

Projeto (3º ano de LCC)

VMS, Simular Visual para a máquina de Stack Virtual VM
Relatório de Desenvolvimento

Adriano Campinho
(a79032)

Vasco Leitão
(a79220)

1 de Julho de 2018

Resumo

Este documento apresenta o relatório desenvolvido no âmbito da Unidade Curricular de Projecto, do curso de Licenciatura em Ciências da Computação, Universidade do Minho, no segundo semestre do ano letivo 2017/2018. Pretendendo-se explicitar o processo de conceção e implementação do mesmo, assim como as escolhas tomadas e as pretendidas resoluções para cada problema exposto no enunciado fornecido, prestando especial atenção ao modo de utilização do simulador;

Supervisores: Pedro Rangel Henriques e José João
Área: Processamento de Linguagens

Conteúdo

1	Introdução	2
2	Análise e Especificação	3
2.1	Descrição Informal da Máquina	3
2.2	Planeamento e objetivos	4
2.3	Especificação do Requisitos	4
2.3.1	Parser	5
2.3.2	Processador	5
2.3.3	Interface Gráfica	5
3	Concepção/desenho da Resolução	6
3.1	Organização do código	6
3.1.1	Estruturas	6
3.2	Fases de Desenvolvimento	7
3.2.1	Parser	7
3.2.2	Processador	7
3.2.3	Interface	7
4	Guia de Utilização e Testes	8
4.1	Modo de utilização	8
4.2	Mensagens de erro	8
4.3	Alternativas, Decisões e Problemas de Implementação	9
4.4	Testes realizados e Resultados	9
5	Conclusão	10

Capítulo 1

Introdução

Nos últimos anos, temos usado nas disciplinas de Processamento de Linguagens (e Compiladores) a máquina virtual VM, uma máquina de stack que suporta inteiros, reais e strings e tem uma heap para permitir um uso dinâmico da memória (necessário por exemplo para implementar listas ligadas). A VM é programada em Assembly e que disponibiliza um conjunto de instruções máquina mínimo e muito compreensível, o que a torna pedagogicamente muito relevante como máquina objeto (destino) em tarefas de compilação.

Para que o uso da VM seja realmente um bom instrumento de trabalho nestes cursos, é fundamental que exista um simulador que permita testar (de preferência passo a passo) o código gerado. Atualmente existe, como é de conhecimento de todos, um Assembler para traduzir o Assembly produzido numa lista de códigos máquina e existe um interpretado que executa esse código e que fornece uma interface visual que permite acompanhar a execução verificando a evolução dos vários blocos de memória e dos registo de controlo da máquina.

Esta versão, codificada em Java, está funcional e pode ser usada, mas tem alguns problemas de implementação, nomeadamente o elevado tempo de execução, quando os programas aumentam um pouco.

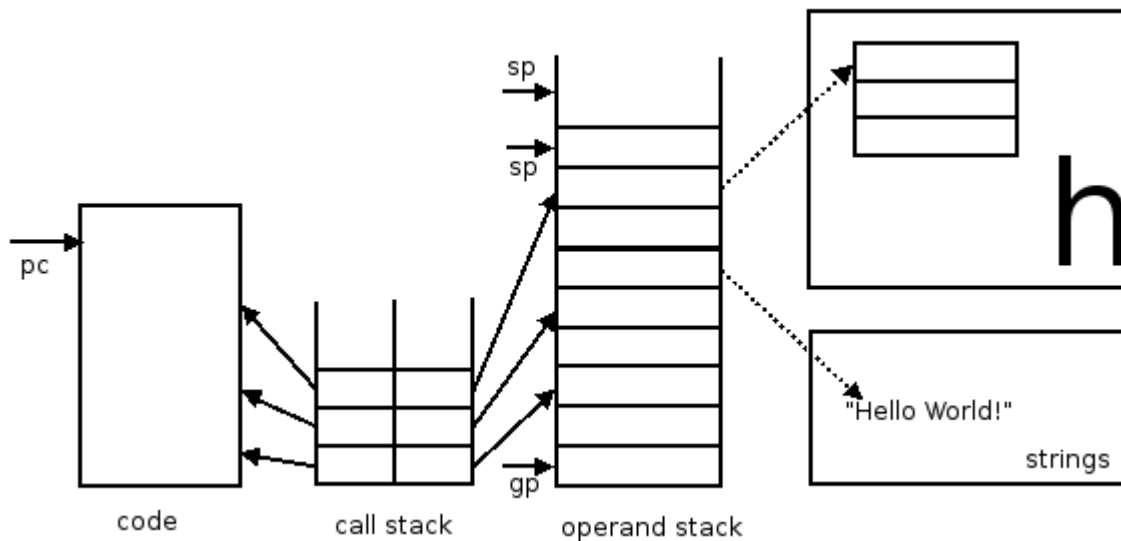
Neste projeto pretende-se que os alunos reconstruam o simulador com interface interativo e visual para a máquina VM desde o início, em Java ou em C (conforme preferência do grupo) produzindo um executável que evite os problemas atuais.

Capítulo 2

Análise e Especificação

2.1 Descrição Informal da Máquina

Trata-se duma máquina de pilhas (por oposição às máquinas de registos), esta é composta duma pilha de execução, duma pilha de chamadas, duma zona de código e de duas heaps:



Code - Esta pilha contém o conjunto das instruções do programa, esta é preenchida no início da execução do programa, mantendo-se igual do princípio ao fim da execução do mesmo;

Operation Stack - Esta pilha contém valores guardados no registo da máquina, estes são todos indexados, podendo ser do tipo inteiro, real e endereço;

Heap - Esta pilha contém cadeias de caracteres (strings), cada uma destas referenciada por endereços, cada bloco, contendo um certo número de caracteres;

Call Stack - Esta pilha contém blocos estruturados, cada um destes referenciado por endereços, cada bloco estruturado contém um certo número de valores (estes são do mesmo tipo dos que se podem encontrar na Operation Stack);

Um endereço pode apontar para quatro tipos de informação: para código, para a pilha, para um bloco estruturado ou para uma string. Ao longo da execução do programa são guardados os seguintes quatro registos:

- **SP** (Stack Pointer) o registo aponta para o topo corrente da pilha/ para a primeira célula livre da pilha.
- **FP** (Frame Pointer) o registo aponta para o endereço de base das variáveis locais.
- **GP** (Global Pointer) o registo contém o endereço de base das variáveis globais.
- **PC** (Program Counter) o registo aponta para a instrução corrente (da zona de código) por executar.

A pilha de chamada permite guardar as chamadas: contém pares de apontadores (i, f). O endereço i guarda o registo de instrução pc e f o registo fp.

Instruções

As instruções são designadas por um nome e podem aceitar um ou dois parâmetros. Estes parâmetros podem ser dos seguintes tipos:

- Constantes inteiras,
- Constantes reais,
- Cadeias de caracteres delimitadas por aspas. Estas cadeias de caracteres seguem as mesmas regras de formatação que as cadeias da linguagem C (em particular no que diz respeito aos caracteres especiais como ; ou \\\),
- Uma etiqueta simbólica designando uma zona no código.

O conjunto das instruções pode ser encontrado no documento original fornecido com as fontes da maquina virtual aqui descrita, destas foram ignoradas as primitivas gráficas, não sendo esta uma implementação de uma versão gráfica da maquina virtual;

2.2 Planeamento e objetivos

O simulador deve ser implementado de forma a funcionar num maior numero de ambientes possível: deve ser possível correr a simulação de vários modos possibilitando uma simulação interativa de cada programa. Existindo a possibilidade do jogador interagir com a simulação através da linha de comandos ou através de uma Interface gráfica a ser desenvolvida.

Podemos então dividir o projecto em 3 grandes grupos:

- **Parser:** O parser é responsável por percorrer o input do sistema e devolve os tokens, que correspondem a qualquer instrução, isto é, qualquer conjunto de símbolos que não incluam espaçamentos e mudanças de linha. Por fim, o parser filtra também o texto por forma a remover qualquer ocorrência de espaçamentos e mudanças de linhas. Assim, o parser é responsável por processar o input para que este fique num estado que possa ser digerido pelo simulador, permitindo assim estabelecer uma maior diversidade de formatos de input que, mesmo não sendo totalmente aceites pelo processador, são mais completos e possuem uma componente estética maior.
- **Simulador:** um programa que através do input recebido do parser e, se necessário, do input do utilizador, faça a simulação em si do programa;
- **Interface Gráfica:** uma interface que permita ao utilizador interagir com a execução do simulador, e dando-lhe acesso a informação sobre o estado atual da maquina; Cada vez que o utilizador interatue com a interface (e.g. clicando num botão), esta deve ser atualizada, refletindo as alterações feitas no estado na maquina e na interface.

2.3 Especificação do Requisitos

Esta secção de forma mais especifica os requisitos de cada um dos grupos de desenvolvimento demonstrando os problemas que surgiram durante cada fase:

2.3.1 Parser

O parser deve aceitar programas com a seguinte sintaxe:

```
<code>          ::= <instr>*
<instr>         ::= <ident> :
                  | <instr_atom>
                  | <instr_int> <integer>
                  | pushf <float>
                  | (pushs | err) <string>
                  | check <integer> , <integer>
                  | (jump | jz | pusha) <ident>
<instr_atom>    ::= add | sub | mul | div | mod | not | inf | infeq | sup |
                  | supeq | fadd | fsub | fmul | fdiv | fcos | fsin | |
                  | finf | finfeq | fsup | fsupeq | concat | equal | atoi | atof |
                  | itof | ftoi | stri | strf |
                  | pushsp | pushfp | pushgp | loadn | storen | swap |
                  | writei | writef | writes | read | call | return |
                  | start | nop | stop | allocn | free | dupn | popn
<instr_int>     ::= pushi | pushn | pushg | pushl | load |
                  | dup | pop | storel | storeg | alloc
```

Ainda, o parser devera ignorar tudo o que sirva para embelezar o código do programa (sejam espaços e mais, linhas em branco etc.), assim como devera ser possível a utilização de comentários que deverão ser ignorados pelo mesmo.

2.3.2 Processador

O processador devera receber o programa tratado pelo Parser e executar as instruções produzindo o resultado esperado (descrito no documento referido acima), permitindo ao utilizador controlar a execução do programa;

2.3.3 Interface Gráfica

A interface deve permitir ao utilizador executar as seguintes ações:

- Realizar 1 passo na execução do programa;
- Realizar n passos (passado como argumento) na execução do programa;
- Realizar a execução do programa ate ao fim de uma vez só;
- Carregar um programa para ser simulado;
- Recarregar o programa que esta atualmente a ser simulado;
- Em qualquer momento da execução do programa, mostrar o estado actual da maquina (stacks e pointers);
- Permitir ao utilizador introduzir input para o programa quando necessário;
- Permitir ao utilizador introduzir múltiplas linhas de input de uma vez;
- Carregar um ficheiro com o input para ser lido por um programa;

Capítulo 3

Concepção/desenho da Resolução

3.1 Organização do código

Em termos de organização do código, este foi dividido nos seguintes ficheiros:

- **README** - Este ficheiro contém uma lista de packages que necessitam de ser instalados para o correto funcionamento do simulador;
- **structs/** - Pasta que contém os módulos das estruturas de dados criadas;
- **vmsMan.1** - Este ficheiro é relativo a man page do simulador, este é acessível após a instalação da vms;
- **lex.l** - Ficheiro lex referente ao parser, que contém todas as instruções e convenções necessárias;
- **semantic.c** - Coisas
- **syntax.y** - Coisas
- **interface.c** - Coisas
- **makefile** - Coisas
- **semantic.h** - Coisas
- **vms.c** - Coisas

3.1.1 Estruturas

Em termos de estruturas criadas, estas foram divididas nos seguintes grupos (sendo que cada grupo representa um par de ficheiros .h e .c):

- **array** - Coisas
- **callStack** - Coisas
- **code** - Coisas
- **heap** - Coisas
- **opStack** - Coisas
- **types** - Coisas

3.2 Fases de Desenvolvimento

3.2.1 Parser

Os programas

Convenções lexicais

<digit> ::= 0-9

<alpha> ::= a-z | A-Z

<ident> ::= ...

<comentario> ::= #.* | //.*

<integer> ::= -? <digit>+

<float> ::= -? <digit>+ (.<digit>*?)

3.2.2 Processador

Estruturas de dados etc.

3.2.3 Interface

Não faço ideia

Capítulo 4

Guia de Utilização e Testes

4.1 Modo de utilização

Esta máquina destina-se a uma utilização interactiva do utilizador, para visualizar a execução de um programa. A execução poderá ser conduzida passo a passo, ou ser feita de uma só vez. Sendo feita da seguinte forma:

vms [opção] ficheiro.vm OU vms opção [ficheiro.vm]

O projeto desenvolvido, tem 3 opções de utilização, estas são:

Máquina em Modo Silencioso (sem flag de opção) O programa passado como argumento, e executado de início ao fim, apenas permitindo ao utilizador introduzir input quando pedido;

Máquina em Modo Debug (-d) O modo Debug permite ao utilizador o acesso a mais informação sobre o programa, disponibilizando os seguintes comandos, que permitem uma execução interativa do programa:

- **run:** executa o programa ate ao fim;
- **next 'n':** executa 'n' passos;
- **file 'f':** "carrega"o ficheiro 'f' para ser executado;
- **reload:** recarrega o ficheiro no "local"do ultimo ficheiro executado;
- **quit:** termina a execução do simulador;

Máquina em Modo Interface (-g) O modo Interface

4.2 Mensagens de erro

- **Error: Index out of array -** Provocado por um acesso indevido a zona de código, seja acesso as duas heaps ou a pilha;
- **Error: fp == sp -** Provocado quando ...
- **Error: Opening file -** Provocado quando ocorreu um erro a abrir um ficheiro;
- **Error: You must specify a '-d' or '-g' option, or a program file'man vms' for more information.-** Provocado quando e iniciado o simulador com as opções/flags incorrectas;

- **Error: Invalid type (write / not / equal / padd second arg not int / padd first argument not adress / alocn / free / atox arg is not heap adress / itof / ftoi / stri / strf / loadn / load / dupn / popn / pon / storen / jz / call)** - Provocado quando o(s) valores retirados da pilha não tem o tipo esperado;
- **Segmentation Fault** - Provocado por um acesso indevido a zona de código, as duas heaps ou a pilha;
- **Error *mensagem*** - Provocado por uma instrução err;

4.3 Alternativas, Decisões e Problemas de Implementação

4.4 Testes realizados e Resultados

Mostram-se a seguir alguns testes feitos (valores introduzidos) e os respectivos resultados obtidos:

Capítulo 5

Conclusão

Neste relatório descreveu-se o processo de conceção e implementação do trabalho no âmbito da Unidade Curricular de Projecto, do curso de Licenciatura em Ciências da Computação, Universidade do Minho, no segundo semestre do ano letivo 2017/2018. pedido pelo enunciado, nomeadamente o parser, o simulador e a interface gráfica, além de descrever em detalhe o funcionamento da máquina referida.

Os requisitos impostos pelo enunciado a que este relatório é referente são cumpridos pelos programas desenvolvidos, os quais são apresentados neste relatório. Foram também implementadas várias funcionalidades adicionais.

Apesar de estarmos satisfeitos com o trabalho desenvolvido, há ainda espaço para muitas melhorias desta implementação. Enumeram-se algumas ideias que, por falta de tempo, não foram (ainda) concretizadas:

- Criar um conjunto grande de ficheiros de testes, que sirvam tanto para provar/verificar a implementação da máquina, assim como servir de exemplo da correcta instalação da mesma;
- Fazer melhorias no espaçamento da interface gráfica, permitindo a utilização mais fácil em janelas mais pequenas;

Bibliografia