

Program 4

48 Points

Due: February 15, 2008 at 11:59 p.m. {Changed Feb 11, 2008}

A Cross Reference Generator in C++

This program focuses on using tree data structures to store and information. This assignment is to be done in C++ and the use of binary trees or another approved data structure is required for this assignment.

This assignment requires creating a case-insensitive cross reference generator[1]. The program will take an ASCII text file as input (e.g. Lincoln Gettysburg Address). After reading in the file, the program prints out an alphabetized list of all the words in the file, approximating the following format:

WORD	NUMBER OF OCCURENCES	LINES THAT THE WORD APPEARED ON
------	----------------------	---------------------------------

For example, given the poem “There are holes in The Sky” by Spike Milligan and excluding the title, the cross generator program would print out the tally below.

There are holes in The Sky

There are holes in the sky
 Where the rain gets in
 But they’re ever so small
 That’s why the rain is thin.

Word	Count	Lines
are	1	1
but	1	3
ever	1	3
gets	1	2
holes	1	1
in	2	1, 2
is	1	4
rain	2	2, 4
sky	1	1
small	1	3
so	1	3
that’s	1	4
the	3	1, 2, 4
there	1	1
they’re	1	3
thin	1	4
where	1	2
why	1	4
TOTAL WORDS	18	

What to turn in for Program 4

Official test input files will be made available a few days before the due date. Turn in your assignment using the *turnin* program on the CCC machines. You should turn in a tarred file that includes your source code, a make file and a README file. The README file should provide any information to help the TA or SA test your program for grading purposes.