Test 2 will be Thursday November 3 in class. It will be closed-book and notes, except for one 8.5" x 11" sheet of paper containing any notes that you want. (Yes, you can use both the front and back of this piece of paper.) Plus, you can use your Python Summary handout.

The test will cover Chapters 4 and 5. The following topics (and maybe more) with be covered:

Chapter 4: Recursion
Recursive functions: base-case(s), recursive case(s), tracing recursion via run-time stack or recursion tree, "infinite recursion"
Costs and benefits of recursion
Recursive examples: countDown, OrderedList __str__ method, fibonacci, factorial, binomial coefficient
Divide-and-Conquer technique of solving a problem. Examples: fibonacci, coin-change problem
Backtracking technique of solving a problem: Examples: coin-change problem, maze (textbook)
General concept of dynamic programming solutions for recursive problems that repeatedly solve the same smaller problems over and over. Example fibonacci, coin-change problem, binomial coefficient

Chapter 5: Searching and Sorting
Sequential/Linear search: code and big-oh analysis
Binary Search: code and big-oh analysis
Python List implementation (ListDict) of dictionaries and big-oh analysis
Hashing terminology: hash function, hash table, collision, load factor, chaining/closed-address/external chaining, open-address with some rehashing strategy: linear probing, quadratic probing, primary and secondary clustering
hashing implementation of dictionaries (ChainingDict and OpenAddrHashDict) and their big-oh analysis
General idea of simple sorts
Simple sorts: selection, bubble, insertion sorts and their big-oh analysis
Advanced sorts and their big-oh analysis: heap sort, quick sort and merge sort

Spring 2016 Name: ____________________

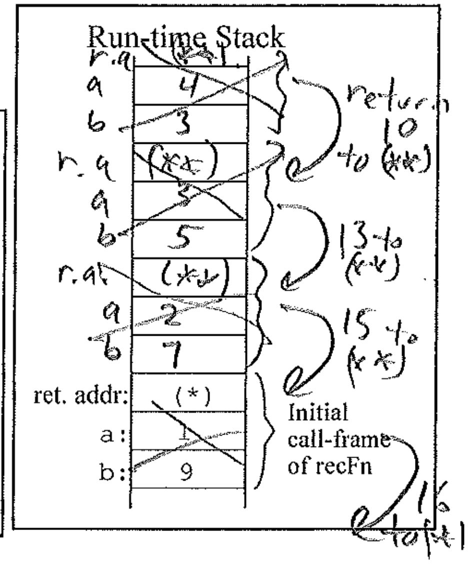
Data Structures - Test 2

Question 1. (10 points) What is printed by the following program? Output:

```
def recFn(a, b):
    print(a, b)
    if a == b:
        return 100
    elif a > b:
        return 10
    else:
        return recFn(a + 1, b - 2) + a
                    (**)

print("result =",recFn(1, 9))
                  (*)
```

1 9
2 7
3 5
4 3
result = 16



Question 2. (10 points) Write a recursive Python function to compute the following mathematical function, G(n):

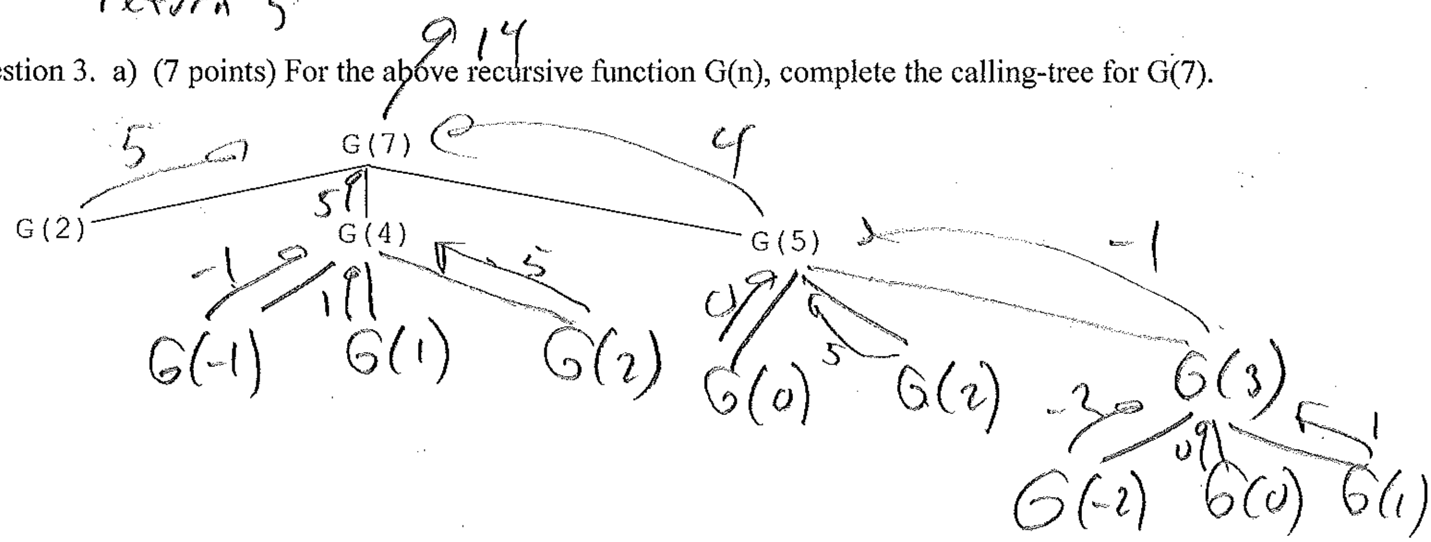
$G(n) = n$ for all value of $n \le 1$

$G(2) = 5$ if $n = 2$

$G(n) = G(n-5) + G(n-3) + G(n-2)$ for all n values > 2.

```
def G(n):
    if n > 2:
        return G(n-5)+G(n-3)+G(n-2)
    elif n <= 1:
        return n
    else: #n=2
        return 5
```

Question 3. a) (7 points) For the above recursive function G(n), complete the calling-tree for G(7).



b) (2 point) What is the value of G(7)? 14

c) (1 point) What is the maximum height of the run-time stack when calculating G(7) recursively? 4

Question 3. The insertion sort code discussed in class is:

```
def insertionSort(myList):
    for firstUnsortedIndex in range(1,len(myList)):
        itemToInsert = myList[firstUnsortedIndex]
        testIndex = firstUnsortedIndex - 1
        while testIndex >= 0 and myList[testIndex] > itemToInsert:
            myList[testIndex+1] = myList[testIndex]
            testIndex = testIndex - 1

        myList[testIndex + 1] = itemToInsert
```

0 n 1 0 2 3 (n-2) (n-1)

Consider the following `insertMergeSort` code which calls the above `insertionSort` code twice with copies of each half of the array, and then merges the two sorted halves back together using the merge code from merge sort. $n \cdot \frac{(n-1)}{2} \approx \frac{n^2}{2}$

```
def insertMergeSort(aList):
    halfSize = len(aList) // 2
    lefthalf = aList[ : halfSize]
    righthalf = aList[halfSize : ]
    insertionSort(lefthalf)
    insertionSort(righthalf)

    #### BELOW IS THE MERGE CODE FROM MERGE SORT ####
    i=0  # index into lefthalf
    j=0  # index into righthalf
    k=0  # index into aList
    while i<len(lefthalf) and j<len(righthalf):  # compare and copy until one half runs out
        if lefthalf[i]<righthalf[j]:
            aList[k]=lefthalf[i]
            i=i+1
        else:
            aList[k]=righthalf[j]
            j=j+1
        k=k+1

    while i<len(lefthalf):       # copy the remaining items from lefthalf if any
        aList[k]=lefthalf[i]
        i=i+1
        k=k+1

    while j<len(righthalf):       # copy the remaining items from righthalf if any
        aList[k]=righthalf[j]
        j=j+1
        k=k+1
```

$\frac{n^2}{4} \cdot 2$ $\frac{n^2}{2}$ n aList n 0   lefthalf righthalf   sorted n aList (n-1)

Consider the following timing of `insertionSort` vs. `insertMergeSort` on lists of 10000 elements.

Initial arrangement of list before sorting	`insertionSort` - at the top of page	`insertMergeSort` - modified version in middle of the page
Sorted in descending order: 10000, 9999, ..., 2, 1	14.3 seconds	7.1 seconds
Already in ascending order: 1, 2, ..., 9999, 10000	0.005 seconds	0.009 seconds
Randomly ordered list of 10000 numbers	7.4 seconds	3.6 seconds

a) (10 points) Explain why `insertMergeSort`(modified version in middle of page) out performs the original `insertionSort`.

b) (10 points) In either version, why does sorting the randomly order list take about halve the time of sorting the initially descending ordered list?

Question 4. (20 points) In insertion sort the inner-loop takes the "first unsorted item" (25 at index 6 in the below example) and "inserts" it into the sorted part of the list "at the correct spot."

Sorted Part						Unsorted Part						
0	1	2	3	4	5	6	7	8				
10	20	35	40	45	60	25	50	90	• • •			

In class we discussed the following insertion sort code which sorts in ascending order (smallest to largest) and builds the sorted part on the left-hand side of the list, i.e.:

```
def insertionSort(myList):
    for firstUnsortedIndex in range(1, len(myList)):
        itemToInsert = myList[firstUnsortedIndex]
        testIndex = firstUnsortedIndex - 1
        while testIndex >= 0 and myList[testIndex] > itemToInsert:
            myList[testIndex+1] = myList[testIndex]
            testIndex = testIndex - 1

        myList[testIndex + 1] = itemToInsert
```

last len(myList)-2, -1, -1
last
<= len(myList)-1

For this question write a variation of the above insertion sort that:

- sorts in **descending order** (largest to smallest)
- builds the **sorted part on the right-hand side** of the list, i.e.,

Unsorted Part								Sorted Part						
0	1	2	3	4	5	6	7	8						
			• • •			25	40	90	60	50	45	35	20	10

40
len(myList)-1
Sorted

```
def insertionSortVariation(myList):
```

Question 5. Two common rehashing strategies for open-address hashing are linear probing and quadratic probing:

quadratic probing	Check the square of the attempt-number away for an available slot, i.e., [home address + ((rehash attempt #)^2 +(rehash attempt #))/2] % (hash table size), where the hash table size is a power of 2. Integer division is used above

a) (8 points) Insert "Paul Gray" and then "Sarah Diesburg" using Linear (on left) and Quadratic (on right) probing.

Hash Table with Linear Probing

0	Ben Schafer
1	Paul Gray
2	Sarah Diesburg
3	Philip East
4	
5	
6	Mark Fienup
7	John Doe

Hash function

hash(John Doe) = 7
hash(Philip East) = 3
hash(Mark Fienup) = 6
hash(Ben Schafer) = 0
hash(Paul Gray) = 6
hash(Sarah Diesburg) = 7

Hash Table with Quad. Probing

0	Ben Schafer
1	Paul Gray
2	Sarah Diesburg
3	Philip East
4	
5	
6	Mark Fienup
7	John Doe

$\frac{1^2+1}{2}$ $\left(\frac{2^2+2}{2}+6\right)\%8$ $\left(\frac{2^2+2}{2}+7\right)\%8$

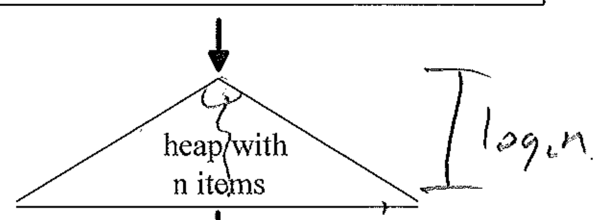
b) (7 points) Explain why both linear and quadratic probing both suffer from primary clustering?

Both base rehashing pattern on home addr. + rehash # so things hashing to same home addr. will follow same rehash pattern

Question 6. Recall the general idea of Heap sort which uses a min-heap (class `BinHeap` with methods: `BinHeap()`, `insert(item)`, `delMin()`, `isEmpty()`, `size()`)) to sort a list.

Generl idea of Heap sort: myList | unsorted list with n items |

1. Create an empty heap
2. Insert all n list items into heap



3. delMin heap items back to list in sorted order

myList | sorted list with n items | O O

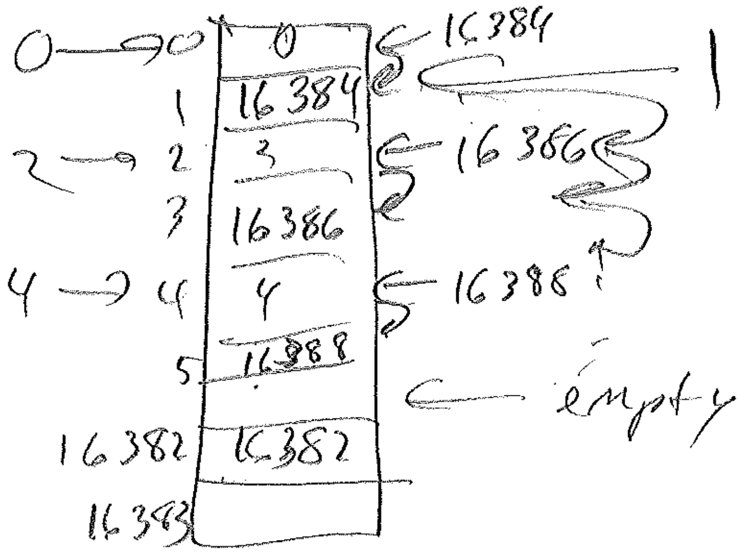
a) (10 points) Complete the code for `heapSort` so that it **sorts in descending order**

```
from bin_heap import BinHeap
def heapSort(myList):
   myHeap = BinHeap()  # Create an empty heap
for item in myList
    myHeap.insert(item)
for index in range(len(myList)-1, -1, -1) - O(n)
    myList[index] = myHeap.delMin() O(log2 n)
```

b) (5 points) Determine the overall O( ) for your heap sort and briefly justify your answer.

$O(n \log_2 n)$

linear



% 16384

Chaining Dec?