

Compilador da linguagem Legol para linguagem C

Aluno: Leandro da Cruz Farias¹
Orientador: Prof. Dr. Raimundo Corrêa de Oliveira¹

Trabalho de Conclusão de Curso II - Engenharia de Computação

¹Escola Superior de Tecnologia - Universidade do Estado do Amazonas (UEA)

Dezembro - 2014

Sumário

- Introdução
 - Justificativa
 - Objetivos
 - Objetivo Geral
 - Objetivos Específicos
 - Metodologia
- Compilador
 - Estrutura de um compilador
 - Análise léxica
 - Gerador do analisador léxico
 - Análise sintática
 - Gerador do analisador sintático
 - Análise semântica
 - Tratamento de erros
 - Geração de código

Sumário

- Compilador da linguagem Legol
 - Analisador léxico
 - Regras léxicas
 - Geração do analisador léxico
 - Analisador sintático
 - Regras sintáticas
 - Geração do analisador sintático
 - Analisador semântico
 - Regras semânticas
 - Implementação do analisador semântico
 - Tratamento de erros
 - Geração de código C
 - Diferenças entre o Portugol e o Legol

Sumário

- Resultados
 - Testes
 - Comparação de tempo de execução
 - Resultados
- Conclusão
 - Dificuldades encontradas
 - Trabalhos futuros
 - Disciplinas aplicadas

Justificativa

- Importância da lógica de programação;
- Os problemas com a ferramenta VisuAlg.

Justificativa

- Importância da lógica de programação;
- Os problemas com a ferramenta VisuAlg.

Justificativa

```
algoritmo "semnome"  
// Função :  
// Autor :  
// Data : 19/12/2014  
// Seção de Declarações  
var  
a: inteiro  
b: inteiro  
inicio  
a <- 4  
b <- 5  
  
se (a > 6 e b > 7) entao  
    escreva("Sucesso.")  
fimse  
fimalgoritmo
```

	A	I	4
	B	I	5

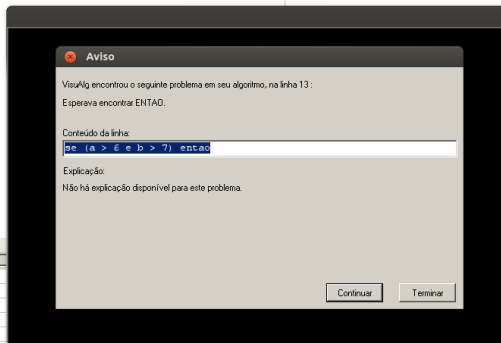


Figura 1: Mensagem de erro sem lógica.

Justificativa

```
algoritmo "Exemplo50"
var
resultado: logico

inicio
resultado <- ("1" + "e") = verdadeiro
escreva(resultado)
finalgoritmo
```

Nome	Tipo	Valor
RESULTADO	L	VERDADEIRO

Início da execução
VERDADEIRO
Fim da execução.

Figura 2: Erro de execução no VisuAlg.

Objetivo Geral

- Desenvolver um compilador que faça a tradução de algoritmos escritos em linguagem Legol para algoritmos em linguagem C.

Objetivos Específicos

- Identificar as palavras reservadas das principais estruturas de programação do Legol;
- Definir os padrões léxicos do Legol para gerar o analisador léxico;
- Definir as regras gramaticais do Legol para gerar o analisador sintático;
- Implementar as ações do analisador semântico;
- Implementar o tratamento de erros nas fases de análise;
- Implementar a geração de código escrito na linguagem alvo.

Objetivos Específicos

- Identificar as palavras reservadas das principais estruturas de programação do Legol;
- Definir os padrões léxicos do Legol para gerar o analisador léxico;
- Definir as regras gramaticais do Legol para gerar o analisador sintático;
- Implementar as ações do analisador semântico;
- Implementar o tratamento de erros nas fases de análise;
- Implementar a geração de código escrito na linguagem alvo.

Objetivos Específicos

- Identificar as palavras reservadas das principais estruturas de programação do Legol;
- Definir os padrões léxicos do Legol para gerar o analisador léxico;
- Definir as regras gramaticais do Legol para gerar o analisador sintático;
- Implementar as ações do analisador semântico;
- Implementar o tratamento de erros nas fases de análise;
- Implementar a geração de código escrito na linguagem alvo.

Objetivos Específicos

- Identificar as palavras reservadas das principais estruturas de programação do Legol;
- Definir os padrões léxicos do Legol para gerar o analisador léxico;
- Definir as regras gramaticais do Legol para gerar o analisador sintático;
- Implementar as ações do analisador semântico;
- Implementar o tratamento de erros nas fases de análise;
- Implementar a geração de código escrito na linguagem alvo.

Objetivos Específicos

- Identificar as palavras reservadas das principais estruturas de programação do Legol;
- Definir os padrões léxicos do Legol para gerar o analisador léxico;
- Definir as regras gramaticais do Legol para gerar o analisador sintático;
- Implementar as ações do analisador semântico;
- Implementar o tratamento de erros nas fases de análise;
- Implementar a geração de código escrito na linguagem alvo.

Objetivos Específicos

- Identificar as palavras reservadas das principais estruturas de programação do Legol;
- Definir os padrões léxicos do Legol para gerar o analisador léxico;
- Definir as regras gramaticais do Legol para gerar o analisador sintático;
- Implementar as ações do analisador semântico;
- Implementar o tratamento de erros nas fases de análise;
- Implementar a geração de código escrito na linguagem alvo.

Metodologia

- Modelo iterativo e incremental;
- Ferramentas:
 - Flex como gerador do analisador léxico;
 - Bison como gerador do analisador sintático.
 - Auxilia no tratamento de erros, análise semântica e geração de código.

Metodologia

- Modelo iterativo e incremental;
- Ferramentas:
 - Flex como gerador do analisador léxico;
 - Bison como gerador do analisador sintático.
 - Auxilia no tratamento de erros, análise semântica e geração de código.

Metodologia

- Modelo iterativo e incremental;
- Ferramentas:
 - Flex como gerador do analisador léxico;
 - Bison como gerador do analisador sintático.
 - Auxilia no tratamento de erros, análise semântica e geração de código.

Metodologia

- Modelo iterativo e incremental;
- Ferramentas:
 - Flex como gerador do analisador léxico;
 - Bison como gerador do analisador sintático.
 - Auxilia no tratamento de erros, análise semântica e geração de código.

Estrutura de um compilador

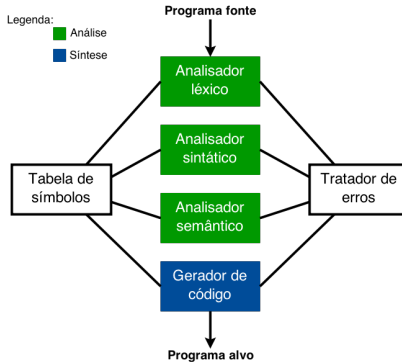


Figura 3: Fases de um compilador.

Gerador do analisador léxico

- O gerador de analisador léxico chamado Flex;
- Os padrões léxicos são definidos por meio de expressões regulares.
 - Exemplo: `STRING` `\ "[^\\ \"\n]*\"`

Gerador do analisador léxico

- O gerador de analisador léxico chamado Flex;
- Os padrões léxicos são definidos por meio de expressões regulares.
 - Exemplo: `STRING` `\ "[^\\ \"\n]*\"`

Gerador do analisador léxico

- O gerador de analisador léxico chamado Flex;
- Os padrões léxicos são definidos por meio de expressões regulares.
 - Exemplo: STRING `\ "[^\\ \"\n]*\"`

Gerador do analisador sintático

- O gerador de analisador sintático chamado Bison.

expressao -> expressao + digito | expressao - digito | digito

digito -> 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

Figura 6: Exemplo de gramática livre de contexto.

Análise semântica

- Verificação de tipos;
- Validação de atribuição de valores a uma variável;
- Índices de variáveis complexas;
- Declaração e uso de variáveis.

Análise semântica

- Verificação de tipos;
- Validação de atribuição de valores a uma variável;
- Índices de variáveis complexas;
- Declaração e uso de variáveis.

Análise semântica

- Verificação de tipos;
- Validação de atribuição de valores a uma variável;
- Índices de variáveis complexas;
- Declaração e uso de variáveis.

Análise semântica

- Verificação de tipos;
- Validação de atribuição de valores a uma variável;
- Índices de variáveis complexas;
- Declaração e uso de variáveis.

Tratamento de erros

- Mensagens de erro claras e que identifiquem onde o erro ocorreu;
- Erro léxico;
- Erro sintático;
- Erro semântico.

Tratamento de erros

- Mensagens de erro claras e que identifiquem onde o erro ocorreu;
- Erro léxico;
- Erro sintático;
- Erro semântico.

Tratamento de erros

- Mensagens de erro claras e que identifiquem onde o erro ocorreu;
- Erro léxico;
- Erro sintático;
- Erro semântico.

Tratamento de erros

- Mensagens de erro claras e que identifiquem onde o erro ocorreu;
- Erro léxico;
- Erro sintático;
- Erro semântico.

Geração de código

- Consulta à tabela de símbolos;
- Comandos semanticamente equivalentes aos que estavam no algoritmo de origem.

Geração de código

- Consulta à tabela de símbolos;
- Comandos semanticamente equivalentes aos que estavam no algoritmo de origem.

Regras léxicas

- Caracteres alfanuméricos e especiais;
- Uso de comentários;
- Todas as palavras reservadas com letras minúsculas;
- Variáveis devem iniciar com letra seguida por número, *underline* ou outra letra;
- *Case sensitive*;

Regras léxicas

- Caracteres alfanuméricos e especiais;
- Uso de comentários;
- Todas as palavras reservadas com letras minúsculas;
- Variáveis devem iniciar com letra seguida por número, *underline* ou outra letra;
- *Case sensitive*;

Regras léxicas

- Caracteres alfanuméricos e especiais;
- Uso de comentários;
- Todas as palavras reservadas com letras minúsculas;
- Variáveis devem iniciar com letra seguida por número, *underline* ou outra letra;
- *Case sensitive*;

Regras léxicas

- Caracteres alfanuméricos e especiais;
- Uso de comentários;
- Todas as palavras reservadas com letras minúsculas;
- Variáveis devem iniciar com letra seguida por número, *underline* ou outra letra;
- *Case sensitive*;

Regras léxicas

- Caracteres alfanuméricos e especiais;
- Uso de comentários;
- Todas as palavras reservadas com letras minúsculas;
- Variáveis devem iniciar com letra seguida por número, *underline* ou outra letra;
- *Case sensitive*;

Geração do analisador léxico

- Definições escritas em linguagem C;
- Declaração das expressões regulares correspondentes aos padrões léxicos;
- Regras das ações léxicas.

Geração do analisador léxico

- Definições escritas em linguagem C;
- Declaração das expressões regulares correspondentes aos padrões léxicos;
- Regras das ações léxicas.

Geração do analisador léxico

- Definições escritas em linguagem C;
- Declaração das expressões regulares correspondentes aos padrões léxicos;
- Regras das ações léxicas.

Regras sintáticas

```
algoritmo "Exemplo42"
var
a : inteiro
resultado : caractere

funcao verifica(a : inteiro) : caractere
inicio
    se (a > 0) entao
        retorne "Positivo."
    senao
        retorne "Negativo."
    fimse
fimfuncao

inicio
    escreva("Informe o valor a ser verificado: ")
    leia(a)
    resultado <- verifica(a)
    escreval(resultado)
finalgoritmo
```

Figura 7: Algoritmo escrito em Legol.

Geração do analisador sintático

- Definições e funções escritas em linguagem C.
- Declarações em linguagem nativa do Bison.
- Produções da gramática livre de contexto.
- Funções auxiliares escritas em linguagem C.

Geração do analisador sintático

- Definições e funções escritas em linguagem C.
- Declarações em linguagem nativa do Bison.
- Produções da gramática livre de contexto.
- Funções auxiliares escritas em linguagem C.

Geração do analisador sintático

- Definições e funções escritas em linguagem C.
- Declarações em linguagem nativa do Bison.
- Produções da gramática livre de contexto.
- Funções auxiliares escritas em linguagem C.

Geração do analisador sintático

- Definições e funções escritas em linguagem C.
- Declarações em linguagem nativa do Bison.
- Produções da gramática livre de contexto.
- Funções auxiliares escritas em linguagem C.

Regras semânticas

- Uso de variável de acordo com o seu escopo;
- Atribuição de valor a uma variável com tipo compatível a ele.
- Índice de variáveis complexas e variável de controle de estruturas de iteração são do tipo *inteiro*;
- Os comandos *se*, *enquanto* e *repita* devem conter expressões com valores lógicos;
- Assinaturas de subprogramas devem ser cumpridas de acordo com as suas respectivas declarações.

Regras semânticas

- Uso de variável de acordo com o seu escopo;
- Atribuição de valor a uma variável com tipo compatível a ele.
- Índice de variáveis complexas e variável de controle de estruturas de iteração são do tipo *inteiro*;
- Os comandos *se*, *enquanto* e *repita* devem conter expressões com valores lógicos;
- Assinaturas de subprogramas devem ser cumpridas de acordo com as suas respectivas declarações.

Regras semânticas

- Uso de variável de acordo com o seu escopo;
- Atribuição de valor a uma variável com tipo compatível a ele.
- Índice de variáveis complexas e variável de controle de estruturas de iteração são do tipo *inteiro*;
- Os comandos *se*, *enquanto* e *repita* devem conter expressões com valores lógicos;
- Assinaturas de subprogramas devem ser cumpridas de acordo com as suas respectivas declarações.

Regras semânticas

- Uso de variável de acordo com o seu escopo;
- Atribuição de valor a uma variável com tipo compatível a ele.
- Índice de variáveis complexas e variável de controle de estruturas de iteração são do tipo *inteiro*;
- Os comandos *se*, *enquanto* e *repita* devem conter expressões com valores lógicos;
- Assinaturas de subprogramas devem ser cumpridas de acordo com as suas respectivas declarações.

Regras semânticas

- Uso de variável de acordo com o seu escopo;
- Atribuição de valor a uma variável com tipo compatível a ele.
- Índice de variáveis complexas e variável de controle de estruturas de iteração são do tipo *inteiro*;
- Os comandos *se*, *enquanto* e *repita* devem conter expressões com valores lógicos;
- Assinaturas de subprogramas devem ser cumpridas de acordo com as suas respectivas declarações.

Implementação do analisador semântico

- Ação de cada produção da gramática livre de contexto;
- Funções em linguagem C para validar a consistência semântica dos comandos identificados.
- Consulta a uma tabela *hash*.

Implementação do analisador semântico

- Ação de cada produção da gramática livre de contexto;
- Funções em linguagem C para validar a consistência semântica dos comandos identificados.
- Consulta a uma tabela *hash*.

Implementação do analisador semântico

- Ação de cada produção da gramática livre de contexto;
- Funções em linguagem C para validar a consistência semântica dos comandos identificados.
- Consulta a uma tabela *hash*.

Tratamento de erros

- O caractere ponto (.), como expressão regular, identifica um erro léxico;
- Valor terminal error na gramática livre de contexto captura um erro sintático;
- Funções em linguagem C garantem mensagens de erro semântico.

Tratamento de erros

- O caractere ponto (.), como expressão regular, identifica um erro léxico;
- Valor terminal error na gramática livre de contexto captura um erro sintático;
- Funções em linguagem C garantem mensagens de erro semântico.

Tratamento de erros

- O caractere ponto (.), como expressão regular, identifica um erro léxico;
- Valor terminal error na gramática livre de contexto captura um erro sintático;
- Funções em linguagem C garantem mensagens de erro semântico.

Geração de código C

- Tabela *hash* é consultada para obter os atributos das variáveis, funções e procedimentos;
- Tipo lógico na linguagem C;
- Definição do tipo String;
- Comando *if* composto ao traduzir o comando *escolha*;

Geração de código C

- Tabela *hash* é consultada para obter os atributos das variáveis, funções e procedimentos;
- Tipo lógico na linguagem C;
- Definição do tipo String;
- Comando *if* composto ao traduzir o comando *escolha*;

Geração de código C

- Tabela *hash* é consultada para obter os atributos das variáveis, funções e procedimentos;
- Tipo lógico na linguagem C;
- Definição do tipo String;
- Comando *if* composto ao traduzir o comando *escolha*;

Geração de código C

- Tabela *hash* é consultada para obter os atributos das variáveis, funções e procedimentos;
- Tipo lógico na linguagem C;
- Definição do tipo String;
- Comando *if* composto ao traduzir o comando *escolha*;

Geração de código C

- Tabela *hash* é consultada para obter os atributos das variáveis, funções e procedimentos;
- Tipo lógico na linguagem C;
- Definição do tipo String;
- Comando *if* composto ao traduzir o comando *escolha*;

Geração de código C

```
algoritmo "Exemplo"
var
a:caractere
b:caractere
c:caractere

inicio
a <- "Compilador "
b <- "Legol"
c <- "Compilador C"

se ((a+b) > c) entao
    escreva("Concatenação é maior.")
fimse

finalgoritmo
```

Figura 8: Exemplo de operações com cadeias de caractere.

Geração de código C

- Uso de uma estrutura de dados do tipo fila;
- Após a confirmação da inexistência de erros, o algoritmo em linguagem C é escrito em um arquivo com extensão .c.

Geração de código C

- Uso de uma estrutura de dados do tipo fila;
- Após a confirmação da inexistência de erros, o algoritmo em linguagem C é escrito em um arquivo com extensão .c.

Diferenças entre o Portugol e o Legol

```
algoritmo "Portugol"  
var  
  valor1, valor2: inteiro  
  
inicio  
  valor1 <- 5  
  VALOR2 <- 5  
  escreva(VALOR1+vaLoR2)  
fimalgoritmo
```

```
algoritmo "Legol"  
var  
  valor1: inteiro  
  valor2: inteiro  
  
inicio  
  valor1 <- 5  
  valor2 <- 5  
  escreva(valor1+valor2)  
fimalgoritmo
```

Figura 9: Diferenças entre o Portugol e o Legol na declaração de variáveis e no uso delas no corpo do algoritmo.

Diferenças entre o Portugol e o Legol

```
algoritmo "Portugol"
var

procedimento imprimir_media(valor1,valor2:real; texto:caractere)
var

inicio
  escreva(texto, ((valor1+valor2)/2))
fimprocedimento

nota1, nota2:real
nome:caractere

inicio
  leia(nota1)
  leia(nota2)
  leia(nome)

  imprimir_media(nota1,nota2,nome)
finalgoritmo
```

```
algoritmo "Legol"
var
  nota1:real
  nota2:real
  nome:caractere

procedimento imprimir_media(valor1:real, valor2:real, texto:caractere)
inicio
  escreva(texto, ((valor1+valor2)/2))
fimprocedimento

inicio
  leia(nota1)
  leia(nota2)
  leia(nome)

  imprimir_media(nota1,nota2,nome)
finalgoritmo
```

Figura 10: Diferenças entre o Portugol e o Legol na declaração de parâmetros em subprogramas, na ordem de declarações de variáveis e subprogramas e no uso da palavra reservada *var*.

Diferenças entre o Portugol e o Legol

```
algoritmo "Portugol"
var
time:caractere

inicio
  leia(TIME)

  escolha (time)
    caso "Fluminense", "Botafogo"
      escreva("Time do Rio de Janeiro.")
  fimescolha
fimalgoritmo
```

```
algoritmo "Legol"
var
time:caractere

inicio
  leia(time)

  escolha (time)
    caso "Fluminense"
      escreva("Time do Rio de Janeiro.")
    caso "Botafogo"
      escreva("Time do Rio de Janeiro.")
  fimescolha
fimalgoritmo
```

Figura 11: Diferenças entre o Portugol e o Legol em relação à estrutura do comando *escolha*.

Testes

- Cinquenta algoritmos compilados pelo compilador Legol;
- Compilador GCC compilou os algoritmos em C.

Testes

- Cinquenta algoritmos compilados pelo compilador Legol;
- Compilador GCC compilou os algoritmos em C.

Comparação de tempo de execução

- Tempo de execução médio de um algoritmo no VisuAlg: 118,76 ms;
- Tempo de execução médio de um algoritmo compilado pelo compilador do Legol e pelo GCC: 1 ms.

Comparação de tempo de execução

- Tempo de execução médio de um algoritmo no VisuAlg: 118,76 ms;
- Tempo de execução médio de um algoritmo compilado pelo compilador do Legol e pelo GCC: 1 ms.

Resultados

- Artefatos:
 - Código fonte para gerar um analisador léxico com a ferramenta Flex;
 - Código fonte para gerar um analisador sintático com a ferramenta Bison, em que o analisador semântico, o tratamento de erros e a geração de código encontram-se implementados por meio de ações permitidas às produções da gramática livre de contexto.
 - Implementação de uma tabela *hash*.
 - Implementação de uma estrutura de dados do tipo fila.

Resultados

- Artefatos:
 - Código fonte para gerar um analisador léxico com a ferramenta Flex;
 - Código fonte para gerar um analisador sintático com a ferramenta Bison, em que o analisador semântico, o tratamento de erros e a geração de código encontram-se implementados por meio de ações permitidas às produções da gramática livre de contexto.
 - Implementação de uma tabela *hash*.
 - Implementação de uma estrutura de dados do tipo fila.

Resultados

- Artefatos:
 - Código fonte para gerar um analisador léxico com a ferramenta Flex;
 - Código fonte para gerar um analisador sintático com a ferramenta Bison, em que o analisador semântico, o tratamento de erros e a geração de código encontram-se implementados por meio de ações permitidas às produções da gramática livre de contexto.
 - Implementação de uma tabela *hash*.
 - Implementação de uma estrutura de dados do tipo fila.

Resultados

- Artefatos:
 - Código fonte para gerar um analisador léxico com a ferramenta Flex;
 - Código fonte para gerar um analisador sintático com a ferramenta Bison, em que o analisador semântico, o tratamento de erros e a geração de código encontram-se implementados por meio de ações permitidas às produções da gramática livre de contexto.
 - Implementação de uma tabela *hash*.
 - Implementação de uma estrutura de dados do tipo fila.

Dificuldades encontradas

- Ambiguidades;
- Mensagens de erro repetidas;
- Alocação de memória para cadeias de caractere.

Dificuldades encontradas

- Ambiguidades;
- Mensagens de erro repetidas;
- Alocação de memória para cadeias de caractere.

Dificuldades encontradas

- Ambiguidades;
- Mensagens de erro repetidas;
- Alocação de memória para cadeias de caractere.

Trabalhos futuros

- Construção de uma ferramenta que permita a edição de algoritmos em Legol e a compilação dos mesmos;
- Compilar os algoritmos em Legol para a linguagem Assembly;
- Permitir a depuração do código Legol pelo usuário.

Trabalhos futuros

- Construção de uma ferramenta que permita a edição de algoritmos em Legol e a compilação dos mesmos;
- Compilar os algoritmos em Legol para a linguagem Assembly;
- Permitir a depuração do código Legol pelo usuário.

Trabalhos futuros

- Construção de uma ferramenta que permita a edição de algoritmos em Legol e a compilação dos mesmos;
- Compilar os algoritmos em Legol para a linguagem Assembly;
- Permitir a depuração do código Legol pelo usuário.

Disciplinas aplicadas

- Linguagem de Programação I
- Linguagem de Programação II
- Estrutura de Dados
- Pesquisa e Ordenação de Dados
- Compiladores

Disciplinas aplicadas

- Linguagem de Programação I
- Linguagem de Programação II
- Estrutura de Dados
- Pesquisa e Ordenação de Dados
- Compiladores

Disciplinas aplicadas

- Linguagem de Programação I
- Linguagem de Programação II
- Estrutura de Dados
- Pesquisa e Ordenação de Dados
- Compiladores

Disciplinas aplicadas

- Linguagem de Programação I
- Linguagem de Programação II
- Estrutura de Dados
- Pesquisa e Ordenação de Dados
- Compiladores

Disciplinas aplicadas

- Linguagem de Programação I
- Linguagem de Programação II
- Estrutura de Dados
- Pesquisa e Ordenação de Dados
- Compiladores

Referências

- Aho, A. V., Lam, M. S., Sethi, R., and Ullman, J. D. (2008).
Compiladores: princípios, técnicas e ferramentas. Pearson
Addison-Wesley, second edition. Tradução Daniel Vieira; Revisão
técnica Mariza Bigonha.
- Levine, J. R. (2009). flex & bison. O'Reilly Media, first edition.