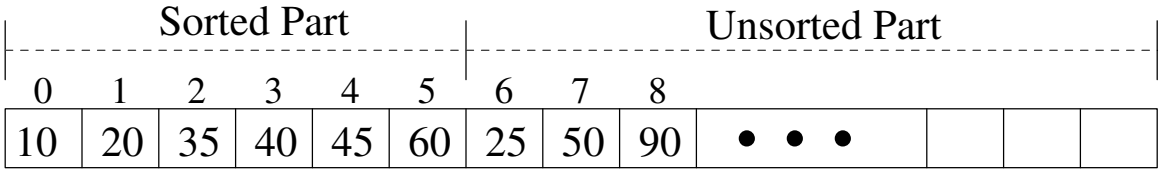
Homework #1 Data Structures
Due: Sept. 3, 2010 (Friday at 3 PM)

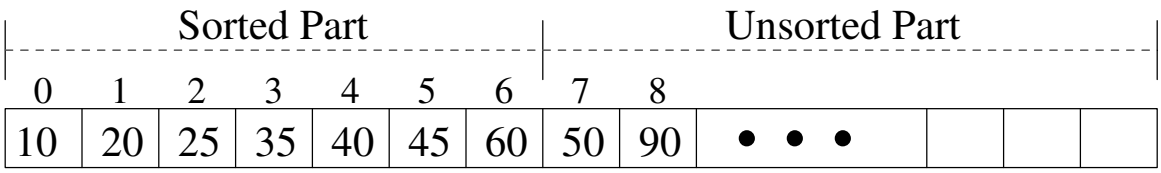
Homework #1 is a pencil-and-paper assignment involving algorithm/program analysis. Answer the questions for each of the following algorithms.

1. Another simple sort is called insertion sort. Recall that in a simple sort:
- the outer loop keeps track of the dividing line between the sorted and unsorted part with the sorted part growing by one in size each iteration of the outer loop.
    - the inner loop's job is to do the work to extend the sorted part's size by one.

After several iterations of insertion sort's outer loop, an list might look like:



In insertion sort the inner-loop takes the "first unsorted element" (25 at index 6 in the above example) and "inserts" it into the sorted part of the list "at the correct spot." After 25 is inserted into the sorted part, the list would look like:



Code for insertion is given below:

```
def insertionSort(myList):
    """Rearranges the items in myList so they are in ascending order"""

    for firstUnsortedIndex in xrange(1,len(myList)):
        itemToInsert = myList[firstUnsortedIndex]
        # Scan the sorted part from the right side
        # Shift items to the right while you have not scanned past the left
        # end of the list and you have not found the spot to insert
        testIndex = firstUnsortedIndex - 1

        while testIndex >= 0 and myList[testIndex] > itemToInsert:
            myList[testIndex+1] = myList[testIndex]
            testIndex = testIndex - 1

        # Insert the itemToInsert at the correct spot
        myList[testIndex + 1] = itemToInsert
```

a) To help you analyze the above algorithm to determine its **worst-case** theta notation, $\Theta(\ )$, complete the following table with the **maximum number** of times the inner-loop executes for each iteration of the outer-loop.

`firstUnsortedIndex value`	**1**	**2**	**3**	**. . .**	**len(myList)-2**	**len(myList)-1**
Maximum number of times inner-loop executes				**. . .**		

b) Write a summation formula for the total number of times that the inner-loop executes.

c) What is the overall worst-case $\Theta(\ )$ notation for insertion sort?

d) What is the worst-case $\Theta(\ )$ notation for the number of element moves?

e) What is the worst-case $\Theta(\ )$ notation for the number of element comparisons?

f) What is the best-case $\Theta(\ )$ notation for the number of element moves?

g) What is the best-case $\Theta(\ )$ notation for the number of element comparisons?

2. Consider the following alternative coding of insertion sort which utilizes an insert function.

```
def insert(myList, itemToInsert, lastSortedIndex):
    """  Inserts itemToInsert into myList's sorted part at the
         correct spot"""
    # Scan the sorted part from the right side
    # Shift items to the right while you have not scanned past the left
    # end of the list and you have not found the spot to insert
    testIndex = lastSortedIndex
    while testIndex >= 0 and myList[testIndex] > itemToInsert:
        myList[testIndex+1] = myList[testIndex]
        testIndex = testIndex - 1

    # Insert the itemToInsert at the correct spot
    myList[testIndex + 1] = itemToInsert


def insertionSort(myList):
    """Rearranges the items in myList so they are in ascending order"""
    for firstUnsortedIndex in xrange(1,len(myList)):
        insert(myList, myList[firstUnsortedIndex], firstUnsortedIndex-1)
```

a) Since the insertionSort function only calls insert, does this improve the worst-case $\Theta(\ )$ notation. Explain your answer.

b) What implications does your answer in part (a) have for analyzing large programs that are split into many functions?

3. Consider a third alternative of insertion sort (no code provided) which utilizes the fact that the "sorted" part of the list is sorted. The previous two versions of insertion sort linearly/sequentially scanned the sorted part of the list for the insertion point. In this problem consider the performance improvements achieved by performing a binary search on the sorted part of the list to find the spot to insert the "first unsorted elemen

Answer the following questions about this **third alternative of insertion sort** (you do NOT need to write the code for this algorithm):

a) What is the worst-case $\Theta(\ )$ notation for the number of element moves? (justify your answer)

b) What is the worst-case $\Theta(\ )$ notation for the number of element comparisons? (justify your answer)

c) What is the overall worst-case $\Theta(\ )$ notation for this third insertion sort? (justify your answer)

d) What is the best-case $\Theta(\ )$ notation for the number of element moves? (justify your answer)

e) What is the best-case $\Theta(\ )$ notation for the number of element comparisons? (justify your answer)

f) What is the overall best-case $\Theta(\ )$ notation for this third insertion sort? (justify your answer)