2020春《大数据高级数据结构设计与实践》作业二 主讲教师: 史建焘

姓名: 许天骁 学号: 1171000405 班号: 1703109

Tire树

1. 简介

英文字典树每一个内部结点都有26个子结点,树的高度为最长字符串长度一棵子树代表具有相同前缀的关键码的集合。例如"an"子树代表具有相同前缀an-的关键码集合{and, ant}

- 2. 实现原理
 - 1. 插入

首先根据插入纪录的关键码找到需要插入的结点位置。如果该结点是叶结点,那么就将为其分裂出两个子结点,分别存储这个纪录和以前的那个纪录。如果是内部结点,则在那个分支上应该是空的,所以直接为该分支建立一个新的叶结点即可。

2. 查找

对待查询的的字符与树中的子节点进行比较,第一个字符与树根相比,如果不为空,则查询树根的所有子树并且和第二个字符相比。以此类推,如果进行到某一层,所有字符均已被匹配,且当前层存在标记,则查询成功。若在任意一层对应位置的字符的子树为空或者,最后一层不存在标记,则查询失败。

3. 删除

根据插入纪录的关键码找到需要删除的结点位置。如果一个被删除结点的父结点没有其他的儿子,那么就需要合并。否则只需要将此分支设置为空即可。

实验实现

此次实验,使用了个人完成的tire树进行。

1. 实验任务1

```
D:\LEARNING\CODES\DSJSJJG\work2\project1.exe

code
exist
above
exist
abc
not exist
tire
not exist
tree
exist
pat
not exist
```

2. 实验任务2







```
■ D:\LEARNING\CODES\DSJSJJG\work2\project2.exe
there 3
they 1
this 5
time 6
times 1
to 11
top 1
towards 1
trees 1
unlike 1
until 1
 up 6
up o
upon 1
valley 1
very 3
view 1
vision 1
warm 1
 was 10
was 10
waterfalls 1
were 1
what 2
where 1
whined 2
why 1
with 7
worked 1
world 1
would 2
ves 1
 yes 1
  ■ D:\LEARNING\CODES\DSJSJJG\work2\project2.exe
time 6
times 1
 to 11
top 1
top 1
towards 1
trees 1
unlike 1
until 1
up 6
upon 1
valley 1
very 3
view 1
vision 1
warm 1
 warm 1
was 10
 waterfalls 1
wateriall
were 1
what 2
where 1
whined 2
why 1
with 7
worked 1
world 1
world 1
yes 1
you 4
your 3
```

源代码

```
tire.h
#include <string>
struct node
{
    node *next[26];
    bool isStr;
    int count = 0;
    node()
    {
        for (int i = 0; i < 26; i++)
            next[i] = NULL;
    }
    void insert(std::string s)
        node *p = this;
        for (int i = 0; i < s.size(); i++)</pre>
            if (p->next[s[i] - 'a'] == NULL)
                node *temp = new node;
                temp->isStr = false;
                p->next[s[i] - 'a'] = temp;
            p = p->next[s[i] - 'a'];
        p->isStr = true;
        p->count++;
    bool find(std::string s)
    {
        node *p = this;
        for (int i = 0; i < s.size(); i++)</pre>
        {
            if (p->next[s[i] - 'a'] == NULL)
            {
                return false;
            }
            else
                p = p->next[s[i] - 'a'];
            }
        if (p->isStr)
            return true;
        else
            return false;
    }
    bool remove(std::string s)
    {
```

```
node *p = this;
        for (int i = 0; i < s.size(); i++)</pre>
        {
             if (p->next[s[i] - 'a'] == NULL)
             {
                 return false;
             }
            else
             {
                 p = p->next[s[i] - 'a'];
             }
        }
        if (p->isStr)
             p->isStr = false;
            p \rightarrow count = 0;
             return true;
        }
        else
        {
             return false;
        }
    }
};
project1:
#include "tire.h"
#include <iostream>
#include <fstream>
#include <string>
void init(node &root)
{
    std::ifstream wordlist("wordlist1.txt");
    std::string line;
    while(getline(wordlist, line))
        root.insert(line);
    }
}
int main()
{
    node T;
    init(T);
    while(1)
    {
        std::string op;
        std::cin >> op;
        if(op=="q")
        {
             return 0;
```

```
}
        else
        {
            if(T.find(op))
            {
                 std::cout << "exist" << std::endl;</pre>
             }
            else
            {
                 std::cout << "not exist" << std::endl;</pre>
            }
        }
    }
}
project2:
#include "tire.h"
#include <iostream>
#include <fstream>
#include <string>
void init(node &root)
    std::ifstream file("article1.txt");
    std::string str;
    while (file >> str)
    {
        root.insert(str);
}
void show(node *root, std::string now)
    for (int i = 0; i < 26; i++)
        if (root->next[i] == NULL)
            continue;
        if (root->next[i]->isStr)
            std::cout << now + (std::string(1, char('a' + i))) << " " << root->next[i]->count <</pre>
        show(root->next[i], now + (std::string(1, char('a' + i))));
    }
}
int main()
{
    node T;
    init(T);
    show(&T, "");
    getchar();
}
```