

# BÀI BÁO CÁO THUẬT TOÁN DSA – WEEK 6

***Nguyễn Đình Trung Kiên – 24120195***

## **1. readCompanyList:**

- Sử dụng một string tạm thời để đọc từng dòng từ file với *ifstream*.
- Từ dòng vừa đọc được và lưu vào string *tmpRead*, ta dùng stringstream để đưa string đó thành một luồng và xử lý.
- Ta tách luồng từ stream được đưa vào thành các string con và lưu vào các phần thông tin của công ty bằng cách sử dụng một biến *Company* tạm thời.
- Sau khi đã xử lý xong một dòng (thông tin của một công ty) và đưa vào biến *Company* thì chèn *Company* đó vào vector kết quả.
- Tiếp tục quá trình trên đến khi không còn dòng nào trong file, ta trả về vector các công ty đã được lấy thông tin.

## **2. hashString:**

- Ta sử dụng vòng lặp xuất phát từ 0 với điều kiện dừng là biến chạy bé hơn chiều dài xâu (trường hợp chiều dài xâu không vượt quá 20) và bé hơn 20 (chiều dài xâu từ 20 ký tự trở lên). Khi đó biến chạy sẽ dừng vòng lặp khi đã duyệt đủ 20 ký tự cuối cùng (đối với xâu có từ 20 ký tự trở lên) hoặc đã duyệt hết toàn bộ xâu (không đủ 20 ký tự).
- Ở mỗi lần duyệt, ta biến đổi từ biến chạy *i* (xuất phát từ 0) để duyệt lần lượt được các ký tự trong số 20 ký tự cuối cùng của xâu bằng  $i + len - 20$  (với *len* là chiều dài xâu). Ta tính vị trí duyệt tới bằng phép tính trên nếu  $len \geq 20$  thì mới cập nhật đến đó, nếu không thì ta xét ngay tại vị trí là *i* vì khi đó chiều dài xâu sẽ không vượt quá 20.
- Ở mỗi bước duyệt, ta tính giá trị Hash theo công thức, tuy nhiên cần dùng phép chia lấy dư ở mỗi bước (đặc biệt là phép tính  $31^i$  vì số tính được tương đối lớn).
- Khi đã duyệt xong thì trả về kết quả (vẫn cần dùng phép modulo với 2000 thêm lần nữa vì ta đã modulo từng phần tử bên trong tổng này rồi).

## **3. insert:**

- Từ tên của công ty cần chèn, ta tìm vị trí chèn công ty trong bảng băm dựa vào hàm *hashString*.

- Nếu tại vị trí tìm được chưa có thông tin, ta chèn trực tiếp thông tin vào vị trí đó.
- Nếu vị trí tìm được đã có thông tin, ta đã gặp phải collision và sẽ sử dụng phương pháp Linear Probing để giải quyết, từ vị trí tìm được ta duyệt tiếp tục cả mảng để tìm ra một vị trí trống đầu tiên rồi chèn thông tin vào đó.

#### **4. createHashTable:**

- Ta duyệt các phần tử của vector các công ty đã được truyền vào, ở mỗi công ty ta sử dụng thuật toán chèn ở hàm *insert* để chèn thông tin của công ty đó vào bảng băm.
- Duyệt hết tất cả các công ty trong vector thì ta sẽ có được bảng băm hoàn chỉnh và trả về con trỏ đến bảng băm đó.

#### **5. search:**

- Từ tên của công ty được đưa vào, ta sử dụng hàm hash để tìm ra vị trí của công ty đó nếu đã được chèn vào.
- Tại vị trí tính được, nếu ta tìm thấy tên công ty cần tìm tại đó thì trả về con trỏ đến công ty đó trong hash table.
- Nếu tại vị trí đó không có công ty nào thì lập tức trả về NULL vì không tồn tại công ty đó trong hash table.
- Nếu tại vị trí đó đã có thông tin của một công ty nào đó và không phải công ty có tên cần tìm, ta tiếp tục duyệt xuôi theo mảng đến khi tìm thấy hoặc gặp một vị trí trống:
  - + Nếu tìm thấy thì trả về con trỏ đến công ty đó.
  - + Nếu không tìm thấy (gặp một ô trống trước khi tìm thấy) thì trả về con trỏ NULL. Bởi khi dùng hàm *insert* để thêm các công ty vào Hash Table thì ta luôn chắc chắn được rằng giữa vị trí cuối cùng của công ty với vị trí ban đầu tính được sẽ không có ô trống nào, vì ta luôn tìm ô trống để đặt công ty vào hash table.

#### **6. Hàm bổ sung (printCompany):**

- Sử dụng để in thông tin của một công ty theo mẫu.

**MINH CHỨNG BÀI LÀM ĐÃ ĐƯỢC UPLOAD LÊN GITHUB:**