PROJECT1
Chinese-English Dictionary based on Binary Search Tree
October 10, 2023

1 Summary

To better understand the tree data structures learned in this course, the project designed for you to use red-black tree and B-tree to implement an English-Chinese dictionary.

2 Percentage Point

Item		Cost	Description
Red-Black tree(20%)	Initialization	4%	Initialize the tree with init.txt
	Delete	4%	Delete the words with delete.txt
	Insert	4%	Insert new words with insert.txt
	Search	4%	All correct search
	Preorder print	4%	Print the tree in right format to rbt.txt
B-Tree(20%)	Initialization	4%	The same as read-black tree
	Delete	4%	-
	Insert	4%	-
	Search	4%	-
	Preorder print	4%	-
General		5%	Correct implementation
UI		15%	Friendly and easy to use
Document		30%	Full and detail. Reasonable analysis
Coding style		5%	Proper comment
bonus		5%	Creative thought

3 Requirements

3.1 Functions

Implement red-black tree and B-tree by yourself. (including INSERT, DELETE, SEARCH, PREORDER_PRINT methods)
ATTENTION: The branch number of B-tree should be easily changeable. And the initial branch number is suggested to be 10.

3.1.1 PREORDER_PRINT

Display the tree by PREORDER traversal to a file.
Here we give an example after insert into trees with number 7, 3, 5, 1, 6. We

require you to print like that format. In this project, you are supposed to replace the number with words.
Print result for red-black tree:

level=0 child=0 5(BLACK)
level=1 child=0 3(BLACK)
level=2 child=0 1(RED)
level=3 child=0 null
level=3 child=1 null
level=2 child=1 null
level=1 child=1 7(BLACK)
level=2 child=0 6(RED)
level=3 child=0 null
level=3 child=1 null
level=2 child=1 null

Print result for B-tree (Here we show the tree with branch number 4)

level=0 child=0 /5/
level=1 child=0 /1/3/
level=1 child=1 /6/7/

3.1.2 INSERT/DELETE

1. A batch of words which contain in a file. The first line in the file represents which operation it will do and next lines are data. Each time you operate with one file, please call your **PREORDER_PRINT** method in tree classes to print the tree to rbt.txt / bt.txt.
2. A single word.

3.1.3 SEARCH   3.1.3

1. Some words in range and give their meanings. E.g. we give a query: please search from 'aa' to 'apc', then you give the words between 'aa' and 'apc' as well as their meanings. The boundary values don't have to be exactly words.
2. A single word. Just give its Chinese meanings.
Hint: You know that English words are in lexicographic order. So please use the word as key and build the trees.

3.2 Analysis work

Another work you should do in this project is to compare the operations like insertion and deletion for both trees.
Please call your methods in the following ways for each kind of tree, and analyses each time they spent.

Note: Don't change the files we have provided or disarrange the step order.

1. insert into trees the data in the file 1_initial.txt.
2. delete the data in the file 2_delete.txt. When deleting, you need to check whether the word to be deleted is in the tree. If it's not, don't perform the delete operation.
3. Add the data in the file 3_insert.txt. When inserting, you need to check whether the word to be inserted is in the tree. If it is, don't perform the insert operation.
4. Query a word.
5. Query some words.   4 5

For the first three steps, after each operation on 100 pieces of data, you should record the time used. And finally give a document based on this analysis and testing time should be attached on it.

4   Design

To make your program more flexible, we don't recommend that you input the command files by hard-coding. We prefer that you provide an interface through which the user can import their files.
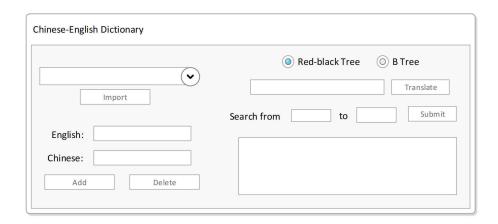Here we provide a UI design graphic

Figure 1:

4.1   The left is the manage part

1. The user can choose a file and execute the corresponding operations (INITIALIZATION)

2. The user can add or delete a single word(delete use English only). (INSERT/DELETE)

4.2 The right-top is choice part

It means which kind of tree you are using to implement the dictionary. The user can make the choice.

4.3 The right-bottom is look-up part

1. Just give a Chinese with the given English(SEARCH)
2. Give the words and the paraphrases in the query range(Inorder Traversal)

5 Points for Attention

(1) For the implementation of these algorithms, you are free to select a programming language of your choice.

(2)Your document should be submitted in electronic format whenever possible. If you have a handwritten document, please ensure that the writing is neat and the layout is well-organized. The document format should be either Word, PDF, or Markdown.

(3)Interviews will be arranged for everyone after the deadline to showcase their work related to the project and UI.The interview will take place during the laboratory class on October 31st. We will conduct interviews in the order listed on the student roster.

(4) Kindly upload the source code files along with their associated documentation in a compressed ZIP format to the elearning system for assessment.

(5) The deadline of this lab is 23:59:59 on October 29, 2023.

(6) The naming format for the file should be "proj1-StudentID-Name," and make sure to compress all the files into a single compressed folder.

(7) If you have any questions please feel free to contact teaching assistants.