- Previous Lecture:
  - Linear Search
  - Bubble Sort, Insertion Sort
- Today's Lecture:
  - "Divide and conquer" strategies
    - Binary search
    - Merge sort
    - Recursion
- Announcements:
  - Discussion this week in classrooms (today and up to 1:10pm tomorrow). Attendance this week is optional, but the posted exercise is required!
  - Prelim 3 will be returned at end of lecture. If your paper isn't here, pick it up from CS1112 consultants in ACCEL during consulting hrs (today 4-10pm; consulting resumes Mon 4pm)
  - Project 6 due Dec 2nd. Part 1 posted; Part 2 to be posted today. Can wait until next Tues lecture to do Part 2.2, but do Part 1 and Part 2.1 by next Tues!

Other efficiency considerations

- Worst case, best case, average case
- Use of subfunction incurs an "overhead"
- Memory use and access

- Example: Rather than directing the *insert* process to a subfunction, have it done "in-line."
- Also, Insertion sort can be done "in-place," i.e., using "only" the memory space of the original vector.


```
function x = insertSort(x)
% Sort vector x in ascending order with insertion sort

n = length(x);
for i= 1:n-1
      % Sort x(1:i+1) given that x(1:i) is sorted
      j= i;
      need2swap= x(j+1) < x(j);
      while  need2swap

          % swap x(j+1) and x(j)
            temp= x(j);
            x(j)= x(j+1);
            x(j+1)= temp;

          j= j-1;
          need2swap= j>0 && x(j+1)<x(j);
      end
end
```

Sorting is done "in-place."
Same is true for Bubble Sort.

An ordered (sorted) list

The Manhattan phone book has 1,000,000+ entries.

How is it possible to locate a name by examining just a tiny, tiny fraction of those entries?

Key idea of "phone book search": repeated halving

To find the page containing Pat Reed's number…

```
while  (Phone book is longer than 1 page)
      Open to the middle page.
      if  "Reed" comes before the first entry,
            Rip and throw away the 2nd half.
      else
            Rip and throw away the 1st half.
      end
end
```


What happens to the phone book length?

```
Original:      3000 pages
After 1 rip:   1500 pages
After 2 rips:   750 pages
After 3 rips:   375 pages
After 4 rips:   188 pages
After 5 rips:    94 pages
       :
After 12 rips:    1 page
```

Binary Search

Repeatedly halving the size of the "search space" is the main idea behind the method of binary search.

An item in a sorted array of length n can be located with just $\log_2 n$ comparisons.


```
% Linear Search
% f is index of first occurrence of value x in vector v.
% f is -1 if x not found.
k= 1;
while  k<=length(v) && v(k)~=x
    k= k + 1;
end
if  k>length(v)
    f= -1; % signal for x not found
else
    f= k;
end
```

n comparisons against the target are needed in worst case, n=length(v).

Binary Search

Repeatedly halving the size of the "search space" is the main idea behind the method of binary search.

An item in a sorted array of length n can be located with just $\log_2 n$ comparisons.

"Savings" is significant!

n	log2(n)
100	7
1000	10
10000	13

Binary search: target x = 70

	1	2	3	4	5	6	7	8	9	10	11	12
v	12	15	33	35	42	45	51	62	73	75	86	98

L: 1
Mid: 6
R: 12

v(Mid) <= x

So throw away the left half…

Binary search: target x = 70

	1	2	3	4	5	6	7	8	9	10	11	12
v	12	15	33	35	42	45	51	62	73	75	86	98

L: 6
Mid: 9
R: 12

x < v(Mid)

So throw away the right half…

Binary search: target x = 70

	1	2	3	4	5	6	7	8	9	10	11	12
v	12	15	33	35	42	45	51	62	73	75	86	98

L: 6
Mid: 7
R: 9

v(Mid) <= x

So throw away the left half…

Binary search: target x = 70

	1	2	3	4	5	6	7	8	9	10	11	12
v	12	15	33	35	42	45	51	62	73	75	86	98

L: 7
Mid: 8
R: 9

v(Mid) <= x

So throw away the left half…

Binary search: target x = 70

	1	2	3	4	5	6	7	8	9	10	11	12
v	12	15	33	35	42	45	51	62	73	75	86	98

L: 8
Mid: 8
R: 9

Done because R-L = 1


```
function L = binarySearch(x, v)
% Find position after which to insert x. v(1)<…<v(end).
% L is the index such that v(L) <= x < v(L+1);
% L=0 if x<v(1).  If x>v(end), L=length(v) but x~=v(L).

% Maintain a search window [L..R] such that v(L)<=x<v(R).
% Since x may not be in v, initially set ...
L=0;  R=length(v)+1;

% Keep halving [L..R] until R-L is 1,
%   always keeping  v(L) <= x < v(R)
while  R ~= L+1
    m= floor((L+R)/2);  % middle of search window
    if


    else


    end
end
```

Binary search is efficient, but how do we sort a vector in the first place so that we can use binary search?

- Many different algorithms out there…
- We saw bubble sort and insertion sort
- Let's look at **merge sort**
- An example of the "divide and conquer" approach
- We'll compare their efficiency later

Merge sort: Motivation

If I have two helpers, I'd…

- Give each helper half the array to sort
- Then I get back the sorted subarrays and merge them.

What if those two helpers each had two sub-helpers?

And the sub-helpers each had two sub-sub-helpers? And…

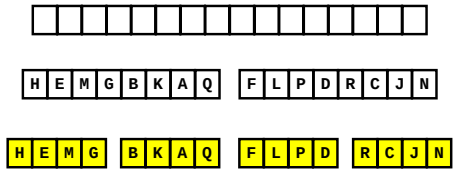
Subdivide the sorting task

H	E	M	G	B	K	A	Q	F	L	P	D	R	C	J	N

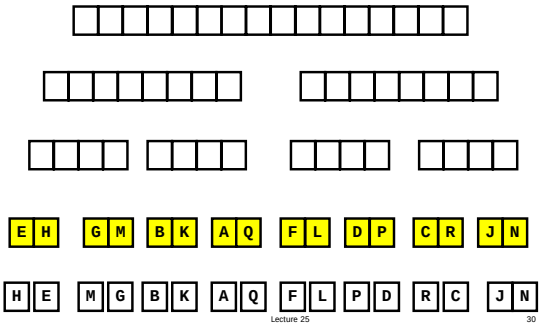
H	E	M	G	B	K	A	Q

F	L	P	D	R	C	J	N

Subdivide again



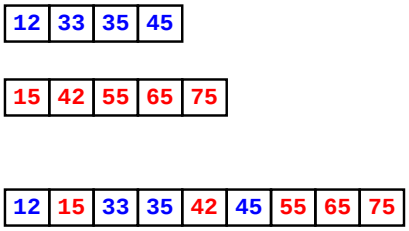
Now merge



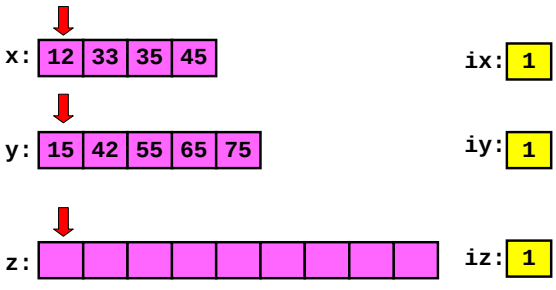

```
function y = mergeSort(x)
% x is a vector.  y is a vector
% consisting of the values in x
% sorted from smallest to largest.

n = length(x);
if n==1
      y = x;
else
      m  = floor(n/2);
      yL = mergeSortL(x(1:m));
      yR = mergeSortR(x(m+1:n));
      y  = merge(yL,yR);
end
```


The central sub-problem is the merging of two sorted arrays into one single sorted array

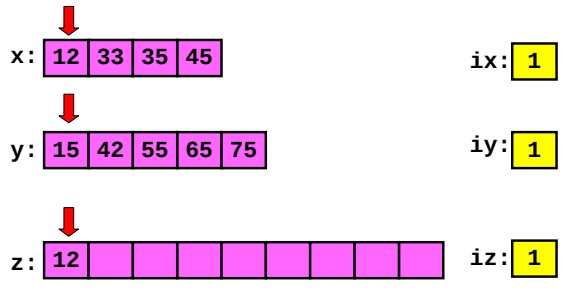


Merge

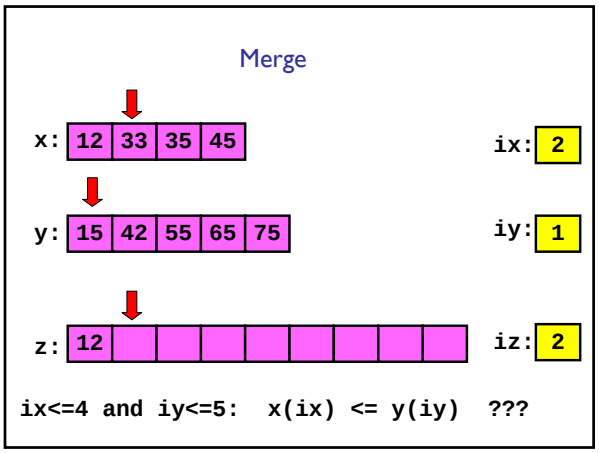


ix<=4 and iy<=5: x(ix) <= y(iy) ???

Merge



ix<=4 and iy<=5: x(ix) <= y(iy) YES




```
function z = merge(x,y)
nx = length(x); ny = length(y);
z = zeros(1, nx+ny);
ix = 1; iy = 1; iz = 1;
while ix<=nx && iy<=ny




end
% Deal with remaining values in x or y
```

```
function y = mergeSort(x)
% x is a vector.  y is a vector
% consisting of the values in x
% sorted from smallest to largest.

n = length(x);
if n==1
      y = x;
else
      m  = floor(n/2);
      yL = mergeSortL(x(1:m));
      yR = mergeSortR(x(m+1:n));
      y  = merge(yL,yR);
end
```



```
function y = mergeSortL(x)
% x is a vector.  y is a vector
% consisting of the values in x
% sorted from smallest to largest.

n = length(x);
if n==1
      y = x;
else
      m  = floor(n/2);
      yL = mergeSortL_L(x(1:m));
      yR = mergeSortL_R(x(m+1:n));
      y  = merge(yL,yR);
end
```



```
function y = mergeSort(x)
% x is a vector.  y is a vector
% consisting of the values in x
% sorted from smallest to largest.

n = length(x);
if n==1
      y = x;
else
      m  = floor(n/2);
      yL = mergeSort(x(1:m));
      yR = mergeSort(x(m+1:n));
      y  = merge(yL,yR);
end
```



```
function y=mergeSort(x)
n=length(x);
if n==1
      y=x;
else
      m=floor(n/2);
      yL=mergeSort(x(1:m));
      yR=mergeSort(x(m+1:n));
      y=merge(yL,yR);
end
```