

## 1.2.1 Перекрестные ссылки. Бинарное дерево поиска.

3530904/90002 Мэн Цзянин

### 1. Общая постановка задачи

- 1) Для разрабатываемого словаря реализовать основные операции:
  - INSERT (ключ, значение) – добавить запись с указанным ключом и значением
  - SEARCH (ключ)- найти запись с указанным ключом
  - DELETE (ключ)- удалить запись с указанным ключом
- 2) Предусмотреть обработку и инициализацию исключительных ситуаций, связанных, например, с проверкой значения полей перед инициализацией и присваиванием.
- 3) Программа должна быть написана в соответствии со стандартом программирования: C++  
Programming Style Guidelines (<http://geosoft.no/development/cppstyle.html>).
- 4) Тесты должны учитывать как допустимые, так и не допустимые последовательности входных данных.

### 2. Требования

Разработать и реализовать алгоритм формирования перекрестных ссылок:

- прочитать текст и вывести его с добавлением последовательных номеров строк;
- собрать все слова, встречающиеся в тексте;
- сформировать таблицу, в которой все слова будут расположены в алфавитном порядке и для каждого слова будет указан список строк его нахождения (по возрастанию номеров строк)

Для реализации задания использовать бинарное дерево поиска, узел которого может содержать:

- Ключ – слово
- Информационная часть – ссылка на список, содержащий номера строк

### 3. Спецификация

- 1) Файл должен быть открыт
  - i. В противном случае выводим ошибку, содержащую текст:  
«cant\_open\_the\_file»
- 2) root, который в бинарном дереве поиска нельзя было nullptr
- 3) все данные должны иметь правильный тип
  - i. Поля Номер строки(row), столбца(col), счетчик(count) и длина слова(length) должны иметь тип данных unsigned int
  - ii. Номер строки, столбца и длина слова должны больше чем 1
- 4) Если пустая строка пропускается, счетчик не накапливается

- 5) Для повторяющихся слов
- i. В одной статье может быть несколько повторяющихся слов, и номера строк и столбцов этих слов должны храниться в определенном контейнере
  - ii. При печати функции(`printWordAndInfo`), которая выводит слова, должен быть предоставлен внешний интерфейс, чтобы пользователь мог выбрать, следует ли печатать повторяющиеся слова
    - a) Если `«bool repetitive_word = true»`: выводит все повторяющиеся слова вместе с их информацией
    - b) Если `«bool repetitive_word = false»`: выводит это слово, номер строк и столбцов, которые появляются в первый раз
  - б) Для одного и того же слова, но с большой буквы или с “ ’s ”. Все буквы должны быть преобразованы в маленькую букву.

**Пример:**

Hello	→	hello
World	→	world
student's	→	student
Student's	→	student

Рис.1

## 4. Структура данных

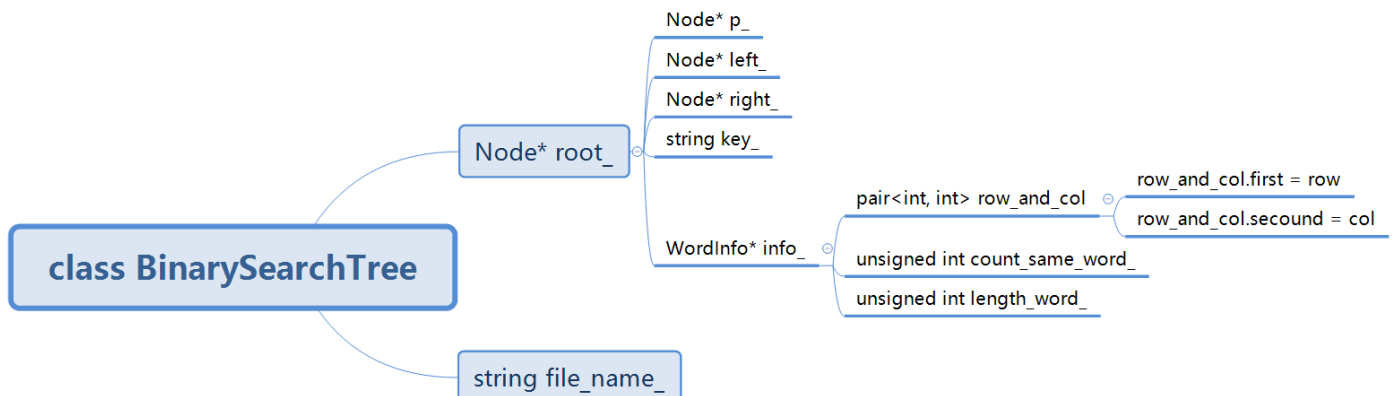


Рис.2

## 5. Тест план

\*для тестового примера используется файл с название 1.txt. Текст (1.txt) представлена для правильного ввода

Ввод:

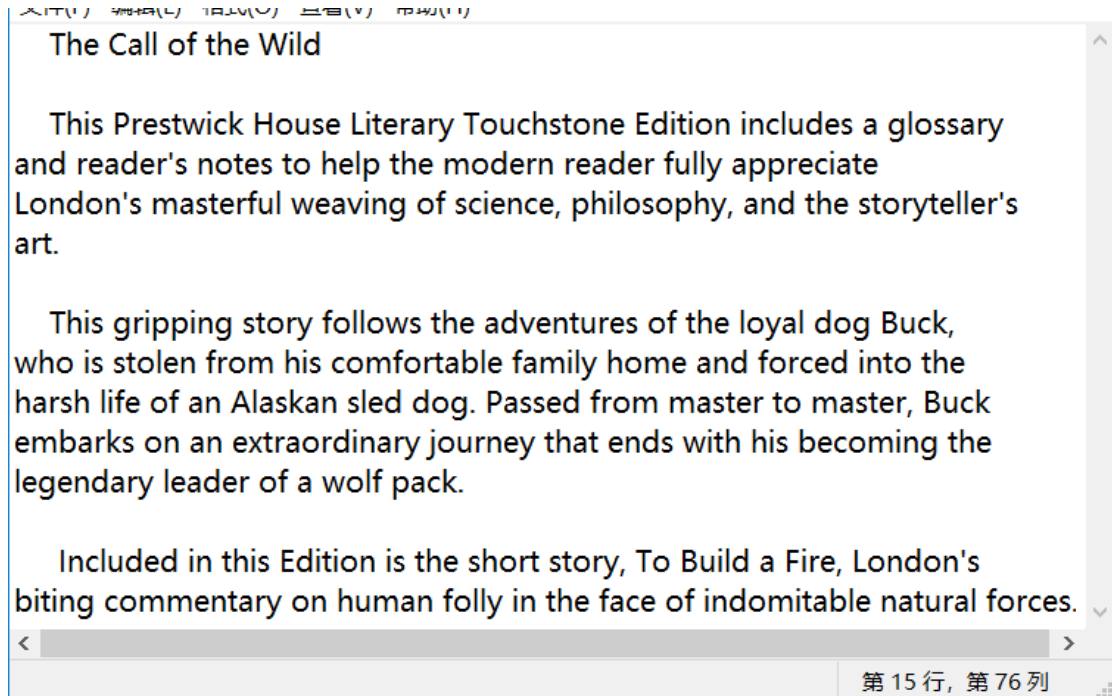


Рис.3 (1.txt)

Вывод (Пример):

- 1 - The Call of the Wild 【Words in this row : 5】
- 2 -
- 3 - This Prestwick House Literary Touchstone Edition includes a glossary 【Words in this row : 9】
- 4 - and reader's notes to help the modern reader fully appreciate 【Words in this row : 10】
- 5 - London's masterful weaving of science, philosophy, and the storyteller's 【Words in this row : 9】
- 6 - art. 【Words in this row : 1】
- 7 -
- 8 - This gripping story follows the adventures of the loyal dog Buck, 【Words in this row : 11】
- 9 - who is stolen from his comfortable family home and forced into the 【Words in this row : 12】
- 10 - harsh life of an Alaskan sled dog. Passed from master to master, Buck 【Words in this row : 13】
- 11 - embarks on an extraordinary journey that ends with his becoming the 【Words in this row : 11】
- 12 - legendary leader of a wolf pack. 【Words in this row : 6】
- 13 -
- 14 - Included in this Edition is the short story, To Build a Fire, London's 【Words in this row : 13】
- 15 - biting commentary on human folly in the face of indomitable natural forces. 【Words in this row : 12】

=====  
Amount of words in the article : 112  
Amount of row in the article : 15  
=====|

Рис.4 (Часть 1 - Пример)

Word	Row	Col	Count	Length
a	3	8	3	1
adventures	8	6	1	10
alaskan	10	5	1	7
an	10	4	2	2
and	4	1	3	3
appreciate	4	10	1	10
....				
Пример				
....				
touchstone	3	5	1	10
weaving	5	3	1	7
who	9	1	1	3
wild	1	5	1	4
with	11	8	1	4
wolf	12	5	1	4

Рис.5 (Часть 2 - Пример)

## 6. Для некорректного ввода данных:

Спецификация	Ожидаемый вывод		Данные	Ожидаемый вывод		Операция	Ожидаемый вывод	
2.t	ERROR --- cant_open_the_file ---		root_ = nullptr	ERROR - nullptr_root		Неправильный выбор	No corresponding operation !	
1. txt						Нет инициализации	no_data	
1.c						Неверные данные при search	Don't have word [...] !	
1txt			word in null	ERROR - null_word		Неверные данные при delete	delete fail, don't have word [...]	
1								
txt								

Таблица 1