Pontifícia Universidade Católica
do Rio de Janeiro

Hisham Hashem Muhammad

Estudo sobre APIs de linguagens de script

PUC-Rio - Certificação Digital Nº 0410837/CB

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós–graduação em Informática do Departamento de Informática da PUC–Rio

Orientador: Prof. Roberto Ierusalmschy

Rio de Janeiro
Agosto de 2006

Pontifícia Universidade Católica
do Rio de Janeiro

Hisham Hashem Muhammad

Estudo sobre APIs de linguagens de script

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós–graduação em Informática do Departamento de Informática do Centro Técnico Científico da PUC–Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Prof. Roberto Ierusalmschy
Orientador
Departamento de Informática — PUC–Rio

Prof. Noemi de la Rocque Rodriguez
Departamento de Informática — PUC–Rio

Prof. Renato Fontoura de Gusmão Cerqueira
Departamento de Informática — PUC–Rio

Prof. Luiz Henrique de Figueiredo
IMPA

Prof. José Eugenio Leal
Coordenador Setorial do Centro Técnico Científico — PUC–Rio

Rio de Janeiro, 28 de Agosto de 2006

PUC-Rio - Certificação Digital Nº 0410837/CB

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

Hisham Hashem Muhammad

PUC-Rio - Certificação Digital Nº 0410837/CB

Ficha Catalográfica

Muhammad, Hisham Hashem

Estudo sobre APIs de linguagens de script / Hisham Hashem Muhammad; orientador: Roberto Ierusalmschy. — Rio de Janeiro : PUC–Rio, Departamento de Informática, 2006.

v., 123 f: il. ; 29,7 cm

1. Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui referências bibliográficas.

1. Informática – Tese. 2. Linguagens de Programação. 3. Linguagens de Script. 4. Interfaces para Programação de Aplicações. I. Ierusalmschy, Roberto. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

Agradecimentos

Agradeço ao meu orientador, Prof. Roberto Ierusalimschy, pela confiança, e aos professores do Departamento de Informática com quem tive a oportunidade de estudar, pelo aprendizado.

Ao CNPq e à PUC–Rio, pelos auxílios concedidos.

Ao amigo Alexandre Wolf pela grande ajuda quando da minha chegada no Rio.

Aos colegas da PUC, em especial ao pessoal do LabLua – Fabio, Sérgio e Alexandra – que fizeram deste um ambiente ótimo de se trabalhar, e ao Leonardo pela parceria na reta final de noites em claro enquanto terminávamos nossas dissertações.

Aos colegas de apartamento, Renato e Sabrina, pela acolhida quando do meu retorno dos Estados Unidos e pelo bom ano e meio em que moramos juntos.

Ao Sport Club Internacional, pela inesquecível campanha da conquista da América, que me fez sentir mais próximo de casa durante esse ano.

E principalmente aos que mesmo longe estiveram sempre perto.

PUC-Rio - Certificação Digital Nº 0410837/CB

Resumo

Muhammad, Hisham Hashem; Ierusalmschy, Roberto. **Estudo sobre APIs de linguagens de script**. Rio de Janeiro, 2006. 123p. Dissertação de Mestrado — Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Um cenário comum atualmente é o de aplicações desenvolvidas usando duas linguagens de programação a fim de otimizar partes onde o desempenho é crítico e permitir extensibilidade através de *scripts* escritos pelo usuário. Há várias formas de se obter esse tipo de interoperabilidade; idealmente, entretanto, uma linguagem deve prover uma interface de acesso externo (*foreign language interface*, FLI) que permita ao programador receber e enviar tanto chamadas como dados para outra linguagem.

Este trabalho discute as principais questões envolvendo o projeto de APIs para integração de ambientes de execução de linguagens em aplicações C. Apresentamos os principais problemas enfrentados na interação entre código executando em um ambiente com características inerentemente dinâmicas como o de uma linguagem de script com código C. Comparamos aqui as abordagens empregadas por cinco linguagens no tratamento da comunicação entre os espaços de dados de C e do ambiente de execução embutido e as conseqüências destas abordagens no gerenciamento de memória, bem como no compartilhamento de código entre a aplicação C e o da linguagem de script.

Ilustramos as diferenças das APIs destas linguagens e o impacto destas no código resultante de uma aplicação C através de um estudo de caso. Diferentes linguagens de script são embutidas como plugins de uma mesma biblioteca, que por sua vez expõe a aplicações clientes uma API genérica de scripting. Assim, o código de cada plugin permite observar de forma clara e isolada os procedimentos adotados em cada linguagem para chamada de funções, registro de funções C e conversão de dados entre os ambientes.

Palavras–chave

Linguagens de Programação. Linguagens de Script. Interfaces para Programação de Aplicações.

PUC-Rio - Certificação Digital Nº 0410837/CB

PUC-Rio - Certificação Digital Nº 0410837/CB

Abstract

Muhammad, Hisham Hashem; Ierusalmschy, Roberto. **A study on APIs for scripting languages**. Rio de Janeiro, 2006. 123p. MsC Thesis — Department of Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Applications written in two programming languages, in order to optimize parts where performance is critical or to obtain extensibility through user-written scripts, are commonplace nowadays. There are several ways to obtain this kind of interoperability; ideally, however, a language should provide a foreign language interface (FLI), allowing the programmer to send and receive both data and function calls to the external language.

This work discusses the main issues involving the design of APIs for the integration of language environments within C applications. We present the main problems faced in the interaction between code executed in an environment with inherently dynamic characteristics such as a scripting language and C code. We compare the approaches employed by five languages when handling communication between the data spaces of C and the embedded runtime environment and the consequences of these approaches in memory management, as well as sharing of code between the C application and that from the scripting language.

We illustrate the differences of the APIs of those languages and their impact in the resulting code of a C application through a case study. Different scripting languages were embedded as plugins for a library, which on its turn exposes to client applications a generic scripting API. This way, the code of each plugin allows us to observe in a clear and isolated way the procedures adopted by each language for function calls, registration of C functions and conversion of data between the environments.

Keywords

Programming Languages. Scripting Languages. Application Programming Interfaces.

Sumário

1 Introdução **10**
1.1 Objetivo 11
1.2 Estrutura do texto 12

2 Interação entre linguagens de programação **13**
2.1 Tradução de código 13
2.2 Compartilhamento de máquinas virtuais 14
2.3 Modelos de objetos independentes de linguagem 15
2.4 C como linguagem intermediária 16
2.5 Interfaces com C 17
2.6 Linguagens de script 19

3 APIs de linguagens de script **24**
3.1 Transferência de dados 24
3.2 Coleta de lixo 47
3.3 Chamada de funções a partir de C 60
3.4 Registro de funções C 71

4 Estudo de caso: LibScript **83**
4.1 LibScript 83
4.2 Implementação dos plugins 91
4.3 Conclusões 109

5 Conclusões **111**

Referências Bibliográficas **115**

A API de LibScript **120**
A.1 Inicialização e Término 120
A.2 Registro de Funções 120
A.3 Buffer de parâmetros 120
A.4 Executando Código 121
A.5 API Exportada por Plugins 122

PUC-Rio - Certificação Digital Nº 0410837/CB

Lista de figuras

3.1 Classe Java contendo um método implementado externamente 38
3.2 Código C implementando um método Java 39
3.3 Acesso possivelmente inválido a uma referência a um `PyObject` em código C 49
3.4 Rotina para concatenar os elementos de um array representados como strings. 55

4.1 Visão geral da arquitetura de LibScript 84
4.2 Ambientes virtuais em LibScript 85
4.3 API para execução de código em LibScript 87

PUC-Rio - Certificação Digital Nº 0410837/CB

PUC-Rio - Certificação Digital Nº 0410837/CB

"Dreams do come true."

Anthony Quinn, *Grauman's Chinese Theater, 1968.*