Heaps

Jaak Vilo
2010 Spring

Priority queue

- Insert $Q, x$
- Retrieve $x$ from $Q$ s.t. $x.value$ is min (or max)
- Sorted linked list:
  - $O(n)$ to insert $x$ into right place
  - $O(1)$ access-min, $O(1)$ delete-min

Binary heap

- Complete – missing nodes only at the lowest level
- Heap property – on any path the parent has higher priority than child
- Typically: min-heaps

Priority queue

- Insert ($Q, x$)
- pop $Q$

Complete Binary Trees

Array Storage

- Fill the array following a breadth-first traversal:

Heap/Priority queue

- Find min/Delete; Insert;
- Decrease key (change value of the key)
- Merge two heaps ...

Binomial heaps:

- Performance: All of the following operations work in $O(\log n)$ time on a binomial heap with $n$ elements:
  - Insert a new element to the heap
  - Find the element with minimum key
  - Delete the element with minimum key from the heap
  - Decrease key of a given element
  - Delete a given element from the heap
  - Merge two given heaps to one heap
  - Finding the element with minimum key can also be done in $O(1)$ by using an additional pointer to the minimum.
Some links


Binomial trees

Binomial trees, Fibonacci heaps, and applications

Dan Feldman
Lemma 20.1

- For the binomial tree $B_k$,
  1. there are $2^k$ nodes,
  2. the height of the tree is $k$,
  3. there are exactly $\binom{k}{i}$ nodes at depth $i$ for $i = 0, 1, \ldots, k$, and
  4. the root has degree $k$, which is greater than that of any other node; moreover if the children of the root are numbered from left to right by $k - 1, k - 2, \ldots, 0$, child $i$ is the root of a subtree $B_i$.

Properties of binomial trees

1) $|B_k| = 2^k$
2) degree(root($B_k$)) $\leq k$
3) depth($B_k$) $\leq k$

=> The degree and depth of a binomial tree with at most $n$ nodes is at most $\log(n)$.

Define the rank of $B_k$ to be $k$

Binomial heaps (def)

A collection of binomial trees with at most one of every rank.
Items at the nodes, heap ordered.

Figure 20.3 A binomial heap $H$ with $n = 13$ nodes. (a) The heap consists of binomial trees $B_2, B_3$, and $B_4$, which have 1, 4, and 6 nodes respectively, totaling $n = 13$ nodes. Since each binomial tree is min-heap ordered, the key of any node is smaller than the key of its parent. Also shown is the root list, which is a linked list of roots in order of increasing degree. (b) A more detailed representation of binomial heap $H$. Each binomial tree is stored in the left-child, right-sibling representation, and each node stores its degree.

Possible rep: Doubly link roots and children of every node. Parent pointers needed for delete.

Figure 20.4 The binomial tree $B_k$ with nodes labeled in binary by a postorder walk.
Binomial heaps (operations)

Operations are defined via a basic operation, called linking, of binomial trees:
Produce a $B_k$ from two $B_{k-1}$, keep heap order.

Binomial heaps (ops cont.)

Basic operation is $\text{meld}(h_1, h_2)$:
Like addition of binary numbers.

```
    B_3  B_4  B_2  B_1
h1:   B_4  B_3  B_1  B_0  +  
      B_4  B_3  B_0
h2:   B_3  B_4  B_2  B_1

    B_3  B_4  B_2
```

The execution of BINOMIAL-HEAP-UNION.(a)
Binomial heaps $H_1$ and $H_2$.

Delete min

Find min (=1)
Extract tree
Split tree, reverse
Merge/meld
Decrease key \( (y=26 \Rightarrow y=7) \)

Binomial heaps (ops cont.)

Findmin(h): obvious
Insert(x,h): meld a new heap with a single B0 containing x, with h
delete(min)(h): Chop off the minimal root. Meld the subtrees with h. Update minimum pointer if needed.
delete(x,h): Bubble up and continue like delete-min
decrease-key(x,h,\( \delta \)) : Bubble up, update min ptr if needed

All operations take \( O(\log n) \) time on the worst case, except find-min(h) that takes \( O(1) \) time.

Amortized analysis

We are interested in the worst case running time of a sequence of operations.

Example: binary counter

single operation -- increment

```
Increment(A)
1. i=0
2. while i<A.len and A[i]==1
3. A[i] = 0
4. i++
5. if i < A.len
6. A[i] = 1
```

In the amortized running time analysis we pretend that very fast operations take a little bit longer than they actually do.

This additional time is then later subtracted from the actual running time of slow operations.

The amount of time saved for later use is measured at any given moment by a potential function.

Incrementing binary counter

<table>
<thead>
<tr>
<th>value</th>
<th>bits</th>
<th>Total cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increment(A)</td>
<td>0</td>
<td>0 0 0 0 0 0 0 0</td>
</tr>
<tr>
<td>1. i=0</td>
<td>1</td>
<td>0 0 0 0 0 0 0 1</td>
</tr>
<tr>
<td>2. while i&lt;A.len and A[i]==1</td>
<td>2</td>
<td>0 0 0 0 0 0 0 3</td>
</tr>
<tr>
<td>3. A[i] = 0</td>
<td>3</td>
<td>0 0 0 0 0 0 0 4</td>
</tr>
<tr>
<td>4. i++</td>
<td>4</td>
<td>0 0 0 0 0 0 0 7</td>
</tr>
<tr>
<td>5. if i &lt; A.len</td>
<td>5</td>
<td>0 0 0 0 0 0 0 8</td>
</tr>
<tr>
<td>6. A[i] = 1</td>
<td>6</td>
<td>0 0 0 0 0 0 0 10</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>0 0 0 0 0 0 0 11</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>0 0 0 0 0 0 0 15</td>
</tr>
<tr>
<td>9</td>
<td>9</td>
<td>0 0 0 0 0 0 0 16</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>0 0 0 0 0 0 0 18</td>
</tr>
<tr>
<td>11</td>
<td>11</td>
<td>0 0 0 0 0 0 0 19</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>0 0 0 0 0 0 0 22</td>
</tr>
<tr>
<td>13</td>
<td>13</td>
<td>0 0 0 0 0 0 0 23</td>
</tr>
<tr>
<td>14</td>
<td>14</td>
<td>0 0 0 0 0 0 0 25</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>0 0 0 0 0 0 0 26</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>0 0 0 0 0 0 0 31</td>
</tr>
</tbody>
</table>

Amortized analysis (Cont.)

On the worst case increment takes \( O(k) \).
\( k = \# \text{digits} \)

What is the complexity of a sequence of increments (on the worst case)?

Define a potential of the counter:

\[ \Phi(c) = ? \]

Amortized(increment) = actual(increment) + \( \Delta \Phi \)
Amortized analysis (Cont.)

Amortized(increment\textsubscript{1}) = actual(increment\textsubscript{1}) + \Phi\textsubscript{1} - \Phi\textsubscript{0} 
Amortized(increment\textsubscript{i}) = actual(increment\textsubscript{i}) + \Phi\textsubscript{i} - \Phi\textsubscript{i-1} 

\ldots 

Amortized(increment\textsubscript{n}) = actual(increment\textsubscript{n}) + \Phi\textsubscript{n} - \Phi\textsubscript{n-1} 

\sum_{i} Amortized(increment\textsubscript{i}) = \sum_{i} actual(increment\textsubscript{i}) + \Phi\textsubscript{n} - \Phi\textsubscript{0} 

\sum_{i} Amortized(increment\textsubscript{i}) \geq \sum_{i} actual(increment\textsubscript{i}) 

if \Phi\textsubscript{n} - \Phi\textsubscript{0} \geq 0

Amortized analysis (Cont.)

Define a potential of the counter:

\[ \Phi (c) = \#\text{(ones)} \]

Amortized(increment) = actual(increment) + \Delta\Phi

Amortized(increment) = 1 + \#(1 \Rightarrow 0) + 1 - \#(1 \Rightarrow 0) = O(1)

==\> Sequence of \( n \) increments takes \( O(n) \) time

Binomial heaps - amortized ana.

\[ \Phi \text{ (collection of heaps)} = \# \text{(trees)} \]

Amortized cost of insert \( O(1) \)

Amortized cost of other operations still \( O(\log n) \)

Binomial heaps + lazy meld

Allow more than one tree of each rank.

Meld \((h1,h2)\):

- Concatenate the lists of binomial trees.
- Update the minimum pointer to be the smaller of the minimums

\( O(1) \) worst case and amortized.

Binomial heaps + lazy meld

As long as we do not do a delete-min our heaps are just doubly linked lists:

\[ 9 5 9 11 4 6 \]

Delete-min : Chop off the minimum root, add its children to the list of trees.

Successive linking: Traverse the forest keep linking trees of the same rank, maintain a pointer to the minimum root.

Binomial heaps + lazy meld

Possible implementation of delete-min is using an array indexed by rank to keep at most one binomial tree of each rank that we already traversed.

Once we encounter a second tree of some rank we link them and keep linking until we do not have two trees of the same rank. We record the resulting tree in the array

\[ \text{Amortized(delete-min)} = \]

\[ = (\#\text{links} + \text{max-rank}) - \#\text{links} \]

\[ = O(\log(n)) \]
Fibonacci heaps (Fredman & Tarjan 84)

Want to do \texttt{decrease-key}(x,h,\delta) faster than delete+insert.

Ideally in \textbf{O}(1) time.

Why?

---

Dijkstra’s shortest path algorithm

Let \( G = (V,E) \) be a weighted (weights are non-negative) undirected graph, let \( s \in V \). Want to find the distance (length of the shortest path), \( d(s,v) \) from \( s \) to every other vertex.

---

Application #2 : Prim’s algorithm for MST

Start with \( T \) a singleton vertex.

Grow a tree by repeating the following step:

Add the minimum cost edge connecting a vertex in \( T \) to a vertex out of \( T \).

---

Application #2 : Prim’s algorithm for MST

Maintain the vertices out of \( T \) but adjacent to \( T \) in a heap.

The key of a vertex \( v \) is the weight of the lightest edge \((v,w)\) where \( w \) is in the tree.

Iteration: Do a delete-min. Let \( v \) be the minimum vertex and \((v,w)\) the lightest edge as above. Add \((v,w)\) to \( T \). For each edge \((w,u)\) where \( u \notin T \).

if \( \text{key}(u) = \infty \) insert \( u \) into the heap with \( \text{key}(u) = w(w,u) \)

if \( w(w,u) < \text{key}(u) \) decrease the key of \( u \) to be \( w(w,u) \).

With regular heaps \( \textbf{O}(m \log(n)) \).

With F-heaps \( \textbf{O}(n \log(n) + m) \).

---

Insert (left from root)
Finding the minimum node

- The minimum node of a Fibonacci heap $H$ is given by the pointer $\text{min}[H]$, so we can find the minimum node in $O(1)$ actual time. Because the potential of $H$ does not change, the amortized cost of this operation is equal to its $O(1)$ actual cost.

Figure 21.3 The action of FIB-HEAP-EXTRACT-MIN.

Fibonacci heaps (cont.)

Decrease-key $(x, h, \delta)$: indeed cuts the subtree rooted by $x$ if necessary as we showed.

In addition we maintain a mark bit for every node. When we cut the subtree rooted by $x$ we check the mark bit of $p(x)$. If it is set then we cut $p(x)$ too. We continue this way until either we reach an unmarked node in which case we mark it, or we reach the root.

This mechanism is called cascading cuts.

Suggested implementation for decrease-key$(x, h, \delta)$:
If $x$ with its new key is smaller than its parent, cut the subtree rooted at $x$ and add it to the forest. Update the minimum pointer if necessary.
Two calls of FIB-HEAP-DECREASE-KEY.

(a) The initial Fibonacci heap. (b) The node with key 46 has its key decreased to 15. The node becomes a root, and its parent (with key 24), which had previously been unmarked, becomes marked. (c) The node with key 35 has its key decreased to 5. The node becomes a root, and its parent, which is unmarked, becomes marked. (d) A cascading cut occurs, since the node with key 26 is marked. The node is cut from its parent and made an unmarked root. (e) Another cascading cut occurs, since the node with key 24 is marked. The node is cut from its parent and made an unmarked root. The cascading cuts stop at this point, since the node with key 7 is a root. (f) The result of the FIB-HEAP-DECREASE-KEY operation is shown in part (f), with min[H] pointing to the new minimum node.

Decrease-key (cont.)

Does it work?

Obs1: Trees need not be binomial trees any more.

Do we need the trees to be binomial?

Where have we used it?

In the analysis of delete-min we used the fact that at most log(n) new trees are added to the forest. This was obvious since trees were binomial and contained at most n nodes.

Fibonacci heaps (cont.)

We shall allow non-binomial trees, but will keep the degrees logarithmic in the number of nodes.

Rank of a tree = degree of the root.

Delete-min: do successive linking of trees of the same rank and update the minimum pointer as before.

Insert and meld also work as before.

Fibonacci heaps (delete)

Delete(x,h) : Cut the subtree rooted at x and then proceed with cascading cuts as for decrease key.

Chop off x from being the root of its subtree and add the subtrees rooted by its children to the forest

If x is the minimum node do successive linking.
• The potential of a Fibonacci heap is given by
  
  \[ \text{Potential} = t + 2m \]
  
  where \( t \) is the number of trees in the Fibonacci heap, and \( m \) is the number of marked nodes. A node is marked if at least one of its children was cut since this node was made a child of another node (all roots are unmarked).

\[ \Phi(\text{coll} \text{ec} \text{tion of heaps}) = \#(\text{trees}) + 2\#(\text{marked nodes}) \]

Actual\( (\text{decrease-key}) = O(1) + \#(\text{cascading cuts}) \)

\[ \Delta\Phi(\text{decrease-key}) = O(1) - \#(\text{cascading cuts}) \]

\[ \Rightarrow \text{amortized}(\text{decrease-key}) = O(1) \]

Fibonacci heaps (analysis)

What about delete and delete-min?

Cascading cuts and successive linking will pay for themselves. The only question is what is the maximum degree of a node?

How many trees are being added into the forest when we chop off a root?

Fibonacci heaps (analysis)

Lemma 1: Let \( x \) be any node in an F-heap. Arrange the children of \( x \) in the order they were linked to \( x \), from earliest to latest. Then the \( i \)-th child of \( x \) has rank at least \( i-2 \).

Proof:

When the \( i \)-th node was linked it must have had at least \( i-1 \) children.

Fibonacci heaps (analysis)

Corollary 1: A node \( x \) of rank \( k \) in a F-heap has at least \( \Phi^k \) descendants, where \( \Phi = (1 + \sqrt{5})/2 \) is the golden ratio.

Proof:

Let \( s_k \) be the minimum number of descendants of a node of rank \( k \) in a F-heap.

By Lemma 1 \( s_k = \sum_{i=0}^{k} s_i + 2 \)

\[ s_0 = 1, s_1 = 2 \]

Fibonacci heaps (analysis)

Figure 26.1: (a) A Fibonacci heap consisting of five heap-ordered trees and 14 nodes. The dashed line indicates the root \( x \). The minimum node of the heap is the node containing the key 3. The three marked nodes are shaded. The potential of the particular Fibonacci heap is \( 5 + 2 \times 3 = 11 \).

(b) A more complete representation showing pointers to (up arrows), child (down arrows), and self and right (side-ways arrows). These details are omitted in the remaining figures in this chapter, since all the information shown here can be determined from what appears in part (a).
Fibonacci heaps (analysis)

Proof (cont):
Fibonacci numbers satisfy
\[ F_{k+2} = 2 \sum_{i=2} F_i + 2, \quad k \geq 2, \quad F_2=1 \]
so by induction \( s_k \geq F_{k+2} \)
It is well known that \( F_{k+2} \geq \phi^k \)

It follows that the maximum degree \( k \) in a F-heap
with \( n \) nodes is such that
\[ \phi^k \leq \sqrt{5} \]
so \( k \leq \log(n) / \log(\phi) = 1.4404 \log(n) \)

---

Make-Fibonacci-Heap
\[
\begin{align*}
n[H] &:= 0 \\
\text{min}[H] &:= \text{NIL} \\
\text{return } H
\end{align*}
\]

Fibonacci-Heap-Minimum
\[
\begin{align*}
\text{return } \text{min}[H]
\end{align*}
\]

Fibonacci-Heap-Link
\[
\begin{align*}
\text{remove } y \text{ from the root list of } H \\
\text{make } y \text{ a child of } x \\
\text{degree}[x] &:= \text{degree}[x] + 1 \\
\text{mark}[y] &:= \text{FALSE}
\end{align*}
\]

CONSOLIDATE(H)
\[
\begin{align*}
&\text{for } i = 0 \text{ to } D(n[H]) \\
&\quad \text{do } A[i] := \text{NIL} \\
&\text{for each node } w \text{ in the root list of } H \\
&\quad \text{do } x := w \\
&\quad \text{d} := \text{degree}(x) \\
&\quad \text{while } A[d] != \text{NIL} \\
&\quad \text{do } y := A[d] \\
&\quad \text{if key}[x] > key[y] \\
&\quad \text{then exchange } x \leftrightarrow y \\
&\quad \text{Fibonacci-Heap-Link}(H, y, x) \\
&\quad A[d] := \text{NIL} \\
&\quad d := d + 1 \\
&\quad A[d] := x
\end{align*}
\]

Fibonacci-Heap-Union
\[
\begin{align*}
&\text{H1} \\
&\text{H2} \\
&\text{H} := \text{Make-Fibonacci-Heap}() \\
&\text{n}[H] := \text{n}[H1] + \text{n}[H2] \\
&\text{return } H
\end{align*}
\]

Fibonacci-Heap-Insert
\[
\begin{align*}
&\text{degree}[x] := 0 \\
&p[x] := \text{NIL} \\
&\text{child}[x] := \text{NIL} \\
&\text{left}[x] := x \\
&\text{right}[x] := \text{FALSE} \\
&\text{mark}[x] := \text{FALSE} \\
&\text{concatenate the root list containing } x \text{ with root list } H \\
&\text{if } \text{min}[H] = \text{NIL} \text{ or } \text{key}[x] < \text{key}[\text{min}[H]] \\
&\text{then } \text{min}[H] := x \\
&\text{n}[H] := \text{n}[H]+1
\end{align*}
\]
Fibonacci-Heap-Extract-Min\((H)\)
\[ z := \text{min}[H] \]
\[ \text{if } x = \text{NIL} \]
\[ \text{then for each child } x \text{ of } z \]
\[ \text{do add } x \text{ to the root list of } H \]
\[ p[x] := \text{NIL} \]
\[ \text{remove } z \text{ from the root list of } H \]
\[ \text{if } z = \text{right}[z] \]
\[ \text{then } \text{min}[H] := \text{NIL} \]
\[ \text{else } \text{min}[H] := \text{right}[z] \]
\[ \text{CONsolidate}(H) \]
\[ n[H] := n[H]-1 \]
\[ \text{return } z \]

Fibonacci-Heap-Decrease-Key\((H, x, k)\)
\[ \text{if } k > \text{key}[x] \]
\[ \text{then error "new key is greater than current key"} \]
\[ \text{key}[x] := k \]
\[ y := p[x] \]
\[ \text{if } y = \text{NIL} \text{ and key}[x] < \text{key}[y] \]
\[ \text{then } \text{CUT}(H, x, y) \]
\[ \text{CASCAding-CUT}(H, y) \]
\[ \text{if } \text{key}[x] < \text{key}[\text{min}[H]] \]
\[ \text{then } \text{min}[H] := x \]

CUT\((H, x, y)\)
Remove \(x\) from the child list of \(y\), decrementing degree\([y]\)
Add \(x\) to the root list of \(H\)
\[ p[x] := \text{NIL} \]
\[ \text{mark}[x] := \text{FALSE} \]

CASCADING-CUT\((H, y)\)
\[ z := p[y] \]
\[ \text{if } z = \text{NIL} \]
\[ \text{then if } \text{mark}[y] = \text{FALSE} \]
\[ \text{then } \text{mark}[y] := \text{TRUE} \]
\[ \text{else } \text{CUT}(H, y, z) \]
\[ \text{CASCADING-CUT}(H, z) \]

C code
- http://www.cs.unc.edu/~bbb/foe/binheaps/ihp.h
- http://www.cs.unc.edu/~bbb/Binomial_Heaps

Visualization:
van Emde Boas tree

- A van Emde Boas tree (or van Emde Boas priority queue), also known as a vEB tree, is a tree data structure which implements an associative array with \( m \)-bit integer keys. It performs all operations in \( O(\log m) \) time. Notice that \( m \) is the size of the keys — therefore \( O(\log m) \) is \( O(\log \log n) \) in a full tree, exponentially better than a self-balancing binary search tree. They also have good space efficiency when they contain a large number of elements, as discussed below. They were invented by a team led by Peter van Emde Boas in 1977.[1]

PQ – Structure (Cont’d)

- Subset \( S \subseteq \{1, \ldots, n\} \) Representation
  - Mark leaves in \( S \) and all nodes on the paths from root to the marked leaves.

PQ – (Cont’d)

Sketches of Algorithms

- Insert(\( i \))
- Delete(\( i \))
- Member(\( i \))
- Min
- Predecessor
PQ – (Cont’d)
Sketches of Algorithms

• Insert(i)
• Delete(i)
• Member(i)
• Min
• Predecessor