CPE 101, slides adapted from UW course

Lecture 15:
Linear & Binary Search

Concepts This Lecture

Searching an array

Linear search

Binary search

Comparing algorithm performance

Searching

Searching = looking for something

Searching an array is particularly common

- Goal: determine if a particular value is in the array

We'll see that more than one algorithm will work

Searching as a Function

Specification:

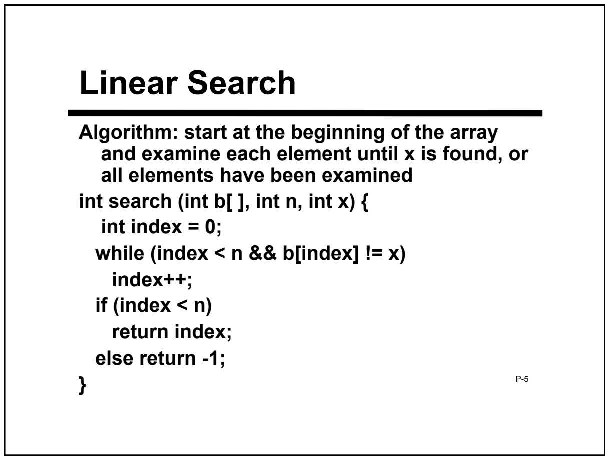
Let b be the array to be searched, n is the size of the array, and x is value we search to find.

If x appears in b[0..n-1], return its index, i.e., return k such that b[k]==x. If x not found, return –1

None of the parameters are changed by the function

Function outline:

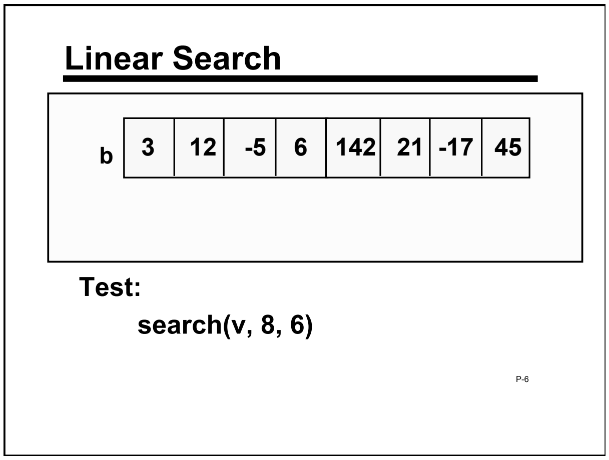
```
int search (int b[ ], int n, int x) {
...
}
```



Linear Search

Algorithm: start at the beginning of the array and examine each element until x is found, or all elements have been examined

```
int search (int b[ ], int n, int x) {
   int index = 0;
  while (index < n && b[index] != x)
     index++;
  if (index < n)
     return index;
  else return -1;
}
```

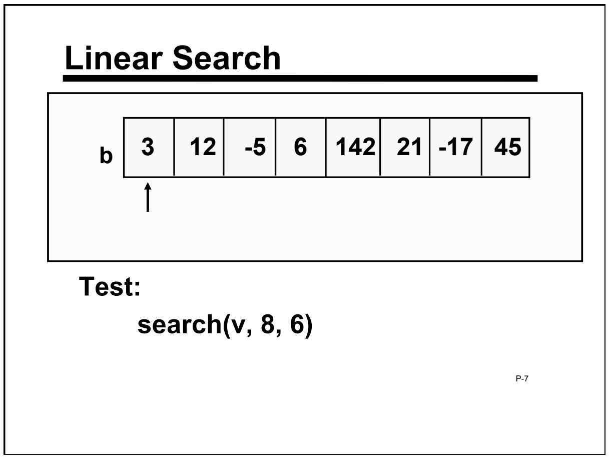


Linear Search

b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 6)

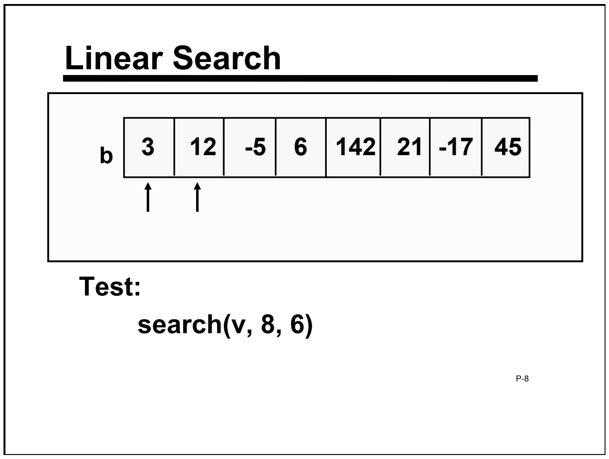


Linear Search

b	3	12	-5	6	142	21	-17	45
	↑							

Test:

search(v, 8, 6)

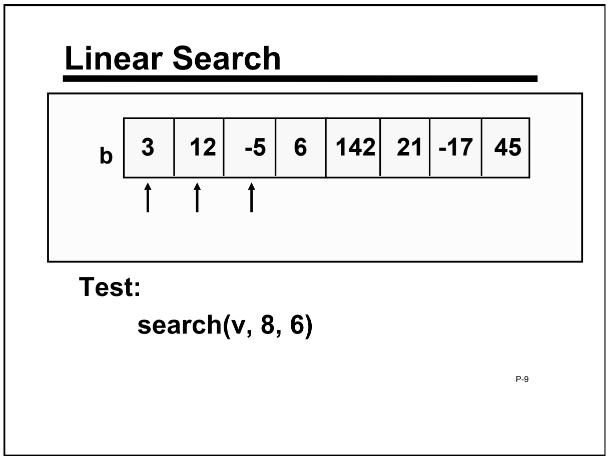


Linear Search

b	3	12	-5	6	142	21	-17	45
	↑	↑						

Test:

search(v, 8, 6)

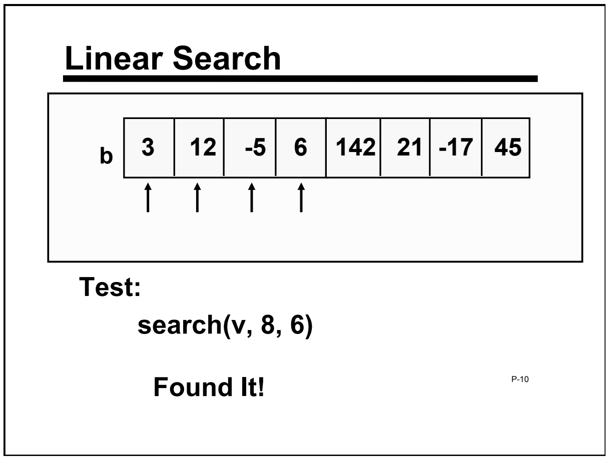


Linear Search

b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 6)



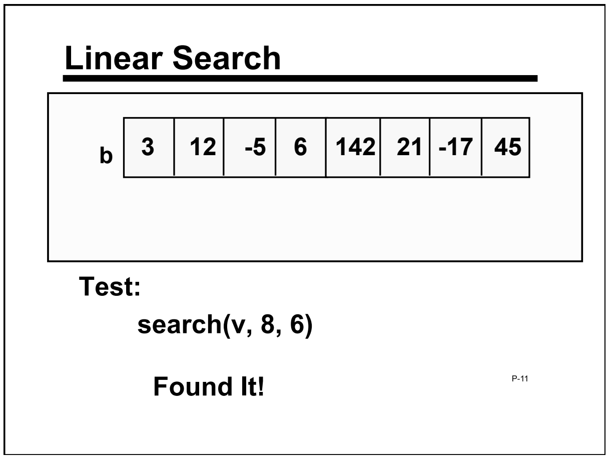
Linear Search

b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 6)

Found It!



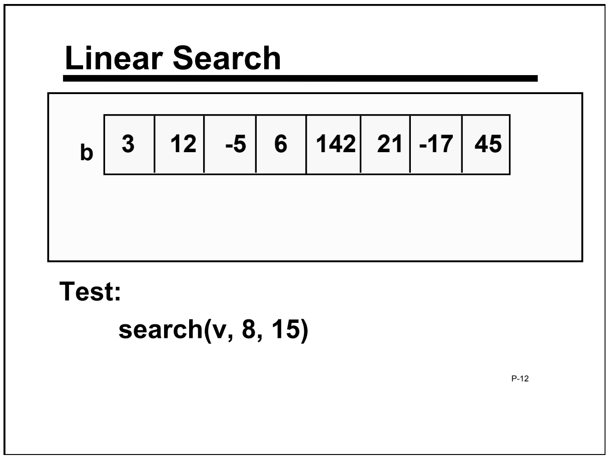
Linear Search

b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 6)

Found It!

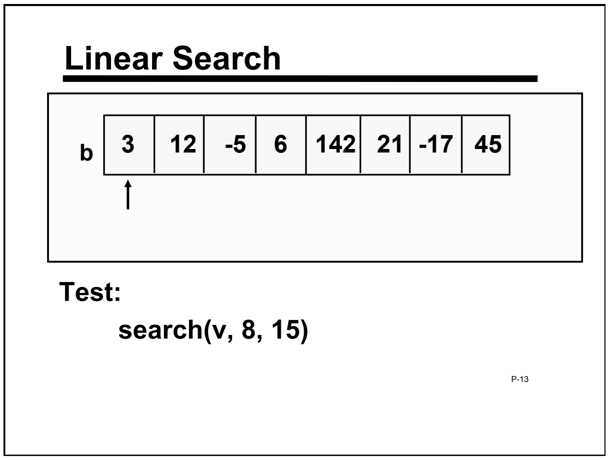


Linear Search

b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 15)



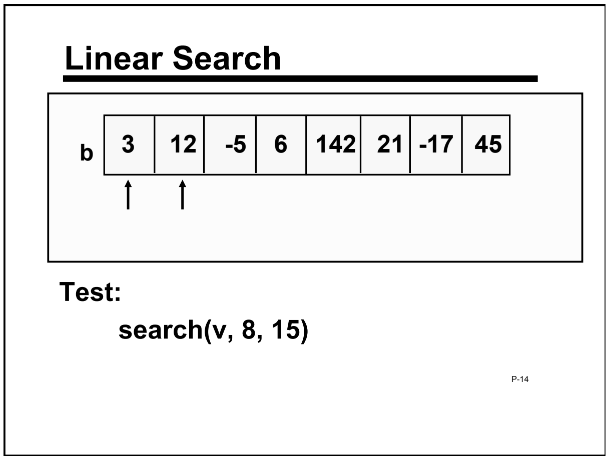
Linear Search

b	3	12	-5	6	142	21	-17	45
	↑							

Test:

search(v, 8, 15)

P-13



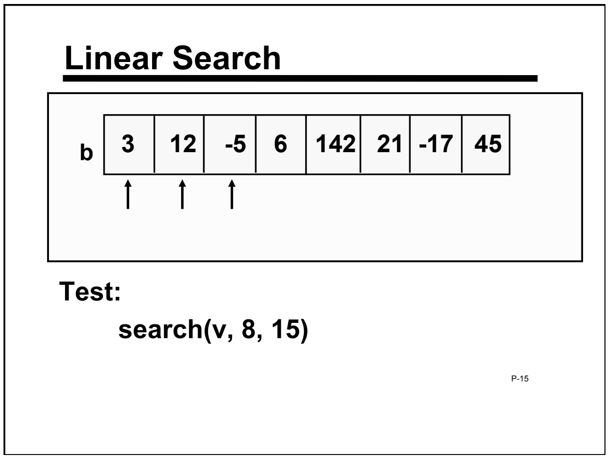
Linear Search

b	3	12	-5	6	142	21	-17	45
	↑	↑						

Test:

search(v, 8, 15)

P-14



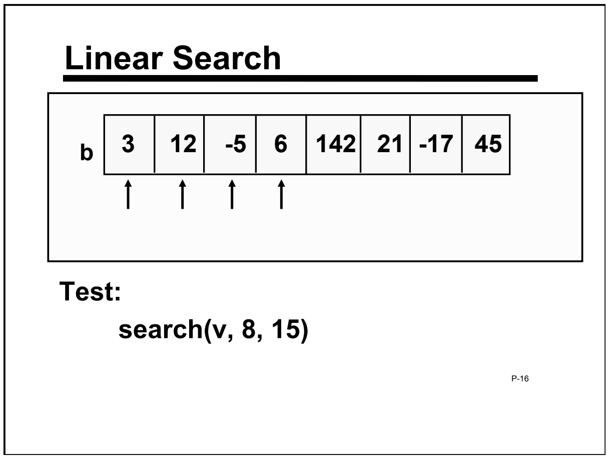
Linear Search

b	3	12	-5	6	142	21	-17	45
	↑	↑	↑					

Test:

search(v, 8, 15)

P-15



Linear Search

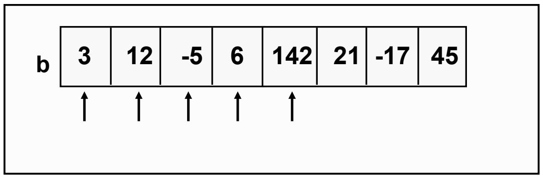
b	3	12	-5	6	142	21	-17	45
	↑	↑	↑	↑				

Test:

search(v, 8, 15)

P-16

Linear Search

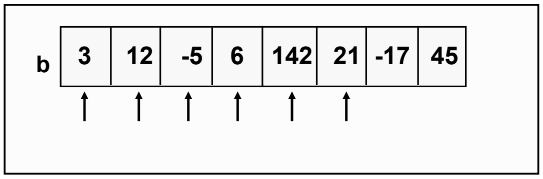


b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 15)

Linear Search

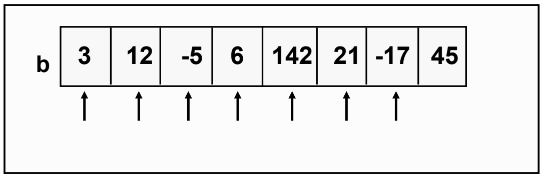


b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 15)

Linear Search

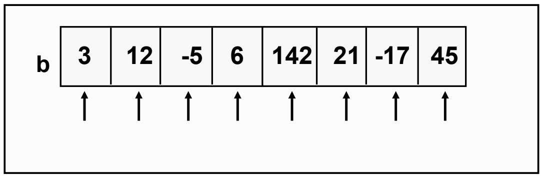


b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 15)

Linear Search

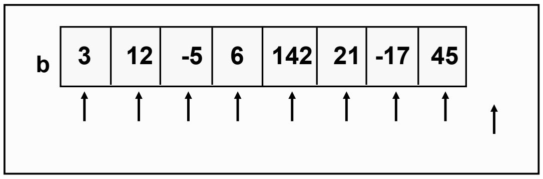


b	3	12	-5	6	142	21	-17	45

Test:

search(v, 8, 15)

Linear Search

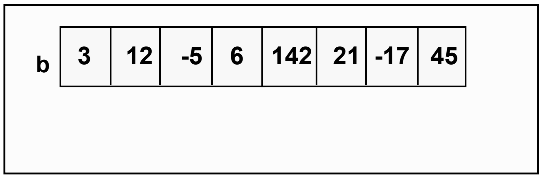


Test:

search(v, 8, 15)

Ran off the end! Not found.

Linear Search

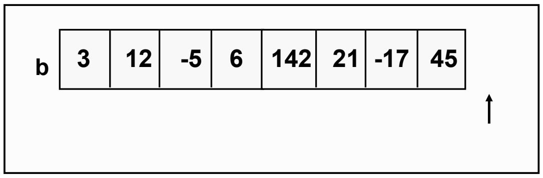


Test:

search(v, 8, 15)

Ran off the end! Not found.

Linear Search



Note: The loop condition is written so *b[index]* is not accessed if *index>=n*.

while (index < n && b[index] != x)

(Why is this true? Why does it matter?)

Can we do better?

Time needed for linear search is proportional to the size of the array.

An alternate algorithm, "Binary search," works if the array is sorted

1. Look for the target in the middle.
2. If you don't find it, you can ignore half of the array, and repeat the process with the other half.

Example: Find first page of pizza listings in the yellow pages

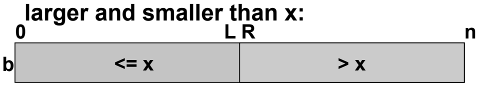
Binary Search Strategy

Binary Search Strategy

What we want: Find split between values larger and smaller than x:

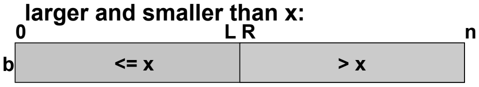
Binary Search Strategy

What we want: Find split between values larger and smaller than x:



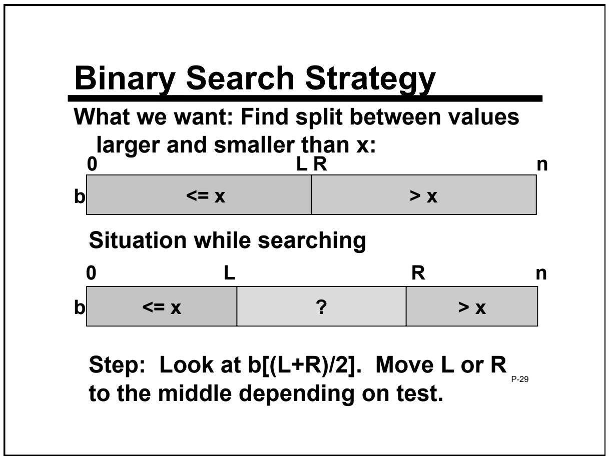
Binary Search Strategy

What we want: Find split between values larger and smaller than x:



Situation while searching

Step: Look at b[(L+R)/2]. Move L or R to the middle depending on test.



Binary Search Strategy

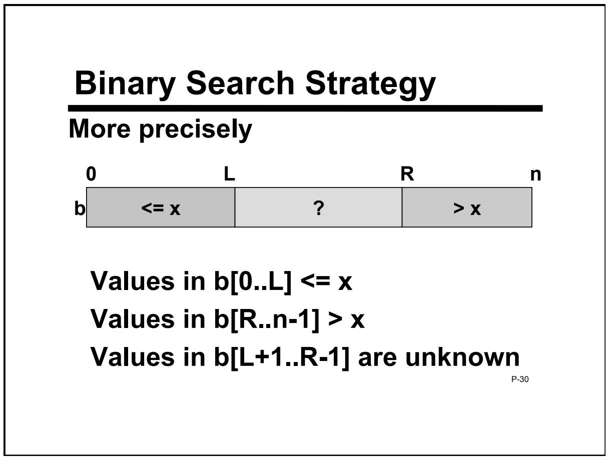
What we want: Find split between values larger and smaller than x:

	0	L R	n
b	<= x	> x	

Situation while searching

	0	L	R	n
b	<= x	?	> x	

Step: Look at b[(L+R)/2]. Move L or R to the middle depending on test.



Binary Search Strategy

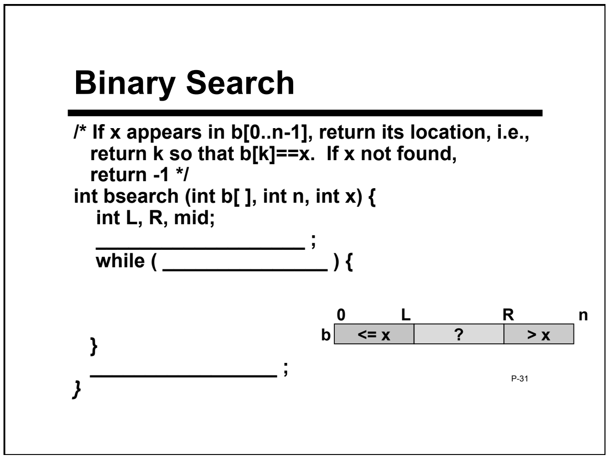
More precisely

	0	L	R	n
b	<= x	?	> x	

Values in b[0..L] <= x

Values in b[R..n-1] > x

Values in b[L+1..R-1] are unknown



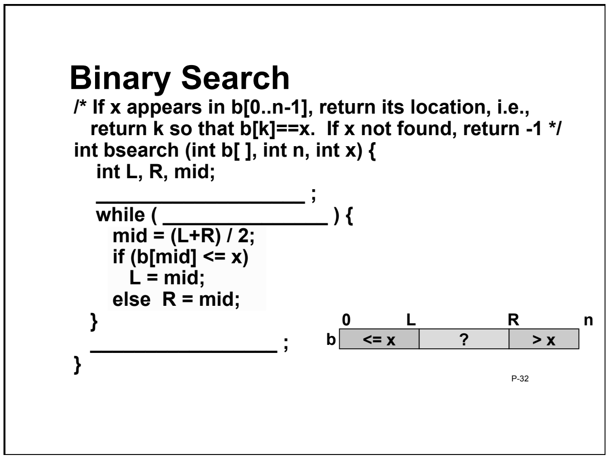
Binary Search

```
/* If x appears in b[0..n-1], return its location, i.e.,
   return k so that b[k]==x.  If x not found,
   return -1 */
int bsearch (int b[ ], int n, int x) {
    int L, R, mid;
    _________________ ;
    while ( _______________ ) {



    }
   _________________ ;
}
```

	0	L	R	n
b	<= x	?	> x	



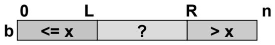
Binary Search

```
/* If x appears in b[0..n-1], return its location, i.e.,
   return k so that b[k]==x.  If x not found, return -1 */
int bsearch (int b[ ], int n, int x) {
    int L, R, mid;
    _______________ ;
    while ( _______________ ) {
      mid = (L+R) / 2;
      if (b[mid] <= x)
         L = mid;
      else  R = mid;
    }
   _________________ ;
}
```

	0	L	R	n
b	<= x	?	> x	

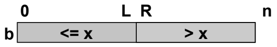
Loop Termination

```
/* If x appears in b[0..n-1], return its location, i.e.,
   return k so that b[k]==x.  If x not found,
   return -1 */
int bsearch (int b[ ], int n, int x) {
    int L, R, mid;
    ________________ ;
    while ( L+1 != R ) {
       mid = (L+R) / 2;
       if (b[mid] <= x)
          L = mid;
       else  R = mid;
   }
   ______________ ;
}
```



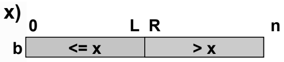
Initialization

```
/* If x appears in b[0..n-1], return its location, i.e.,
   return k so that b[k]==x.  If x not found, return
    -1 */
int bsearch (int b[ ], int n, int x) {
    int L, R, mid;
    L = -1; R = n;
    while ( L+1 != R ) {
       mid = (L+R) / 2;
       if (b[mid] <= x)  L = mid;
       else  R = mid;
   }
   ______________ ;
}
```



Return Result

```
/* If x appears in b[0..n-1], return its location, i.e.,
   return k so that b[k]==x.  If x not found, return -1 */
int bsearch (int b[ ], int n, int x) {
    int L, R, mid;
    L = -1; R = n;
    while ( L+1 != R ) {
       mid = (L+R) / 2;
       if (b[mid] <= x)  L = mid;
       else  R = mid;
   }
   if (L >= 0 && b[L] == x)
       return L
   else return -1;
}
```



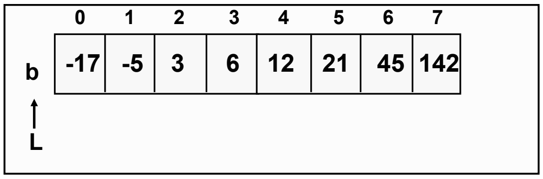
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

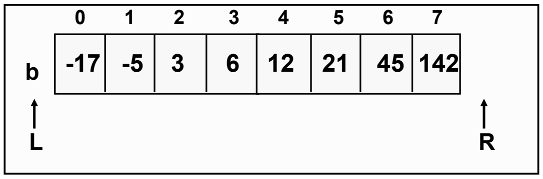


0	1	2	3	4	5	6	7
-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

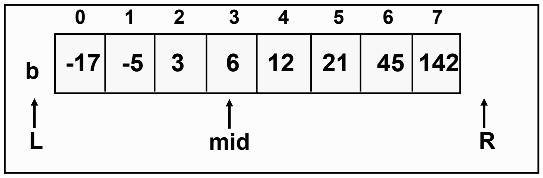


0	1	2	3	4	5	6	7
-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

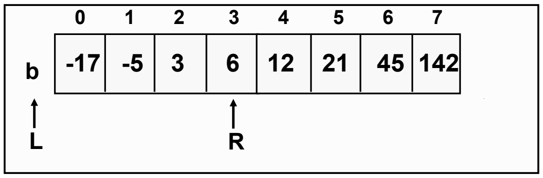


0	1	2	3	4	5	6	7
-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

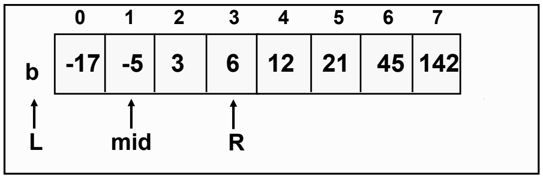


0	1	2	3	4	5	6	7
-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

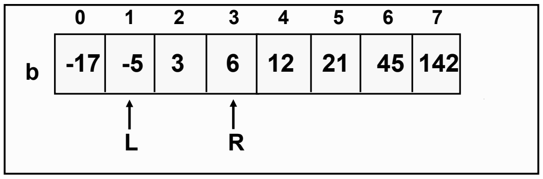


	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

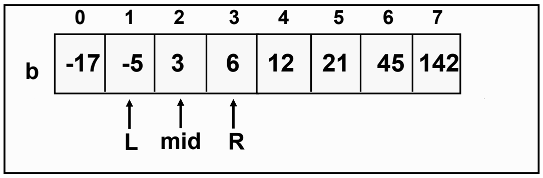


	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Binary Search

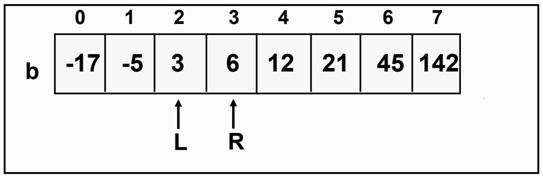


	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

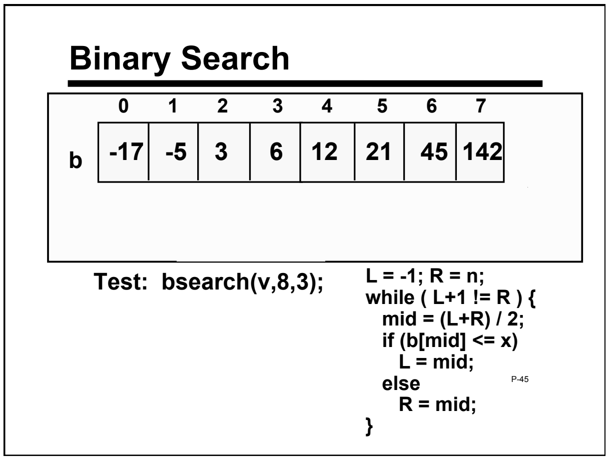
Binary Search



	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```



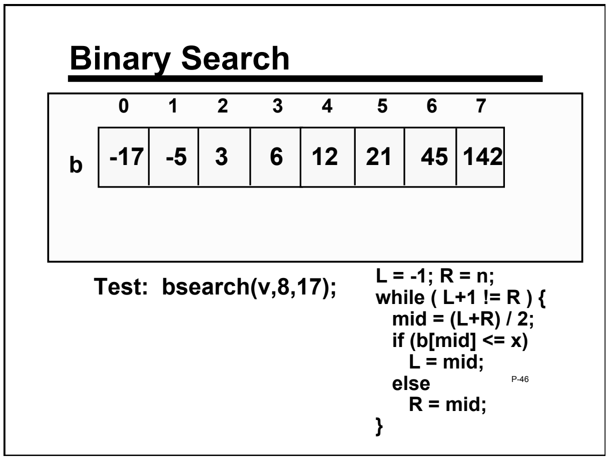
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,3);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-45



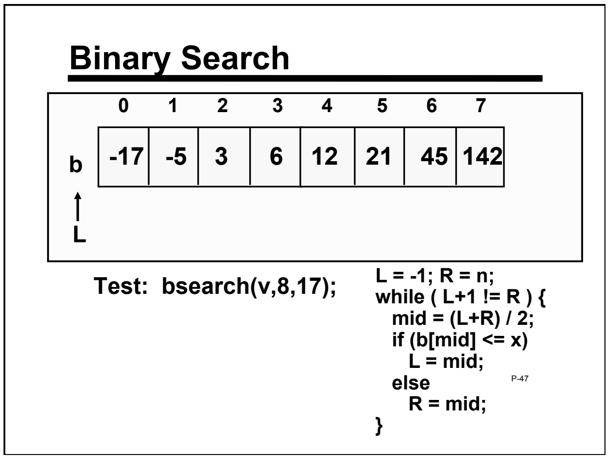
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-46



Binary Search

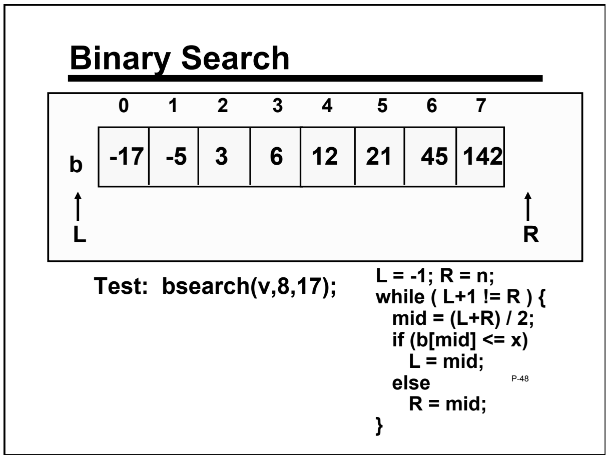
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

↑
L

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-47



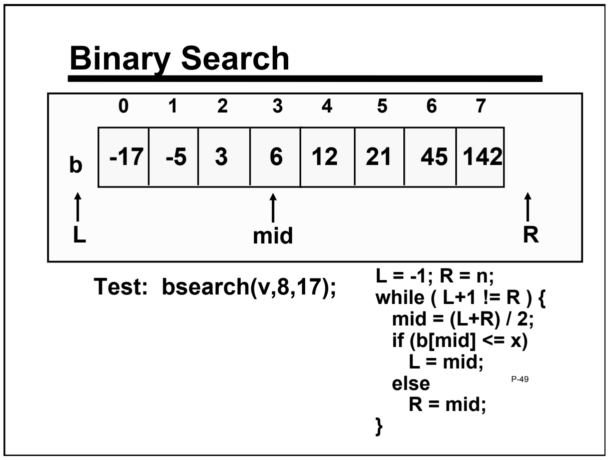
Binary Search

	0	1	2	3	4	5	6	7	
b	-17	-5	3	6	12	21	45	142	
↑ L									↑ R

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-48

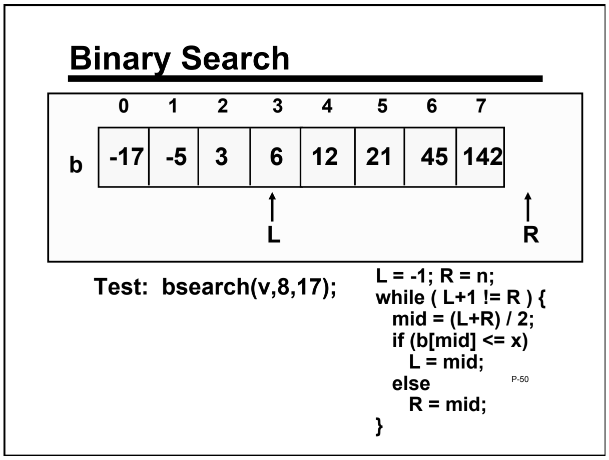


Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

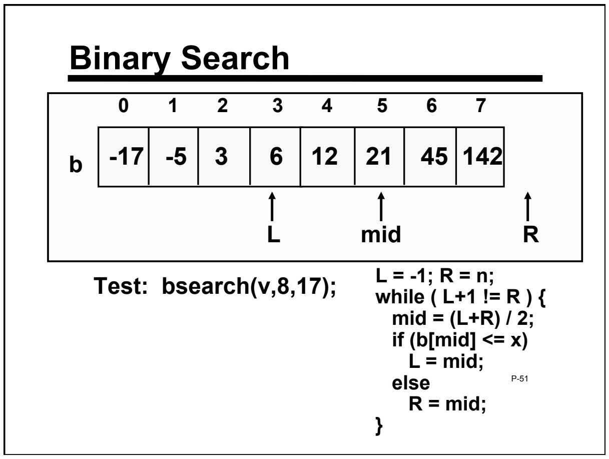


Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

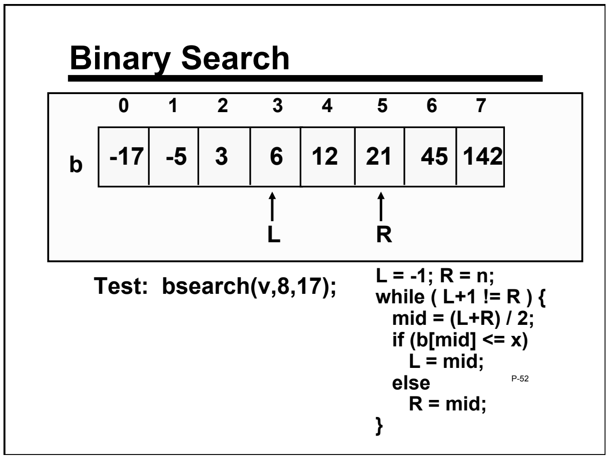


Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

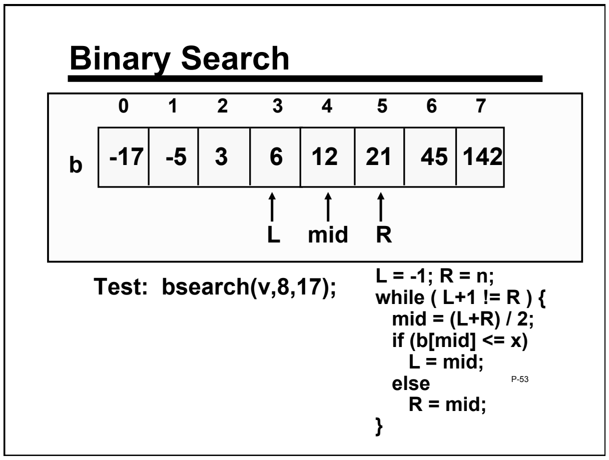


Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```



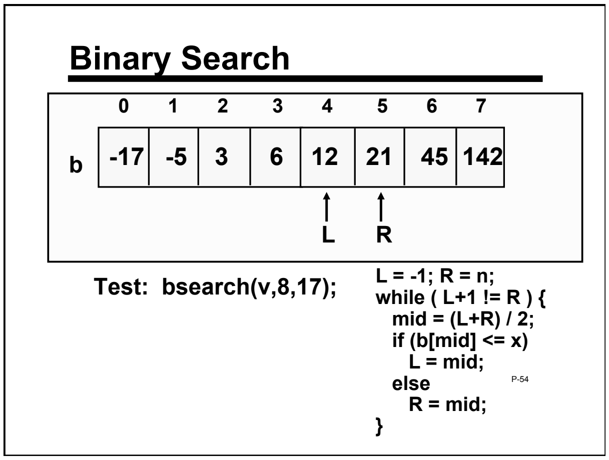
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142
				L	mid	R		

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-53



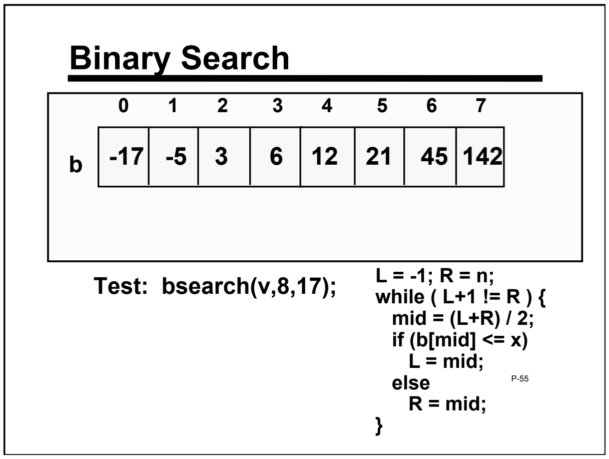
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142
					L	R		

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-54



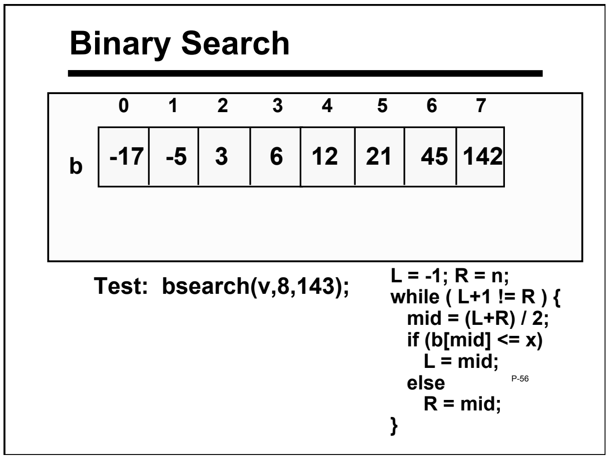
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,17);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-55



Binary Search

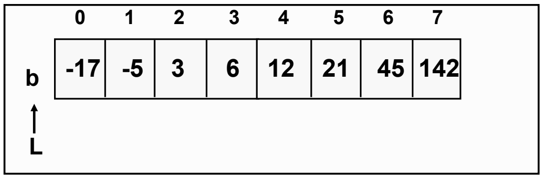
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-56

Binary Search



	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

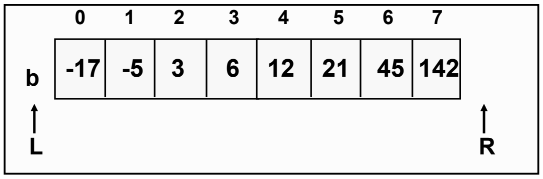
L

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-57

Binary Search



	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

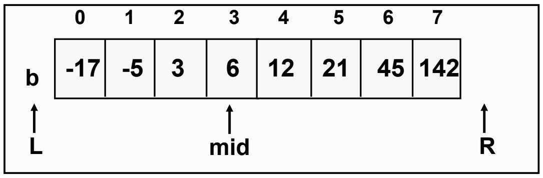
L R

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-58

Binary Search



	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

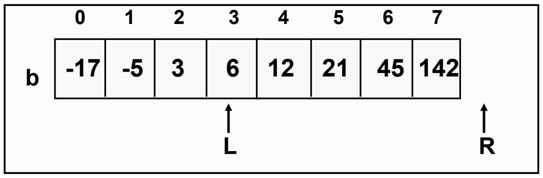
L mid R

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-59

Binary Search



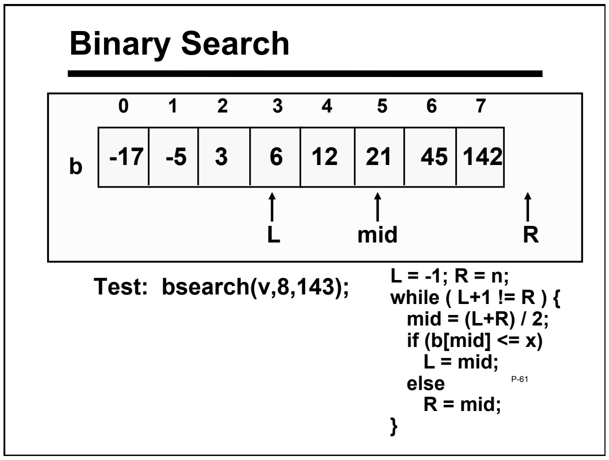
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

L R

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-60



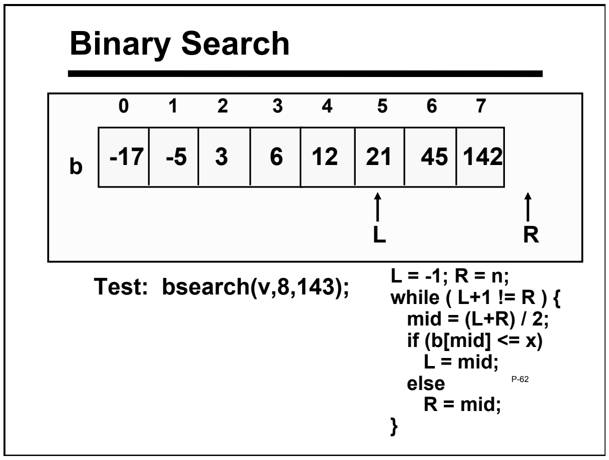
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-61



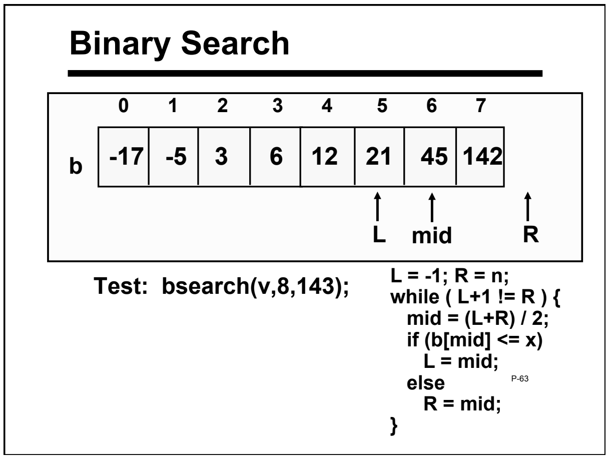
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-62



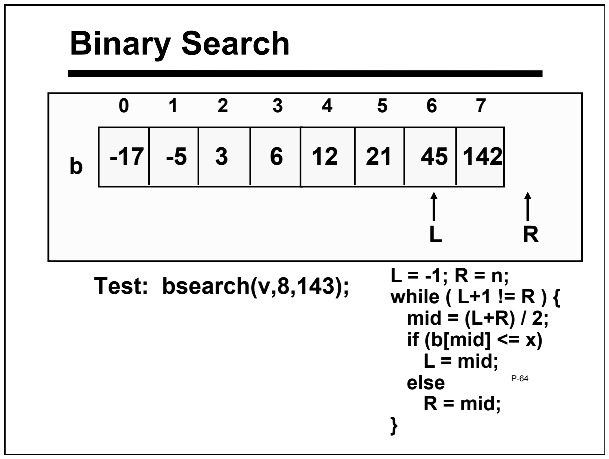
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-63



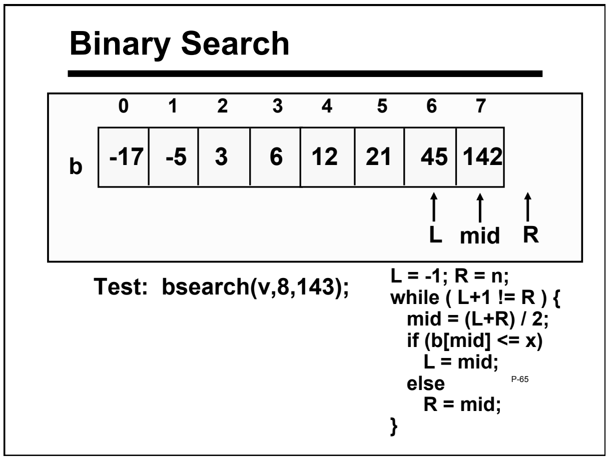
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-64



Binary Search

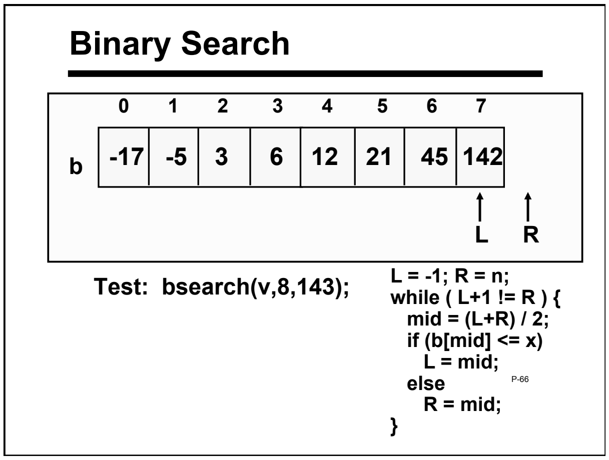
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

L mid R

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-65



Binary Search

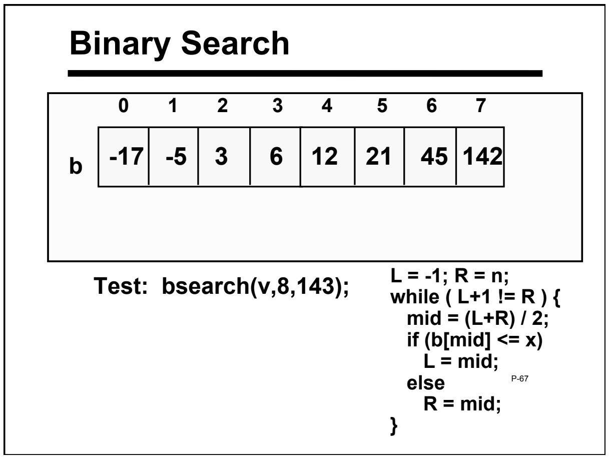
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

L R

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-66



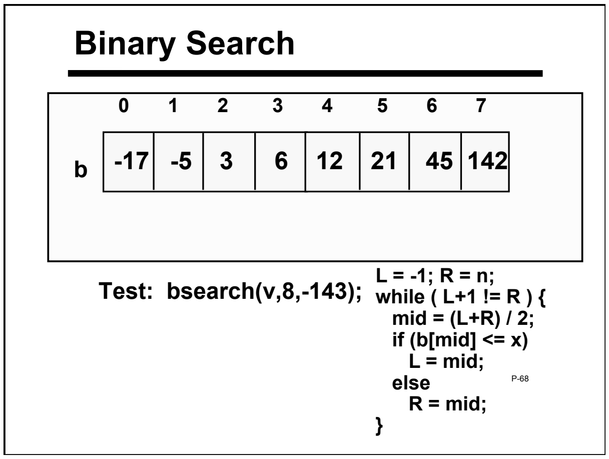
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-67



Binary Search

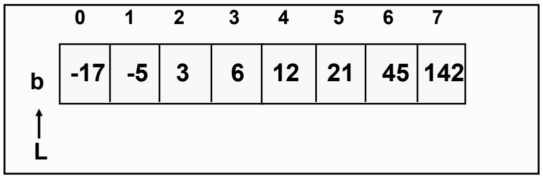
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-68

Binary Search



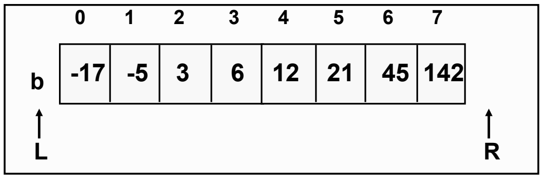
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-69

Binary Search



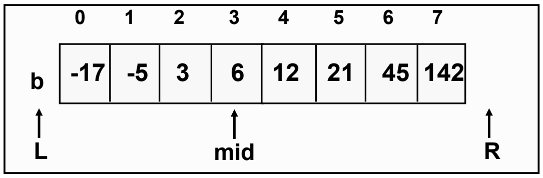
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-70

Binary Search



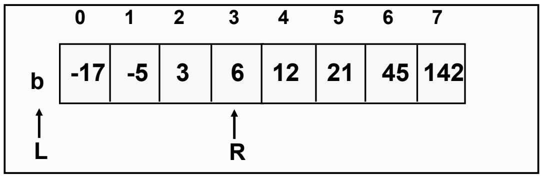
	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-71

Binary Search

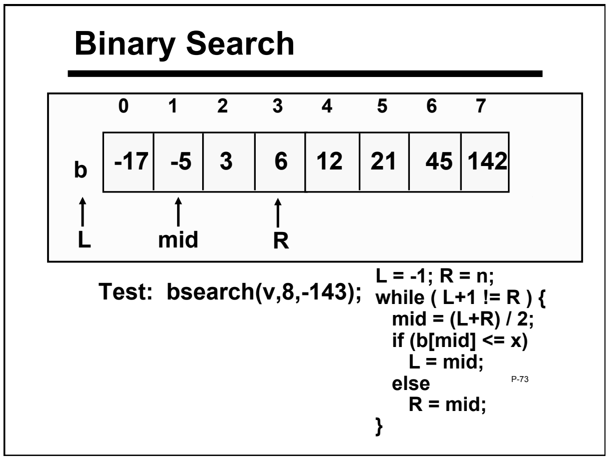


	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
  mid = (L+R) / 2;
  if (b[mid] <= x)
    L = mid;
  else
    R = mid;
}
```

P-72



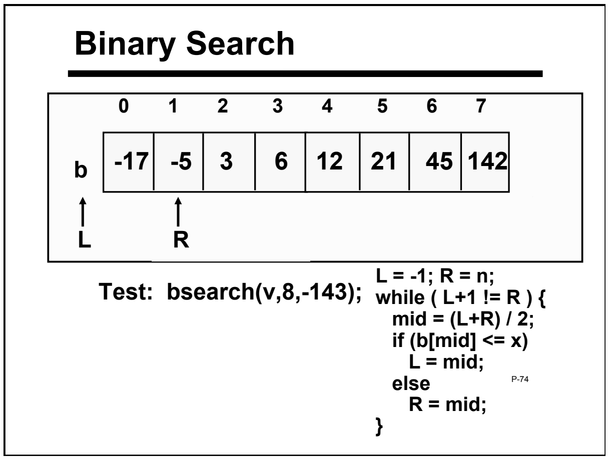
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-73



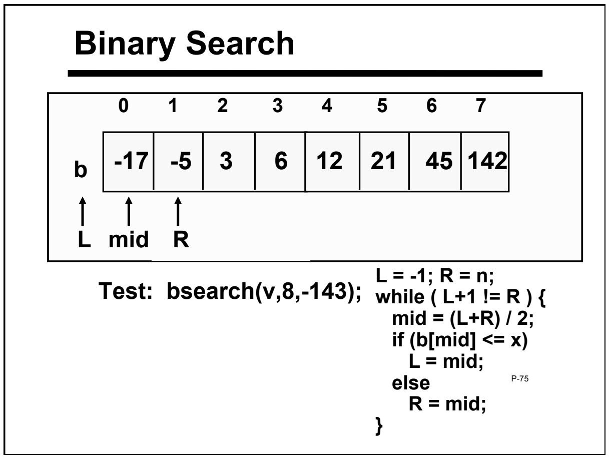
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-74



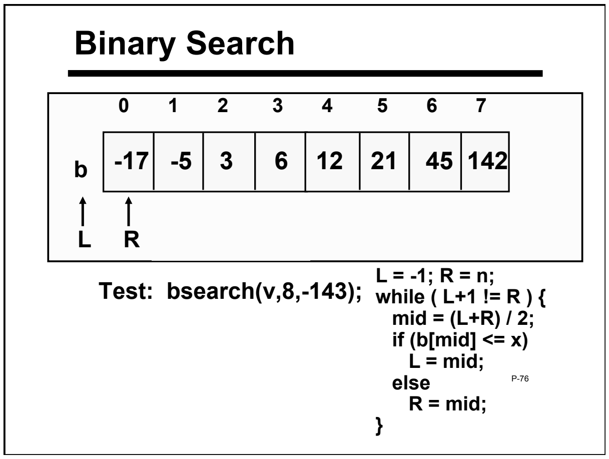
Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-75



Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

P-76

Binary Search

	0	1	2	3	4	5	6	7
b	-17	-5	3	6	12	21	45	142

Test: bsearch(v,8,-143);

```
L = -1; R = n;
while ( L+1 != R ) {
   mid = (L+R) / 2;
   if (b[mid] <= x)
      L = mid;
   else
      R = mid;
}
```

Is it worth the trouble?

Is it worth the trouble?

Suppose you had 1000 elements

Ordinary search would require maybe 500 comparisons on average

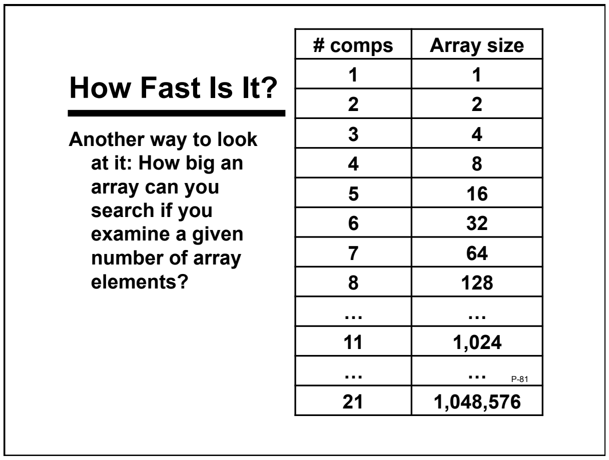
Binary search

- after 1st compare, throw away half, leaving 500 elements to be searched.
- after 2nd compare, throw away half, leaving 250. Then 125, 63, 32, 16, 8, 4, 2, 1 are left.
- After at most 10 steps, you're done!

What if you had 1,000,000 elements??

How Fast Is It?

Another way to look at it: How big an array can you search if you examine a given number of array elements?

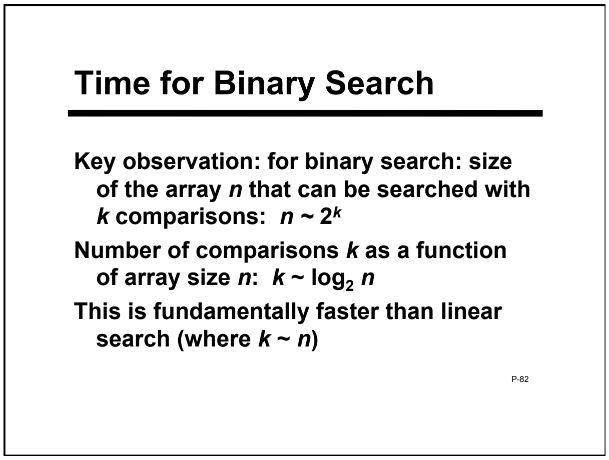


How Fast Is It?

Another way to look at it: How big an array can you search if you examine a given number of array elements?

# comps	Array size
1	1
2	2
3	4
4	8
5	16
6	32
7	64
8	128
…	…
11	1,024
…	…
21	1,048,576

P-81



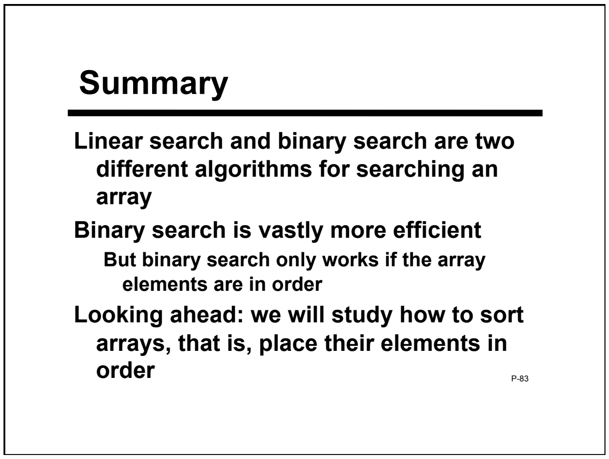
Time for Binary Search

Key observation: for binary search: size of the array *n* that can be searched with *k* comparisons: $n \sim 2^k$

Number of comparisons *k* as a function of array size *n*: $k \sim \log_2 n$

This is fundamentally faster than linear search (where $k \sim n$)

P-82



Summary

Linear search and binary search are two different algorithms for searching an array

Binary search is vastly more efficient

- But binary search only works if the array elements are in order

Looking ahead: we will study how to sort arrays, that is, place their elements in order

P-83