Linguagem de Definição e Geração de Analisadores Sintáticos em Semântica Denotacional Legível

José Leite da Silva Júnior

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Ciência da Computação
Março de 1993

FOLHA DE APROVAÇÃO

Linguagem de Definição e Geração de Analisadores Sintáticos em Semântica Denotacional Legível

JOSÉ LEITE DA SILVA JÚNIOR

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

PROF. ROBERTO DA SILVA BIGONHA - Orientador
DCC - ICEx - UFMG

PROF. JOÃO EDUARDO DE REZENDE DANTAS
DCC - ICEx - UFMG

PROF. CLARINDO ISAÍAS PEREIRA DA SILVA E PÁDUA
DCC - ICEx - UFMG

Belo Horizonte, 29 de março de 1993.

Agradecimentos

Deus não nos propõe senão enigmas
Feôdor Dostoiévski

À CAPES e à UFMG pelo apoio financeiro.

Ao meu orientador, prof. Roberto Bigonha, pelo constante incentivo, ajuda na solução de diversos problemas e orientação precisa.

Aos colegas do grupo de linguagens de programação — Lucas M. Amaral, Walace Rodrigues e Mariza Bigonha — pela ajuda em diversos problemas e pelas informações sempre muito valiosas.

Aos amigos de Belo Horizonte, Daniel, Eveline, Geraldo, Higino, José Américo, Luis Henrique, Márcio e Vivien, pela companhia e apoio.

Aos *hermanos* de Fortaleza, Alfredo, Antônio Leite, Beatriz, Catarina, Edite, Ênio, Erivaldo, Galileu, Herbert, Mário, Paulo, Sobral, Tibério, pelos conselhos, amizade, cartas, telefonemas (mesmo às 2:00 h), cerveja e biscoitos.

A meus pais, Isauda e José Leite, minhas irmãs, Márcia e Nadja, e meus irmãos, Cláudio, Augusto e Marcos, pelo paitrocínio, carinho, presença, amor e saudade.

Sinopse

Este trabalho apresenta uma nova abordagem à definição formal de linguagens de programação. Partindo do fato que os principais problemas encontrados em definições formais são sua baixa legibilidade e dificuldade de compreensão, foi examinada a adequação da aplicação das técnicas de *literate programming* a semântica denotacional.

O resultado da análise de *literate programming* no contexto de semântica formal foi a definição da linguagem de especificação LDS. Essa linguagem une prosa informal e equações semânticas numa mesma estrutura, promovendo um estilo de composição de definições onde a ênfase se encontra na estruturação e apresentação das definições.

A existência de dois modos de exposição (prosa e equações) torna mais evidente o duplo caráter (as duas audiências) das descrições em LDS: processamento pelo computador para obtenção de automática de um compilador e a leitura por pessoas para a comunicação de idéias sobre uma linguagem de programação. Seguindo o modelo de *literate programming*, dois programas foram desenvolvidos para dirigir as definições à audiência desejada.

A segunda parte da dissertação trata da primeira fase do processo de geração automática de compiladores em LDS, i.e., a produção dos analisadores léxico e sintático a partir da definição sintática da linguagem. O algoritmo de geração de analisadores e sua implementação são descritos, juntamente com o esquema de compactação de tabelas de análise usado nos compiladores produzidos. Comentários sobre a linguagem de especificação sintática e sua interação com o algoritmo de geração também são apresentados.

Abstract

This work presents a new approach to the formal definition of programming languages. One assumption underlies the proposed approach: the main problems of formal definitions are their low legibility and difficult of comprehension. To solve these problems, it was verified the suitability of applying the techniques of literate programming to denotational semantics.

The result of the analysis of literate programming in a denotational semantics context was the design of the specification language LDS. This language associates in the same structure informal prose and semantic equations, promoting a style of definition composition where the emphasis is in the definition structuring and presentation.

The two modes of exposition (prose and equations) make more evident the two objectives (two audiences) of LDS descriptions: computer processing, to the automatic generation of compilers, and human reading, to communicate ideas about a programming language. Following the literate programming model, two programs were implemented to direct the definitions to the desired audience.

The second part of this dissertation deals with the first phase of automatic generation of compilers, i.e., the production of lexical analyzers and parsers from the syntatic definition. The algorithm to construct the analyzers and its implementation are described, along with the method for compactation of parsing tables used in the analyzers produced. Comments about the syntatic specification language and its interation with the algorithm are also presented.

Sumário