Data Structures (CS 1520) Fall 2020

Time and Place: 8 - 9:15 AM Tuesday and Thursday in ITT 328 and lab: 8 – 9:50 AM Wednesday in Wright 112 UNI eLearning will be used for this course. It is accessed at https://elearning.uni.edu using your CatID

Instructor: Mark Fienup (mark.fienup@uni.edu) Office: ITT 313 Phone: 319-273-5918 (Home 319-266-5379) Office Hours: M: 8-11, 1:10-2, T: 9:30-10:30, 1:10-2, W: 1:10-2, Th: 9:30-10:30, 1:10-2 Zoom Meeting: https://uni.zoom.us/j/3192735918 Password: UNI

Credit Hours: 4 hours. This course meets the Course Credit Hour Expectation outlined in the Course Catalog. Students should expect to work approximately 2 hours per week outside of class for every course credit hour.

Prerequisite: Intro. to Computing (CS 1510), and pre- or corequisite Discrete Structures (CS 1800)

Student Learning Outcomes: After this course you should be able to:
1. implement efficient data structures include: stacks, queues, lists, heaps, hash tables, trees, and graphs
2. design and implement "medium" sized programs using functional decomposition and be able to select appropriate data structures.

Course Description: Introduction to use and implementation of data and file structures such as sets, hash tables, stacks, trees, queues, heaps, and graphs. Basic algorithm analysis. Searching and sorting. Basic object-oriented analysis, design, and modeling tools.


Assignments: Assignments will consist of weekly laboratory exercises along with concurrent weekly or bi-weekly programming assignments.

Pedagogic Approach: In class, I will tend to break up the lecture with active learning exercises to aid learning. Students benefit by (1) increased depth of understanding, (2) increased comfort and confidence, and (3) increased motivation. This might sound great, but it will require you (and me) to work differently to prepare for class. Before the class, you must read the assigned reading, thought about what I asked you to think about, etc.; otherwise, you won’t be able to effectively participate during class.

Attendance Policy: I encourage you to attend class and lab if you are well and feel safe. I think you learn more by attending. However, I will be posting Panopto videos shortly after each class period on eLearning.

Grading policy: There will be three tests (including the final). Tentative test dates and weighting of course components are:

<table>
<thead>
<tr>
<th>Component</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labs</td>
<td>20%</td>
</tr>
<tr>
<td>Homework Assignments</td>
<td>20%</td>
</tr>
<tr>
<td>In-class Test 1:</td>
<td>20%</td>
</tr>
<tr>
<td>In-class Test 2:</td>
<td>20%</td>
</tr>
<tr>
<td>Final Exam</td>
<td>20%</td>
</tr>
</tbody>
</table>

(Complete within 1 week and submit on eLearning) (Typically 2 weeks in duration and submit on eLearning) (Thursday, September 24) (Thursday, October 29) (Tuesday, November 24 from 8 - 9:50 AM in ITT 328)

Grades will be assigned based on straight percentages off the top student score. If the top student's score is 92%, then the grading scale will be, i.e., 100-82 A, 81.9-72 B, 71.9-62 C, 61.9-52 D, and below 52 F. Plus and minus grades will be assigned for students near cutoff points.

PPE/Personal Protective Equipment: Protecting our campus from COVID-19 depends on all of us acting with care and responsibility. To protect each other and our campus community, we are required to wear masks or face shields that cover our mouths and noses inside all campus buildings, including throughout the duration of class. We are asked to self-screen for COVID-19 symptoms, stay away from others and seek medical attention if we’re not feeling well and/or experience any symptoms such as a fever over 100.4, and to communicate and plan proactively to make up for missed learning. We will maintain physical distancing by sitting in designated areas in the classroom. Failure to follow these requirements can result in students being referred to the student conduct process and faculty being referred to the Associate Provost for Faculty. We take these steps together recognizing that my
mask protects you, your mask protects me, and together wearing masks protects the entire UNI community. Our collective actions will determine our ability to remain together in an in-person learning environment.

**Scholastic Conduct:** You are responsible for being familiar with the University' Academic Ethics Policies (http://www.uni.edu/pres/policies/301.shtml). Copying from other students is expressly forbidden. Doing so on exams or assignments will be penalized every time it is discovered. The penalty can vary from zero credit for the copied items (first offense) up to a failing grade for the course. If an assignment makes you realize you don't understand the material, ask questions designed to improve your understanding, not ones designed to discover how another student solved the assignment. The solutions to assignments should be individual, original work unless otherwise specified. Remember: discussing assignments is good. Copying code or test-question answers is cheating.

Any substantive contribution to your assignment solution by another person or taken from a publication (or the web) should be properly acknowledged in writing. Failure to do so is plagiarism and will necessitate disciplinary action. In addition to the activities we can all agree are cheating (plagiarism, bringing notes to a closed book exam, texting during an exam, etc.), assisting or collaborating on cheating is cheating (e.g., supplying code for another student). Cheating can result in failing the course and/or more severe disciplinary actions.

**The Learning Center @ Rod Library:** All students are encouraged to use The Learning Center @ Rod Library for assistance with writing, math, science, and college reading and learning strategies. Beginning week two, The Learning Center (TLC) operates on a walk-in basis and is open 10:00 am to 10:00 pm Monday through Thursday. For more information, go to https://tlc.uni.edu/tutoring, email TheLearningCenter@uni.edu, call 319-273-6023, or visit the TLC desk located on the main floor of Rod Library. If you are unable to come in during normal tutoring hours, online tutoring is available through Smarthinking. You will need your CATID and passphrase to gain access. To access the Smarthinking platform go to https://tlc.uni.edu/schedule.

**Non-Discrimination:** The University of Northern Iowa does not discriminate in employment or education. Visit https://policies.uni.edu/1303 for additional information.

**Student Accessibility Services:** The University of Northern Iowa (UNI) complies with the Americans with Disabilities Act Amendments Act of 2008 (ADAAA), Section 504 of the Rehabilitation Act of 1973, the Fair Housing Act, and other applicable federal and state laws and regulations that prohibit discrimination on the basis of disability. To request accommodations please contact Student Accessibility Services (SAS), located at ITTC 007, for more information either at (319) 273-2677 or Email accessibilityservices@uni.edu. Visit Student Accessibility Services (https://sas.uni.edu/) for additional information.

### Data Structures Reading/Class Schedule Fall 2020

NOTE: Chapters below are for the online textbook which split chapter 2 in the paper textbook into chapters 2 and 3.

<table>
<thead>
<tr>
<th>Lect #</th>
<th>Tuesday</th>
<th>Thursday</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8/18</td>
<td>8/20</td>
</tr>
<tr>
<td>2</td>
<td>Ch. 1: Python Review and functions</td>
<td>Ch. 2: Classes and Ch. 3: Algorithm Analysis; Big-oh; timing</td>
</tr>
<tr>
<td>3</td>
<td>8/25</td>
<td>8/27</td>
</tr>
<tr>
<td>4</td>
<td>Ch. 3: Performance of Python Built-in data structures</td>
<td>Ch. 4: Linear data structures; stack implementations</td>
</tr>
<tr>
<td>5</td>
<td>9/1</td>
<td>9/3</td>
</tr>
<tr>
<td>6</td>
<td>Stack applications; Queue implementations</td>
<td>Queue applications; Deque</td>
</tr>
<tr>
<td>7</td>
<td>9/8</td>
<td>9/10</td>
</tr>
<tr>
<td>8</td>
<td>Ch. 7.8 (6.6): Priority Queue: binary heap implementation</td>
<td>Unordered Lists implementations</td>
</tr>
<tr>
<td>9</td>
<td>9/15</td>
<td>9/17</td>
</tr>
<tr>
<td>10</td>
<td>Ordered List implementation</td>
<td>Recursion examples: Coin-change problem</td>
</tr>
<tr>
<td>11</td>
<td>Ch. 5: Recursion; Run-time stack</td>
<td>Test 1</td>
</tr>
<tr>
<td>12</td>
<td>9/22</td>
<td>9/24</td>
</tr>
<tr>
<td>13</td>
<td>Review for Test 1</td>
<td>Ch. 6: Searching: linear and binary search Hashing: Open-address with rehashing</td>
</tr>
<tr>
<td>14</td>
<td>9/29</td>
<td>10/1</td>
</tr>
<tr>
<td>15</td>
<td>Backtracking vs. dynamic programming coin-change problem</td>
<td>Ch. 6: Searching: linear and binary search Hashing: Open-address with rehashing</td>
</tr>
<tr>
<td>16</td>
<td>10/6</td>
<td>10/8</td>
</tr>
<tr>
<td>17</td>
<td>Chaining/closed-address hashing Dictionary implementations</td>
<td>Open-address implementation; Simple sorts: bubble, selection, and insertion sorts</td>
</tr>
<tr>
<td>Lect #</td>
<td>Tuesday</td>
<td>Thursday</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>17</td>
<td>10/13</td>
<td>10/15</td>
</tr>
<tr>
<td></td>
<td>Advanced sorts: heap, merge and quick sorts</td>
<td>Ch. 7: tree terminology, binary tree implementation, parse tree application, and tree traversals</td>
</tr>
<tr>
<td>19</td>
<td>10/20</td>
<td>10/22</td>
</tr>
<tr>
<td></td>
<td>Binary Search Tree implementation</td>
<td>Binary Search Tree delete method</td>
</tr>
<tr>
<td>21</td>
<td>10/27</td>
<td>10/29</td>
</tr>
<tr>
<td></td>
<td>Review for Test 2</td>
<td>Test 2</td>
</tr>
<tr>
<td>23</td>
<td>11/3</td>
<td>11/5</td>
</tr>
<tr>
<td></td>
<td>AVL height-balanced trees</td>
<td>File structures vs. In-memory Data Structures B+ trees</td>
</tr>
<tr>
<td>25</td>
<td>11/10</td>
<td>11/12</td>
</tr>
<tr>
<td></td>
<td>Ch. 8: Graph terminology, traversals (BFS &amp; DFS) Graph implementations</td>
<td>Select Graph algorithms: topological sort and Dijkstra’s algorithm</td>
</tr>
<tr>
<td>27</td>
<td>11/17</td>
<td>11/19</td>
</tr>
<tr>
<td></td>
<td>Select Graph algorithms: Prim’s algorithm; Modifications to BinHeap</td>
<td>Review for Final Exam</td>
</tr>
</tbody>
</table>

Final: 8:00 - 9:50 AM Tuesday, November 24 in ITT 328

Data Structures Lab Schedule Fall 2020

<table>
<thead>
<tr>
<th>Lab #</th>
<th>Date</th>
<th>Wednesday (Wright Hall 112)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8/19</td>
<td>Python Review and functions</td>
</tr>
<tr>
<td>2</td>
<td>8/26</td>
<td>Big-oh and Classes</td>
</tr>
<tr>
<td>3</td>
<td>9/2</td>
<td>FIFO Queue Implementations</td>
</tr>
<tr>
<td>4</td>
<td>9/9</td>
<td>Priority Queue Implementations</td>
</tr>
<tr>
<td>5</td>
<td>9/16</td>
<td>Recursion</td>
</tr>
<tr>
<td>6</td>
<td>9/23</td>
<td>No new lab - catch-up session since Test 1 on 9/24</td>
</tr>
<tr>
<td>7</td>
<td>9/30</td>
<td>Divide-and-Conquer vs. Dynamic Programming</td>
</tr>
<tr>
<td>8</td>
<td>10/7</td>
<td>Dictionary Implementations using Hashing</td>
</tr>
<tr>
<td>9</td>
<td>10/14</td>
<td>Sorting</td>
</tr>
<tr>
<td>10</td>
<td>10/21</td>
<td>Binary Search Trees</td>
</tr>
<tr>
<td>11</td>
<td>10/28</td>
<td>No new lab - catch-up session since Test 2 is on 10/28</td>
</tr>
<tr>
<td>12</td>
<td>11/4</td>
<td>AVL Tree put Implementation</td>
</tr>
<tr>
<td>13</td>
<td>11/11</td>
<td>Graph implementation and Traversals</td>
</tr>
<tr>
<td>14</td>
<td>11/18</td>
<td>Graph Algorithms: Prim’s and Dijkstra’s Algorithm; Improvements to BinHeap</td>
</tr>
</tbody>
</table>

Final: 8:00 - 9:50 AM Tuesday November 24 in ITT 328 so NO lab Wednesday November 25

Data Structures Homework Schedule Fall 2020

<table>
<thead>
<tr>
<th>HW #</th>
<th>Start</th>
<th>Due Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8/18</td>
<td>Saturday 8/29 at 11:59 PM</td>
<td>Math Tutor Program</td>
</tr>
<tr>
<td>2</td>
<td>8/30</td>
<td>Saturday 9/12 at 11:59 PM</td>
<td>JUMBLE Solver Program</td>
</tr>
<tr>
<td>3</td>
<td>9/13</td>
<td>Saturday 9/26 at 11:59 PM</td>
<td>Cursor-based List Class</td>
</tr>
<tr>
<td>4</td>
<td>9/27</td>
<td>Saturday 10/10 at 11:59 PM</td>
<td>Text-Editor Program</td>
</tr>
<tr>
<td>5</td>
<td>10/11</td>
<td>Saturday 10/24 at 11:59 PM</td>
<td>Improved Insertion and Heap Sorts</td>
</tr>
<tr>
<td>6</td>
<td>10/25</td>
<td>Saturday 11/14 at 11:59 PM</td>
<td>Word-Concordance and Dictionary Comparison</td>
</tr>
</tbody>
</table>

All late labs and homeworks must be submitted by NOON on Wednesday November 25!
**Python Summary**

### Print Function:

The `print()` function takes a list of values to print and writes them to the output, e.g.,

```python
print('cat', 5, 'dog')
print()  # blank line
print('pi is about', 3.14)
```

<table>
<thead>
<tr>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>cat 5 dog</td>
</tr>
<tr>
<td>pi is about 3.14</td>
</tr>
</tbody>
</table>

Optional keyword arguments can be used to replace the defaults: space-character (" ") as a separator, the new-line-character ("
") as the ending character, and output file of the console (`sys.stdout`). The syntax with default parameters explicitly shown is:

```python
print(value, ..., sep=' ', end='
', file=sys.stdout)
```

<table>
<thead>
<tr>
<th><strong>Print Function</strong></th>
<th><strong>Expected Output</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><code>print('cat', 5, end='')</code></td>
<td>cat 5</td>
</tr>
<tr>
<td><code>print('horse')</code></td>
<td>horse</td>
</tr>
<tr>
<td><code>print('cow')</code></td>
<td>cow</td>
</tr>
<tr>
<td><code>print ('cat', 5, 'dog', sep='23', end='#')</code></td>
<td>cat23523dog#</td>
</tr>
<tr>
<td><code>print ('&lt;', end='')</code></td>
<td>&lt;cat&gt;, &lt;5&gt;, &lt;dog&gt;</td>
</tr>
</tbody>
</table>
| `print ('cat', 5, 'dog', sep='>', < ', end='>
')` | error since keyword arguments must be at the end of the parameter list |
| `print ('cat', 5, 'dog', sep='23', 'horse')` | |

### String Formatting:

Inside a string we use formatting placeholders (e.g., `%d`, `%s`, `%5.2f`), follow the string with the format operator (%), and a tuple supplying values for corresponding placeholders. For example:

```python
print("Name: %s Age: %d GFA: %.2f" % ("Bob", 20, 3.14))
```

outputs the line:

| Name: Bob | Age: 20 | GFA: 3.14 |

%s means left-justify string in 10 spaces, %d means right-justify decimal/int in 7 spaces, and %.2f left-justify float using the minimum spaces but with 2 decimal places. Another example with a character and exponent:

```python
print("character: %c float with exponent %e" % (',', 123.456))
```

outputs the line:

| character: $ | float with exponent 1.234560e+02 |

### Assignment Statement:

The assignment statement creates a variable in memory and sets its value. The syntax is:

<variable identifier> = <constant or expression value>, where identifiers must start with a letter or underscore ("_") and then can be followed by letters, underscores, or digits.

<table>
<thead>
<tr>
<th>Assignment Statements and Print Functions</th>
<th>Expected Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>a = 123</code></td>
<td>a is 123</td>
</tr>
<tr>
<td><code>b = a</code></td>
<td>b is 123</td>
</tr>
<tr>
<td><code>a = a + 1</code></td>
<td>c is catfish</td>
</tr>
<tr>
<td><code>d = c</code></td>
<td>d is cat</td>
</tr>
<tr>
<td><code>c = 'cat'</code></td>
<td>e = ['cat', 'dog', 'cow']</td>
</tr>
<tr>
<td><code>d = c + 'fish'</code></td>
<td>f = ['cat', 'dog', 'cow']</td>
</tr>
<tr>
<td><code>print('a is', a)</code></td>
<td></td>
</tr>
<tr>
<td><code>print ('b is', b)</code></td>
<td></td>
</tr>
<tr>
<td><code>print('c is', c)</code></td>
<td></td>
</tr>
<tr>
<td><code>print('d is', d)</code></td>
<td></td>
</tr>
<tr>
<td><code>print('e is', e)</code></td>
<td></td>
</tr>
<tr>
<td><code>print('f is', f)</code></td>
<td></td>
</tr>
</tbody>
</table>

NOTE: This last example deals with assigning lists. In `e = ['cat', 'dog']` the variable `e` is assigned a reference/pointer to the list, so `f = e` assigns `f` a reference to the same list. There is only a single list! Thus, when we append 'cow' to the list using `e`'s reference, it is also printed in `f`'s list because it's the same list.

The first two examples deal with integers and strings which are immutable (i.e., unchangeable). New immutable values are created with new references being assigned. After `b = a` both variables reference 123. When `a = a + 1` executes, a new integer constant of 124 is created and its reference is assign to variable `a`. Variable `b` still references 123.
**Input Function:** the `input()` function reads a line from the keyboard (`sys.stdin`) and returns it as a string with the trailing new-line stripped. To input numeric values, the string needs to be explicitly cast (e.g. `eval(input())`). For example, we can input and echo the user’s name and age.

```python
name = input("Enter your name: ")
age = eval(input("Enter your age: "))
print('Hi', name, end='!')
print('Your age is', age)

Enter your name: Bob
Enter your age: 10
Hi Bob! Your age is 10
```

**Control Statements:** the body of control statements are indented and there is NO other "end" ("}") marker

**if statements:** An if statement allows code to be executed or not based on the result of a comparison. If the condition evaluates to True, then the statements of the **indented body** is executed. If the condition is False, then the body is skipped. The syntax of if statements is:

```python
if <condition>:
    statement_1
    statement_2
    statement_3

elif <condition2>:
    statement
    statement
else:
    statement_1
    statement_2
```

Typically, the condition involves comparing “stuff” using relational operators (\(<\), \(\geq\), \(\leq\), \(\neq\), \(!=\)).

Complex conditions might involve several comparisons combined using Boolean operators: not, or, and. For example, we might want to print “Your grade is B.” if the variable score is less than 90, but greater than or equal to 80.

```python
if score < 90 and score >= 80:
    print("Your grade is B.")
```

The precedence for mathematical operators, Boolean operators, and comparisons are given in the table.

<table>
<thead>
<tr>
<th>Operator(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>** (exponential)</td>
</tr>
<tr>
<td>+, - (unary pos. \&amp; neg.)</td>
</tr>
<tr>
<td>*, /, % (rem), // (integer div)</td>
</tr>
<tr>
<td>+, - (add, sub)</td>
</tr>
<tr>
<td>&lt;, &gt;, ==, &lt;=, &gt;=, !=, is, is not</td>
</tr>
<tr>
<td>not and or</td>
</tr>
<tr>
<td>= (assignment)</td>
</tr>
</tbody>
</table>

**for loop:** the **for loop** iterates once for each item in some sequence type (i.e. list, tuple, string).

```python
for value in [1, 3, 9, 7]:
    print(value)
```

```python
for character in 'house':
    print( character )
```

The for loop iterates over an iterable data-structure object (list, string, dictionary) or a range object created by the built-in `range` function which generate each value one at a time for each iteration of the loop. The syntax of: `range([start,] end, [step])`, where [ ] are used to denote optional parameters. Examples:

- `range(5)` generates the sequence of values: `0, 1, 2, 3, 4`
- `range(2, 7)` generates the sequence of values: `2, 3, 4, 5, 6`
- `range(10, 2, -1)` generates the sequence of values: `10, 9, 8, 7, 6, 5, 4, 3`

For example:

```python
for count in range(1,6):
    print(count, end= " ")
print("\nDone")
```

```
1 2 3 4 5
Done
```
**Python Summary**

**while loop:** A while statement allows code to be executed repeated (zero or more times) as long as the condition evaluates to True. The syntax of a while statement is:

```
while <condition>:
    statement_1
    statement_2
    statement_3
```

An *infinite loop* is one that would loop forever. (FYI, in a Python shell ctrl-c (^c) can be used to kill the running program.) Most infinite loops are caused by programmer error, but sometimes they are intentional. The following “sentinel-controlled” code uses an infinite loop and a `break` statement that immediately causes control to exit the loop.

```python
total = 0
counter = 0
while True:  # an infinite loop
    score = eval(input("Enter a score (or negative value to exit): "))
    if score < 0:
        break
    total += score
    counter += 1
print("Average is", total/counter)
```

**Strings:** Strings in Python are sequential collections of only characters. Strings are *immutable* (i.e., cannot be changed), so new strings are generated by string operations. Operations on strings (or any sequence collection) include:

<table>
<thead>
<tr>
<th>Operation</th>
<th>Operator</th>
<th>Explanation</th>
<th>Example</th>
<th>Result of Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indexing</td>
<td>[ &lt;index&gt; ]</td>
<td>Access the element specified by the index</td>
<td>myString[1]</td>
<td>'e'</td>
</tr>
<tr>
<td>Slicing</td>
<td>[ : ]</td>
<td>Extract a part of the string</td>
<td>myString[1:5]</td>
<td>'ello'</td>
</tr>
<tr>
<td>Concatenation</td>
<td>+</td>
<td>Combine strings together</td>
<td>myString + aString</td>
<td>'Hello!!!cat'</td>
</tr>
<tr>
<td>Repetition</td>
<td>*</td>
<td>Concatenate a repeated number of times</td>
<td>aString * 3</td>
<td>'catcatcatcat'</td>
</tr>
<tr>
<td>Membership</td>
<td>in</td>
<td>Ask whether a substring is in a string</td>
<td>'ell' in myString</td>
<td>True</td>
</tr>
<tr>
<td><strong>…<code>len</code></strong></td>
<td><code>len(string)</code></td>
<td>How many items are in the string?</td>
<td><code>len(myString)</code></td>
<td>8</td>
</tr>
</tbody>
</table>

Indexing of strings starts with 0 on the left end, and -1 on the right end:

```
  1111
  01234567890123
cheer = 'GO Panthers!!!'
```

Omitted indexes in a slice means “from the end.” For example, `cheer[:4]` generates ‘GO P’.

Omitted indexes in a slice means “from the end.” For example, `cheer[-4:]` generates ‘s!!!’.
String objects also have the following methods: (the string module can be imported to provide more operations.)

<table>
<thead>
<tr>
<th>Method</th>
<th>Usage</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>center</td>
<td>myString.center(w)</td>
<td>Returns a string with myString centered in a field of size w</td>
</tr>
<tr>
<td>ljust</td>
<td>myString.ljust(w)</td>
<td>Returns a string with myString left-justified in a field of size w</td>
</tr>
<tr>
<td>rjust</td>
<td>myString.rjust(w)</td>
<td>Returns a string with myString right-justified in a field of size w</td>
</tr>
<tr>
<td>upper</td>
<td>myString.upper()</td>
<td>Returns a string with myString in all upper-case characters</td>
</tr>
<tr>
<td>lower</td>
<td>myString.lower()</td>
<td>Returns a string with myString in all lower-case characters</td>
</tr>
<tr>
<td>strip</td>
<td>myString.strip()</td>
<td>Returns a string with leading and trailing whitespace (space, tab, new-line) chars. removed. An optional string parameter can be used to supply characters to strip instead of whitespace.</td>
</tr>
<tr>
<td>count</td>
<td>myString.count(sub)</td>
<td>Returns number of occurrences of sub in myString (Optional parameters: myString.count(sub [, start [, end ]])</td>
</tr>
<tr>
<td>endswith</td>
<td>myString.endswith(sub)</td>
<td>Returns True if myString ends with the substring sub; otherwise it returns False</td>
</tr>
<tr>
<td>startswith</td>
<td>myString.startswith(sub)</td>
<td>Returns True if myString starts with the substring sub; otherwise it returns False</td>
</tr>
<tr>
<td>isdigit</td>
<td>myString.isdigit()</td>
<td>Returns True if myString contains only digits; otherwise it returns False</td>
</tr>
<tr>
<td>isalpha</td>
<td>myString.isalpha()</td>
<td>Returns True if myString contains only letters; otherwise it returns False</td>
</tr>
<tr>
<td>split</td>
<td>myString.split()</td>
<td>Returns a list of substrings of myString splits at whitespace characters. An optional string parameter can supply characters to split on.</td>
</tr>
<tr>
<td>find</td>
<td>myString.find(sub)</td>
<td>Returns the starting index of the first occurrence of sub. (Optional parameters: myString.find(sub [, start [, end ]])</td>
</tr>
<tr>
<td>replace</td>
<td>myString.replace(old,new)</td>
<td>Returns a string with all occurrences of substring old replaced by substring new. An additional integer parameter can specify the number of replacements to perform, e.g., myString.replace(old,new, 3)</td>
</tr>
</tbody>
</table>

Lists: A Python list is also a sequence collection, but a list can contain items of any type (e.g., character, strings, integers, floats, other lists, etc.), and lists are mutable. Lists are represented by comma-separated values enclosed in square brackets ("[", "]"). Operations on lists (or any sequence collection, e.g., strings) include:

<table>
<thead>
<tr>
<th>Operation</th>
<th>Operator</th>
<th>Explanation</th>
<th>Example</th>
<th>Result of Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slicing</td>
<td>[ :: ]</td>
<td>Extract a part of the list</td>
<td>myList[ 1:3 ]</td>
<td>[6, 7]</td>
</tr>
<tr>
<td>Concatenation</td>
<td>+</td>
<td>Combine lists together</td>
<td>myList + ListB</td>
<td>[5, 6, 7, 8, 8, 9 ]</td>
</tr>
<tr>
<td>Repetition</td>
<td>*</td>
<td>Concatenate a repeated number of times</td>
<td>ListB * 3</td>
<td>[8, 9, 8, 9, 8, 9]</td>
</tr>
<tr>
<td>Membership</td>
<td>in</td>
<td>Ask whether an item is in a list</td>
<td>3 in myList</td>
<td>False</td>
</tr>
<tr>
<td>Length</td>
<td>len(list)</td>
<td>How many items are in the list?</td>
<td>len( myList )</td>
<td>4</td>
</tr>
</tbody>
</table>
The following list methods are provided by Python:

<table>
<thead>
<tr>
<th>Method</th>
<th>Usage</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>append</td>
<td>myList.append(item)</td>
<td>Adds item to the end of myList</td>
</tr>
<tr>
<td>extend</td>
<td>myList.extend(otherList)</td>
<td>Extends myList by adding all items in otherList to myList's end</td>
</tr>
<tr>
<td>insert</td>
<td>myList.insert(i, item)</td>
<td>Insert item in myList at index i</td>
</tr>
<tr>
<td>pop</td>
<td>myList.pop()</td>
<td>Remove and return the last item in myList</td>
</tr>
<tr>
<td>pop(i)</td>
<td>myList.pop(i)</td>
<td>Remove and return the i'th item in myList</td>
</tr>
<tr>
<td>del</td>
<td>del myList[i]</td>
<td>Deletes the item in i'th position of myList</td>
</tr>
<tr>
<td>remove</td>
<td>myList.remove(item)</td>
<td>Removes the first occurrence of item in myList</td>
</tr>
<tr>
<td>index</td>
<td>myList.index(item)</td>
<td>Returns the index of the first occurrence of item in myList</td>
</tr>
<tr>
<td>count</td>
<td>myList.count(item)</td>
<td>Returns the number of occurrences of item in myList</td>
</tr>
<tr>
<td>sort</td>
<td>myList.sort()</td>
<td>Modifies myList to be sorted</td>
</tr>
<tr>
<td>reverse</td>
<td>myList.reverse()</td>
<td>Modifies myList to be in reverse order</td>
</tr>
</tbody>
</table>

* Note: raises an IndexError if the index i is not in the list  
** Note: raises a ValueError if the item is not in the list

**Tuples:** A tuple is another sequence data type, so the sequence operations of indexing, slicing, concatenation, repetition, membership (in), and len() work on tuples too. Tuples are very similar to lists, i.e., comma-separated items enclosed in parentheses. The main difference is that tuples are immutable (cannot be modified).

Create two tuples as:

```python
student1 = ('Bob', 123456, 'Jr', 3.12)  
student2 = 'Sally', 654321, 'Fr', 0.0
```

In addition to indexing, "fields" of a tuple can be unpacked using a single assignment statement as:

```python
name, idnum, rank, gpa = student1
```

(NOTE: This allows multiple values to be returned from a function)

**Dictionaries:** A dictionary is an unordered set of key-value pairs (written as key:value). Keys must be unique and immutable (e.g., numerics, strings, tuples of immutable objects). Dictionaries are typically used to look up the value corresponding to a specified key. Dictionaries can be written as comma-separated key:value pairs enclosed in curly braces. For example,

```python
phoneNumbers = {'fienup':35918,'gray':35917,'east':32939,'drake':35811,'schafer':32187}
```

Access to individual key:value pairs looks syntactically like a sequence lookup using a key instead of an index. For example, `phoneNumbers['east']` returns 32939, and a new key:value pair can be added by `phoneNumbers['wallingford'] = 35919`. Additional, methods on dictionaries are:

<table>
<thead>
<tr>
<th>Method</th>
<th>Usage</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>keys</td>
<td>myDictionary.keys()</td>
<td>Returns keys in an iterable dict_keys object</td>
</tr>
<tr>
<td>values</td>
<td>myDictionary.values()</td>
<td>Returns values in an iterable dict_values object</td>
</tr>
<tr>
<td>items</td>
<td>myDictionary.items()</td>
<td>Returns key:value tuples in an iterable dict_items object</td>
</tr>
<tr>
<td>get</td>
<td>value = myDictionary[myKey]</td>
<td>Returns the value associated with myKey; otherwise raises a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KeyError if myKey is not in the dictionary</td>
</tr>
<tr>
<td>set</td>
<td>myDictionary[myKey] = value</td>
<td>Inserts a key-value entry if myKey does not exist or replaces</td>
</tr>
<tr>
<td></td>
<td></td>
<td>the old value with value if myKey exists.</td>
</tr>
<tr>
<td>get</td>
<td>myDictionary.get(myKey)</td>
<td>Returns the value associated with myKey; otherwise None</td>
</tr>
<tr>
<td>get</td>
<td>myDictionary.get(myKey, alt)</td>
<td>Returns the value associated with myKey; otherwise alt</td>
</tr>
<tr>
<td>in</td>
<td>myKey in myDictionary</td>
<td>Returns True if myKey is in myDictionary; otherwise False</td>
</tr>
<tr>
<td>del</td>
<td>del myDictionary[myKey]</td>
<td>Deletes the key:value pair whose key is myKey</td>
</tr>
</tbody>
</table>
Text Files: Below is a summary of the important text-file operations in Python.

<table>
<thead>
<tr>
<th>General syntax</th>
<th>Example</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>open(filename)</td>
<td><code>f = open('data.txt', 'w')</code></td>
<td>Modes: ‘r’ read only; ‘w’ write only; ‘a’ append; ‘r+’ both reading and writing. Default mode is ‘r’</td>
</tr>
<tr>
<td>open(filename, mode)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f.close()</td>
<td><code>f.close()</code></td>
<td>Close the file to free up system resources.</td>
</tr>
<tr>
<td>loop over the file</td>
<td><code>for line in f:</code></td>
<td>Memory efficient, fast and simple code to loop over each line in the file.</td>
</tr>
<tr>
<td>object</td>
<td><code>print (line)</code></td>
<td></td>
</tr>
<tr>
<td>f.readline()</td>
<td><code>nextLine = f.readline()</code></td>
<td>Returns the next line from the file. The newline (‘\n’) character is left at the end of the string.</td>
</tr>
<tr>
<td>f.write(string)</td>
<td><code>f.write('cats and dogs\n')</code></td>
<td>Writes the string to the file.</td>
</tr>
<tr>
<td>f.read()</td>
<td><code>all = f.read()</code></td>
<td>Returns the whole file as a single string.</td>
</tr>
<tr>
<td>f.read(size)</td>
<td><code>chunk = f.read(100)</code></td>
<td>Returns a string of at most 100 (size) bytes. If the file has been completely read, an empty string is returned.</td>
</tr>
<tr>
<td>f.readlines()</td>
<td><code>allLines = f.readlines()</code></td>
<td>Returns a list containing all the lines of the file.</td>
</tr>
<tr>
<td>f.readlines(size)</td>
<td><code>someLines = f.readlines(5000)</code></td>
<td>Returns the next 5000 bytes of line. Only complete lines will be returned.</td>
</tr>
</tbody>
</table>

Below is a summary of the important file-system functions from the os and os.path modules in Python.

### os Module File-system Functions

<table>
<thead>
<tr>
<th>General syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>getcwd()</td>
<td>Returns the complete path of the current working directory</td>
</tr>
<tr>
<td>chdir(path)</td>
<td>Changes the current working directory to path</td>
</tr>
<tr>
<td>listdir(path)</td>
<td>Returns a list of the names in directory named path</td>
</tr>
<tr>
<td>mkdir(path)</td>
<td>Creates a new directory named path and places it in the current working directory</td>
</tr>
<tr>
<td>rmdir(path)</td>
<td>Removes the directory named path from the current working directory</td>
</tr>
<tr>
<td>remove(path)</td>
<td>Removes the file named path from the current working directory</td>
</tr>
<tr>
<td>rename(old, new)</td>
<td>Renames the file or directory named old to new</td>
</tr>
</tbody>
</table>

### os.path Module File-system Functions

<table>
<thead>
<tr>
<th>General syntax</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>exists(path)</td>
<td>Returns True if path exists and False otherwise</td>
</tr>
<tr>
<td>isdir(path)</td>
<td>Returns True if path is a directory and False otherwise</td>
</tr>
<tr>
<td>.isfile(path)</td>
<td>Returns True if path is a file and False otherwise</td>
</tr>
<tr>
<td>getsize(path)</td>
<td>Returns the size in bytes of the object named path</td>
</tr>
</tbody>
</table>

NOTE: The initial “current working directory” is the directory where the program is located. Typically, it is useful to access files relative to the “current working directory” instead of specifying an absolute (complete) path. You can use the strings:
- `'.'` to specify the current working directory, e.g., `currentDirectoryList = os.listdir('.' )`  
- `'/.'` to specify the parent of current working directory, e.g., `os.chdir('/.')` which changes the current working directory to the parent directory
Functions:
A function is a procedural abstract, i.e., a named body of code that performs some task when it is called/invoked. Often a function will have one or more parameter that allows it to perform a more general (variable) task. For example, the cube function below can be called with any numeric value with the corresponding cube of that number being returned.

```python
# Function to calculate the cube of a number
def cube(num):
    num_squared = num * num
    return num_squared * num

# call the function
value = 2
print('The value', value, 'raised to the power 3 is', cube(value))
print('The value 3 raised to the power 3 is', cube(3))
```

Terminology:
- a formal parameter is the name of the variable used in the function definition. It receives a value when the function is called. In the function cube, num is the formal parameter. Formal parameters are only known inside of the function definition. The section of a program where a variable is known is called its scope, so the scope of a formal parameter (and other local variable defined in the function such as num_squared) is limited to the function in which it is defined.
- an actual parameter/argument is the value used in the function call that is sent to the function. In the call to function cube, the variable value supplies the actual parameter value of 2.
- a global variable is created outside all functions and is known throughout the whole program file, e.g. value.

It is helpful to understand the “rules of the game” when a function is called. Memory is used to store the current program and the data associated with it. The memory used to store the data is divided as shown below.
- Global memory is used to store the global variables (and constants).
- The heap is used to store dynamically allocated objects as the program runs, e.g. lists, strings, ints, objects
- The run-time stack is used to store call-frames (or activation records) that get pushed on the stack when a function is called, and popped off the stack when a function returns.

When a function is called the section of code doing the calling is temporarily suspended, and a new call-frame gets pushed on top of the stack before execution of the function body. The call-frame contains the following information about the function being called:
- the return address -- the spot in code where the call to the function occurred. This is needed so execution (control) can return there when the end of the function is reached or a return statement executes.
- room to store the formal parameters used by the function. In Python, parameters are passed-by-value which means that the value of each actual parameter in the function call is assigned to the corresponding formal parameter in the function definition before the function starts executing. However, the memory location for actual parameters for strings, lists, dictionaries, tuples, atomic objects contain only references to the heap
- room to store the local variables defined in the function (these are probably references to objects in the heap)

When a function returns, execution resumes at the function call (which is specified by the return address). A function typically sends back a value to the call by specifying an expression after return in the return statement. In Python if no expression is specified returned, then the special object None is returned.
def play(myInt, myLongInt, myList, myString):
    print('START OF play Function')
    print('myInt=',myInt,'myLongInt=',myLongInt)
    print('myList=',myList,'myString=',myString)
    myInt += 1
    myLongInt += 1
    myList.append(1)
    myString += 'a'
    print('END OF play Function')
    print('myInt=',myInt,'myLongInt=',myLongInt)
    print('myList=',myList,'myString=',myString)
    return

anInt=10  # 1st STATEMENT EXECUTED, play loaded in IDLE
aLongInt = 123456789012345678901234567890L
aList = range(5)
aString = 'hello'
print('BEFORE CALL')
print('anInt=',anInt,'aLongInt=',aLongInt)
print('aList=',aList,'aString=',aString)
play(anInt, aLongInt, aList, aString)
print('AFTER CALL')  (#)
print('anInt=',anInt,'aLongInt=',aLongInt)
print('aList=',aList,'aString=',aString)

Output of complete program:

```python
>>> BEFORE CALL
anInt= 10 aLongInt= 123456789012345678901234567890
aList= [0, 1, 2, 3, 4] aString= hello
START OF play Function
myInt= 10 myLongInt= 123456789012345678901234567890
myList= [0, 1, 2, 3, 4] myString= hello
END OF play Function
myInt= 11 myLongInt= 123456789012345678901234567891
myList= [0, 1, 2, 3, 4, 1] myString= helloa
AFTER CALL
anInt= 10 aLongInt= 123456789012345678901234567890
aList= [0, 1, 2, 3, 4, 1] aString= hello
```
Classes: A class definition is like a blueprint (recipe) for each of the objects of that class. A class specifies a set of data attributes and methods for the objects of that class:

- The values of the data attributes of a given object make up its state
- The behavior of an object depends on its current state and on the methods that manipulate this state
- The set of a class's methods is called its interface

The general syntax of class definition is:

```python
class MyClass ( ( superClass1 [, superClass2 ]* ) ) :
    '''Document comment which becomes the __doc__ attribute for the class'''
    def __init__(self, [param [, param]*) :
        '''Document comment for constructor method with self be referencing to the object itself'''
        #__init__ body

    # defs of other class methods and assignments to class attributes

# end class MyClass
```

```
***
File: simple_die.py
Description: This module defines a six-sided Die class.
***

from random import randint

class Die(object):
    """This class represents a six-sided die.""

def __init__(self):
    """The initial face of the die.""
    self.__currentRoll = randint(1, 6)

def roll(self):
    """Resets the die's value to a random number between 1 and 6.""
    self.__currentRoll = randint(1, 6)

def getRoll(self):
    """Returns the face value of the die.""
    return self.__currentRoll

def __str__(self):
    """Returns the string representation of the die.""
    return str(self.__currentRoll)
```

Consider the following script to test the Die class and its associated output:

```
# testDie.py - script to test Die class
from simple_die import Die

die1 = Die()
die2 = Die()
print('die1 =', die1)  #calls __str__
print('die2 =', die2)
print()
print('die1.getRoll() = ', die1.getRoll())
print('die2.getRoll() = ', die2.getRoll())
die1.roll()
print('die1.getRoll() = ', die1.getRoll())
print('str(die1): ' + str(die1))
print('die1 + die2: ', die1.getRoll() + die2.getRoll())

>>> die1 = 2
die2 = 5
die1.getRoll() = 2
die2.getRoll() = 5
die1.getRoll() = 3
str(die1): 3
die1 + die2: 8
```
Classes in Python have the following characteristics:

- **all class attributes (data attributes and methods) are public by default, unless your identifier starts with a single underscores, e.g., self._currentRoll.**
- **all data types are objects, so they can be used as inherited base classes**
- **objects are passed by reference when used as parameters to functions**
  - all classes have a set of standard methods provided, but may not work properly (str, doc, etc.)
  - most built-in operators (+, -, *, <, >, ==, etc.) can be redefined for a class. This makes programming with objects a lot more intuitive. For example suppose we have two Die objects: die1 & die2, and we want to add up their combined rolls. We could use accessor methods to do this:
    
    ```python
diceTotal = die1.getRoll() + die2.getRoll()
```
  
  Here, the getRoll method returns an integer (type int), so the ‘+’ operator being used above is the one for ints. But, it might be nice to “overload” the + operator by defining an __add__ method as part of the Die class, so the programmer could add dice directly as in:
    
    ```python
diceTotal = die1 + die2
```

The three most important features of **Object-Oriented Programming (OOP)** to simplify programs and make them maintainable are:

1. **encapsulation** - restricts access to an object's data to access only by its methods
   - helps to prevent indiscriminate changes that might cause an invalid object state (e.g., 6-side die with a of roll 8)

2. **inheritance** - allows one class (the subclass) to pickup data attributes and methods of other class(es) (the parents)
   - helps code reuse since the subclass can extend its parent class(es) by adding addition data attributes and/or methods, or overriding (through polymorphism) a parent's methods

3. **polymorphism** - allows methods in several different classes to have the same names, but be tailored for each class
   - helps reduce the need to learn new names for standard operations (or invent strange names to make them unique)

Consider using inheritance to extend the Die class to a generalized AdvancedDie class that can have any number of sides. The interface for the AdvancedDie class are:

<table>
<thead>
<tr>
<th>Method</th>
<th>Example Usage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>init</strong></td>
<td>myDie = AdvancedDie(8)</td>
<td>Constructs a die with a specified number of sides and randomly rolls it (Default of 6 sides if no argument supplied)</td>
</tr>
<tr>
<td>getRoll</td>
<td>myDie.getRoll()</td>
<td>Returns the current roll of the die (inherited from Die class)</td>
</tr>
<tr>
<td>getSides</td>
<td>myDie.getSides()</td>
<td>Returns the number of sides on the die (did not exist in Die class)</td>
</tr>
<tr>
<td>roll</td>
<td>myDie.roll()</td>
<td>Rerolls the die randomly (By overriding the roll method of Die, an AdvancedDie can generate a value based on its # of sides)</td>
</tr>
<tr>
<td><strong>eq</strong></td>
<td>if myDie == otherDie:</td>
<td>Allows == operations to work correctly for AdvancedDie objects.</td>
</tr>
<tr>
<td><strong>lt</strong></td>
<td>if myDie &lt; otherDie:</td>
<td>Allows &lt; operations to work correctly for AdvancedDie objects.</td>
</tr>
<tr>
<td><strong>gt</strong></td>
<td>if myDie &gt; otherDie:</td>
<td>Allows &gt; operations to work correctly for AdvancedDie objects.</td>
</tr>
<tr>
<td><strong>add</strong></td>
<td>sum = myDie + otherDie</td>
<td>Allows the direct addition of AdvancedDie objects, and returns the integer sum of their current roll values.</td>
</tr>
<tr>
<td><strong>str</strong></td>
<td>Directly as:</td>
<td>Returns a string representation for the AdvancedDie. By overriding the <strong>str</strong> method of the Die class, so the “print” statement will work correctly with an AdvancedDie.</td>
</tr>
</tbody>
</table>

myDie.__str__() or indirectly as: print myDie
Consider the following script and associated output:

```python
# testAdvancedDie.py - script to test
AdvancedDie class
from advanced_die import AdvancedDie

die1 = AdvancedDie(100)
die2 = AdvancedDie(100)
die3 = AdvancedDie()

print('die1 =', die1)  # calls __str__
print('die2 =', die2)
print('die3 =', die3)

print('die1.getRoll() = ', die1.getRoll())
print('die1.getSides() = ', die1.getSides())
die1.roll()
print('die1.getRoll() = ', die1.getRoll())
print('die2.getRoll() = ', die2.getRoll())
print('die1 == die2:', die1==die2)
print('die1 < die2:', die1<die2)
print('die1 > die2:', die1>die2)
print('die1 != die2:', die1!=die2)
print('str(die1): ' + str(die1))
print('die1 + die2:', die1 + die2)

help(AdvancedDie)
```

die1 = Number of Sides=100 Roll=32
die2 = Number of Sides=100 Roll=76
die3 = Number of Sides=6 Roll=5
die1.getRoll() = 32
die1.getSides() = 100

die1.getRoll() = 70
die2.getRoll() = 76
die1 == die2: False
die1 < die2: True
die1 > die2: False
die1 != die2: True
str(die1): Number of Sides=100 Roll=70
die1 + die2: 146

Help on class AdvancedDie in module advanced_die:

class AdvancedDie(simple_die.Die):
    | Advanced die class that allows for any number of sides
    | Method resolution order:
    | AdvancedDie
    |   simple_die.Die
    |     __builtin__ object
    | Methods defined here:

Notice that the testAdvancedDie.py script needed to import AdvancedDie, but not the Die class.

The initial AdvancedDie objects created can be visualized as:

![Diagram of Global Memory and Initial AdvancedDie objects]

<table>
<thead>
<tr>
<th>Global Memory</th>
<th>Initial AdvancedDie objects</th>
</tr>
</thead>
<tbody>
<tr>
<td>die1</td>
<td>_numSides: 100, _currentRoll: 32</td>
</tr>
<tr>
<td>die2</td>
<td>_numSides: 100, _currentRoll: 76</td>
</tr>
<tr>
<td>die3</td>
<td>_numSides: 6, _currentRoll: 5</td>
</tr>
</tbody>
</table>
The AdvancedDie class that inherits from the Die superclass.

```python
class AdvancedDie(Die):
    """Advanced die class that allows for any number of sides"""

    def __init__(self, sides = 6):
        """Constructor for any sided Die that takes an the number of sides
        as a parameter; if no parameter given then default is 6-sided."""
        Die.__init__(self)  # call Die parent class constructor
        self._numSides = sides
        self._currentRoll = randint(1, self._numSides)

    def roll(self):
        """Causes a die to roll itself — overrides Die class roll""
        self._currentRoll = randint(1, self._numSides)

    def __eq__(self, rhs_Die):
        """Overrides default '__eq__' operator to allow for deep comparison of Dice"
        return self._currentRoll == rhs_Die._currentRoll

    def __lt__(self, rhs_Die):
        """Overrides default '__lt__' operator to allow for deep comparison of Dice"
        return self._currentRoll < rhs_Die._currentRoll

    def __gt__(self, rhs_Die):
        """Overrides default '__gt__' operator to allow for deep comparison of Dice"
        return self._currentRoll > rhs_Die._currentRoll

    def __str__(self):
        """Returns the string representation of the AdvancedDie.""
        return 'Number of Sides='+str(self._numSides)+' Roll='+str(self._currentRoll)

    def __add__(self, rhs_Die):
        """Returns the sum of two dice rolls"
        return self._currentRoll + rhs_Die._currentRoll

    def getSides(self):
        """Returns the number of sides on the die.""
        return self._numSides
```

A better AdvancedDie class would check and enforce preconditions by raising appropriate exceptions (e.g., Exception, TypeError, IndexError, KeyError, ValueError, etc.). For example the `__init__`:

```python
class AdvancedDie(Die):
    def __init__(self, sides = 6):
        """Constructor for any-sided Die that takes the number of sides as a parameter;
        if no parameter is given then default is 6-sided.
        Precondition: sides is an integer >= 1
        Postcondition: current roll of the die between 1 and sides
        Raises: TypeError if sides in not an integer and ValueError if sides < 1"
        # Check preconditions
        if not isinstance(sides, int):
            raise TypeError("Die sides must be an integer!")
        if sides < 1:
            raise ValueError("Die sides must be a positive integer!")
        Die.__init__(self)  # call Die parent class constructor
        self._numSides = sides
        self._currentRoll = randint(1, self._numSides)
```

A better AdvancedDie class would check and enforce preconditions by raising appropriate exceptions (e.g., Exception, TypeError, IndexError, KeyError, ValueError, etc.). For example the `__init__`:
1. The print function has optional *keyword arguments* which can be listed last that modify it behavior. The print function syntax: `print(value,...,sep=' ',end='
', file=sys.stdout)

a) Predict the expected output of each of the following.

<table>
<thead>
<tr>
<th>Program</th>
<th>Expected Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>print('cat',5,'dog')</code></td>
<td></td>
</tr>
<tr>
<td><code>print()</code></td>
<td></td>
</tr>
<tr>
<td><code>print('cat',5,end='')</code></td>
<td></td>
</tr>
<tr>
<td><code>print('horse')</code></td>
<td></td>
</tr>
<tr>
<td><code>print('cow')</code></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Program</th>
<th>Expected Output</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>print ('cat',5,'dog',end='#',sep='23')</code></td>
<td></td>
</tr>
<tr>
<td><code>print ('cat',5,'dog',sep='23','horse')</code></td>
<td></td>
</tr>
<tr>
<td><code>print ('cat',5,'dog',sep='*3')</code></td>
<td></td>
</tr>
</tbody>
</table>

2. Review of assignment statements. Predict the output of the following programs

```python
a = 123
b = a
da += 1
print ('a is', a)
print ('b is', b)
c = ['cat', 'dog']
d = c
(c.append('cow'))
print('c is', c)
print('d is', d)
c = 'cat'
d = c
c += 'fish'
print('c is', c)
print('d is', d)
```

Most simple programs have a similar functional-decomposition design pattern (IPO - Input, Process, Output):

```
<table>
<thead>
<tr>
<th>Input</th>
<th>Process</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>welcome user</td>
<td>compute answer</td>
<td>display results</td>
</tr>
<tr>
<td>initialize variables</td>
<td></td>
<td></td>
</tr>
<tr>
<td>prompt for input</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

```python
def main():
    label, values = getInput()
    total = sum(values)
    displayResults(label, total)

def getInput():
    """ Get label and list of values to sum."""
    label = input("What are we summing? ")
    numberOfValues=int(input("How many values are there? "))
    values = []
    for i in range(numberOfValues):
        values.append(eval(input("Enter the next number: ")))
    return label, values

def displayResults(label, total):
    """ Display sum of values. """
    print("The sum of", label, "values is", total)
main()  # starts the main function running
```

What are we summing? money
How many values are there? 4
Enter the next number: 10
Enter the next number: 20
Enter the next number: 30
Enter the next number: 50
The sum of money values is 110
When a function is called, a call-frame is pushed onto the run-time stack part of memory. A call-frame contains information about the function: (1) return address -- where function was called from, (2) formal parameters, and (3) local variables -- temporary values created inside the function.

```python
def main():
    label, values = getInput()
    total = sum(values)  # sum
    displayResults(label, total)

def getInput():
    """ Get label and list of values to sum."""
    label = input("What are we summing? ")
    numberOfValues = int(input("How many values are there? "))
    values = []
    for i in range(numberOfValues):
        values.append(eval(input("Enter the next number: ")))
    return label, values

def displayResults(label, total):
    """ Display sum of values. """
    print("The sum of", label, "values is", total)
```

Sequence of snapshots of the run-time stack during execution of program:

At start of main

<table>
<thead>
<tr>
<th>r.a.</th>
<th>main's call-frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>label</td>
<td>?</td>
</tr>
<tr>
<td>numberOfValues</td>
<td>?</td>
</tr>
<tr>
<td>values</td>
<td>?</td>
</tr>
<tr>
<td>total</td>
<td>?</td>
</tr>
</tbody>
</table>

At start of getInput

<table>
<thead>
<tr>
<th>r.a.</th>
<th>getInput's call-frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>label</td>
<td>?</td>
</tr>
<tr>
<td>numberOfValues</td>
<td>?</td>
</tr>
<tr>
<td>values</td>
<td>?</td>
</tr>
<tr>
<td>i</td>
<td>?</td>
</tr>
</tbody>
</table>

After returning to (***) from getInput

<table>
<thead>
<tr>
<th>r.a.</th>
<th>&quot;money&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>label</td>
<td>[10, 20, 30, 50]</td>
</tr>
<tr>
<td>numberOfValues</td>
<td>4</td>
</tr>
<tr>
<td>values</td>
<td>4</td>
</tr>
<tr>
<td>total</td>
<td>?</td>
</tr>
</tbody>
</table>

At start of displayResults

<table>
<thead>
<tr>
<th>r.a.</th>
<th>displayResults's call-frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>label</td>
<td>&quot;money&quot;</td>
</tr>
<tr>
<td>values</td>
<td>[10, 20, 30, 50]</td>
</tr>
<tr>
<td>total</td>
<td>110</td>
</tr>
</tbody>
</table>

After returning to (****) from displayResults

<table>
<thead>
<tr>
<th>r.a.</th>
<th>&quot;money&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>label</td>
<td>[10, 20, 30, 50]</td>
</tr>
<tr>
<td>values</td>
<td>[10, 20, 30, 50]</td>
</tr>
<tr>
<td>total</td>
<td>110</td>
</tr>
</tbody>
</table>

After returning to (*) from main
3. Design a program to roll two 6-sided dice 1,000 times to determine the percentage of each outcome (i.e., sum of both dice). Report the outcome(s) with the highest percentage.
   a) How would you solve this problem “by hand” without a computer?

   b) What built-in Python module/function can be used to simulate a 6-sided die?

   c) What built-in data structure(s) could be used to tally the 1,000 roll outcomes?

   d) Customize the functional-decomposition diagram for the dice problem by briefly describing what each function does and what parameters are passed.
Objective: To practice writing Python code using functions.

To start the lab: Download and unzip the file lab1.zip on eLearning.

Part A: In the folder lab1, open the diceOutcomes.py program in IDLE. (Right-click on diceOutcomes.py | Edit with IDLE) It contains a partial program we started to discuss in class to solve the problem:

"Write a program to roll two 6-sided dice 1,000 times to determine the percentage of each outcome (i.e., sum of both dice). Report the outcome(s) with the highest percentage."

I decided to functional-decomposition this problem as:

```
main - provides an outline of program by calling top-level functions
welcomeAndInputRolls - Displays welcome message for the user. Gets and returns the number of dice rolls from the user.
calculateFrequentRolls - Rolls the dice the correct number of times, tallies the outcomes, and returns a list of outcomes with the highest count and highest count.
rollAndTallyOutcomes - Rolls the dice the correct number of times and tallies the outcomes. Returns a list of tallies with the index being the outcome.
max - built-in function to return the largest item in an iterable data structure like a list.
findOutcomes - Returns a list of outcomes with the highest count.
displayResults - Displays the outcome(s) with the highest percentage.
```

Consider running the program with only 10 dice rolls instead of 1,000. The program output with some extra debugging prints showing the two Python lists used: outcomeCounts and mostFrequentRolls.

```
This programs rolls two 6-sided dice many times to determine the outcome(s) with the highest percentage.
How many times would you like to roll the pair of dice? 10
outcomeCounts: [0, 0, 1, 0, 2, 1, 0, 3, 0, 3, 0, 0]
mostFrequentRolls: [7, 10] and highestCount: 3
The highest percentage is 30.0 for outcome(s): 7 10
```

Your task for lab 1 is to complete the code for the rollAndTallyOutcomes and findOutcomes functions.

EXTRA CREDIT -- Part B: Rewrite the program using a dictionary instead of a list for the outcomeCounts. Your dictionary will have the outcome for the key and its corresponding tally as its value.

After you have working code, zip the lab1 folder and submit it on eLearning.
Data Structures: Homework #1
Due: August 29 (Sat. at 11:59 PM)

You are to write a Math Tutor program for small children. Your program's interaction should look something like: (Student input shown in bold.)

Welcome to the Math Tutor

Enter your name: Bob<Enter>
Enter your grade (1, 2, or 3): Third<Enter>
Sorry invalid grade, please re-enter your grade using only (1, 2, or 3): 3<Enter>

Directions: Please answer the following 5 problems Bob.

Problem #1: 395
+ 137

Answer: 422<Enter>

Sorry Bob, but your answer is incorrect. The correct answer is 532.

Problem #2: 427
+ 344

Answer: 771<Enter>

Great Bob, your answer is correct!!!

Summary: 4 out of 5 problems correct.

Your program should randomly select the difficulty of the problems according to the following table.

<table>
<thead>
<tr>
<th>Grade Level</th>
<th>Range of the Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (first grade)</td>
<td>0 to 9</td>
</tr>
<tr>
<td>2 (second grade)</td>
<td>10 to 99</td>
</tr>
<tr>
<td>3 (third grade)</td>
<td>100 to 999</td>
</tr>
</tbody>
</table>

Your program also needs to use functions (See lab 1 for an example of how I want your program). It should:

- be designed using the functional-decomposition (top-down) design before you start to write code! You will need to turn in a design document (see lab 1 example below for details).
- name your program mathTutor.py inside a folder hw1 containing it,
- alignment the numbers to be added vertically aligned on the one’s place
- check for a valid grade selection. Anything other than '1', '2', or '3' should print an appropriate error message and ask for the user to re-enter the grade. See example_programs_hw1.zip on the eLearning system (Content | Unit #1 | Homework #1 subfolder).

Extra credit features: (You don’t need to do all features to get partial extra credit)

- Randomly pick the arithmetic operation based on the grade level. First grade should only be addition. Second grade can be addition or subtraction, but if subtraction is to be performed the "upper" number should be larger than the "lower" number. Third grade should allow addition, subtraction, and multiplication. Third grade subtraction should also use numbers such that the "upper" number is larger than the "lower" number.
- In a file whose name is Log.txt, append a line of text summarizing each user, include the user’s name, grade-level, and score.
- Provide more helpful feedback to the user. For example, in the first example in the above interaction you might suggest "review the concept of carrying".
When you write your program, be sure you:

- think about the functional-decomposition (top-down) design before you start to write code! You will need to turn in a design document design.doc (or design.pdf, or design.txt, or design.rtf). This document describes the design of your program including a functional-decomposition diagram showing parameters and returned values, and text describing each function (see lab 1 description and program lab1/diceOutcome.py for examples)
- use meaningful variable names with good style (i.e., useCamelCase)
- use comments (""" Multi-line Comment """") at the start of the program and immediately after each function definition describing what they do (see lab 1 diceOutcomes.py program for an example)
- use a main function (see lab1 diceOutcomes.py program) located at the top of program with a call to it at the bottom to start execution
- do not use global variables, except for global constants where appropriate with good style (ALL_CAPS_AND_UNDERSCORES). Put your global constants after your initial comments describing the program and before your main function definition so they can be found and changed easily in future versions of your program. (see DIE_SIDES in lab 1 diceOutcomes.py program for an example)

On eLearning (Content | Unit #1 | Homework #1 subfolder), submit a single .zip file, hw1.zip containing the following:

- mathTutor.py (your Python program)
- design.docx (or hand-written and scanned as design.pdf, or photo design.jpg, or design.rtf) a document describing the design of your program including a functional-decomposition diagram showing parameters and returned values, and text describing each function (see lab 1 description)

(If you miss the deadline, you can still submit it without a late penalty. However, I there will be a homework 2, etc. and you don’t want to get too far behind!)

Recall the Lab 1 functional-decomposition that you can model for this homework’s design document. Your design diagram should include the flow of information between your functions.

```
main
  numberOfRolls
  displayWelcomeAndInputRolls
    numberOfRolls
    |mostFrequentRolls
    |highestCount
  |mostFrequentRolls
  |highestCount
  displayResults

calculateFrequentRolls
  |mostFrequentRolls
  |highestCount
  outcomeCounts
  |highestCount
  max
  rollAndTallyOutcomes
    outcomeCounts
    |highestCount
    |mostFrequentRolls
    |highestCount
  findOutcomes
```

For each function, include a brief description of what it does (NOT the code/algorithim):

- main - provides an outline of program by calling top-level functions
- displayWelcomeAndInputRolls - Displays welcome message for the user. Gets and returns the number of dice rolls from the user.
- calculateFrequentRolls - Rolls the dice the correct number of times, tallies the outcomes, and returns a list of outcomes with the highest count and highest count.
- rollAndTallyOutcomes - Rolls the dice the correct number of times and tallies the outcomes. Returns a list of tallies with the index being the outcome.
- max - built-in function to return the largest item in an iterable data structure like a list.
- findOutcomes - Returns a list of outcomes with the highest count.
- displayResults - Displays the outcome(s) with the highest percentage.
1. An alternative to functional-decomposition design is to use object-oriented design (OOD). For the following program, what objects would be useful and what methods (operations on the objects) should each support? Write a program to roll two 6-sided dice 1,000 times to determine the percentage of each outcome (i.e., sum both dice). Report the outcome(s) with the highest percentage.” (You only need consider the program’s OOD)

2. Consider the Die and AdvancedDie classes from the Python Summary handout.
   a) What data attributes of AdvancedDie are inherited from the parent Die class?

   b) What new data attributes are added as part of the subclass AdvancedDie?

   c) Which Die class methods are used directly for an AdvancedDie object?

   d) Which Die class methods are redefined/overridden by the AdvancedDie object?

   e) Which methods are new to the AdvancedDie class and not in the Die class?

   f) If die1 and die2 are AdvancedDie objects, then the statement “if die1 == die2:” invokes the __eq__ method of AdvancedDie with die1 “passed” as self and die2 passed as rhs_die.

   ```python
def __eq__(self, rhs_die):
    """Overrides default '__eq__' operator to allow for deep comparison of dice""
    return self._currentRoll == rhs_die._currentRoll
```

What would the code be for AdvancedDie __le__ method to allow for the “if die1 <= die2:” statement?

   g) Good software engineering practice is to include precondition and postcondition comments on each method/function where the:
   - precondition - indicates what must be true for the method to work correctly. Typically, the precondition describes the valid values of the parameters. If the precondition is not satisfied, the method does not need to work correctly!
   - postcondition - describes the expected state after the method has executed

Consider the AdvancedDie constructor:

```python
class AdvancedDie(Die):
    """Advanced die class that allows for any number of sides""
    def __init__(self, sides = 6):
        """Constructor for any sided Die that takes as input the number of sides as a parameter; if no parameter given then default is 6-sided.""
        Die.__init__(self) # call Die parent class constructor
        self._numSides = sides
        self._currentRoll = randint(1, self._numSides)
```

What precondition and postcondition comments should we add?

   h) If a method/function has a precondition that is not met when invoked (e.g., die1 = AdvancedDie("six")), why should the method raise an error?
3. General “Algorithmic-Complexity Analysis” terminology:

- **problem** - question we seek an answer for, e.g., "What is the sum of all the items in a list/array?"
- **parameters** - variables with unspecified values
- **problem instance** - assignment of values to parameters, i.e., the specific input to the problem

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>myList</td>
<td>5</td>
<td>10</td>
<td>2</td>
<td>15</td>
<td>20</td>
<td>1</td>
<td>11</td>
</tr>
</tbody>
</table>

(sum: ?)

- **algorithm** - step-by-step procedure for producing a solution
- **basic operation** - fundamental operation in the algorithm (i.e., operation done the most) Generally, we want to derive a function for the number of times that the basic operation is performed related to the **problem size**.
- **problem size** - input size. For algorithms involving lists/arrays, the problem size is the number of elements ("n").

**Big-oh notation (O())** - As the size of a problem grows (i.e., more data), how will our program's run-time grow.

Consider the following **sumList** function.

```python
def sumList(myList):
    """Returns the sum of all items in myList""
    total = 0
    for item in myList:
        total = total + item
    return total
```

a) What is the basic operation of **sumList** (i.e., operation done the most)?

b) What is the problem size of **sumList**?

c) If n is 10000 and **sumList** takes 10 seconds, how long would you expect **sumList** to take for n of 20000?

d) What is the big-oh notation for **sumList**?

4. Consider the following **someLoops** function.

```python
def someLoops(n):
    total = 0
    for i in range(n):
        for j in range(n):
            total = total + i + j
    return total
```

![Execution flow diagram]

a) What is the basic operation of **someLoops** (i.e., operation done the most)?

b) How many times will the basic operation execute as a function of n?

c) What is the big-oh notation for **someLoops**?

d) If we input n of 10000 and **someLoops** takes 10 seconds, how long would you expect **someLoops** to take for n of 20000?
1. Draw the graph for `sumList (O(n))` and `someLoops (O(n^2))` from the previous lecture.

```
import time

def main():
    n = eval(input("Enter size of list: "))
    alist = list(range(1, n+1))
    start = time.perf_counter()  # <<< time.clock() is deprecated
    sum = sumSomeListItems(alist)
    end = time.perf_counter()    # <<< time.clock() is deprecated
    print("Time to sum the list was %.9f seconds" % (end-start))

def sumSomeListItems(myList):
    """Returns the sum of some items in myList""
    total = 0
    index = len(myList) - 1
    while index > 0:
        total = total + myList[index]
        index = index // 2
    return total

main()
```

a) What is the problem size of `sumSomeListItems`?

b) If we input `n` of 10,000 and `sumSomeListItems` takes 10 seconds, how long would you expect `sumSomeListItems` to take for `n` of 20,000?

(Hint: For `n` of 20,000, how many more times would the loop execute than for `n` of 10,000?)

c) What is the big-oh notation for `sumSomeListItems`?

d) Add the execution-time graph for `sumSomeListItems` to the graph.
3.

```python
i = 1
while i <= n:
    for j in range(n):
        # something of O(1)
    # end for
    i = i * 2
# end while
```

a) Analyze the above algorithm to determine its big-oh notation, $O()$.

b) If $n$ of 10,000, takes 10 seconds, how long would you expect the above code to take for $n$ of 20,000?

c) Add the execution-time graph for the above code to the graph.

4. Most programming languages have a built-in array data structure to store a collection of same-type items. Arrays are implemented in RAM memory as a contiguous block of memory locations. Consider an array $X$ that contains the odd integers:

<table>
<thead>
<tr>
<th>address</th>
<th>Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>1</td>
</tr>
<tr>
<td>4004</td>
<td>3</td>
</tr>
<tr>
<td>4008</td>
<td>5</td>
</tr>
<tr>
<td>4012</td>
<td>7</td>
</tr>
<tr>
<td>4016</td>
<td>9</td>
</tr>
<tr>
<td>4020</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a) Any array element can be accessed randomly by calculating its address. For example, address of $X[5] = 4000 + 5 \times 4 = 4020$. What is the general formula for calculating the address of the $i$th element in an array?

b) What is the big-oh notation for accessing the $i$th element?

c) A Python list uses an array of references (pointers) to list items in their implementation of a list. For example, a list of strings containing the alphabet:

```
0 1 2 3 (len()-1)
\downarrow \downarrow \downarrow \downarrow
'a' 'b' 'c' ... 'z'
```

Since a Python list can contain heterogeneous data, how does storing references in the list aid implementation?
5. Arrays in most HLLs are static in size (i.e., cannot grow at run-time), so arrays are constructed to hold the “maximum” number of items. For example, an array with 1,000 slots might only contain 3 items:

| size: 3 | scores: 20 10 30 999 |

a) The physical size of the array is the number of slots in the array. What is the physical size of scores?
b) The logical size of the array is the number of items actually in the array. What is the logical size of scores?
c) The load factor is fraction of the array being used. What is the load factor of scores?
d) What is the \( O() \) for “appending” a new score to the “right end” of the array?
e) What is the \( O() \) for adding a new score to the “left end” of the array?
f) What is the average \( O() \) for adding a new score to the array?
g) During run-time if an array fills up and we want to add another item, the program can usually:
- Create a bigger array than the one that filled up
- Copy all the items from the old array to the bigger array
- Add the new item
- Delete the smaller array to free up its memory
When creating the bigger array, how much bigger than the old array should it be?

h) What is the \( O() \) of moving to a larger array?

6. Consider the following list methods in Python:

<table>
<thead>
<tr>
<th>Method</th>
<th>Usage</th>
<th>Average ( O() ) for myList containing ( n ) items</th>
</tr>
</thead>
<tbody>
<tr>
<td>index []</td>
<td>itemValue = myList[i]</td>
<td></td>
</tr>
<tr>
<td>append</td>
<td>myList.append(item)</td>
<td></td>
</tr>
<tr>
<td>extend</td>
<td>myList.extend(otherList)</td>
<td></td>
</tr>
<tr>
<td>insert</td>
<td>myList.insert(i, item)</td>
<td></td>
</tr>
<tr>
<td>pop</td>
<td>myList.pop( )</td>
<td></td>
</tr>
<tr>
<td>pop(i)</td>
<td>myList.pop(i)</td>
<td></td>
</tr>
<tr>
<td>del</td>
<td>del myList[i]</td>
<td></td>
</tr>
<tr>
<td>remove</td>
<td>myList.remove(item)</td>
<td></td>
</tr>
<tr>
<td>index</td>
<td>myList.index(item)</td>
<td></td>
</tr>
<tr>
<td>iteration</td>
<td>for item in myList:</td>
<td></td>
</tr>
<tr>
<td>reverse</td>
<td>myList.reverse( )</td>
<td></td>
</tr>
</tbody>
</table>

Dictionary Operations:

<table>
<thead>
<tr>
<th>Method</th>
<th>Usage</th>
<th>Explanation</th>
<th>Average ( O() ) for ( n ) keys</th>
</tr>
</thead>
<tbody>
<tr>
<td>get item</td>
<td>myDictionary.get(myKey) value = myDictionary[myKey]</td>
<td>Returns the value associated with myKey; otherwise None</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>set item</td>
<td>myDictionary[myKey]=value</td>
<td>Change or add myKey: value pair</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>in</td>
<td>myKey in myDictionary</td>
<td>Returns True if myKey is in myDictionary; otherwise False</td>
<td>( O(1) )</td>
</tr>
<tr>
<td>del</td>
<td>del myDictionary[myKey]</td>
<td>Deletes the mykey: value pair</td>
<td>( O(1) )</td>
</tr>
</tbody>
</table>
Objective: To get a feel for big-oh notation by analyzing algorithms as well as timing them (Part A), and gain some experience writing Python classes.

Informal Big-oh (and Big-Theta) Definition: As the size of a computational problem grows (i.e., more data), we expect our program to run longer, but this run-time growth is not necessarily linear. Big-oh notation gives us an idea how our program’s run-time will grow with respect to its problem size on larger data.

This might seem like a lot of mathematical mumbo-jumbo, but knowing an algorithms big-oh notation can help us predict its run-time on large problem sizes. While running a large size problem, we might want to know if we have time for a quick lunch, a long lunch, a long nap, go home for the day, take a week of vacation, pack-up the desk because the boss will fire you for a slow algorithm, etc.

For example, consider the following algorithm:

```python
result = 0
for r in range(n):
    for c in range(n):
        for d in range(n/2):
            result = result + d
    # end for
# end for
```

Clearly, the body of the inner-most loop (the “result = result + d” statement) will execute \( n^3 / 2 \) times, so this algorithm is “big-oh” of \( n \)-cubed, \( O(n^3) \). Thus, the execution-time formula, \( T(n) \), with-respect-to \( n \) is:

\[
T(n) = cn^3 + \text{(slower growing terms)}
\]

For large values of \( n \), the execution time as a function of \( n \), \( T(n) \approx cn^3 \), where \( c \) is the constant of proportionality on the fastest growing term (the machine dependent time related to how long it takes to execute the inner-most loop once). If we know that \( T(10,000) = 1 \text{ second} = c \times 10,000^3 \), then we can calculate \( c \) and use it to predict what \( T(1,000,000) \). First approximate \( c \approx \frac{T(n)}{n^3} = 1 \text{ second} / 10,000^3 = 1 \text{ second} / 10^{12} = 10^{-12} \text{ seconds} \).

Since we are running the algorithm on the same machine, \( c \) is unchanged for the larger problem. Thus,

\[
T(1,000,000) \approx c \times 1,000,000^3 = c \times 10^{18} = 10^{-12} \text{ seconds} \times 10^{18} = 10^6 \text{ seconds} \text{ or about 11.6 days. (A couple weeks of vacation is appropriate!)}
\]

To start the lab: Download and unzip the file lab2.zip from eLearning.

Part A: In the folder lab2, open the timeStuff.py program in IDLE. (Right-click on timeStuff.py | Edit with IDLE) Start it running in IDLE by selecting Run | Run Module from the menu. While it is running, answer the following questions about each of the algorithms in timeStuff.py.

a) What is the big-oh of Algorithm 0?

Algorithm 0:

```python
result = 0
for i in range(10000000):
    result = result + i
```

b) What is the big-oh of Algorithm 1?

Algorithm 1:

```python
result = 0
for i in range(n):
    result = result + i
# end for
```
c) What is the big-oh of Algorithm 2?

Algorithm 2:
```python
result = 0
for r in range(n):
    c = n
    while c > 1:
        result = result + c
        c = c // 2
    # end while
# end for
```

d) What is the big-oh of Algorithm 3?

Algorithm 3:
```python
result = 0
for r in range(n):
    for c in range(n):
        result = result + c
    # end for
# end for
```

e) What is the big-oh of Algorithm 4?

Algorithm 4:
```python
result = 0
for r in range(n):
    for c in range(n):
        for d in range(n*n*n):
            result = result + d
        # end for
    # end for
# end for
```

f) What is the big-oh of Algorithm 5?

Algorithm 5:
```python
result = 0
i = 0
while i < 2**n:
    result = result + i
    i += 1
# end while
```
g) Complete the following timing table from the output of timeStuff.py.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Execution Time in Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n = 0</td>
</tr>
<tr>
<td>Algorithm 0</td>
<td></td>
</tr>
<tr>
<td>Algorithm 1</td>
<td></td>
</tr>
<tr>
<td>Algorithm 2</td>
<td></td>
</tr>
<tr>
<td>Algorithm 3</td>
<td></td>
</tr>
<tr>
<td>Algorithm 4</td>
<td></td>
</tr>
<tr>
<td>Algorithm 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

h) For Algorithm 5, use the timing for n = 20 to compute the constant of proportionality, c, on the fastest growing term.

i) Using the constant of proportionality computed in (h), predict the run-time of Algorithm 5 for n = 30.

j) How does your prediction in (i) compare to the actual time from (g)?

(Note: Part B is on the backside of this sheet)
Part B: The lab2.zip file also contains:
- A simple Die class (in the simple_die.py module) for a six-sided die.
- An AdvancedDie class (in the module advanced_die.py module) for a die which can be constructed with any number of sides. The AdvancedDie class inherits from the Die class.
- An averageOutcome.py program that computes the average outcome (i.e., average total on the pair of dice) on a pair of 10-side dice over 1,000 rolls. Unfortunately, it uses randInt(1,10)+randInt(1,10).

a) Modify the averageOutcome.py program so that it uses the AdvancedDie class by:
- creating two 10-sided AdvancedDie objects (remember to “from advanced_die import AdvancedDie”)
- rolls the pair of dice 1,000 times to compute the average outcome

(Note: most of the program will remain unchanged. See Python Summary Page 11 in the course packet (or lab2/pythonSummary.pdf) to see examples of the AdvancedDie class usage)

Part C: (NOTE: Part C does not need to use averageOutcome.py program)

a) Implement a new subclass MoreAdvancedDie in a new file more_advanced_die.py which inherits from the AdvancedDie class, and includes only a new setRoll method. The setRoll method can be used for testing certain dice games, because the new method setRoll takes as a parameter a roll value that is used to set a die’s current roll to a specified value. We can test the new MoreAdvancedDie class after loading its definition in the Python shell by using the Run | Run Module IDLE menu selection:

```python
>>> myDie = MoreAdvancedDie(10) # create a 10 sided MoreAdvancedDie object
>>> print(myDie) # print the current roll
>>> myDie.setRoll(3) # sets myDie's current roll to 3
>>> print(myDie) # print the new current roll which should be 3
```

When implementing the MoreAdvancedDie class you must:
- Include documentation with the MoreAdvancedDie class (comments right below its class line)
- Include documentation with the setRoll method including preconditions and postconditions
- Enforce the setRoll method’s preconditions by raising appropriate exceptions.

(see the AdvancedDie class’s __init__ method as an examples of raising exceptions)

We can test that exceptions are raised by setRoll method after loading the MoreAdvancedDie class into the Python shell by:

```python
>>> myDie = MoreAdvancedDie(10) # create a 10 sided MoreAdvancedDie object
>>> myDie.setRoll(3) # sets myDie’s current roll to 3
>>> myDie.setRoll(0) # should raise a ValueError
>>> myDie.setRoll(11) # should raise a ValueError
>>> myDie.setRoll(3.14) # should raise a TypeError
>>> myDie.setRoll("six") # should raise a TypeError
```

b) View your programmer-authored documentation for the MoreAdvancedDie class by typing help(MoreAdvancedDie) at the IDLE shell prompt (“>>>”) after selecting Run | Run Module in the file more_advanced_die.py which contains the MoreAdvancedDie class.

After you have working code, zip the lab1 folder and submit it on eLearning.

If you do not get done today, then submit it by next week’s lab period. Today save your lab 2 files (USB drive, etc.) and remember to log off.

If you have extra time, this would be a good chance to work on Homework #1!
1. An "abstract" view of the stack:

Using an array implementation would look something like:

<table>
<thead>
<tr>
<th>items:</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>(max-1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>c</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>top:</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>max:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

Complete the big-oh notation for the following stack methods assuming an array implementation: ("n" is the # items)

<table>
<thead>
<tr>
<th></th>
<th>push(item)</th>
<th>pop()</th>
<th>peek()</th>
<th>size()</th>
<th>isEmpty()</th>
<th>isFull()</th>
<th>Constructor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big-oh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Since Python does not have a (directly accessible) built-in array, we can use a list.

```python
class Stack:
    def __init__(self):
        self.items = []

    def isEmpty(self):
        return self.items == []

    def push(self, item):
        self.items.append(item)

    def pop(self):
        return self.items.pop()

    def peek(self):
        return self.items[len(self.items)-1]

    def size(self):
        return len(self.items)
```

Since Python uses an array of references (pointers) to list items in their implementation of a list.

"Abstract" Stack Object list Object

```
Stack
<table>
<thead>
<tr>
<th>c</th>
<th>top</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>bottom</td>
</tr>
</tbody>
</table>
```

a) Complete the big-oh notation for the stack methods assuming this Python list implementation: ("n" is the # items)

<table>
<thead>
<tr>
<th></th>
<th>push(item)</th>
<th>pop()</th>
<th>peek()</th>
<th>size()</th>
<th>isEmpty()</th>
<th>init</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big-oh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Which operations should have what preconditions?
3. The text's alternative stack implementation also using a Python list is:

```python
class Stack:
    def __init__(self):
        self.items = []

    def isEmpty(self):
        return self.items == []

    def push(self, item):
        self.items.insert(0, item)

    def pop(self):
        return self.items.pop(0)

    def peek(self):
        return self.items[0]

    def size(self):
        return len(self.items)
```

Since an array is used to implement a Python list, the alternate Stack implementation using a list:

"Abstract" Stack Object

```
items: 0 1 2 3
```

"alternate" Stack Object

```
c b a
```

a) Complete the big-oh notation for the "alternate" Stack methods: ("n" is the # items)

<table>
<thead>
<tr>
<th>Big-oh</th>
<th>push(item)</th>
<th>pop()</th>
<th>peek()</th>
<th>size()</th>
<th>isEmpty()</th>
<th>init</th>
</tr>
</thead>
</table>

4. How could we use a stack to check if a word is a palindrome (e.g., radar, toot)?

5. How could we check to see if we have a balanced string of nested symbols? ("(([]){()}[])")
Data Structures  Homework #2  Due: Sept. 12 (Sat. at 11:59 PM)

You are to write a program to help solve “JUMBLE” puzzles. A “JUMBLE” puzzle is a series of scrambled words which needs to be unscrambled to real words. For example, the follow is series of 6 scrambled words in a file named jumble.txt and the corresponding expected output file named solution.txt.

<table>
<thead>
<tr>
<th>REUVC</th>
<th>REUVC unscrambles to ['CURVE']</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOHOP</td>
<td>TOHOP unscrambles to ['PHOTO']</td>
</tr>
<tr>
<td>FNELEN</td>
<td>FNELEN unscrambles to ['FENNEL']</td>
</tr>
<tr>
<td>EERSDC</td>
<td>EERSDC unscrambles to ['SECEDE']</td>
</tr>
<tr>
<td>XYZ</td>
<td>XYZ could not be unscrambled!</td>
</tr>
<tr>
<td>OTP</td>
<td>OTP unscrambles to ['OPT', 'POT', 'TOP']</td>
</tr>
<tr>
<td>RVOSTE</td>
<td>RVOSTE unscrambles to ['STOVER', 'STROVE', 'TROVES', 'VOTERS']</td>
</tr>
</tbody>
</table>

The file hw2/dictionary.txt contains an alphabetical list of “all” real words that JUMBLE words can unscramble to. A couple hints about how I would solve this problem reasonably efficiently:

- I would sort the letters in each word to get a unique string for comparisons. For example FNELEN would become EEFLNN, and FENNEL would become a matching EEFLNN.
- Before reading the JUMBLE words, I would preprocess the dictionary.txt file to create a Python dictionary whose keys are the sorted letters of words in the dictionary file with corresponding values being the list of “matching” dictionary words. For example in this dictionary will be entries: EEFLNN: ['FENNEL'], OPT: ['OPT', 'POT', 'TOP'], EORSTV: ['STOVER', 'STROVE', 'TROVES', 'VOTERS'], and AEKXNS: ['SKEANS', 'SNAKES', 'SNEAKS'].

Your program can assume an input file named jumble.txt containing one jumble word per line, and it should generate an output file named solution.txt formatted similar to the above example.

When you write your program, be sure you:

- think about the functional-decomposition (top-down) design before you start to write code! You will need to turn in a design document design.doc (or design.pdf, or design.txt, or design.rtf). This document describes the design of your program including a functional-decomposition diagram showing parameters and returned values, and text describing each function (see lab 1 description and program lab1/diceOutcome.py for examples)
- use meaningful variable names with good style (i.e., useCamelCase)
- use comments (""" Multi-line Comment """") at the start of the program and immediately after each function definition describing what they do (see lab 1 diceOutcomes.py program for an example)
- use a main function (see lab1 diceOutcomes.py program) located at the top of program with a call to it at the bottom to start execution
- do not use global variables, except for global constants where appropriate with good style (ALL_CAPS_AND_UNDERSCORES). Put your global constants after your initial comments describing the program and before your main function definition so they can be found and changed easily in future versions of your program. (see DIE_SIDES in lab1 diceOutcomes.py program for an example)

On eLearning (Course Content | Unit #1 | Homework #2 subfolder), submit a single .zip file, hw2.zip containing the following:

- jumbleSolver.py (your Python program)
- dictionary.txt (my supplied dictionary file)
- design.docx (or hand-written and scanned as design.pdf, or photo design.jpg, or design.rtf) a document describing the design of your program including a functional-decomposition diagram showing parameters and returned values, and text describing each function (see lab 1 description)

(If you miss the deadline, you can still submit it without a late penalty. However, I there will be a homework 3, etc. and you don’t want to get too far behind!)
1. The `Node` class (in `node.py`) is used to dynamically create storage for a new item added to the stack. The `LinkedStack` class (in `linked_stack.py`) uses this `Node` class. Conceptually, a `LinkedStack` object would look like:

```
  class Node:
    def __init__(self, initdata):
      self.data = initdata
      self.next = None

    def getData(self):
      return self.data

    def getNext(self):
      return self.next

    def setData(self, newdata):
      self.data = newdata

    def setNext(self, newnext):
      self.next = newnext
```

```
  class LinkedStack(object):
    *** Link-based stack implementation.***

    def __init__(self):
      self.top = None
      self.size = 0

    def push(self, newItem):
      ***Inserts newItem at top of stack.***
      def pop(self):
        ***Removes and returns the item at top of the stack.
        Precondition: the stack is not empty.***
      def peek(self):
        ***Returns the item at top of the stack.
        Precondition: the stack is not empty.***
        return self.top.getData()
      def size(self):
        ***Returns the number of items in the stack.***
        return self.size
      def isEmpty(self):
        return self.size == 0
      def __str__(self):
        ***Items strung from top to bottom.***
```

a) Complete the `push`, `pop`, and `__str__` methods.

b) Stack methods big-oh's?
   (Assume "n" items in stack)
   - constructor `__init__`
   - `push(item)`
   - `pop()`
   - `peek()`
   - `size()`
   - `isEmpty()`
   - `str()`
A FIFO queue is basically what we think of as a waiting line.

The operations/methods on a queue object, say myQueue are:

<table>
<thead>
<tr>
<th>Method Call on myQueue object</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>myQueue.dequeue()</td>
<td>Removes and returns the front item in the queue.</td>
</tr>
<tr>
<td>myQueue.enqueue(myItem)</td>
<td>Adds myItem at the rear of the queue.</td>
</tr>
<tr>
<td>myQueue.peek()</td>
<td>Returns the front item in the queue without removing it.</td>
</tr>
<tr>
<td>myQueue.isEmpty()</td>
<td>Returns True if the queue is empty, or False otherwise.</td>
</tr>
<tr>
<td>myQueue.size()</td>
<td>Returns the number of items currently in the queue</td>
</tr>
<tr>
<td>str(myQueue)</td>
<td>Returns the string representation of the queue.</td>
</tr>
</tbody>
</table>

2. Complete the following table by indicating which of the queue operations should have preconditions. Write "none" if a precondition is not needed.

<table>
<thead>
<tr>
<th>Method Call on myQueue object</th>
<th>Precondition(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>myQueue.dequeue()</td>
<td>none</td>
</tr>
<tr>
<td>myQueue.enqueue(myItem)</td>
<td>none</td>
</tr>
<tr>
<td>myQueue.peek()</td>
<td>none</td>
</tr>
<tr>
<td>myQueue.isEmpty()</td>
<td>none</td>
</tr>
<tr>
<td>myQueue.size()</td>
<td>none</td>
</tr>
<tr>
<td>str(myQueue)</td>
<td>none</td>
</tr>
</tbody>
</table>

3. The textbook's Queue implementation use a Python list:

```python
class Queue:
    def __init__(self):
        self.items = []

    def isEmpty(self):
        return self.items == []

    def enqueue(self, item):
        self.items.insert(0, item)

    def dequeue(self):
        return self.items.pop()

    def peek(self):
        return self.items[0]

    def size(self):
        return len(self.items)

    def __str__(self):
```

a) Complete the _peek, and _str_ methods

b) What are the Queue methods big-oh's? (Assume “n” items in the queue)

- constructor __init__:
- isEmpty()
- enqueue(item)
- dequeue()
- peek()
- size()
- str()
4. An alternate queue implementation using a linked structure (LinkedQueue class) would look like:

"Abstract Queue"

\[
\begin{array}{ccc}
\text{'w'} & \text{'x'} & \text{'y'} \\
\text{front} & \text{rear} \\
\end{array}
\]

LinkedQueue Object

- \_front: \\
- \_size: 3 \\
- \_rear: \\

Node Objects

\[\text{data} \rightarrow \text{next} \rightarrow \text{data} \rightarrow \text{next} \rightarrow \text{data} \rightarrow \text{next}\]

a) Draw on the picture and number the steps for the enqueue method of the "normal" case (non-empty queue)

b) Write the enqueue method code for the "normal" case:

c) Starting with the empty queue below, draw the resulting picture after your "normal" case code executes.

empty LinkedQueue Object

- \_front: \\
- \_size: 0 \\
- \_rear: \\

d) Fix your "normal" case code to handle the "special case" of an empty queue.

e) Draw on the above picture and number the steps for the dequeue method of the "normal" case (non-empty queue)
f) Write the `dequeue` method code for the "normal" case:


g) What "special case(s)" does the `dequeue` method code need to handle?

h) Draw the picture for each special case and number the steps for the `dequeue` method in the "special" case(s)

i) Combine the "normal" and special case(s) code for a complete `dequeue` method.

j) Complete the big-oh notation for the `LinkedQueue` methods: ("n" is the # items)

<table>
<thead>
<tr>
<th></th>
<th>init</th>
<th>enqueue(item)</th>
<th>dequeue()</th>
<th>peek()</th>
<th>size()</th>
<th>isEmpty()</th>
<th>str</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big-oh</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data Structures (CS 1520)  Lecture 6  Name: 

A Deque (pronounced “Deck”) is a linear data structure which behaves like a double-ended queue, i.e., it allows adding or removing items from either the front or the rear of the Deque.

![Deque Diagram]

1. One possible implementation of a Deque would be to use a Python list to store the Deque items such that
   - the rear item is always stored at index 0,
   - the front item is always stored at the highest index (or -1)

```
Deque Object
items: [ ]

List Object
0 1 2 3
'a' 'b' 'c' 'd'
(rear) (front)
```

```
class Deque(object):
    def __init__(self):
```

a) Complete the `__init__` method and determine the big-oh, \( O() \), for each Deque operation, assuming the above implementation. Let \( n \) be the number of items in the Deque.

<table>
<thead>
<tr>
<th>isEmpty</th>
<th>addFront</th>
<th>removeFront</th>
<th>addRear</th>
<th>removeRear</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Write the methods for the `addRear` and `removeRear` operation.

```python
def addRear(self, newItem):
def removeRear(self):
```

2. An alternative implementation of a Deque would be a linked implementation as in:

![LinkedDeque Diagram]

```
class LinkedDeque(object):
    def __init__(self):
```

a) Complete the `__init__` method and determine the big-oh, \( O() \), for each Deque operation assuming the above linked implementation. Let \( n \) be the number of items in the Deque.

<table>
<thead>
<tr>
<th>isEmpty</th>
<th>addFront</th>
<th>removeFront</th>
<th>addRear</th>
<th>removeRear</th>
<th>size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) Suggest an improvement to the above linked implementation of the Deque to speed up some of its operations.
from node import Node

class Node2Way(Node):
    def __init__(self, initdata):
        Node.__init__(self, initdata)
        self.previous = None

    def getPrevious(self):
        return self.previous

    def setPrevious(self, newprevious):
        self.previous = newprevious

3. An alternative implementation of a Deque would be a doubly-linked implementation as in:

DoublyLinkedDeque Object

_ahead:

_previous:

_size: 4

a) Determine the big-oh, $O()$, for each Deque operation assuming the above doubly-linked implementation. Let $n$ be the number of items in the Deque.

<table>
<thead>
<tr>
<th>isEmpty</th>
<th>addFront</th>
<th>removeFront</th>
<th>addRear</th>
<th>removeRear</th>
<th>size</th>
</tr>
</thead>
</table>

4. A priority queue has the same operations as a regular queue, except the items are NOT returned in the FIFO (first-in, first-out) order. Instead, each item has a priority that determines the order they are removed. A hospital emergency room operates like a priority queue -- the person with the most serious injury has highest priority even if they just arrived.

a) Suppose that we have a priority queue with integer priorities such that the smallest integer corresponds to the highest priority. For the following priority queue, which item would be dequeued next?

priority queue:

40 10 79

30 13

5

b) To implement a priority queue, we could use an unordered Python list. If we did, what would be the big-oh notation for each of the following methods: (justify your answer)

- enqueue:
- dequeue:

c) To implement a priority queue, we could use a Python list ordered by priorities in descending order. If we did, what would be the big-oh notation for each of the following methods: (justify your answer)

- enqueue:
- dequeue
1. Section 6.6 discusses a very “non-intuitive”, but powerful list/array-based approach to implement a priority queue, call a binary heap. The list/array is used to store a complete binary tree (a full tree with any additional leaves as far left as possible) with the items being arranged by heap-order property, i.e., each node is \( \leq \) either of its children. An example of a min heap “viewed” an a complete binary tree would be:

```
6
 /\  
[1][2] /
  /\  
[3][4] /
  /\  
[5][6] /
  /\  
[7][8]
```

```
Python List actually used to store heap items
```

```
0  [6] 15 10 114 20 20 50 300 125 117
```

a) For the above heap, the list/array indexes are indicated in []'s. For a node at index \( i \), what is the index of:

- its left child if it exists:
- its right child if it exists:
- its parent if it exists:

b) What would the above heap look like after inserting 13 and then 3? (show the changes on above tree)

General Idea of insert(newItem):

- append newItem to the end of the list (easy to do, but violates heap-order property)
- restore the heap-order property by repeatedly swapping the newItem with its parent until it percolates to correct spot

c) What is the big-oh notation for inserting a new item in the heap?

d) Complete the code for the percUp method used by insert.

class BinHeap:
    def __init__(self):
        self.heapList = [0]
        self.currentSize = 0
    def percUp(self, currentIdx):
        parentIndex =
        while
```

def insert(self, k):
    self.heapList.append(k)
    self.currentSize = self.currentSize + 1
    self.percUp(self.currentSize)
```
2. Now let us consider the `delMin` operation that removes and returns the minimum item.

   ![Heap Diagram]

   Python List actually used to store heap items
   
<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>6</td>
<td>15</td>
<td>10</td>
<td>114</td>
<td>20</td>
<td>20</td>
<td>50</td>
<td>300</td>
<td>125</td>
<td>117</td>
</tr>
</tbody>
</table>

a) What item would `delMin` remove and return from the above heap?

b) What is the quickest way to fill the hole left by `delMin`?

c) What new problem does this cause?

General Idea of `delMin()`:
- remember the minimum value so it can be returned later (easy to find - at index 1)
- copy the last item in the list to the root, delete it from the right end, decrement size
- restore the heap-order property by repeatedly swapping this item with its smallest child until it percolates down to the correct spot
- return the minimum value

d) What would the above heap look like after `delMin`? (show the changes on above tree)

e) Complete the code for the `percDown` method used by `delMin`.

```python
class BinHeap:
    ...
    ...

    def minChild(self, i):
        if i * 2 + 1 > self.currentSize:  # if only left child
            return i * 2
        else:
            if self.heapList[i * 2] < self.heapList[i * 2 + 1]:
                return i * 2
            else:
                return i * 2 + 1

    def delMin(self):
        retval = self.heapList[1]
        self.currentSize = self.currentSize - 1
        self.heapList.pop()
        self.percDown(1)
        return retval

def percDown(self, currentindex):
    ...
```

e) Complete the code for the `percDown` method used by `delMin`.

f) What is the big-oh notation for `delMin`?
Once we have a working BinHeap, then implementing the PriorityQueue class using a BinHeap is a piece of cake:

```python
class PriorityQueue:
    def __init__(self):
        self._heap = BinHeap()

    def isEmpty(self):
        return self._heap.isEmpty()

    def enqueue(self, item):
        self._heap.insert(item)

    def dequeue(self):
        return self._heap.delMin()

    def size(self):
        return self._heap.size()

    def __str__(self):
        return str(self._heap)
```

>>> q = PriorityQueue()
>>> print(q)
[]
>>> q.enqueue(5)
>>> q.enqueue(1)
>>> q.enqueue(7)
>>> print(q)
[1, 5, 7]
>>> q.dequeue()
1
>>> print(q)
[5, 7]

3. A “list” is a generic term for a sequence of items in a linear arrangement. Unlike stacks, queues and deques access to list items is not limited to either end, but can be from any position in the list. The general terminology of a list is illustrated by:

```
"Abstract view of a list"

[0] [1] [2] [3] ← index/position in the list

head

'a' 'b' 'c'

tail
```

There are three broad categories of list operations that are possible:

- **Index-based operations** - the list is manipulated by specifying an index location, e.g.,
  ```python
  myList.insert(3, item)  # insert item at index 3 in myList
  ```

- **Content-based operations** - the list is manipulated by specifying some content (i.e., item value), e.g.,
  ```python
  myList.remove(item)  # removes the item from the list based on its value
  ```

- **Cursor-based operations** - a cursor (current position) can be moved around the list, and it is used to identify list items to be manipulated, e.g.,
  ```python
  myList.first()  # sets the cursor to the head item of the list
  myList.next()  # moves the cursor one position toward the tail of the list
  myList.remove()  # deletes the second item in the list because that's where the cursor is currently located
  ```

The following table summarizes the operations from the three basic categories on a list, L:

<table>
<thead>
<tr>
<th>Index-based operations</th>
<th>Content-based operations</th>
<th>Cursor-based operations</th>
</tr>
</thead>
<tbody>
<tr>
<td>L.insert(index, item)</td>
<td>L.add(item)</td>
<td>L.hasNext()</td>
</tr>
<tr>
<td>item = L[index]</td>
<td>L.remove(item)</td>
<td>L.next()</td>
</tr>
<tr>
<td>L[index] = newValue</td>
<td>L.search(item) #return Boolean</td>
<td>L.hasPrevious()</td>
</tr>
<tr>
<td>L.pop(index)</td>
<td>i = L.index(item).</td>
<td>L.previous()</td>
</tr>
</tbody>
</table>

Built-in Python lists are unordered with a mixture of index-based and content-based operations. We know they are implemented using a contiguous block of memory (i.e., an array). The textbook talks about an unordered list ADT, and a sorted list ADT which is more content-based. Both are implemented using a singly-linked list.

a) Why would a singly-linked list be a bad choice for implementing a cursor-based list ADT?
Objective: To understand priority queue implementations in Python including being able to determine the big-oh of each operation.

To start the lab: Download and unzip the lab4.zip file from eLearning.

Part A: Python list implementations of a Priority Queue

a) Suppose that we have a priority queue with integer priorities such that the smallest integer corresponds to the highest priority. For the following priority queue, which item would be dequeued next?

```
priority queue:

40  10  60  35
30  25
```

b) The `ListPriorityQueue` implementation in `lab4/list_priority_queue.py` uses an unordered Python list.

```
ListPriorityQueue Object
_items: [30, 40, 60, 25, 35]
```

What would be the big-oh notation for each of the following methods: (justify your answer)

- enqueue:
- dequeue:

c) The `SortedListPriorityQueue` implementation in `lab4/sorted_list_priority_queue.py` uses a Python list order by priorities in descending order.

```
SortedListPriorityQueue Object
_items: [60, 40, 35, 30, 25, 10]
```

What would be the big-oh notation for each of the following methods: (justify your answer)

- enqueue:
- dequeue

d) Why would it be a bad idea to implement a priority queue using a Python list order by priorities in reverse (ascending) order? (HINT: What is the big-oh notations for enqueue and dequeue?)

```
SortedListPriorityQueue Object
_items: [10, 25, 30, 35, 40, 60]
```

49
Part B: (Lecture 7 and) Section 6.6 discusses a very “non-intuitive”, but powerful list/array-based approach to implement a priority queue, call a binary heap. The list/array is used to store a complete binary tree (a full tree with any additional leaves as far left as possible) with the items being arranged by heap-order property, i.e., each node is ≤ either of its children. An example of a min heap “viewed” as a complete binary tree would be:

![Binary Heap Diagram]

Python List actually used to store heap items

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
<td>23</td>
<td>17</td>
<td>34</td>
<td>25</td>
<td>60</td>
<td>90</td>
<td>120</td>
<td>44</td>
<td>28</td>
<td>31</td>
<td>96</td>
<td>84</td>
<td>98</td>
<td></td>
</tr>
</tbody>
</table>

a) For the above heap, the list/array indexes are indicated in [ ]s. For a node at index i, what is the index of:
- its left child if it exists:
- its right child if it exists:
- its parent if it exists:

Recall the General Idea of insert(newItem):
- append newItem to the end of the list (easy to do, but violates heap-order property)
- restore the heap-order property by repeatedly swapping the newItem with its parent until it *percolates up* to the correct spot

b) What would the above heap look like after inserting 18 and then 27? (show the changes on above tree)

c) What is the big-oh notation for inserting a new item in the heap?

Now let us consider the *delMin* operation that removes and returns the minimum item. Recall the General Idea of *delMin*):
- remember the minimum value so it can be returned later (easy to find - at index 1)
- copy the last item in the list to the root, delete it from the right end, decrement size
- restore the heap-order property by repeatedly swapping this item with its smallest child until it *percolates down* to the correct spot
- return the minimum value

![Updated Binary Heap Diagram]

Python List actually used to store heap items

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>12</td>
<td>23</td>
<td>17</td>
<td>34</td>
<td>25</td>
<td>60</td>
<td>90</td>
<td>120</td>
<td>44</td>
<td>28</td>
<td>31</td>
<td>96</td>
<td>84</td>
<td>98</td>
<td></td>
</tr>
</tbody>
</table>

d) What would the above heap look like after *delMin*? (show the changes on above tree)
Part C: (a) Run the lab4/timePriorityQueues.py program that enqueues 20,000 random integers followed by dequeueing all 20,000 integers from various priority queues discussed above. Complete the following timing table from the output of timePriorityQueues.py.

<table>
<thead>
<tr>
<th>Priority Queue Implementation</th>
<th>Execution Time in Seconds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Enqueuing 20,000 Random ints</td>
</tr>
<tr>
<td>Unsorted Python list</td>
<td></td>
</tr>
<tr>
<td>Sorted Python list in descending order</td>
<td></td>
</tr>
<tr>
<td>“Reverse” sorted Python list in ascending order</td>
<td></td>
</tr>
<tr>
<td>Binary heap stored in a Python list</td>
<td></td>
</tr>
</tbody>
</table>

b) Why does it take more time to enqueue 20,000 items in the “unsorted” Python list version than dequeue 20,000 in the sorted Python list version?

c) Why does it take more time to dequeue 20,000 items in the heap version than enqueue 20,000 in the heap version?

d) Why is the heap implementation of the priority queue considered “better” than the other three?

After you have answered all the questions, submit your answers on eLearning.

If you do not get done today, then submit it by next week’s lab period.

If you have extra time, this would be a good chance to work on Homework #2!
1. The textbook's unordered list ADT uses a singly-linked list implementation. I added the _size and _tail attributes:

UnorderedList Object

- _head
- _size: 4
- _tail

![Linked List Diagram]

a) The search(targetItem) method searches for targetItem in the list. It returns True if targetItem is in the list; otherwise it returns False. Complete the search(targetItem) method code:

```python
class UnorderedList:
    def search(self, targetItem):
```

b) The textbook's unordered list ADT does not allow duplicate items, so operations `add(item)`, `append(item)`, and `insert(pos, item)` would have what precondition?

c) Complete the append(item) method including a check of it's precondition(s)?

```python
def append(self, item):
```

d) Why do you suppose I added a _tail attribute?
e) The textbook's `remove(item)` and `index(item)` operations "Assume the item is present in the list." Thus, they would have a precondition like "Item is in the list." When writing a program using an UnorderedList object (say `myGroceryList = UnorderedList()`), how would the programmer check if the precondition is satisfied?

```python
itemToRemove = input("Enter the item to remove from the Grocery list: ")
if
    myGroceryList.remove(itemToRemove)
```

f) The `remove(item)` and `index(item)` methods both need to look for the `item`. What is inefficient in this whole process?

g) Modify the `search(targetItem)` method code in (a) to set additional data attributes to aid the implementation of the `remove(item)` and `index(item)` methods.

h) Write the `index(item)` method including a check of its precondition(s).

```python
def index(self, item):
```

i) Write the `remove(item)` method including a check of its precondition(s).

```python
def remove(self, item):
```
Objective: To gain experience implementing linked data structures by implementing a cursor-based list using doubly-linked nodes.

To start the homework: Download and extract the file hw3.zip from eLearning at Course Content | Unit 1 | Homework #3. The hw3.zip file contains:
- the Node class (in the node.py module) and the Node2Way class (in the node2way.py module)
- the skeleton CursorBasedList class (in the cursor_based_list.py module) which you will complete
- the cursorBasedListTester.py file that you can use to interactively test your CursorBasedList class.

Recall that in a cursor-base list a cursor (indicating the current item) can be moved around the list with the cursor being used to identify the region of the list to be manipulated. We will insert and removing items relative to the current item. A current item must always be defined as long as the list is not empty.

<table>
<thead>
<tr>
<th>Cursor-based operations</th>
<th>Description of operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>L.getCurrent()</td>
<td>Precondition: the list is not empty. Returns the current item without removing it or changing the current position.</td>
</tr>
<tr>
<td>L.hasNext()</td>
<td>Precondition: the list is not empty. Returns True if the current item has a next item; otherwise return False.</td>
</tr>
<tr>
<td>L.next()</td>
<td>Precondition: hasNext returns True. Postcondition: The current item has moved right one item</td>
</tr>
<tr>
<td>L.hasPrevious()</td>
<td>Precondition: the list is not empty. Returns True if the current item has a previous item; otherwise return False.</td>
</tr>
<tr>
<td>L.previous()</td>
<td>Precondition: hasPrevious returns True. Postcondition: The current item has moved left one item</td>
</tr>
<tr>
<td>L.first()</td>
<td>Precondition: the list is not empty. Makes the first item the current item.</td>
</tr>
<tr>
<td>L.last()</td>
<td>Precondition: the list is not empty. Makes the last item the current item.</td>
</tr>
<tr>
<td>L.insertAfter(item)</td>
<td>Inserts item after the current item, or as the only item if the list is empty. The new item is the current item.</td>
</tr>
<tr>
<td>L.insertBefore(item)</td>
<td>Inserts item before the current item, or as the only item if the list is empty. The new item is the current item.</td>
</tr>
<tr>
<td>L.replace(newValue)</td>
<td>Precondition: the list is not empty. Replaces the current item by the newValue.</td>
</tr>
<tr>
<td>L.remove()</td>
<td>Precondition: the list is not empty. Removes and returns the current item. Making the next item the current item if one exists; otherwise the tail item in the list is the current item unless the list is now empty.</td>
</tr>
</tbody>
</table>

The cursor_based_list.py file contains a skeleton CursorBasedList class. **You MUST uses a doubly-linked list implementation with a header node and a trailer node.** All “real” list items will be inserted between the header and trailer nodes to reduce the number of “special cases” (e.g., inserting first item in an empty list, deleting the last item from the list, etc.). An empty list looks like:

![Diagram of an empty list with header and trailer nodes]

Use the provided cursorBasedListTester.py program to test your list.
On eLearning (Course Content | Unit #1 | Homework #3 subfolder), submit a single .zip file, hw3.zip containing the following:
- the Node class (in the node.py module) and the Node2Way class (in the node2way.py module)
- the completed CursorBasedList class (in the cursor_based_list.py module)
- the cursorBasedListTester.py file

Note: No design document needed for this homework.

(If you miss the deadline, you can still submit it without a late penalty. However, I there will be a homework 4, etc. and you don’t want to get too far behind!)
Test 1 Review Topics

Test 1 will be Thursday September 24th in class. It will be closed-book and notes, except for one 8.5" x 11" sheet of paper (you can use front and back) containing any notes that you want AND the Python Summary handout (in front of you packet). The test will cover the following topics (and maybe more).

Chapter 1. Introduction
Python control structures: if, while, for,
Python built-in data structures: list, dictionary, string
Preconditions, postconditions, and raising exceptions to enforce the precondition
Defining classes (e.g., Die) including inheritance, overriding methods

Chapter 2. Algorithm Analysis
Machine dependent measures of performance: program running time and instruction count
Machine independent measures of performance: big-oh, orders of complexity: constant O(1), logarithmic O(log n), linear O(n), "n log n"/log linear O(n log n), quadratic O(n^2), cubic O(n^3), exponential O(2^n)
Complexity analysis of an algorithm to determine its big-oh notation
Implementation of Python lists as an array of object references with implications on operation big-oh (e.g., pop() is O(1) while pop(0) is O(n), etc.)

Chapter 3. Basic Data Structures
General concept of a stack: LIFO, top and bottom
Stack Operations: pop, push, peek, size, isEmpty, __str__
Stack Implementations: Python list to store stack items and linked list of Nodes to store stack items including big-oh of operations
Stack Applications: general idea of using a stack to do parentheses matching and palindrome checking

General concept of a queue: FIFO, front and rear
Queue Operations: enqueue, dequeue, peek, size, isEmpty, __str__
Queue Implementations: Python list to store queue items and linked list of Nodes to store stack items including big-oh of operations

General concept of a deque: double ended queue, front and rear
Deque Operations: addFront, addRear, removeFront, removeRear, size, isEmpty, __str__
Deque Implementations: Python list to store deque items, singly-linked list of Nodes to store deque items, and doubly-linked list of Nodes (e.g, Node2Way) including big-oh of operations
Deque Applications: general idea of using deque to do palindrome checking

General concept of a list: head, tail, index
Categories of List operations: index-based, content-based, cursor-based
Unordered List operations and implementation with a singly-linked list of Nodes including big-oh of operations
Ordered List operations and implementation with a singly-linked list of Nodes including big-oh of operations

Section 6.6. Priority Queue with Binary Heaps
General concept of a priority queue: remove highest priority next
Priority Queue Operations: enqueue, dequeue, peek, size, isEmpty, __str__
Priority Queue Implementations: Python list unordered, Python list ordered by priority, Binary Heap including big-oh of operations
Binary Heap implementation: insert, findMin, delMin, isEmpty, size, buildHeap including big-oh of operations, and binary-tree diagrams after insert and delMin operations
Spring 2020

Data Structures - Test 1

Question 1. (4 points) Consider the following Python code.

```python
for i in range(n*n):
    j = 1
    while j < n:
        print(i, j)
        j *= 2

for k in range(n):
    print(k)
```

What is the big-oh notation $O(\cdot)$ for this code segment in terms of $n$?

Question 2. (4 points) Consider the following Python code.

```python
for i in range(n):
    for j in range(n):
        for k in range(n):
            print(i, j, k)

h = 1
while h < 2**n:  # 2 raised to the n-th power
    print(h)
    h += 1
```

What is the big-oh notation $O(\cdot)$ for this code segment in terms of $n$?

Question 3. (4 points) Consider the following Python code.

```python
def main(n):
    for i in range(n):
        doSomething(n)

def doSomething(n):
    for j in range(n*n):
        doMore(n)
        doMore(n)

def doMore(n):
    for k in range(n*n):
        print(k)
```

What is the big-oh notation $O(\cdot)$ for this code segment in terms of $n$?

Question 4. (6 points) Suppose a $O(n^4)$ algorithm takes 10 second when $n = 1000$. How long should the algorithm run when $n = 10,000$?

Question 5. (7 points) Why should a method/function having a "precondition" raise an exception if the precondition is violated?
Question 6. A FIFO (First-In-First-Out) queue allows adding a new item at the rear using an enqueue operation, and removing an item from the front using a dequeue operation. One possible implementation of a queue would be to use a built-in Python list to store the queue items such that

- the **rear** item is always stored at index 0,
- the **front** item is always at index len(self._items) - 1, or -1

![Queue Object Diagram](image)

**Python List Object**

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>'d'</td>
<td>'c'</td>
<td>'b'</td>
<td>'a'</td>
</tr>
</tbody>
</table>

**rear**

**front**

a) (6 points) Complete the average big-oh $O()$, for each Queue operation, assuming the above implementation. Let $n$ be the number of items in the queue.

<table>
<thead>
<tr>
<th>Operation</th>
<th>$O()$</th>
</tr>
</thead>
<tbody>
<tr>
<td>isEmpty</td>
<td></td>
</tr>
<tr>
<td>enqueue(item)</td>
<td></td>
</tr>
<tr>
<td>dequeue</td>
<td></td>
</tr>
<tr>
<td>peek (returns front item without removing it)</td>
<td></td>
</tr>
<tr>
<td><strong>str</strong></td>
<td></td>
</tr>
<tr>
<td>size</td>
<td></td>
</tr>
</tbody>
</table>

b) (9 points) Complete the method for the dequeue operation, **including the precondition check to raise an exception if it is violated**.

```python
def dequeue(self):
    """Removes and returns the Front item of the Queue
    Precondition: the Queue is not empty.
    Postcondition: Front item is removed from the Queue and returned"
    ... 
```

Suppose we add a new method `peekRear` which returns the value of the rear item without changing the queue.

```python
def peekRear(self):
    """Returns the value of the rear item without changing the queue
    Precondition: the Queue is not empty.
    Postcondition: Value of rear item is returned from the Queue"
    ... 
```

if self.isEmpty():
    raise Exception("Cannot peekRear on an empty queue!")

valueToReturn = self._items.pop(0)
self._items.insert(0, valueToReturn)
return valueToReturn
```

c) (5 points) The above `peekRear` code works correctly, but suggest an improvement to the code that makes it more efficient (you can modify the above code with your changes).
Question 7. (4 points) Suppose we want to implement a priority queue with integer priorities such that the smallest integer corresponds to the highest priority. One possible implementation would use a completely unorder Python list such as:

```
ListPriorityQueue Object
 ITEMS: 0 1 2 3 4 5
30 40 10 60 25 35
```

What would be the big-oh notation for each of the following methods:
- enqueue:
- dequeue:

Question 8. Consider an alternate binary heap approach to implement a priority queue. A Python list is used to store a complete binary tree (a full tree with any additional leaves as far left as possible) with the items being arranged by heap-order property, i.e., each node is ≤ either of its children. An example of a min heap “viewed” as a complete binary tree would be:

```
14 [1]
   /
  [2]  [3]
  19  23
  /
52  34  26
  /
[8]  [9]  [10]  [11]  [12]
110 57  98  61  46
```

Python List actually used to store heap items
```
0 1 2 3 4 5 6 7 8 9 10 11 12
14 19 23 52 34 26 30 110 57 98 61 46
```

a) (7 points) What would the above heap look like after inserting 17 and then 28 (show changes on above tree)

Now consider the binary heap’s `delMin` operation that removes and returns the minimum item.

```
14 [1]
   /
  [2]  [3]
  19  23
  /
52  34  26
  /
[8]  [9]  [10]  [11]  [12]
110 57  98  61  46
```

b) (1 point)  What item would `delMin` remove and return from the above heap?

c) (7 points) What would the heap look like after `delMin`? (show the changes on the tree that’s just above)

d) (2 points) What would be the O() of a single `insert` or `delMin`, where n is the # of items in the heap?

e) (6 points) Explain why the average `insert` operation is faster than the average `delMin` operation.
Question 9. The `Node2Way` and `Node` classes can be used to dynamically create storage for each new item added to a Stack using a doubly-linked implementation as in:

DoublyLinkedStack Object

```
bottom: [a]
_top: [b]
_size: 4
```

a) (6 points) Complete the average big-oh $O()$, for each stack operation, assuming the above implementation. Let $n$ be the number of items in the stack.

<table>
<thead>
<tr>
<th>push(item)</th>
<th>pop()</th>
<th>peek()</th>
<th>size()</th>
<th>isEmpty()</th>
<th><strong>str</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

b) (17 points) Complete the `push` and `__str__` methods for the above DoublyLinkedStack implementation.

```python
class DoublyLinkedStack(object):
    """ Doubly-linked list based stack implementation. """
    def __init__(self):
        self._bottom = None
        self._top = None
        self._size = 0
    def push(self, newItem):
        """ Adds the newItem to the top of the stack. 
        Precondition: none """
    def __str__(self):
        """ Returns a string of stack from top to bottom. 
        Precondition: none """
        resultStr = "(top) "
        temp = self._top
        while temp:
            resultStr += str(temp.data) + " -> "
            temp = temp.next
        return resultStr
```

```python
class Node:
    def __init__(self, initdata):
        self.data = initdata
        self.next = None
        self.previous = None
    def getData(self):
        return self.data
    def getNext(self):
        return self.next
    def setData(self, newData):
        self.data = newData
    def setNext(self, newNext):
        self.next = newNext
```

```python
class Node2Way(Node):
    def __init__(self, initdata):
        Node.__init__(self, initdata)
        self.previous = None
    def getPrevious(self):
        return self.previous
    def setPrevious(self, newPrevious):
        self.previous = newPrevious
```

from node import Node

c) (5 points) Explain how we could use singly-linked nodes (i.e., only Node objects with data and next) to implement the stack without degrading performance (i.e., causing some stack operations to have worse big-oh notations)? Justify your answer.