Priority Queues and Heaps

What is one area for growth that you identified as a takeaway from completing the diagnostic?

(put your answers the chat)
Roadmap

Object-Oriented Programming

arrays
dynamic memory management
linked data structures

real-world algorithms
recursive problem-solving
Life after CS106B!

C++ basics
User/client
vectors + grids
stacks + queues
sets + maps

Core Tools
testing
algorithmic analysis

Diagnostic
Roadmap

C++ basics

User/client

- vectors + grids
- stacks + queues
- sets + maps

Core Tools

- testing
- algorithmic analysis
- recursive problem-solving

Object-Oriented Programming

Implementation

- arrays
- dynamic memory management
- linked data structures

Life after CS106B!

- real-world algorithms

Diagnostic
Today’s questions

How do we implement algorithms for prioritizing data?

How can we make use of multiple levels of abstraction to build better ADTs?
Today’s topics

1. Review (OurVector)
2. Human prioritization algorithms
3. Priority Queues
4. Binary Heaps
Review

[implementing OurVector]
What is OurVector?

- **Goal:** Implement own version of the Stanford C++ Vector

- **Scope Constraints:**
  - We will only implement a subset of the functionality that the Stanford Vector provides.
  - OurVector can only store integers and is not be configurable to store other types.
OurVector Header File

class OurVector {
    public:
        OurVector();
        ~OurVector();
        void add(int value);
        void insert(int index, int value);
        int get(int index);
        void remove(int index);
        int size();
        bool isEmpty();
    private:
        int* elements;
        int allocatedCapacity;
        int numItems;
};
Review: **OurVector** internal state

---

### Client Code

```java
// client code
OurVector vec;
vec.add(106);
vec.add(42);
vec.add(-3);
vec.add(27);
vec.remove(1);
vec.insert(0, 198);
```

---

<table>
<thead>
<tr>
<th>elements</th>
<th>Capacity</th>
<th>numItems</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x1234abef</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>198</th>
<th>106</th>
<th>-3</th>
<th>27</th>
<th>?</th>
<th>?</th>
<th>?</th>
<th>?</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>
Dynamic Array Growth
Implementing ADT Classes

- The first step of implementing an ADT class (as with any class) is answering the three important questions regarding its public interface, private member variables, and initialization procedures.

- Most ADT classes will need to store their data in an underlying array. The organizational patterns of data in that array may vary, so it is important to illustrate and visualize the contents and any operations that may be done.

- The paradigm of "growable" arrays allows for fast and flexible containers with dynamic resizing capabilities that enable storage of large amounts of data.
Implementing ADT Classes

- The first step of implementing an ADT class (as with any class) is answering the three important questions regarding its public interface, private member variables, and initialization procedures.

- Most ADT classes will need to store their data in an underlying array. The organizational patterns of data in that array may vary, so it is important to illustrate and visualize the contents and any operations that may be done.

- The paradigm of "growable" arrays allows for fast and flexible containers with dynamic resizing capabilities that enable storage of large amounts of data.

What about more complex ADTs?
Human Prioritization Algorithms
Multiple Levels of Abstraction
Levels of abstraction

Abstraction boundary for the user

Abstract Data Structures
Data Organization Strategies
Fundamental C++
Data Storage
Levels of abstraction

What is the interface for the user?  
(Vector, Sets, Queues, Grids, etc.)

Abstract Data Structures

Data Organization Strategies

Fundamental C++ Data Storage
Levels of abstraction

What is the interface for the user? (Priority Queue)

What you’ll focus on for Assignment 5

Abstract Data Structures

Data Organization Strategies

Fundamental C++

Data Storage
Levels of abstraction

What is the interface for the user? (Priority Queue)

How is our data organized? (sorted array, binary heap)

Abstract Data Structures

Data Organization Strategies

Fundamental C++ Data Storage
Levels of abstraction

What is the interface for the user?
(Priority Queue)

How is our data organized?
(sorted array, binary heap)

What stores our data?
(arrays, linked lists, etc.)
Levels of abstraction

What is the interface for the user?
(Priority Queue)

How is our data organized?
(sorted array, binary heap)

What stores our data?
(arrays)
Levels of abstraction

What is the interface for the user? (Priority Queue)

How is our data organized? (sorted array, binary heap)

What stores our data? (arrays)

Abstract Data Structures

Data Organization Strategies

Fundamental C++ Data Storage
Priority Queues
What is a priority queue?

- A queue that orders its elements based on a provided “priority”
What is a priority queue?

- A queue that orders its elements based on a provided “priority”
- Like regular queues, you cannot index into them to get an item at a particular position.
What is a priority queue?

- A queue that orders its elements based on a provided “priority”

- Like regular queues, you cannot index into them to get an item at a particular position.

- Useful for maintaining data sorted based on priorities
  - Emergency room (ER) waiting rooms
What is a priority queue?

- A queue that orders its elements based on a provided “priority”

- Like regular queues, you cannot index into them to get an item at a particular position.

- Useful for maintaining data sorted based on priorities
  - Emergency room (ER) waiting rooms
  - Different airline boarding groups (families and first class passengers, frequent flyers, boarding group A, boarding group B, etc.)
What is a priority queue?

- A queue that orders its elements based on a provided “priority”
- Like regular queues, you cannot index into them to get an item at a particular position.
- Useful for maintaining data sorted based on priorities
  - Emergency room (ER) waiting rooms
  - Different airline boarding groups (families and first class passengers, frequent flyers, boarding group A, boarding group B, etc.)

Individual data points can have the same priority!
What is a priority queue?

- A queue that orders its elements based on a provided “priority”

- Like regular queues, you cannot index into them to get an item at a particular position.

- Useful for maintaining data sorted based on priorities
  - Emergency room (ER) waiting rooms
  - Different airline boarding groups (families and first class passengers, frequent flyers, boarding group A, boarding group B, etc.)
  - Filtering data to get the top X results (e.g. most popular Google searches or fastest times for the Women’s 800m freestyle swimming event)
Three fundamental operations

- **enqueue**(priority, elem): inserts elem with given priority
- **dequeue()**: removes the element with the highest priority from the queue
- **peek()**: returns the element with the highest priority in the queue without removing it
Less fundamental operations

- **size()**: returns the number of elements in the queue
- **isEmpty()**: returns true if there are no elements in the queue, false otherwise
- **clear()**: empties the queue
How do we design `PriorityQueue`?

1. **Member functions:** *What public interface should `PriorityQueue` support? What functions might a client want to call?*

2. **Member variables:** *What private information will we need to store in order to keep track of the data stored in `PriorityQueue`?*

3. **Constructor:** *How are the member variables initialized when a new instance of `PriorityQueue` is created?*
How do we design **PriorityQueue**?

1. **Member functions**: What public interface should **PriorityQueue** support? What functions might a client want to call?

2. Member variables: What private information will we need to store in order to keep track of the data stored in **PriorityQueue**?

3. Constructor: How are the member variables initialized when a new instance of **PriorityQueue** is created?

We’ll provide the public interface...
How do we design `PriorityQueue`?

1. **Member functions:** What public interface should `PriorityQueue` support? What functions might a client want to call?

2. **Member variables:** What private information will we need to store in order to keep track of the data stored in `PriorityQueue`?

3. **Constructor:** How are the member variables initialized when a new instance of `PriorityQueue` is created?

You get to decide on the implementation details!
How do we implement `PriorityQueue`?

- We want to be able to access the element that has the highest priority in constant-time (i.e. `peek()`).
How do we implement **PriorityQueue**?

- We want to be able to access the element that has the highest priority in constant-time (i.e. `peek()`).

- **Idea**: We can keep a sorted array where the elements are in order of their priority (highest priority is at the end of the array)!
  - Dequeue will be fast – just get the last element in the array.
  - But every time we enqueue something, we have to adjust the entire array...
How do we implement **PriorityQueue**?

- We want to be able to access the element that has the highest priority in constant-time (i.e. `peek()`).

- **Idea**: We can keep a sorted array where the elements are in order of their priority (highest priority is at the end of the array)!
  - Dequeue will be fast – just get the last element in the array.
  - But every time we **enqueue** something, we have to adjust the entire array…

You’ll get to implement this on the assignment!
How do we implement **PriorityQueue**?

- We want to be able to access the element that has the highest priority in constant-time (i.e. `peek()`).

- **Idea**: We can keep a sorted array where the elements are in order of their priority (highest priority is at the end of the array)!
  - Dequeue will be fast – just get the last element in the array.
  - But every time we enqueue something, we have to adjust the entire array...

- Can we do better?
How do we implement **PriorityQueue**?

- We want to be able to access the element that has the highest priority in constant-time (i.e. `peek()`).

- **Idea**: We can keep a sorted array where the elements are in order of their priority (highest priority is at the end of the array)!
  - Dequeue will be fast – just get the last element in the array.
  - But every time we enqueue something, we have to adjust the entire array...

- Can we do better? (yes!)

> There are multiple possible implementations for the same ADT!
Levels of abstraction

What is the interface for the user?  
(Priority Queue)

How is our data organized?  
(sorted array, **binary heap**)

What stores our data?  
(arrays)
Announcements
Announcements

● Diagnostic regrade requests AND mid-quarter check-ins are due **tonight by 11:59pm PDT**.
  ○ If you opt into mid-quarter check-ins, you must also sign up to meet with your section leader via the IG scheduling feature on Paperless.

● The final project proposals are due next **Wednesday, August 4 at 11:59pm PDT**. Please read the full guidelines on the course website.

● Assignment 5 was released on Wednesday and is due next **Friday, August 6 at 11:59pm PDT**.
  ○ YEAH hours will be on **Sunday, August 1 at 11:30am PDT**. More details coming on Ed!
Binary Heaps
What is a binary heap?

- A heap is a tree-based structure that satisfies the *heap property* that parents have a higher priority than any of their children.
What is a binary heap?

- A heap is a tree-based structure that satisfies the *heap property* that *parents* have a higher priority than any of their children.

- Additional properties
  - **Binary**: Two children per parent (but no implied orderings between siblings)
What is a binary heap?

• A heap is a tree-based structure that satisfies the heap property that parents have a higher priority than any of their children.

• Additional properties
  ○ Binary: Two children per parent (but no implied orderings between siblings)
  ○ Completely filled (each parent must have 2 children) except for the bottom level, which gets populated from left to right
What is a binary heap?

- A heap is a tree-based structure that satisfies the heap property that parents have a higher priority than any of their children.

- Additional properties
  - **Binary**: Two children per parent (but no implied orderings between siblings)
  - **Completely filled**: (each parents must have 2 children) except for the bottom level, which gets populated from left to right

- Two types → which we use depends on what we define as a “higher” priority
  - Min-heap: smaller numbers = higher priority (closer to the root)
  - Max-heap: larger numbers = higher priority (closer to the root)
What is a binary heap?

- A heap is a tree-based structure that satisfies the heap property that parents have a higher priority than any of their children.

- Additional properties
  - **Binary**: Two children per parent (but no implied orderings between siblings)
  - **Completely filled**: (each parent must have 2 children) except for the bottom level, which gets populated from left to right

- Two types ➔ which we use depends on what we define as a “higher” priority
  - **Min-heap**: smaller numbers = higher priority (closer to the root)
  - **Max-heap**: larger numbers = higher priority (closer to the root)
Spot the Valid Min-Heap

Heap 1

Heap 2
Spot the Valid Min-Heap

Heap 1

- Root node: "a", 4
- Children:
  - "b", 6
    - "d", 7
    - "e", 9
  - "c", 8

Heap 2

- Root node: "a", 4
- Children:
  - "b", 6
    - "d", 7
    - "e", 5
Spot the Valid Min-Heap

Heap 1

- {"a", 4}
  - {"b", 6}
    - {"d", 7}
    - {"e", 9}
  - {"c", 8}

Heap 2

- {"a", 4}
  - {"b", 6}
    - {"d", 7}
    - {"e", 5}
  - {"c", 8}
Spot the Valid Min-Heap

Poll: Which of these heaps is a valid min-heap?
Spot the Valid Min-Heap

Heap 1

Heap 2
Spot the Valid Min-Heap

Heap 1

{"a", 4}
{"b", 6}
{"c", 8}
{"d", 7}
{"e", 9}

Heap 2

{"a", 4}
{"b", 6}
{"c", 8}
{"d", 7}
{"e", 5}

This element is not smaller than both its children!
Spot the Valid Min-Heap (Round 2)

Heap 1

Heap 2
Spot the Valid Min-Heap (Round 2)

Poll: Which of these heaps is a valid min-heap?

Heap 1

Heap 2
Spot the Valid Min-Heap (Round 2)

Heap 1

Heap 2
Spot the Valid Min-Heap (Round 2)

Heap 1

"a", 4

"b", 6

Heap 2

"a", 4

"b", 6

"c", 8

"d", 9

"e", 10

This level of the heap is not complete
Binary heaps and implementation
Binary heaps and implementation

What is the interface for the user? (Priority Queue)

How is our data organized? (sorted array, binary heap)

What stores our data? (arrays)
Binary heaps and implementation

What is the interface for the user? (Priority Queue)

How is our data organized? (binary heap)

What stores our data? (arrays)

Abstract Data Structures

Data Organization Strategies

Fundamental C++ Data Storage
Binary heaps + implementation

- Binary heaps are both another way to implement `PriorityQueue` and also an abstraction on top of arrays!
Binary heaps + implementation

- Binary heaps are both another way to implement `PriorityQueue` and also an abstraction on top of arrays!

- Later, we will see a different approach to storing tree structures, but for heaps (which look like trees), the best solution is actually a simple array.
  - The reason for this is because of the **complete** nature of the structure, with all levels filled from left to right.
Binary heaps + implementation

- Binary heaps are both another way to implement `PriorityQueue` and also an abstraction on top of arrays!

- Later, we will see a different approach to storing tree structures, but for heaps (which look like trees), the best solution is actually a simple array.
  - The reason for this is because of the `complete` nature of the structure, with all levels filled from left to right.

*How are parents and children in the tree related in the array?*
Binary heaps + implementation

![Binary heap diagram]

{"a", 4}  
{"b", 6}  
{"c", 8}  
{"d", 7}  
{"e", 9}

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>
Binary heaps + implementation

Parent index: 0
Left child: 1
Right child: 2
Binary heaps + implementation

Parent index: 1
Left child: 3
Right child: 4
Binary heaps + implementation

Parent: $i$
Left child: $2i + 1$
Right child: $2i + 2$

Parent index: 0
Left child: 1
Right child: 2

Parent index: 1
Left child: 3
Right child: 4
Binary heaps + implementation

Parent: \((i-1)/2\)
Child: \(i\)

Parent index: 0
Left child: 1
Right child: 2

Parent index: 1
Left child: 3
Right child: 4
Binary heaps + implementation

Parent index: 0
Left child: 1
Right child: 2

Parent index: 1
Left child: 3
Right child: 4

Parent index: 2
Left child: 5
Right child: 6
Manipulating heap contents
Heap operations

There are three important operations in a heap:

- **peek()**: return the element with the highest priority (lowest number for a min-heap). This operation does not change the state of the heap at all.

- **enqueue(e)**: insert an element $e$ into the heap. Insertion of this element must result in a heap that still retains the heap property! Accomplishing this will require some clever manipulation.

- **dequeue()**: remove the highest priority (smallest element for a min-heap) from the heap. This changes the state of the heap, and thus we have to do work to restore the heap property.
peek()

- Look at the root of the tree (position 0 in your array)
- O(1)
enqueue()

- How might we go about inserting into a binary heap?

- Example: What if we called `enqueue({"j", 9})` into the heap from before?

- The key is to understand how heaps are built: it is critical that we fill each level from left to right.
enqueue() 

- Start by putting the element into the first empty slot at the bottom level. Similar to how we did with the `OurVector` class, we can say something along the lines of `heap[heapSize] = newElement;`

- Inserting our new element into the first empty slot may have destroyed the heap property so now we have to fix it.

- To do so, we "bubble up" the new element into its correct spot.
enqueue()

- Start by putting the element into the first empty slot at the bottom level. Similar to how we did with the `OurVector` class, we can say something along the lines of `heap[heapSize] = newElement;`

- Inserting our new element into the first empty slot may have destroyed the heap property so now we have to fix it.

- To do so, we "bubble up" the new element into its correct spot.
enqueue()
enqueue()  

Operation: enqueue("j", 9)
enqueue() Operation: enqueue("j", 9)

Problem: The heap property is now violated!
enqueue()

- Start by putting the element into the first empty slot at the bottom level. Similar to how we did with the `OurVector` class, we can say something along the lines of `heap[heapSize] = newElement;`.

- Inserting our new element into the first empty slot may have destroyed the heap property so now we have to fix it.

- To do so, we "bubble up" the new element into its correct spot.
enqueue() 

- Inserting our new element into the first empty slot may have destroyed the heap property so now we have to fix it.

- To do so, we "bubble up" the new element into a spot in the heap that is more fitting of its priority.
  - Look at the newly added element and its parent. Do they have a proper min-heap relationship (that is, is the parent smaller than the child element)?
    - If yes, then we're done, terminate the bubble up process.
    - If not, swap the newly added element with its parent.
  - Repeat the above steps until the process terminates or until the newly added element becomes the root of the heap.
enqueue()
enqueue()

Operation: enqueue("j", 9)

Step 1: Compare current element with its parent.
enqueue() Operation: enqueue("j", 9)

Step 1: Compare current element with its parent.
enqueue()

Operation: enqueue("j", 9)

Step 2: Heap property is violated! We must swap!
enqueue()
enqueue()
enqueue()
enqueue()

Operation: enqueue("j", 9)

Step 1: Compare current element with its parent.
enqueue()

Operation: enqueue("j", 9)

Step 2: Heap property is violated! We must swap!
enqueue()

Operation: enqueue("j", 9)

Step 2: Heap property is violated! We must swap!
enqueue()

Operation: enqueue("j", 9)
enqueue()

Operation: enqueue("j", 9)

Step 1: Compare current element with its parent.
enqueue()

Operation: enqueue("j", 9)

Step 1: Compare current element with its parent.
enqueue()

Operation: enqueue("j", 9)

Step 2: Heap property is satisfied! Terminate bubbling!
enqueue()

Operation: enqueue("j", 9)
enqueue()

- After our "bubble up" process completes, the heap is in a proper state again. Yay! We have now successfully inserted a new element into the heap.

- For a cool animation of this process across many enqueue operations, check out this cool online heap animation.

- What is the runtime complexity of the enqueue operation?
  o In the worst case scenario, we have to bubble up the new element all the way up to the root position.
  o Since there are $n$ total elements, the tree will have $\log n$ levels, which means we would do $\log n$ comparisons and $\log n$ swaps along the way.
  o The overall complexity is $O(\log n)$, which we know is blazingly fast! How cool!
dequeue()

- Remove the minimum element: the root of the tree.
- Replace the root with the “last” element in our tree (last level, farthest right) since we know that location will end up empty.
- Bubble down to regain the heap property!
dequeue()

- Remove the minimum element: the root of the tree.

- Replace the root with the “last” element in our tree (last level, farthest right) since we know that location will end up empty.

- Bubble down to regain the heap property!
dequeue()

- Remove the minimum element: the root of the tree.

- Replace the root with the “last” element in our tree (last level, farthest right) since we know that location will end up empty.

- **Bubble down** to regain the *heap property*!
  - Compare the moved element to its new children.
    - If one of the two children is smaller, swap with that child.
    - If both of the children are smaller, swap with the one that’s smaller.
  - Repeat until you no longer bubble down or there are no more children to compare against.
dequeue()

- Remove the minimum element: the root of the tree.
- Replace the root with the “last” element in our tree (last level, farthest right) since we know that location will end up empty.
- Bubble down to regain the heap property!
- \( O(\log n) \): At worst, you do one comparison at each level of the tree.
dequeue()

- Remove the minimum element: the root of the tree.
- Replace the root with the “last” element in our tree (last level, farthest right) since we know that location will end up empty.
- Bubble down to regain the heap property!
- \( \text{O}(\log n) \): At worst, you do one comparison at each level of the tree.

We have a data structure with only \( \text{O}(\log n) \) and \( \text{O}(1) \) operations!
Summary
Levels of abstraction

What is the interface for the user? (Priority Queue)

How is our data organized? (sorted array, binary heap)

What stores our data? (arrays)

Abstract Data Structures

Data Organization Strategies

Fundamental C++ Data Storage
Summary

- **Priority queues** are queues ordered by **priority** of their elements, where the **highest priority** elements get dequeued first.

- **Binary heaps** are a good way of organizing data when creating a priority queue.
  - Use a min-heap when a smaller number = higher priority (what you’ll use on the assignment) and a max-heap when a larger number = higher priority.

- There can be multiple ways to implement the same abstraction! For both ways of implementing our priority queues, we’ll use **arrays** for data storage.
What’s next?
Memory and Pointers

Man, I suck at this game. Can you give me a few pointers?

I hate you.

0x3A2B2A3A 0x6339392C 0x73636E8E