



Judson Santos Santiago

Aplicações do Flex

Compiladores

Introdução

- O Flex é um **gerador de analisador léxico**
 - Recebe especificações de tokens em formato de **expressões regulares**
 - Traduz as expressões regulares para **diagramas de transição**
 - Transforma os diagramas em **código**
- O **código C/C++** gerado é um analisador léxico
 - Usado para **fornecer tokens** para um **analisador sintático**
 - Pode ser usado de forma independente em **outras aplicações**

Números e Texto

- O Flex pode ser usado para identificar e classificar texto:

```
%{  
  // reconhecimento de números e texto  
  #include <iostream>  
  using std::cout;  
  %}  
  
%option noyywrap  
  
%%  
[ \t\n]+ ; // ignora espaços em branco  
[+-]?([0-9]+|([0-9]*\.[0-9]+)([eE][+-]?[0-9]+)?) cout << "número\n";  
.+ cout << "texto\n";  
%%  
  
...
```

Contagem de Palavras

- O Flex pode contar ocorrências de um padrão:

```
%{  
// faz contagem de caracteres, palavras e linhas de um arquivo  
#include <iostream>  
#include <fstream>  
using namespace std;  
unsigned charCount = 0, wordCount = 0, lineCount = 0;  
%}  
  
%option noyywrap  
word [^ \t\n]+  
eol  \n  
  
%%  
{word} { wordCount++; charCount += YYLeng(); }  
{eol}  { charCount++; lineCount++; }  
.  
%%
```


Múltiplos Arquivos

- O Flex processa **vários arquivos**
 - Temos utilizado uma opção para inibir esse comportamento

```
// processa apenas um arquivo
%option noyywrap
```
- Por padrão, a função **yylex()** chama **yywrap()**
 - A chamada é feita ao encontrar o **caractere EOF**
 - Podemos usar **yywrap** para **abrir outros arquivos**
 - **yylex** continua processando se **yywrap** retornar zero

Múltiplos Arquivos

- Funcionamento da função `yylex`:

```
int yylex()
{
    while (!the_end)
    {
        // processa entrada
        
        // fim de arquivo
        if (peek == EOF)
        {
            the_end = yywrap();
        }
    }
}
```

A função `yylex`
continua processando
se `yywrap` retornar
zero.

Estados

- O Flex pode definir e usar **estados**:
 - Estados são criados na seção de declarações:
 - %x - estados exclusivos
 - %s - estados não-exclusivos

```
%{  
#include <iostream>  
using std::cout;  
%}
```

```
%option noyywrap  
%x COMMENT  
%s LITERAL
```

```
%%
```

Estados

- Estados são ativados através da construção **BEGIN**
 - **INITIAL** é um estado pré-definido

```
%{  
#include <iostream>  
using std::cout;  
%}
```

```
%option noyywrap  
%x COMMENT
```

```
%%  
"/*"      BEGIN COMMENT;  // abertura de comentário  
<COMMENT>.*      ;        // comentário  
<COMMENT>"*/"    BEGIN INITIAL;  // fechamento de comentár  
io  
%%
```

...

Em um estado exclusivo,
apenas as regras marcados
com este estado são levadas
em conta.

Resumo

- O código gerado pelo Flex é um analisador léxico
 - Usado para fornecer tokens para um analisador sintático
 - Pode ser usado de forma independente em outras aplicações
 - É uma ferramenta poderosa para processamento de texto:
 - Tradutores
 - Contadores
 - Conversores
 - Classificadores
 - Etc.