EE360C: Algorithms

Priority Queues

Pedro Santacruz

Department of Electrical and Computer Engineering
University of Texas at Austin

Fall 2014

Motivation: Stable Marriage

The stable marriage algorithm needs a data structure that maintains the dynamically changing set of all free men. The algorithm needs to be able to:

- add elements to the set
- delete elements from the set
- select an element from the set, based on some assigned *priority*

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 of 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 must take $O(\log n)$ time

Candidate Data Structures for Sorting

The data structure we select must support inserting a new element, finding the minimum element, and deleting the minimum element.

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

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

Enter the Priority Queue

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

An Example Application

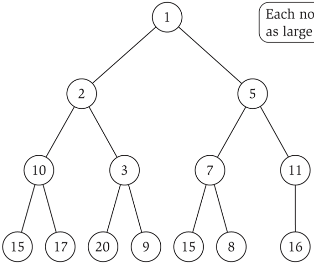
Consider the problem of real-time scheduling of processes on a computer

- each process has a priority
- processes *do not* arrive in order of their priorities
- we need to maintain a set of *active processes* with the ability to quickly extract the one with the highest priority so it can be scheduled
- using a priority queue keyed by process priority, scheduling the highest priority process entails simply finding the one with the lowest priority key

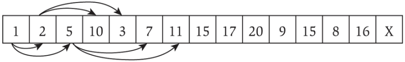
Heaps

- Combine the benefits of both lists and sorted arrays
- Conceptually, a heap is a balanced binary tree
- Heap order: For every element v at 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 a 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 $2i > n$, the number of elements in the heap.

A Heap Example



Each node's key is at least as large as its parent's.

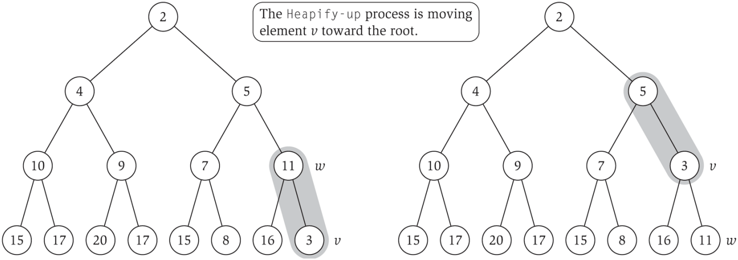


INSERTING AN ELEMENT: `Heapify-up`

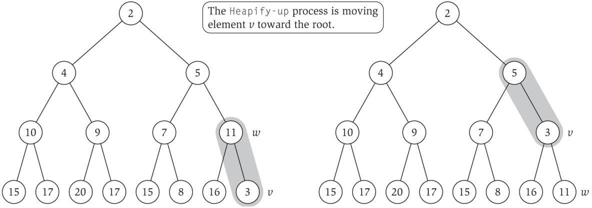
1. Insert a new element at $n+1$
2. Fix the heap order using `Heapify-up`$(H, n+1)$

```
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
```

Heapify-Up EXAMPLE



CORRECTNESS OF `Heapify-Up`



- H is almost a heap with key of $H[i]$ too small if there is a value $\alpha \geq$`key`$(H[i])$ such that increasing `key`$(H[i])$ to α makes H a heap
- Prove by induction on i
  - Base case: $i = 1$
  - Inductive step: if H is almost a heap with key of $H[i]$ too small, after `Heapify-up`(H, i), H is a heap or almost a heap with the key of $H[j]$ too small.
- The running time of `Heapify-up` is $O(\log i)$

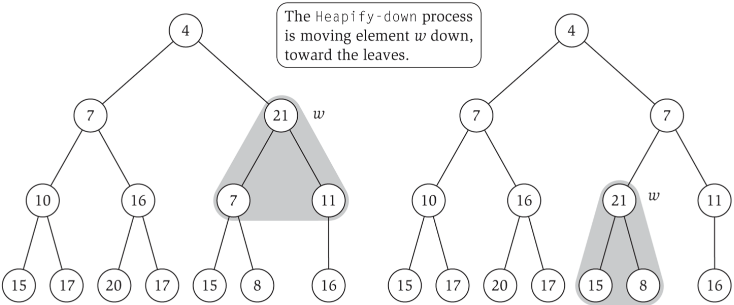
Deleting an Element: Heapify-down

Suppose H has $n+1$ elements

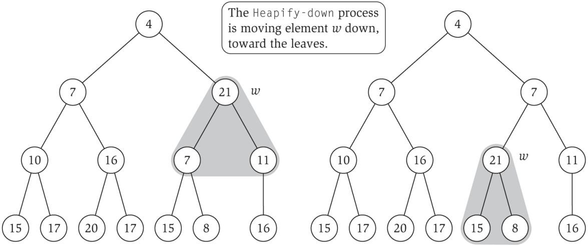
1. Delete element at $H[i]$ by moving element at $H[n+1]$ to $H[i]$
2. If element at $H[i]$ is too small, fix heap order using `Heapify-up`(H, i)
3. If element at $H[i]$ is too large, fix heap order using `Heapify-down`(H, i)

```
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
```

Heapify-down EXAMPLE



Heapify-down CORRECTNESS



- H is almost a heap with key of $H[i]$ too big if there is a value $\alpha \leq$ `key`$(H[i])$ such that decreasing `key`$(H[i])$ to α makes H a heap
- Proof by reverse induction on i
  - Base case: $2i > n$
  - Inductive step: after `Heapify-down`(H, i), H is a heap or almost a heap with the key of $H[j]$ too big
- The running time of `Heapify-down`(H, i) is $O(\log n)$

Sorting with a 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 $O(n \log n)$

Questions

?