Ph.D. Qualification Examination

Algorithms (March 2013)

(1) (20%) Given the string "go go gophers", please construct a Huffman tree that optimally encodes the characters (including the space character) in this string. What is the complexity of the Huffman's algorithm? Is it possible to construct this tree in linear time?

(2) (10%) Please construct a B-tree of order 3 by inserting this sequence of numbers: 10, 5, 20, 25, 15, 30. Briefly discuss why the B-tree is suitable for data search on secondary storage device.

(3) (20%) Consider inserting the keys 10, 22, 31, 4, 15, 28, 17, 88, 59 into a hash table of length $m = 11$ using open addressing with a hash function $h'(k) = k \mod m$. Indicate the result of inserting these keys using linear probing and double hashing. In linear probing, the hash function is used: $h(k, i) = (h'(k) + i) \mod m$. In double hashing, the hash function is used: $h(k, i) = (h'(k) + i \times h_2(k)) \mod m$, where $h_2(k) = 1 + (k \mod (m - 1))$.

(4) (20%) Consider an open-address hash table with uniform hashing. Give upper bounds on the expected number of probes in an unsuccessful search and on the expected number of probes in a successful search when the load factor is 3/4 and when it is 7/8.

(5) (20%) Use the dynamic programming strategy to align the following two sequences with the cost function: match= 2, match with gap= 0, mismatch= -1 (using a table to show your result).

C C G G G T T A C C A
G G A G T T C A

(6) (10%) A certain problem can be solved by an algorithm whose running time is in $O(2^n)$. Which of the following assertions is true?

(a) The problem is tractable.
(b) The problem is intractable.
(c) None of the above.