

**Homework Module 9****Problems**

1. (50 pts) Implement a hash for text. Given a string as input, construct a hash with words as keys, and word counts as values. Your implementation should include:
  - a hash function that has good properties for text
  - storage and collision management using linked lists
  - operations: `insert(key,value)`, `delete(key)`, `increase(key)`, `find(key)`, `list-all-keys`

Output the list of words together with their counts on an output file. For this problem, you cannot use built-in-language datastructures that can index by strings (like hash tables). Use a language that easily implements linked lists, like C/C++.

You can test your code on “Alice’s in Wonderland” by Lewis Carroll, at <ftp://ibiblio.org/pub/docs/books/gutenberg/1/11/11-h/11-h.htm>. The test file used by TA will probably be shorter.

(Extra Credit) Find a way to record not only word counts, but also the positions in text. For each word, besides the count value, built a linked list with positions in the given text. Output this list together with the count.

2. (50points) Implement a red-black tree, including binary-search-tree operations *sort*, *search*, *min*, *max*, *successor*, *predecessor* and specific red-black procedures *rotation*, *insert*, *delete*. The *delete* implementation is Extra Credit (but highly recommended).

Your code should take the input array of numbers from a file and build a red-black tree with this input by sequence of “inserts”. Then interactively ask the user for an operational command like “insert x” or “sort” or “search x” etc, on each of which your code rearranges the tree and if needed produces an output. After each operation also print out the height of the tree.

You can use any mechanism to implement the tree, for example with pointers and struct objects in C++, or with arrays of indices that represent links between parent and children. You cannot use any tree built-in structure in any language.

3. (Extra credit 50 pts) The `int[]` bag describes a bag of non-negative integers. A bag is the same thing as a set, only it may contain repeated elements. The order of elements in a bag does not matter. Given two bags *A* and *B*, we say that *A* is a *sub-bag* of *B* if *A* can be obtained by erasing zero or more elements from *B*. The *weight* of a bag is the sum of its elements. Examples: The bags (1, 2, 1, 3, 1) and (3, 1, 1, 1, 2) are the same, but different from the bag (1, 2, 3, 3). Bags (1, 2) and (3, 1, 1) are sub-bags of the bag (1, 2, 1, 3, 1), but bag (1, 2, 2) is not. The weight of the bag (1, 2, 1, 3, 1) is



10000, 10000, 10000, 10000, 10000, 10000, 10000, 10000, 10000, 10000, 10000, 10000,  
10000, 10000, 10000, 10000, 10000, 10000, 10000}

Return 9: 6555