Seoul National University

M1522.000900 Data Structure

Fall 2017, Kang

Homework 7: Searching (Chapter 9)

Due: November 30, 02:00 PM

**Reminders**

* The points of this homework add up to 100.
* Like all homeworks, this has to be done individually.
* Lead T.A.: Beunguk Ahn ([beunguk.ahn@gmail.com](mailto:beunguk.ahn@gmail.com))
* Please type your answers in English. Illegible handwriting may get no points, at the discretion of the graders.
* If you have a question about assignments, please upload your question in eTL.
* If you want to use slip days or consider late submission with penalties, please note that you are allowed one week to submit your assignment after the due date.

Remember that:

1. Whenever you are making an assumption, please state it clearly

-

**Question 1**

Jump search is a search algorithm which skips some of the items of an array to improve the performance. Let be the number of items in the arrawy. Then, the optimal size of jump is . Answer the following questions [10 points]:

1. What is the number of comparisons when we search the given sorted array for the element of ?

**[** … **]** (a list of elements in power series)

ㅋㅈㅋ

1. What is the number of comparisons when we search the given sorted array for the element of 234?

**[**0 4 5 10 23 30 43 46 60 65 65 100 134 234 456 459**]**

**Question 2**

A student ID has the format of “XXXX-XXXXX” or “OO-OOOOO,” but you change its format to integer as “XXXXXXXXX”, “19OOOOOOO”. For example, student IDs “2013-12345” and “97-54321” are converted into “201312345” and “199754321.” You call this format ***N\_SID*** (Numeric Student ID), and it is used as a key for a hash function.

The minimum value of *N\_SID* is 199700000.  
The maximum value of *N\_SID* is 201799999.

Answer the following questions.

**Question 2-1**

Assume that you have an empty 10-slot **open hashing table** (the slots are numbered 0 through 9) for which you use the following function. [30 points]

1. Show the final hash table that would result if you insert *N\_SID*s as follows.

[201610005 201610001 199743532 201634532 201723453 200733222].

1. How many collisions occur in the question (a) above?
2. What could be the problem of using this hash function ? Describe it in at most two sentences.

**Question 2-2**

You decide to change the hash function as the following: [20 points]

1. Show the final hash table that would result if you insert *N\_SID*s as follows.

[201610005 201610001 199743532 201634532 201723453 200733222].

1. How many collisions occur in the question (a) above?

**Question 2-3**

Suppose is the probability of at least two of the students sharing a same hash value. Assume N\_SIDs are independently and uniformly distributed at the range of the minimum to the maximum. [20 points]

1. At which , does exceed 0.99 using the hash function in the question 2-2?
2. At which , does exceed 0.99 using the hash function ?

**Question 3**

Consider inserting the keys 10, 21, 34, 45, 4, 88, 35, 91, and 59 into a hash table of length 11 using **open addressing**. (Slots are numbered 0 through 10). We will use the following function as a hash function. [10 points]

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Slot | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Key |  |  |  |  |  |  |  |  |  |  |  |

1. For linear probing, we use for probing. Show the final hash table that would result if you insert the given keys.
2. For quadratic probing, we use for probing. Show the final hash table that would result if you insert the given keys.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Slot | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| Key |  |  |  |  |  |  |  |  |  |  |  |

**Question 4**

**Text compression** is an example that uses Move-to-Front heuristic algorithm. For example, “the car on the left hit the car I left” could be compressed as “the car on 3 left hit 3 5 I 5”. [10 points]

Compress the following quote. Ignore the periods and commas.

**I felt happy because I saw the others were happy and because I knew I should feel happy, but I was not really happy.**