

Báo cáo ngắn về website tra từ điển

1. Frameword

Trang web được xây dựng trên Django là một frameword phát triển website, Django được phát hành năm 2005. Django giúp phát triển ứng dụng web theo mô hình MTV (model-template-views).

2. Front-end

Giao diện chủ yếu viết bằng bootstrap,html,css, javascript.

3. Các chức năng

Chức năng tìm kiếm từ, thêm từ, xóa từ.

4. Phương thức lấy request (lấy input người nhập)

Sử dụng phương thức post để lấy từ người dùng nhập vào và trạng thái, thêm tìm kiếm hoặc xóa.

5. Cấu trúc dữ liệu được sử dụng

Sử dụng cấu trúc dữ liệu Trie để tìm kiếm, xóa, thêm trong từ điển.

Data là một file txt từ vựng và nghĩa được cách nhau bởi dấu “:”.

```
go:di  
sleep:ngu  
eat:an  
president:chu tịch
```

Mỗi node trên cây Trie chứa các thông tin như sau char(kí tự hiện tại của nó), children[] (là một mảng chứa các nút con), index (đây là giá trị để sử dụng seek trong python, để tìm kiếm nghĩa của từ, đồng thời đó là dấu hiệu xem phải là kết thúc một từ hay không)

Hàm thêm từ vào cây Trie: Bắt đầu tại nút gốc là root, Ban đầu sẽ duyệt từng từ của từ cần thêm. Giả sử ta duyệt tới ký tự chr, kiểm tra xem giá trị char của các node con tại node đang được duyệt tới có char bằng với chr không. Nếu có thì ta sẽ đệ quy xuống node con đó ngược lại ta sẽ thêm node để thêm ký tự chr này vào. Cứ tiếp tục như vậy cho đến khi hết ký tự trong từ cần thêm.

Hàm tìm kiếm: bắt đầu tìm kiếm tại nút gốc, Ban đầu duyệt từng ký tự của từ cần tìm kiếm từ trái sang phải, giả sử đó là ký tự chr, nếu ký tự chr trùng

khớp bất kì giá trị char của nút con trong node đang duyệt thì ta sẽ đệ quy vào trong nút con đó. Cứ tiếp tục tìm kiếm cho đến khi duyệt hết các kí tự trong từ cần tìm, hoặc không thể đệ quy xuống tiếp. Thì sẽ kết thúc phần tìm kiếm.

Hàm `delete_word`: Bắt đầu duyệt từ node root. Duyệt từ trái sang phải của từ cần xóa. Giả sử duyệt tới ký tự `chr`, nếu giá trị char của node con tại node đang duyệt bằng `chr` thì tiếp tục đệ quy để duyệt. Ngược lại từ đó không có trong từ điển ta sẽ dừng lại. Tiếp tục duyệt đến khi hết các ký tự trong từ cần xóa. Kiểm tra xem node đó có phải là kết thúc một từ hay không (kiểm tra `index` khác `-1` hay không) nếu có thì sẽ thay đổi lại giá trị `index=-1`.

Sau khi thêm từ vào cây ta sử dụng thư viện `pickle.dumps` để lưu lại cấu trúc cây, khi cần thiết sẽ sử dụng lệnh `pickle.load` để đọc lại và sử dụng (không tốn thời gian build lại cây). Tuy nhiên vì đang làm trên web nên em đã setup biến môi trường, nếu cây đã build sẽ không build lại. File pickle tạo ra để lưu trữ lại quá trình chạy website ta thêm sửa như thế nào trên cây này. Khi cần có thể load và sử dụng lại.

6. Cách cài đặt và sử dụng:

- Máy cần có python 3.
- Cài đặt framework django qua lệnh `pip3 install django-admin` (chú ý tùy thuộc vào biến môi trường , `pip3` là `pip` ở python3 phân biệt với python2)
- Cd vào thư mục dic nơi chứa thư mục (`manage.py`) để khởi động web qua lệnh : `Python3 manage.py runserver`
- Chú ý: Mọi warning về cú pháp không ảnh hưởng tới việc chạy chương trình.

Vd 

Đây là warning về việc import do máy không nhận được file. Tuy nhiên không ảnh hưởng lúc khởi động server.

- Cuối cùng truy cập vào trang web , tùy vào máy sẽ cấp phát địa chỉ riêng. Thông thường sẽ là <http://127.0.0.1:8000/>
- Dưới đây là địa chỉ cấp phát để truy cập server.

```
System check identified no issues (0 silenced).  
November 07, 2019 - 16:55:27  
Django version 2.2.6, using settings 'dic.settings'  
Starting development server at http://127.0.0.1:8000/  
Quit the server with CTRL-BREAK.
```

- Tham khảo chi tiết tại github:
- https://github.com/tks1998/build_dictionary
- Các file cài đặt thuật toán trong thư mục /dictionary/dic/Find_dic gồm file process.py chứa cây Trie, file data.txt chứa data, views.py nhận request từ server.