Rochester Institute of Technology
Golisano College of Computing and Information Sciences
School of Information

R·I·T

ISTE-222 Assignment 3: Algorithm Analysis

The Count of Monte Carlo

Problem Statement:

The purpose of this assignment is to study the behavior of search algorithms across various sets of inputs using a technique called Monte Carlo analysis.

Monte Carlo analysis involves subjecting a system to random input perturbations in order to see how it responds across a wide variety of inputs. For this assignment, you will be studying how the average number of operations of three search algorithms varies across a collection of randomly generated data sets. Specifically, you will be analyzing linear search, binary search, and interpolation search in terms of the average number of operations each algorithm performs over a series of data sets.

ClassicSearchesA3.java contains a class member variable called "ops" that is incremented each time one of the algorithms has to perform an operation. An "operation" will consist of comparison(s) and/or recomputing indices.

Your assignment is to randomly generate a collection of data sets to calculate the average number of operations each algorithm performs to find **all** items in a data set as well as searching for items that are **not** in the data set. This will ultimately show how many operations on average that each search algorithm performs for both successful and unsuccessful searches. You will then analyze these results in terms of how they compare with the theoretical average case complexities of each of the three search algorithms being studied.

Rochester Institute of Technology
Golisano College of Computing and Information Sciences
School of Information

R·I·T

Requirements/Specifications:

This assignment involves coding as well as analyzing the results produced by having run the code. You will be generating 100 data sets of 1,000 elements each with random values ranging from 1 to 10,000, which you will use to evaluate linear search, binary search, and interpolation search.

For each data set, you will execute a search using each algorithm to find every element contained in that data set. After a search for a single element using a given algorithm, you'll record the number of operations that the algorithm performed to find that element. Once you've searched for all elements with each algorithm, compute the average number of operations that each algorithm performed across all of the searches for that data set.

Similarly, for each data set, execute a search for 1,000 random values between 1 and 10,000 that are **not** contained in the data set. Record the same data that you did as mentioned above for the successful element searches.

At this point, you'll have 100 averages for each of the three algorithms that denote the average number of operations required for a **successful** search as well as 100 averages for each of the three algorithms that denote the average number of operations required for an **unsuccessful** search.

From those 100 averages for successful and unsuccessful searches from each search algorithm, compute the average of those values for each algorithm.

Compare this average for successful searches across all of the data sets with the theoretical average case for linear search, binary search, and interpolation search. Also, compare the average for unsuccessful searches across all of the data sets with the theoretical average case. Use the Assignment 3 – Report Template document to insert your results and write up your analysis.

Your code and a PDF document containing your data and related analysis is due in the Assignment 3 dropbox by the due date specified on myCourses.

ISTE-222 Assignment 3: Algorithm Grading Sheet   Name:

Item	Possible points	Earned points
Data Set Creation (20%)		
100 total data sets are created	5	
Data sets contain 1,000 elements each	5	
Values in the data sets range from 1 to 10,000	10	
Data Collection (60%)		
Searches for all 1,000 elements contained in a data set	15	
Searches for 1,000 elements **not** contained in a data set	30	
Average number of operations for a data set is properly computed	10	
Average across all successful/unsuccessful searches is properly computed	5	
Report (20%)		
Theoretical average case values are computed correctly	10	
Meaningful analysis of both successful and unsuccessful searches provided	10	
Total:	100	

Comments: