

Instituto Politécnico de Beja

Escola Superior de Tecnologia e Gestão



C Extensions to Collecting Data About Universities

Programming Languages

Bruno Moreira



2011

Docente: José Jasnau Caeiro

Aluno: Bruno Alexandre da Silva Moreira

Número: 6170

Curso: Eng^a Informática 2º ano



ÍNDICE

Resumo.....	3
Programa.....	4
Definição geral de funcionalidades	4
Extracção de listagens	4
Extracção de Estatísticas	4
Extracção de ficheiros	4
Medição do tempo de execução de cada função	4
Descrição de ficheiros	5
rebides.c	5
functions.c, functions.h	5
rebides.i	6
rebides_wrap.c	6
MakeFile	6
main.py	6
Conclusão	7
Bibliografia	8
Anexos	9
Estatísticas.....	9
Listagem	9
Listagem2	10

RESUMO

O trabalho tem como objectivo o tratamento de dados relativos a docentes e instituições entre os anos 2000 e 2009 utilizando a linguagem C.

Os dados acima mencionados foram obtidos no trabalho anterior utilizando para tal a linguagem Python. Estes dados estão disponíveis online na base de dados rebides.

Os dados acima mencionados, foram carregados em base de dados SQLite3.

O tratamento dos dados, consiste em efectuar listagens e estatísticas, desenvolvendo para tal um módulo em C com a aplicação SWIG, utilizando as bibliotecas sqlite3 e Python C API, de forma que este possa ser executado a partir de um qualquer programa em linguagem Python.

PROGRAMA

DEFINIÇÃO GERAL DE FUNCIONALIDADES

Neste trabalho foi-nos solicitado que o programa disponibilizasse algumas funcionalidades, nomeadamente:

EXTRACÇÃO DE LISTAGENS

Foi-nos solicitado que, com base nos dados acima mencionados, fossem disponibilizadas algumas listagens.

EXTRACÇÃO DE ESTATÍSTICAS

Foi-nos solicitado que, com base nos dados acima mencionados, fossem disponibilizadas algumas estatísticas.

EXTRACÇÃO DE FICHEIROS

Foi-nos também solicitado que as estatísticas pudessem ser gravadas num único ficheiro.

O referido ficheiro tem o seguinte formato:

```
*****
*  título
*****
[ano]      [coluna1]      [filhos_coluna1]
                        [filhos_filhos_coluna1]
                        ...
[ano_seguinte]
...
```

MEDIÇÃO DO TEMPO DE EXECUÇÃO DE CADA FUNÇÃO

O tempo de execução de cada função deverá ser medido utilizando para tal a ferramenta gprof.

DESCRIÇÃO DE FICHEIROS

REBIDES.C

Este é o ficheiro principal, sendo que este tem as funções que geram os ficheiros com base nas consultas à base de dados. Todas as funções responsáveis por essas consultas encontram-se no ficheiro functions.c.

FUNCTIONS.C, FUNCTIONS.H

Estes ficheiros contêm as funções que geram toda a informação, quer na consulta à base de dados, construção de listagens e escrita em ficheiros.

Abaixo é descrita a função de cada uma delas:

- void append_string_to_file(char* pString);
 - escreve uma string no final de um ficheiro
- void write_list_to_file(PyObject* pList);
 - escreve uma listagem num ficheiro
- int counterCallback(void **pCounter, int argc, char **argv, char **azColName);
 - trata um valor numérico vindo de uma consulta sql
- int statsListCallback(void **pList, int argc, char **argv, char **azColName);
 - trata uma linha de uma consulta sql, sendo que o último valor é numérico
- int listCallback(void **pList, int argc, char **argv, char **azColName);
 - trata uma linha de uma consulta sql, sendo que todas as colunas são strings
- int intcmp(int p1, int p2);
 - compara 2 valores interiores para decidir qual dos dois é maior
- int poStringCompare(PyObject ** p1, PyObject ** p2);
 - compara cada coluna de dois tuplos entre si alfabeticamente
- int poFinalizedIntCompare(PyObject ** p1, PyObject ** p2);
 - compara cada coluna de dois tuplos entre si alfabeticamente, comparando a última coluna numericamente
- void qSortPyList(PyObject *pList, function f);
 - organiza duas listas com o algoritmo qsort
- void queryDataBase(char* pQuery, function f, void* pArgument);
 - consulta a base de dados com base em queries sql
- PyObject* count_teachers(int iYear);
 - conta o número de professores por ano
- PyObject* count_teachers_per_establishment(int iYear);
 - conta o número de professores por estabelecimento por ano
- PyObject* count_teachers_per_degree(int iYear);
 - conta o número de professores por grau
- PyObject* count_teachers_per_degree_establishment(int iYear);
 - conta o numero de professores por grau por estabelecimento por ano
- PyObject* list_establishments_per_year(int iYear);



- gera uma listagem de estabelecimentos por ano
- PyObject* list_teachers_per_degree_year(int iYear);
 - gera uma listagem de professores por grau
- PyObject* list_teachers_leaving_institution_year(int iYear);
 - gera uma listagem de professores que saem de uma instituição por ano
- PyObject* list_new_teachers_institution_year(int iYear);
 - gera uma listagem de novos professores por instituição por ano
- void generate_list_header(char * cTitle);
 - gera o cabeçalho das listagens
- void generate_year_header(int iYear);
 - gera o cabeçalho do ano

REBIDES.I

Este ficheiro contém as funções que ficam disponíveis para serem usadas no python.

REBIDES_WRAP.C

Este ficheiro é gerado automaticamente pelo compilador SWIG e contém a transformação das funções designadas em rebides.i por forma a que estas possam ser interpretadas pelo Python.

MAKEFILE

Este ficheiro permite de uma forma fácil e estruturada compilar o programa sem ter que, em primeiro lugar escrever tudo sempre que há essa necessidade. E em segundo lugar compilar apenas os ficheiros que necessitam de compilação, não desperdiçando assim tempo a compilar o que já se encontra compilado.

MAIN.PY

Este ficheiro de Python chama as funções compiladas em C, para demonstrar o seu funcionamento.

CONCLUSÃO

Achei este trabalho muito interessante, pois nunca tinha experimentado relacionar funções de duas linguagens completamente distintas.

A utilização do MakeFile foi uma alegre surpresa, pois não só acelerou o desenvolvimento, como me deu a conhecer um novo método de compilação que desconhecia até à data.

Achei no entanto que a medição dos tempos gasto por cada função utilizando a ferramenta gprof, desadequada no meu caso, pois como utilizei PyObjects para todas as funcionalidades, e tendo em conta que o gprof não interpreta estes objectos, a sua utilização torna-se muito trabalhosa, necessitando de uma reprogramação quase por completo. Razão pela qual não foi implementada neste projecto.

Penso no entanto que os conhecimentos adquiridos durante este trabalho poderão facilmente ser utilizados em futuros desenvolvimentos, quer no meio académico quer no meio profissional.

BIBLIOGRAFIA

- [i] ActiveState, QuickSort PyObjects, 2002
<http://code.activestate.com/recipes/65210-implementing-c-function-callbacks-using-python/history/2/>
- [ii] Python/C API Reference Manual, Abril de 2011
<http://docs.python.org/c-api/concrete.html>
- [iii] QLite3 Quickstart example, 2011
<http://www.sqlite.org/quickstart.html>

ANEXOS

ESTATÍSTICAS

```
*****
total number of teachers in the higher
*****
[2000]
    31879
[2001]
    28441
[2002]
    26210
[2003]
    24202
[2004]
    22992
[2005]
    21717
[2006]
    19932
[2007]
    18629
[2008]
    17713
[2009]
    16888
```

Ilustração 1 – Listagem de número total de docentes por ano

LISTAGEM

```
*****
list of establishments per year
*****
[2000]
    Academia Militar
    Academia Nacional Superior de Orquestra
    Academia da Força Aérea
    Academia da Força Aérea - Escola Superior de Tecnologias Militares Aeronáuticas
    Conservatório Superior de Música de Gaia
    DINENSINO-Ensino, Desenvolvimento e Cooperação, CRL (Beja)
    DINENSINO-Ensino, Desenvolvimento e Cooperação, CRL (Setúbal)
    Escola Naval
    Escola Náutica Infante D. Henrique
    Escola Superior Artística do Porto
    Escola Superior Artística do Porto (Guimarães)
    Escola Superior Gallaecia
    Escola Superior Politécnica do Exército
    Escola Superior de Actividades Imobiliárias
    Escola Superior de Artes Decorativas
    Escola Superior de Artes e Design
    Escola Superior de Design
    Escola Superior de Educadores de Infância Maria Ulrich
    Escola Superior de Educação Jean Piaget - Nordeste
    Escola Superior de Educação Jean Piaget de Almada
    Escola Superior de Educação Jean Piaget de Arcozelo
    Escola Superior de Educação Jean Piaget de Arcozelo (Viseu)
    Escola Superior de Educação de Almeida Garrett
    Escola Superior de Educação de Fafe
```

Ilustração 2 – Listagem de estabelecimentos por ano

LISTAGEM2

```
*****
list of new personnel in the institution per year
*****
[2000]
    Academia Militar
        Alberto Jorge Crispim Gomes
        Alexandre Almeida Mendes Borga
        Alfredo Rui Francisco do C. Gonçalves Pereira
        Américo Fernando Carreira Martins
        Américo Fernando Levita Fernandes
        Ana Bela Ribeiro da Costa Santos Bravo
        Aniceto Henrique Afonso
        António Carlos Raínha Perry da Câmara
        António Gervásio Martins de Almeida Leite
        António José Barreiro Telo
        António Lourenço Guedes
        António Luciano Leite Videira
        António Luciano de Oliveira Alves Vieira
        António Maia Farinha Cadete
        António Manuel Nunes Cavaca
```

Ilustração 3 – Listagem de novos docentes na instituição por ano