"There's nothing in your head the sorting hat can't see. So try me on and I will tell you where you ought to be."

-The Sorting Hat, *Harry Potter and the Sorcerer's Stone*
Sorting and Searching

- Fundamental problems in computer science and programming
- Sorting done to make searching easier
- Multiple different algorithms to solve the same problem
  - How do we know which algorithm is "better"?
- Look at searching first
- Examples use arrays of ints to illustrate algorithms
Searching

Google

recursive backtracking

Advanced Search

Preferences

Language Tools

Grep in Project

Searching for "automatic"

11 matching lines:

1. valid, and will automatically terminate your rights under this License.
2. Program, the recipient automatically receives a license from the
3. Then, the excellent work should now automatically have
4. * Added a new option to automatically convert "fuzzy" entries into
5. in "automatically convert fuzzy; entries into

Google Search

I'm Feeling Lucky
Searching

- Given a list of data find the location of a particular value or report that value is not present
- linear search
  - intuitive approach?
  - start at first item
  - is it the one I am looking for?
  - if not go to next item
  - repeat until found or all items checked
- If items not sorted or unsortable this approach is necessary
Linear Search

/*  
pre: list != null  
post: return the index of the first occurrence  
of target in list or -1 if target not present in  
list  
*/

public int linearSearch(int[] list, int target) {
    for(int i = 0; i < list.length; i++)
        if( list[i] == target )
            return i;
    return -1;
}
Linear Search, Generic

/*
 * pre: list != null
 * post: return the index of the first occurrence
 * of target in list or -1 if target not present in
 * list
 */

public int linearSearch(Object[] list, Object target) {
    for(int i = 0; i < list.length; i++)
        if(target.equals(list[i]))
            return i;
    return -1;
}

T(N)? Big O? Best case, worst case, average case?
Attendance Question 1

What is the average case Big O of linear search in an array with N items, if an item is present?

A. O(N)
B. O(N^2)
C. O(1)
D. O(logN)
E. O(NlogN)
Searching in a Sorted List

- If items are sorted then we can *divide and conquer*
- dividing your work in half with each step
  - generally a good thing
- The Binary Search on List in Ascending order
  - Start at middle of list
  - is that the item?
  - If not is it less than or greater than the item?
  - less than, move to second half of list
  - greater than, move to first half of list
  - repeat until found or sub list size = 0
Binary Search

Is middle item what we are looking for? If not is it more or less than the target item? (Assume lower)

and so forth…
Binary Search in Action

<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>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
<td>5</td>
<td>7</td>
<td>11</td>
<td>13</td>
<td>17</td>
<td>19</td>
<td>23</td>
<td>29</td>
<td>31</td>
<td>37</td>
<td>41</td>
<td>43</td>
<td>47</td>
<td>53</td>
</tr>
</tbody>
</table>

public static int bsearch(int[] list, int target) {
    int result = -1;
    int low = 0;
    int high = list.length - 1;
    while( result == -1 && low <= high ) {
        int mid = low + ((high - low) / 2);
        if( list[mid] == target )
            result = mid;
        else if( list[mid] < target)
            low = mid + 1;
        else
            high = mid - 1;
    }
    return result;
}

// mid = ( low + high ) / 2; // may overflow!!!
// or mid = (low + high) >>> 1; using bitwise op
Trace When Key == 3
Trace When Key == 30

Variables of Interest?
Attendance Question 2

What is the worst case Big O of binary search in an array with N items, if an item is present?

A. O(N)
B. O(N^2)
C. O(1)
D. O(logN)
E. O(NlogN)
public static <T extends Comparable<? super T>> int bsearch(T[] list, T target) {

    int result = -1;
    int low = 0;
    int high = list.length - 1;
    while (result == -1 && low <= high) {
        int mid = low + ((high - low) / 2);
        int compareResult = target.compareTo(list[mid]);
        if (compareResult == 0)
            result = mid;
        else if (compareResult > 0)
            low = mid + 1;
        else
            high = mid - 1; // compareResult < 0
    }
    return result;
}
Recursive Binary Search

```java
public static int bsearch(int[] list, int target) {
    return bsearch(list, target, 0, list.length - 1);
}

public static int bsearch(int[] list, int target, int low, int high) {
    if( low <= high){
        int mid = low + ((high - low) / 2);
        if( list[mid] == target )
            return mid;
        else if( list[mid] > target )
            return bsearch(list, target, low, mid - 1);
        else
            return bsearch(list, target, mid + 1, high);
    } return -1;
}

// is this a recursive backtracking algorithm?
```
Other Searching Algorithms

- Interpolation Search
  – more like what people really do
- Indexed Searching
- Binary Search Trees
- Hash Table Searching
- Grover's Algorithm (Waiting for quantum computers to be built)
- best-first
- A*
### U.S. All-time List - Marathon

As of 4/24/08

#### Women

<table>
<thead>
<tr>
<th>Place</th>
<th>Time</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2:19:36</td>
<td>Deena Kastor nee Drossin</td>
</tr>
<tr>
<td>2</td>
<td>2:21:16</td>
<td>Drossin (2)</td>
</tr>
<tr>
<td>3</td>
<td>2:21:21</td>
<td>Joan Benoit Samuelson</td>
</tr>
<tr>
<td>4</td>
<td>2:21:25</td>
<td>Kastor (3)</td>
</tr>
<tr>
<td>5</td>
<td>2:22:43a</td>
<td>Benoit (2)</td>
</tr>
<tr>
<td>6</td>
<td>2:24:52a</td>
<td>Benoit (3)</td>
</tr>
<tr>
<td>7</td>
<td>2:26:11a</td>
<td>Benoit (4)</td>
</tr>
<tr>
<td>8</td>
<td>2:26:26a</td>
<td>Julie Brown</td>
</tr>
<tr>
<td>9</td>
<td>2:26:40a</td>
<td>Kim Jones</td>
</tr>
</tbody>
</table>
Sorting

- A fundamental application for computers
- Done to make finding data (searching) faster
- Many different algorithms for sorting
- One of the difficulties with sorting is working with a fixed size storage container (array)
  - if resize, that is expensive (slow)
- The "simple" sorts run in quadratic time $O(N^2)$
  - bubble sort
  - selection sort
  - insertion sort
Selection sort

Algorithm
- Search through the list and find the smallest element
- swap the smallest element with the first element
- repeat starting at second element and find the second smallest element

```java
public static void selectionSort(int[] list) {
    for(int i = 0; i < list.length - 1; i++) {
        int min = i;
        for(int j = i + 1; j < list.length; j++)
            if( list[j] < list[min] )
                min = j;
        int temp = list[i];
        list[i] = list[min];
        list[min] = temp;
    }
}
```
Selection Sort in Practice

What is the $T(N)$, _actual_ number of statements executed, of the selection sort code, given a list of $N$ elements? What is the Big O?
Generic Selection Sort

public static <T extends Comparable<? super T>>
    void selectionSort(T[] list) {

        for(int i = 0; i < list.length - 1; i++) {
            int min = i;
            for(int j = i + 1; j < list.length; j++)
                if(list[min].compareTo(list[j]) > 0 )
                    min = j;
            T temp = list[i];
            list[i] = list[min];
            list[min] = temp;
        }
    }

Insertion Sort

- Another of the $O(N^2)$ sorts
- The first item is sorted
- Compare the second item to the first
  - if smaller swap
- Third item, compare to item next to it
  - need to swap
  - after swap compare again
- And so forth...
Insertion Sort Code

public void insertionSort(int[] list) {
    for(int i = 1; i < list.length; i++) {
        int temp = list[i];
        int j = i;
        while( j > 0 && temp < list[j - 1]){
            // swap elements
            list[j] = list[j - 1];
            list[j - 1] = temp;
            j--;
        }
    }
}

¬ Best case, worst case, average case Big O?
Comparing Algorithms

Which algorithm do you think will be faster given random data, selection sort or insertion sort?

A. Insertion Sort
B. Selection Sort
C. About the same