

BÀI TẬP THỰC HÀNH 06: BẢNG BĂM

1 Nội dung

Tập tin "MST.txt" là file chứa thông tin của 5000 Công ty ở Việt Nam. Tập tin này có một phần nội dung như sau:

```
1 Ten cong ty|MST|Địa chỉ
2 CONG TY TNHH BEE VIET NAM|0108927262|Số 8 - K8, Khu nhà ở liên kết trung tâm 75, Tổng cục II, Bộ Quốc Phòng, thôn Lai Xá, Xã Kim Chung, Huyện
Hoài Đức, Thành phố Hà Nội
3 CONG TY CỔ PHẦN THƯƠNG MẠI CHAU DUC PHAT|3502406778|Số 266 Ấp Phước Trung, Xã Tam Phước, Huyện Long Điền, Tỉnh Bà Rịa - Vũng Tàu
4 CONG TY CỔ PHẦN XÂY DỰNG ĐẦU TƯ PHÁT TRIỂN DI SẢN SAO VIỆT|0315938079|30/18 Trường Sa, Phường 17, Quận Bình Thạnh, Thành phố Hồ Chí Minh
5 CONG TY TNHH MTV THANH TÂN HOANG LONG|0315938103|2/47 Dương Thanh Lộc 31, Khu Phố 3C, Phường Thanh Lộc, Quận 12, Thành phố Hồ Chí Minh
6 CONG TY TNHH NÔNG NGHIỆP CÔNG NGHỆ CAO MIỀN ĐÔNG VIỆT|3401194911|Thôn 5, Xã Tân Phúc, Huyện Ham Tân, Tỉnh Bình Thuận
7 CONG TY TNHH XÂY DỰNG VINH GIA PHAT|3603671412|Số 171, Xóm 4, Khu 2, Ấp Bau Ca, Xã Trung Hòa, Huyện Trang Bom, Tỉnh Đồng Nai
8 CONG TY TNHH THƯƠNG MẠI DỊCH VỤ PHU LONG RIVERSIDE|3401194929|243 Huỳnh Thúc Khang, KP1, Phường Mũi Né, Thành phố Phan Thiết, Tỉnh Bình Thuận
9 CONG TY TNHH HANH TRANG PHAT|3603671155|To 5, Ấp Thanh Bình, Xã Lộc An, Huyện Long Thành, Tỉnh Đồng Nai
10 CONG TY TNHH THƯƠNG MẠI DỊCH VỤ THIẾT BỊ M.K.K|0315932380|154/1/34 Công Lộ, Phường 15, Quận Tân Bình, Thành phố Hồ Chí Minh
11 CONG TY TNHH TM - DV - NHÀ HÀNG HẢI SẢN KÝ QUANG|0315933352|Số 526 Dương Phạm Văn Đông, Phường 13, Quận Bình Thạnh, Thành phố Hồ Chí Minh
12 CONG TY TNHH MTV KIM LONG|1201613551|Số 170 Nguyễn Minh Dương, Ấp 1, Xã Đạo Thạnh, Thành phố Mỹ Tho, Tỉnh Tiền Giang
13 CONG TY TNHH SONA AGENCY VIET NAM|0108926660|Số 333 Bạch Mai, Phường Bạch Mai, Quận Hai Bà Trưng, Thành phố Hà Nội
```

Trong đó:

- Dòng đầu tiên thể hiện tên các trường thông tin có trong tập tin.
- Những dòng tiếp theo thể hiện thông tin của một công ty, bao gồm Tên công ty, mã số thuế của công ty và địa chỉ. Mỗi trường thông tin cách nhau bởi một dấu gạch thẳng (|).

Sinh viên được yêu cầu đọc thông tin của các Công ty từ file trên vào cấu trúc dữ liệu *Company*, sau đó lưu trữ dưới dạng bảng băm.

2 Thực hành

Cho struct *Company* được định nghĩa như sau:

```
struct Company
{
    string name;
    string profit_tax;
    string address;
};
```

Yêu cầu

Sinh viên hãy định nghĩa các hàm sau:

- Hàm đọc dữ liệu từ tập tin:
 - `vector<Company> ReadListCompany(string file_name)`
 - Đầu vào: `file_name` là tên tập tin đầu vào (trong bài tập thực hành này là `MST.txt`).
 - Đầu ra: danh sách các công ty đọc từ tập tin được lưu trong biến có kiểu dữ liệu `vector<Company>`.
- Hàm băm một chuỗi thành một số với chuỗi là tên của công ty:
 - `int HashString(string company_name)`
 - Đầu vào: `company_name` có kiểu dữ liệu `string` là tên của công ty muốn băm thành số (kiểu dữ liệu số nguyên).

- Đầu ra: là giá trị của chuỗi đầu vào đã được băm thành kiểu dữ liệu số nguyên.
- Công thức băm:

$$hash(s) = \sum_{i=0}^{n-1} ((s[i] \times p^i) \bmod m)$$

Trong đó:

- + `s` là 20 ký tự cuối cùng của chuỗi tên đầu vào `s = company_name.substr(company_name.length() - 20)`.
Nếu chuỗi tên đầu vào ít hơn 20 ký tự thì lấy toàn bộ.
- + `s[i]` là mã ASCII của ký tự tại vị trí thứ `i` trong chuỗi `s`.
- + `p = 31`
- + `m = 109 + 9`
- Ví dụ: - Đầu vào: "CTDLGT"
- 20 ký tự cuối cùng của chuỗi: "CTDLGT" (trong trường hợp này không đủ 20 ký tự nên lấy toàn bộ chuỗi).
- Đầu ra: 472750792

3. Hàm tạo một bảng băm từ danh sách các công ty (`vector<Company>`), với số lượng phần tử của bảng băm là 5000:

- `Company* CreateHashTable(vector<Company> list_company)`
- Đầu vào: `list_company` là danh sách các công ty được đọc từ tập tin.
- Đầu ra: bảng băm được tạo từ danh sách đầu vào.
- Gợi ý: sử dụng hàm `HashString` đã định nghĩa.

4. Hàm thêm thông tin 1 công ty vào bảng băm đã có:

- `void Insert(Company* hash_table, Company company)`
- Đầu vào: - `hash_table`: bảng băm hiện tại.
- `company` là công ty muốn thêm vào bảng băm.

5. Hàm tìm kiếm thông tin của 1 công ty trong bảng băm thông qua tên công ty:

- `Company Search(Company* hash_table, string company_name)`
- Đầu vào: - `hash_table` là bảng băm (được lưu trong mảng).
- `company_name` là tên công ty muốn tìm kiếm thông tin.
- Output: công ty tìm được từ bảng băm, có kiểu dữ liệu `Company`.

3 Quy định nộp bài

- Sinh viên nộp toàn bộ mã nguồn liên quan thông qua tập tin `MSSV.zip` hoặc `MSSV.rar`.
- Mỗi phần cần được đặt trong thư mục riêng. Tất cả nằm trong thư mục `MSSV` (Lưu ý: chỉ nộp file `.h` và `.cpp`).
- Các bài nộp sai quy định sẽ bị 0 điểm.
- Các bài làm giống nhau sẽ bị 0 điểm môn học.