('+')} \} \\
& \mid expr - term \quad \{ \text{print('-')} \} \\
& \mid term \\

term & \rightarrow term \ast factor \quad \{ \text{print('*')} \} \\
& \mid term / factor \quad \{ \text{print('/')} \} \\
& \mid factor \\
\end{align*}
\]

\[
\begin{align*}
factor & \rightarrow ( expr ) \\
& \mid \text{num} \quad \{ \text{print(num.value)} \} \\
& \mid \text{id} \quad \{ \text{print(id.lexeme)} \} \\
\end{align*}
\]

Figure 2.28: Actions for translating into postfix notation
The Scanner

```java
for (; ; peek = next input character ) {
    if ( peek is a blank or a tab ) do nothing;
    else if ( peek is a newline ) line = line+1;
    else break;
}
```

Figure 2.29: Skipping white space

- What is the purpose of line?
- What is the purpose of peek?