Midterm Exam

Instructions. This exam is closed book, closed notes, and no gadgets, but you are allowed **one sheet of notes**. You will have the full class period (50 minutes). If you put answers on the back of a sheet, mark this clearly.

Name (Print): ______________________________

This exam is my own work. I understand that this exam is governed by the **Emory Honor Code**.

Signature: ______________________________

Please leave this blank.

Problem	Topic	Score	Max
1	Short Answer		10
2	Open Hashing		15
3	Strings		20
4	Search Trees		10
5	Fill in the Blank		20
6	Extra Credit		+3
Raw Total			75(+3)
Curved			100
Grade			

Problem 1. (10 points) Short Answer. Answer each question with a few sentences.

1(a). Compare the skiplist to a balanced binary search tree (such as red-black). Give an advantage of the skiplist, and a disadvantage.

1(b). What is the book's distinction between a hash code and a hash function?

Problem 2. (15 points) Open Hashing. In each part, show the result of inserting 12, 32, 45, 53, 62 (in that order) into an empty table T of size 10. We'll use the hash function $h(x) = (x \bmod 10)$. Note that $h(x)$ is the last digit of x, for example $h(12) = 2$.

2(a). Use linear probing.

i	0	1	2	3	4	5	6	7	8	9
$T[i]$										

2(b). Use quadratic probing.

i	0	1	2	3	4	5	6	7	8	9
$T[i]$										

2(c). Use Cuckoo hashing, where $h_1(x) = h(x)$ as above (the last digit), and $h_2(x)$ is the 10's digit of x. For example, $h_2(12) = 1$. (Recall: start by putting x in slot $h_1(x)$.)

i	0	1	2	3	4	5	6	7	8	9
$T[i]$										

Problem 3. (20 points) Strings.

3(a). Draw a trie (compressed or not) storing the strings $\{baaa, aaaa, aabb, aacc, aaba\}$.

3(b). Draw the suffix trie for "`abcabca%`". (The character ordering is `%` < `a` < `b` < `c`.)

3(c). Find the Burrows-Wheeler transform of "`abcabca`", using `%` as the marker (like `hw3`).

3(d). Finish this table, the Knuth-Morris-Pratt failure function for the pattern "abbaabbab".

j	0	1	2	3	4	5	6	7	8
$P[j]$	a	b	b	a	a	b	b	a	b
$f(j)$	0	0	0						

Problem 4. (10 points) Search Trees. Draw a valid (2,4) tree with at least one 2-node, one 3-node, and one 4-node, storing integer keys. (Hint: this can be done with as few as 6 keys.) Next to that, draw a corresponding red-black tree.

Problem 5. (20 points) Fill in the Blank. Put an appropriate word or phrase in each blank below.

To do a rotation in a binary search tree (without parent links), we modify ______________________ links. (A number.)

______________________ is the hashing method taking $O(1)$ worst case time to find or remove a key.

______________________ is the hashing method taking best advantage of modern memory caching systems, for large enough data sets.

______________________ is the hashing method which may safely let the load factor rise above one.

Given a text string T of length n and a query string P of length m, the Boyer-Moore algorithm finds the first occurrence of P in T in worst-case time $O($______________________$)$.

Given T and P as above, plus a precomputed suffix trie for T, we can find the first occurrence of P in T in worst-case time $O($______________________$)$.

Suppose we want to insert a key k into a (2,4) tree, in a "one-pass" (top-down only) way. As we walk down the tree looking for our insertion point, we make sure that on each step, the node we are about to step into is not a(n) ______________________.

In a splay tree of size n, the maximum possible value of the potential function is $\Theta(n \lg n)$. This happens when the tree looks ______________________. (Either a phrase or small picture is OK here.)

Suppose we create an empty splay tree. We insert the integers $1, 2, \ldots, n$ in increasing order, and then we find each of those n integers once again, in increasing order. The total time needed is $O($______________________$)$.

Let A_h be the minimum size (number of keys) of an AVL tree of height h. So $A_1 = 1$ and $A_2 = 2$. It satisfies the recurrence $A_{h+2} =$ ______________________.

Problem 6. (+3 points) Extra Credit. Describe the MAD method, and its purpose.