Priority Queues

T. M. Murali

January 26, 2016

Motivation: Sort a List of Numbers

Sort

INSTANCE: Nonempty list $x_1, x_2, \ldots, x_n$ of integers.

SOLUTION: A permutation $y_1, y_2, \ldots, y_n$ of $x_1, x_2, \ldots, x_n$ such that $y_i \leq y_{i+1}$, for all $1 \leq i < n$.

Motivation: Sort a List of Numbers

Sort

INSTANCE: Nonempty list $x_1, x_2, \ldots, x_n$ of integers.

SOLUTION: A permutation $y_1, y_2, \ldots, y_n$ of $x_1, x_2, \ldots, x_n$ such that $y_i \leq y_{i+1}$, for all $1 \leq i < n$.

- Possible algorithm:
  - Store all the numbers in a data structure D.
  - Repeatedly find the smallest number in D, output it, and remove it.

Motivation: Sort a List of Numbers

Sort

INSTANCE: Nonempty list $x_1, x_2, \ldots, x_n$ of integers.

SOLUTION: A permutation $y_1, y_2, \ldots, y_n$ of $x_1, x_2, \ldots, x_n$ such that $y_i \leq y_{i+1}$, for all $1 \leq i < n$.

- Possible algorithm:
  - Store all the numbers in a data structure D.
  - Repeatedly find the smallest number in D, output it, and remove it.
- To get $O(n \log n)$ running time, each "find minimum" step and each "remove" step must take $O(\log n)$ time.

Candidate Data Structures for Sorting

- Possible algorithm:
  - Store all the numbers in a data structure D.
  - Repeatedly find the smallest number in D, output it, and remove it.
- Data structure must support insertion of a number, finding minimum, and deleting minimum.

Candidate Data Structures for Sorting

- Possible algorithm:
  - Store all the numbers in a data structure D.
  - Repeatedly find the smallest number in D, output it, and remove it.
- Data structure must support insertion of a number, finding minimum, and deleting minimum.

List

Candidate Data Structures for Sorting

- Possible algorithm:
  - Store all the numbers in a data structure D.
  - Repeatedly find the smallest number in D, output it, and remove it.
- Data structure must support insertion of a number, finding minimum, and deleting minimum.

List
: Insertion and deletion take $O(1)$ time but finding minimum requires scanning the list and takes $\Omega(n)$ time.

Sorted array

Candidate Data Structures for Sorting

- Possible algorithm:
  - Store all the numbers in a data structure D.
  - Repeatedly find the smallest number in D, output it, and remove it.
- Data structure must support insertion of a number, finding minimum, and deleting minimum.

List Insertion and deletion take $O(1)$ time but finding minimum requires scanning the list and takes $\Omega(n)$ time.

Sorted array Finding minimum takes $O(1)$ time but insertion and deletion can take $\Omega(n)$ time in the worst case.

Priority Queue

- Store a set S of elements, where each element v has a priority value $\texttt{key}(v)$.
- Smaller key values $\equiv$ higher priorities.
- Operations supported:
  - find the element with smallest key
  - remove the smallest element
  - insert an element
  - delete an element
  - update the key of an element
- Element deletion and key update require knowledge of the position of the element in the priority queue.

Heaps

- Combine benefits of both lists and sorted arrays.
- Conceptually, a heap is a balanced binary tree.
- *Heap order*: For every element v at a node i, the element w at i's parent satisfies $\texttt{key}(w) \leq \texttt{key}(v)$.

Heaps

- Combine benefits of both lists and sorted arrays.
- Conceptually, a heap is a balanced binary tree.
- *Heap order*: For every element v at a node i, the element w at i's parent satisfies $\texttt{key}(w) \leq \texttt{key}(v)$.
- We can implement a heap in a pointer-based data structure.

Heaps

- Combine benefits of both lists and sorted arrays.
- Conceptually, a heap is a balanced binary tree.
- *Heap order*: For every element v at a node i, the element w at i's parent satisfies $\texttt{key}(w) \leq \texttt{key}(v)$.
- We can implement a heap in a pointer-based data structure.
- Alternatively, assume maximum number N of elements is known in advance.
- Store nodes of the heap in an array.
  - Node at index i has children at indices $2i$ and $2i+1$ and parent at index $\lfloor i/2 \rfloor$.
  - Index 1 is the root.
  - How do you know that a node at index i is a leaf?

Heaps

- Combine benefits of both lists and sorted arrays.
- Conceptually, a heap is a balanced binary tree.
- *Heap order*: For every element v at a node i, the element w at i's parent satisfies $\texttt{key}(w) \leq \texttt{key}(v)$.
- We can implement a heap in a pointer-based data structure.
- Alternatively, assume maximum number N of elements is known in advance.
- Store nodes of the heap in an array.
  - Node at index i has children at indices $2i$ and $2i + 1$ and parent at index $\lfloor i/2 \rfloor$.
  - Index 1 is the root.
  - How do you know that a node at index i is a leaf? If $2i > n$, where n is the current number of elements in the heap.

Example of a Heap

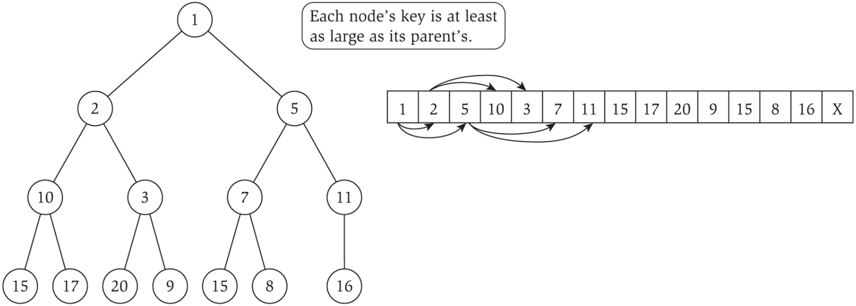

Figure 2.3 Values in a heap shown as a binary tree on the left, and represented as an array on the right. The arrows show the children for the top three nodes in the tree.

Inserting an Element: Heapify-up

1. Insert new element at index $n+1$.
2. Fix heap order using Heapify-up$(H, n+1)$.

```
Heapify-up(H,i):
  If i > 1 then
    let j = parent(i) = ⌊i/2⌋
    If key[H[i]]<key[H[j]] then
      swap the array entries H[i] and H[j]
      Heapify-up(H,j)
    Endif
  Endif
```

Inserting an Element: Heapify-up

1. Insert new element at index $n+1$.
2. Fix heap order using Heapify-up$(H, n+1)$.

```
Heapify-up(H,i):
  If i > 1 then
    let j = parent(i) = ⌊i/2⌋
    If key[H[i]]<key[H[j]] then
      swap the array entries H[i] and H[j]
      Heapify-up(H,j)
    Endif
  Endif
```

- Proof of correctness: read pages 61–62 of your textbook.

Example of Heapify-up

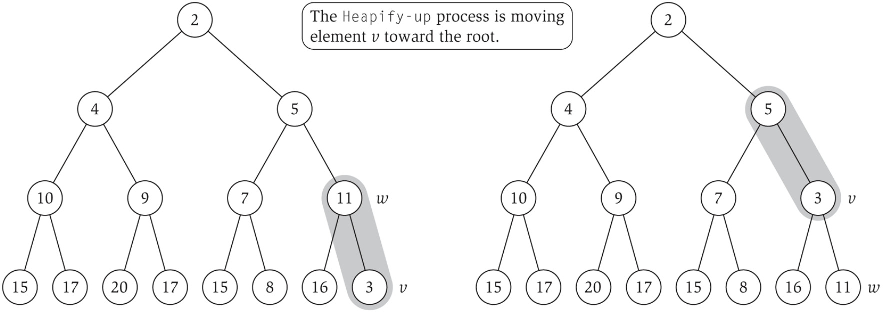

Figure 2.4 The `Heapify-up` process. Key 3 (at position 16) is too small (on the left). After swapping keys 3 and 11, the heap violation moves one step closer to the root of the tree (on the right).

Running time of `Heapify-up`

```
Heapify-up(H,i):
  If i > 1 then
    let j = parent(i) = ⌊i/2⌋
    If key[H[i]]<key[H[j]] then
      swap the array entries H[i] and H[j]
      Heapify-up(H,j)
    Endif
  Endif
```

- Running time of `Heapify-up`(i)

Running time of `Heapify-up`

```
Heapify-up(H,i):
  If i > 1 then
    let j = parent(i) = ⌊i/2⌋
    If key[H[i]]<key[H[j]] then
      swap the array entries H[i] and H[j]
      Heapify-up(H,j)
    Endif
  Endif
```

- Running time of `Heapify-up`(i) is $O(\log i)$.

Running time of `Heapify-up`

```
Heapify-up(H,i):
  If i > 1 then
    let j = parent(i) = ⌊i/2⌋
    If key[H[i]]<key[H[j]] then
      swap the array entries H[i] and H[j]
      Heapify-up(H,j)
    Endif
  Endif
```

- Running time of `Heapify-up`(i) is $O(\log i)$.
- Define $T(i)$ to be the worst-case running time of `Heapify-up`(i) on a heap with i elements.

Running time of `Heapify-up`

```
Heapify-up(H,i):
  If i > 1 then
    let j = parent(i) = ⌊i/2⌋
    If key[H[i]]<key[H[j]] then
      swap the array entries H[i] and H[j]
      Heapify-up(H,j)
    Endif
  Endif
```

- Running time of `Heapify-up`(i) is $O(\log i)$.
- Define $T(i)$ to be the worst-case running time of `Heapify-up`(i) on a heap with i elements.

$$T(i) \leq \begin{cases} T(\lfloor \frac{i}{2} \rfloor) + O(1) & \text{if } i > 1 \\ O(1) & \text{if } i = 1 \end{cases}$$

Deleting an Element: `Heapify-down`

- Suppose H has $n+1$ elements.

1. Delete element at $H[i]$ by moving element at $H[n+1]$ to $H[i]$.
2. If element at $H[i]$ is too small, fix heap order using `Heapify-up(`H, i`)`.
3. If element at $H[i]$ is too large, fix heap order using `Heapify-down(`H, i`)`.

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H, j)
  Endif
```

Deleting an Element: `Heapify-down`

- Suppose H has $n+1$ elements.

1. Delete element at $H[i]$ by moving element at $H[n+1]$ to $H[i]$.
2. If element at $H[i]$ is too small, fix heap order using `Heapify-up`(H, i).
3. If element at $H[i]$ is too large, fix heap order using `Heapify-down`(H, i).

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H,j)
  Endif
```

- Proof of correctness: read pages 63–64 of your textbook.

Example of Heapify-down

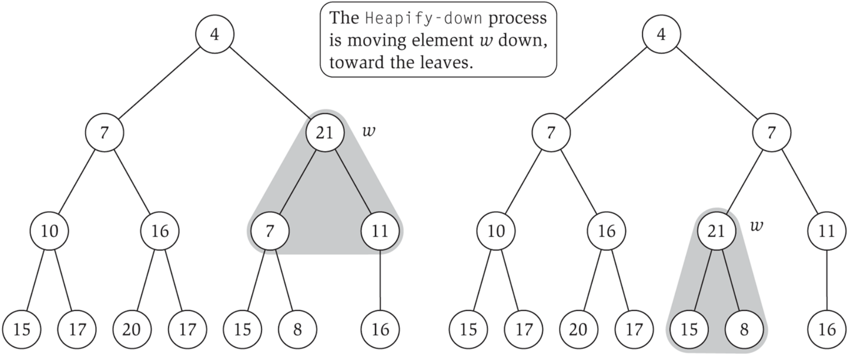

Figure 2.5 The Heapify-down process:. Key 21 (at position 3) is too big (on the left). After swapping keys 21 and 7, the heap violation moves one step closer to the bottom of the tree (on the right).

Running time of `Heapify-down`

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H, j)
  Endif
```

- Recurrence for running time of `Heapify-down`(H, i)

Running time of `Heapify-down`

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H, j)
  Endif
```

- Recurrence for running time of `Heapify-down`(H, i)

$$T(i) = \begin{cases} \max\left(T(2i), T(2i+1)\right) + 1 & \text{if } i > 1 \\ O(1) & \text{if } 2i > n \end{cases}$$

Running time of Heapify-down

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H,j)
  Endif
```

- Recurrence for running time of Heapify-down(H, i)

$$T(i) = \begin{cases} \max\left(T(2i), T(2i+1)\right) + 1 & \text{if } i > 1 \\ O(1) & \text{if } 2i > n \end{cases}$$

- Alternative proof since the recurrence is ugly.

Running time of Heapify-down

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H, j)
  Endif
```

- Recurrence for running time of Heapify-down(H, i)

$$T(i) = \begin{cases} \max\left(T(2i), T(2i+1)\right) + 1 & \text{if } i > 1 \\ O(1) & \text{if } 2i > n \end{cases}$$

- Alternative proof since the recurrence is ugly.
- Every invocation of Heapify-down increases its second argument by a factor of at least two.

Running time of Heapify-down

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H, j)
  Endif
```

- Recurrence for running time of Heapify-down(H, i)

$$T(i) = \begin{cases} \max\left(T(2i), T(2i+1)\right) + 1 & \text{if } i > 1 \\ O(1) & \text{if } 2i > n \end{cases}$$

- Alternative proof since the recurrence is ugly.
- Every invocation of Heapify-down increases its second argument by a factor of at least two.
- After k invocations argument must be at least

Running time of `Heapify-down`

```
Heapify-down(H,i):
  Let n = length(H)
  If 2i > n then
    Terminate with H unchanged
  Else if 2i < n then
    Let left = 2i, and right = 2i + 1
    Let j be the index that minimizes key[H[left]] and key[H[right]]
  Else if 2i = n then
    Let j = 2i
  Endif
  If key[H[j]] < key[H[i]] then
     swap the array entries H[i] and H[j]
     Heapify-down(H, j)
  Endif
```

- Recurrence for running time of `Heapify-down`(H, i)

$$T(i) = \begin{cases} \max\left(T(2i), T(2i+1)\right) + 1 & \text{if } i > 1 \\ O(1) & \text{if } 2i > n \end{cases}$$

- Alternative proof since the recurrence is ugly.
- Every invocation of `Heapify-down` increases its second argument by a factor of at least two.
- After k invocations argument must be at least $i2^k \leq n$, which implies that $k \leq \log_2 n/i$. Therefore running time is $O(\log_2 n/i)$.

Sorting Numbers with the Priority Queue

Sort

INSTANCE: Nonempty list $x_1, x_2, \ldots, x_n$ of integers.

SOLUTION: A permutation $y_1, y_2, \ldots, y_n$ of $x_1, x_2, \ldots, x_n$ such that $y_i \leq y_{i+1}$, for all $1 \leq i < n$.

Sorting Numbers with the Priority Queue

Sort

INSTANCE: Nonempty list $x_1, x_2, \ldots, x_n$ of integers.

SOLUTION: A permutation $y_1, y_2, \ldots, y_n$ of $x_1, x_2, \ldots, x_n$ such that $y_i \leq y_{i+1}$, for all $1 \leq i < n$.

- Final algorithm:
  - Insert each number in a priority queue H.
  - Repeatedly find the smallest number in H, output it, and delete it from H.

Sorting Numbers with the Priority Queue

Sort

INSTANCE: Nonempty list $x_1, x_2, \ldots, x_n$ of integers.

SOLUTION: A permutation $y_1, y_2, \ldots, y_n$ of $x_1, x_2, \ldots, x_n$ such that $y_i \leq y_{i+1}$, for all $1 \leq i < n$.

- Final algorithm:
  - Insert each number in a priority queue H.
  - Repeatedly find the smallest number in H, output it, and delete it from H.
- Each insertion and deletion takes $O(\log n)$ time for a total running time of $O(n \log n)$.