CS 11: Introduction to Computer Science

Pointers, The Heap and Dynamic List Arrays

Mark Sheldon Tufts University

Email: msheldon@cs.tufts.edu

Web: http://www.cs.tufts.edu/~msheldon

Noah Mendelsohn Tufts University

Email: noah@cs.tufts.edu

Web: http://www.cs.tufts.edu/~noah

Goals for this session

- Review pointers & arrays
- Review using "new" and allocating space on the heap
- Applying what we've learned: dynamic array lists
- Preparing to explore object-oriented programming

Unscrambling Pointers and Arrays

What you should get from this interlude

- Often C++ will take the address of an array for you when a pointer is needed
- In C++ you will see subscripting used on pointers (to arrays) as well as on arrays. In all cases it means the same thing.
- You will see arr[] (especially in parameter declarations) and *arr_ptr
 syntax used somewhat interchangeably to declare pointers to arrays.

Pointers, Arrays, and how they relate

int arr[5];

int	int	int	int	int
arr[0]	arr[1]	arr[2]	arr[3]	arr[4]

Pointers, Arrays, and how they relate

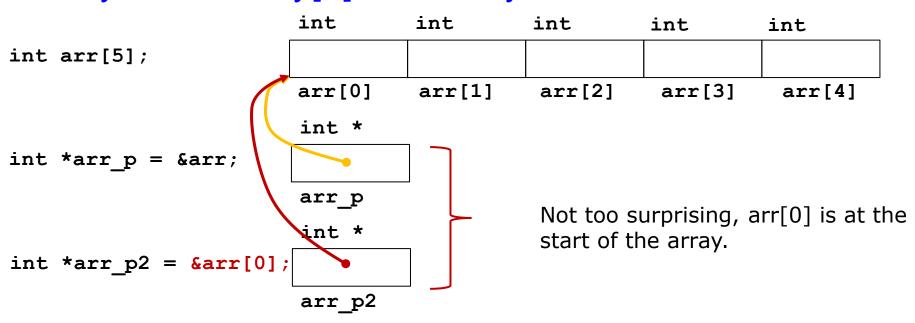
```
int int int int int
int arr[5];

arr[0] arr[1] arr[2] arr[3] arr[4]
int *arr_p = &arr;
arr_p
```

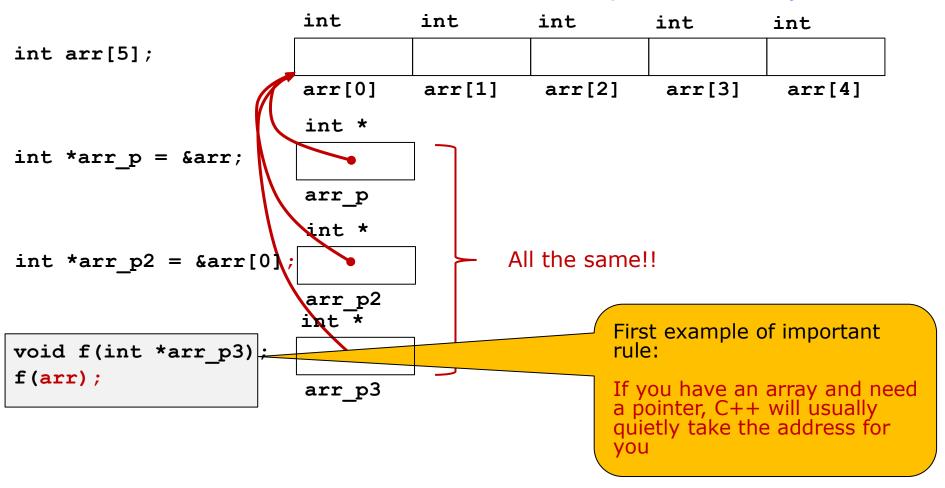
That's pretty clear...

...but C++ for historical reasons plays some tricks. You'll need to at least not be surprised when you see the following....

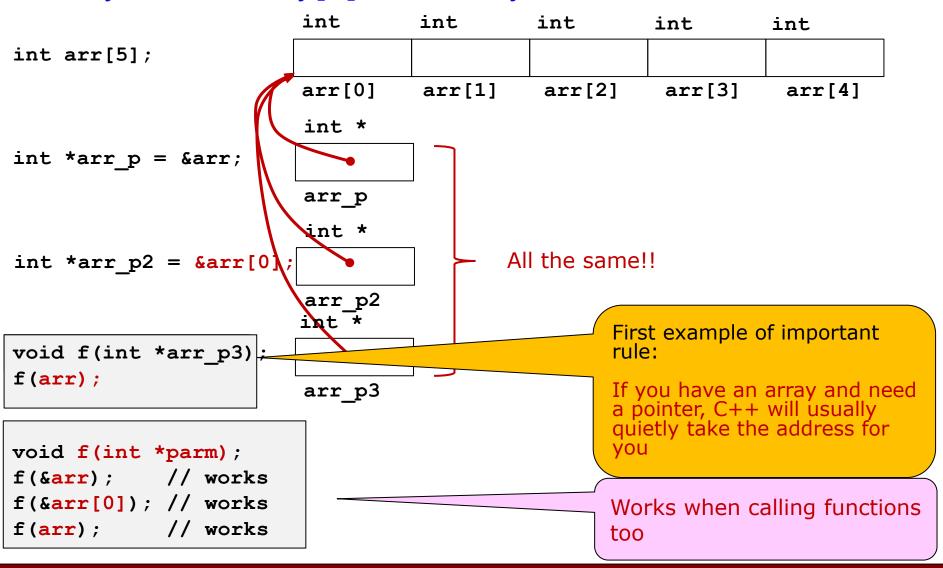
&array and &array[0] are always the same



C++ will often take the address of a pointer for you



&array and &array[0] are always the same



That's a lot of detail...what you mostly need to know is...

<u>In many but not all cases, pointer and array types can be used somewhat interchangeably.</u> Most importantly:

```
void f(int *parm);
void f(int parm[]);
```

Truly equivalent function signatures...so get used to seeing both!

And now we need add one more simple part of the story:

You can subscript pointers...assumed pointing to array

In many but not all cases, pointer and array types can be used somewhat interchangeably. Most importantly:

```
void f(int parm[]) {
   cout << parm[1]; // works, no surprise it's an array
};</pre>
```

```
void f(int *parm) {
   cout << parm[1]; // works exactly the same!!
};</pre>
```

In C++

Subscripting an array and subscripting a pointer to that array are by definition doing exactly the same thing!

Use whichever is more convenient!

You can subscript pointers...assumed pointing to array

In many but not all cases, pointer and array types can be used somewhat interchangeably. Most importantly:

```
void f(int parm[]) {
   cout << parm[1]; // works, no surprise it's an array
};</pre>
```

What you should have gotten from this interlude

- Often C++ will take the address of an array for you when a pointer is needed
- In C++ you will see subscripting used on pointers (to arrays) as well as on arrays. In all cases it means the same thing
- You will see arr[] (especially in parameter declarations) and *arr_ptr
 syntax used somewhat interchangeably to declare pointers to arrays
- arr[5] with an explicit size is different...that's an array of known size, but...
- ...as always, all arrays are passed into functions by reference

Using **new** to get Memory on the Heap An Example Application

I'd like to write a main program like this:

```
int *new array with squares(int size);
void print array(int *array, int length);
int main ()
        int *squares array ptr;
        int len;
        cout << "How many squares do you want to store? ";</pre>
        cin >> len;
        squares array ptr = new array with squares(len);
        print array(squares array ptr, len);
        delete [] squares array ptr;
```

How many squares do you want to store? 5

int	int	int	int	int
0	1	4	9	16
arr[0]	arr[1]	arr[2]	arr[3]	arr[4]

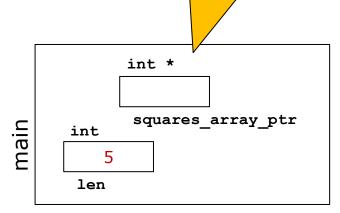
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

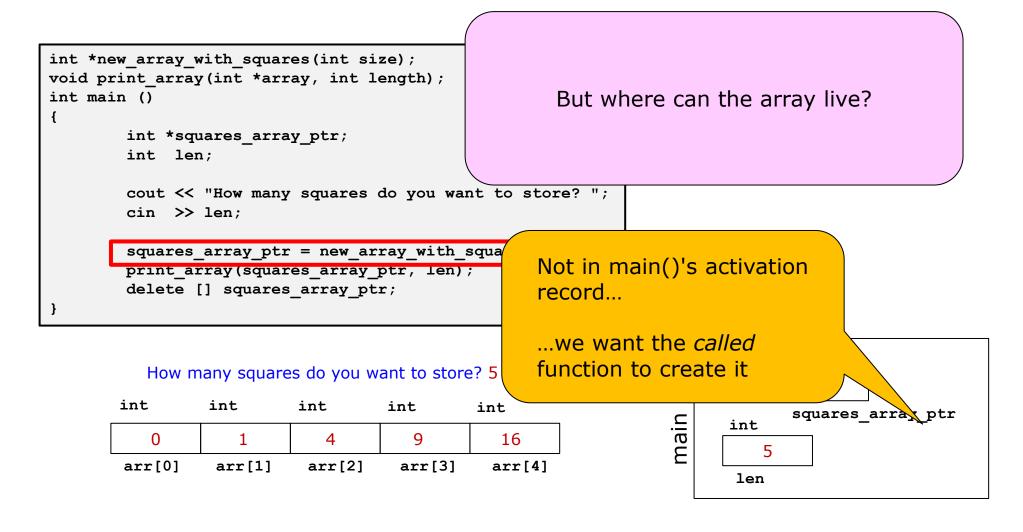
    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

How many squares do you want to store? 5

main()'s local variables live in its activation record, of course.





But where can the array live?

```
int *new array with squares(int size);
void print array(int *array, int length);
int main ()
        int *squares array ptr;
        int len;
                                                                    y_w/squares
        cout << "How many squares do you want to store? ";</pre>
        cin >> len;
        squares array ptr = new array with squares(len);
        print array(squares array ptr, len);
It can't be in the activation record
for new_array_with_squares...
                                                                                 int *
                                           ore? 5
                                              int
                                                                                 squares_array_ptr
                                                                    main
                                                                          int
                                                 16
                                                                              5
        arr[0]
                  arr[1]
                            arr[2]
                                      arr[3]
                                                arr[4]
                                                                           len
```

But where can the array live?

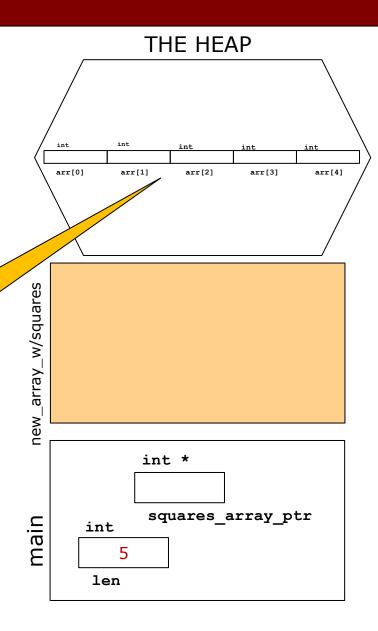
```
int *new array with squares(int size);
void print array(int *array, int length);
int main ()
        int *squares array ptr;
        int len;
                                                                   y_w/squares
        cout << "How many squares do you want to store? ";</pre>
        cin >> len;
        squares array ptr = new array with squares(len);
        print array(squares array ptr, len);
It can't be in the activation record
for new_array_with_squares...
                                                                               int *
                                           ore? 5
...that will go away when the
                                              int
                                                                                squares_array_ptr
                                                                   main
function returns!
                                                                         int
                                                16
                                                                             5
        arr[0]
                  arr[1]
                           arr[2]
                                     arr[3]
                                               arr[4]
                                                                          len
```

```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? cin >> len;

    squares_array_ptr = new array with squares delete [] squares_array_ptr;
}
```

new_array_with_squares will allocate it on the heap using new...

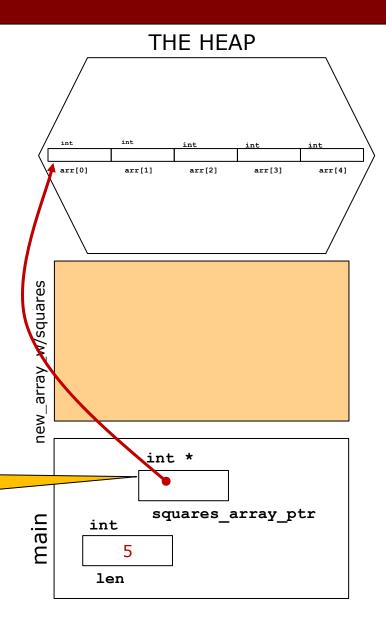


```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

...and will return a pointer that we can keep here.

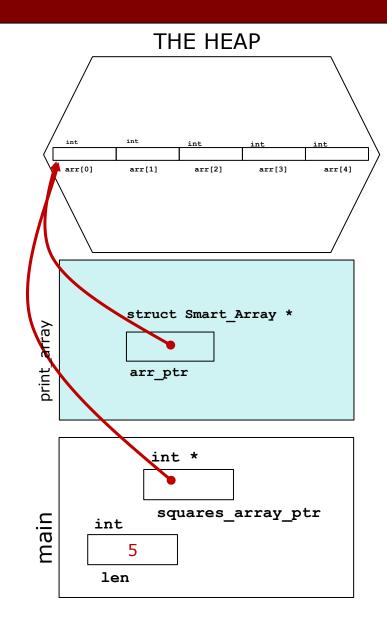


```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares array ptr = new array with squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_tr;
}
```

...so we can pass the pointer into the printing function which can then find the array!



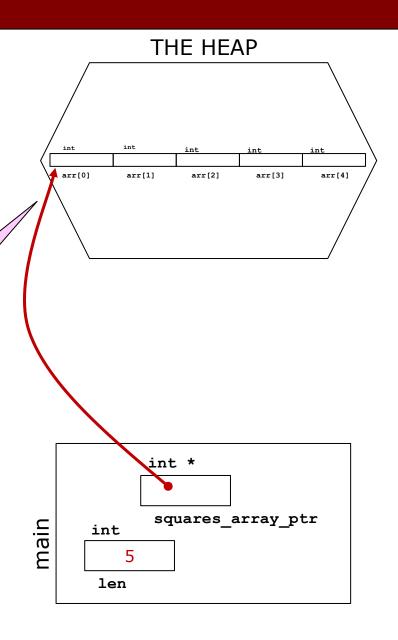
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to storin >> len;

    squares array ptr = new array with sr len);
    print_array(squares_array_ptr, lendelete [] squares_array_ptr;
}
```

Crucially...

Both the array and at least one pointer to it live on after the function that created the array returns!



```
int *new_array with squares(int size);
                                                         arr[0]
                                                               arr[1]
                                                                     arr[2]
                                                                          arr[3]
void print array(int *array, int length);
int main ()
      int *squares array ptr;
          But.. how can new array with squares
          create this array and return the pointer?
                                                                        array ptr
```

THE HEAP

int i;

```
THE HEAP
            arr[1]
                    arr[2]
                             arr[3]
               squares array ptr
main
       int
            5
```

len

int *new array with squares(int size)

// Fill with squares

int *array = new int [size];

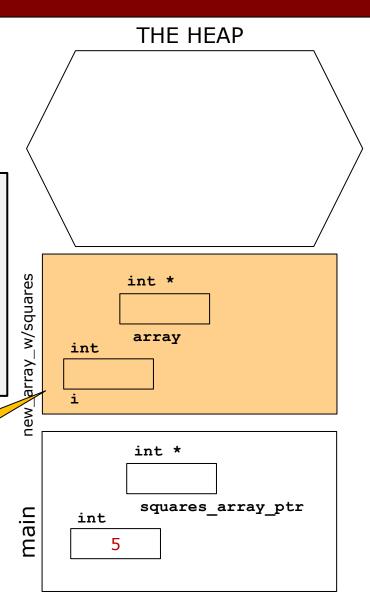
for (i = 0; i < size; i++) {

{

```
int *new_array_with_squares(int size)
{
    int i;
    int *array = new int [size];

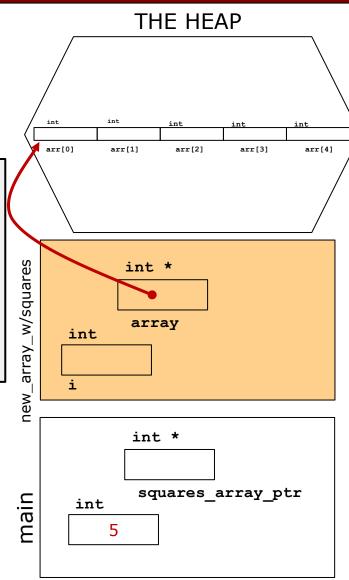
    // Fill with squares
    for (i = 0; i < size; i++) {
        array[i] = i * i; // array[i] gets i squared
    }
    return array;
}</pre>
```

new_array_with_squares' local variables live in its activation record while it is running!



```
int *new_array_with_squares(int size)
{
    int i;
    int *array = new int [size];

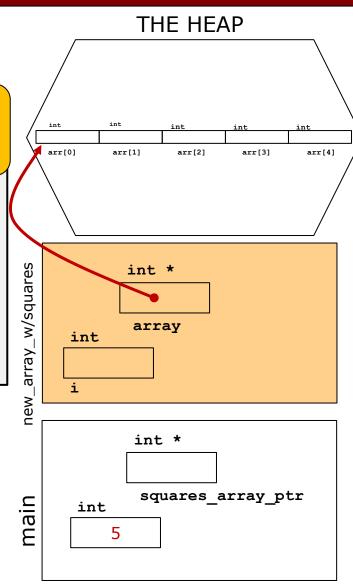
    // Fill with squares
    for (i = 0; i < size; i++) {
        array[i] = i * i; // array[i] gets i squared
    }
    return array;
}</pre>
```



The C++ new operation allocates the array *on the heap!*

```
int *new_array_w
{
    int i;
    int *array = new int [size];

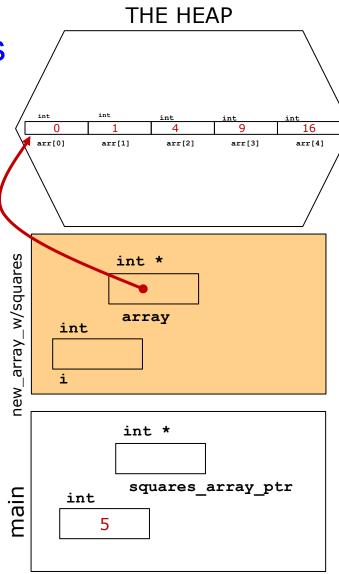
    // Fill with squares
    for (i = 0; i < size; i++) {
        array[i] = i * i; // array[i] gets i squared
    }
    return array;
}</pre>
```



Filling the array with squared numbers

```
int *new_array_with_squares(int size)
{
    int i;
    int *array = new int [size];

    // Fill with squares
    for (i = 0; i < size; i++) {
        array[i] = i * i; // array[i] gets i squared
    }
    return array;
}</pre>
```



How does main find the array?

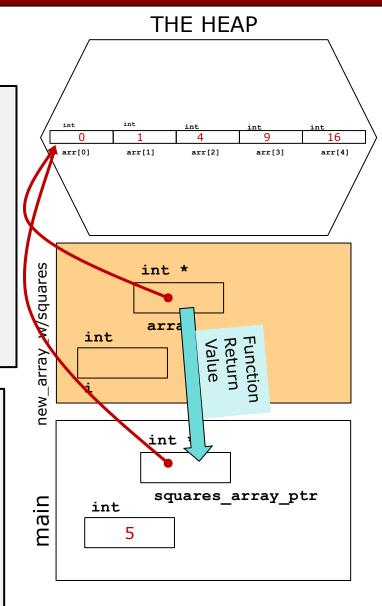
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

```
int *new_array_with_squares(int size)
{
    int i;
    int *array = new int [size];

    // Fill with squares
    for (i = 0; i < size; i++) {
        array[i] = i * i; // array[i] gets i squared
}
return array;
}</pre>
```



How does main find the array?

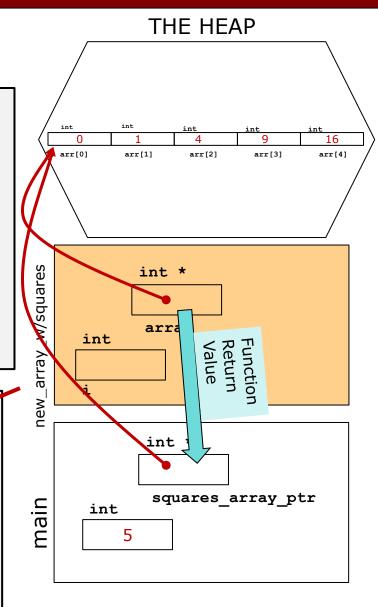
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

```
int *new_array_with_squares(int size)
{
    int i;
    int *array = new int [size];

    // Fill with squares
    for (i = 0; i < size; i+); {
        array[i] = i * i; // array[i] gets i squared
    }
    return array;
}</pre>
```



THE HEAP

Now we can pass the array to a function as always

```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares array ptr = new array with squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

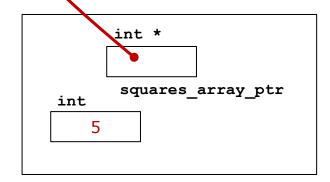
int int int int int

0 1 4 9 16

arr[0] arr[1] arr[2] arr[3] arr[4]

How many squares do you want to store? 5

```
array[0] = 0
array[1] = 1
array[2] = 4
array[3] = 9
array[4] = 16
```



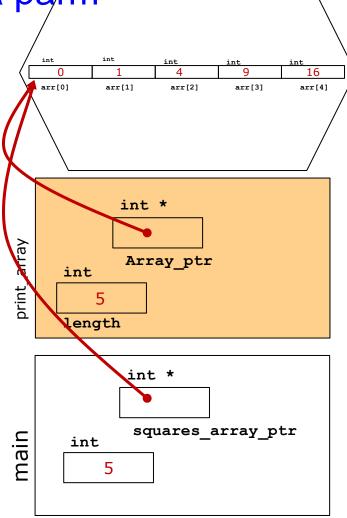
THE HEAP

Print_array gets the array pointer as a parm

```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares array ptr = new array with squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```



THE HEAP

Print_array gets the array pointer as a parm

```
int *new array with squares(int size);
void print array(int *array, int length);
int main ()
                                                                                 arr[1]
                                                                                         arr[2]
                                                                                                arr[3]
        int *squares array ptr;
        int len;
                                             Note that although arr_ptr is typed as a pointer,
                                             print array subscripts it.
        cout << "How many squares do you</pre>
        cin >> len;
        squares array ptr = new array with square
        print_array(squares_array ptr, len);
                                                                                   Array ptr
        delete [] squares array ptr;
                                                                             int
                                                                             length
void print array(int *arr ptr, int length
                                                                                    int *
        int i;
        for (i = 0; i < length; i++//
                                                                                    squares array ptr
                 cout << <u>"array[" << i << "] = "</u>
                                                                             int
                      << arr ptr[i] << endl;</pre>
```

...and use the array until it's deleted!

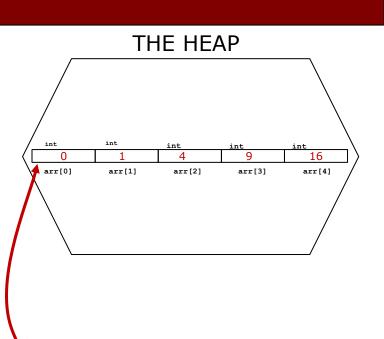
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

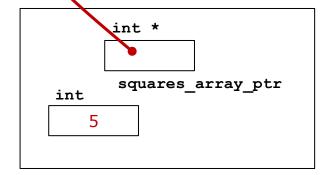
    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print array(squares array ptr, len);
    delete [] squares_array_ptr;
}
```

How many squares do you want to store? 5

```
array[0] = 0
array[1] = 1
array[2] = 4
array[3] = 9
array[4] = 16
```





...and use the array until it's deleted!

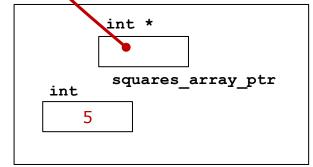
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print array(squares array ptr, len);
    delete [] squares_array_ptr;
}
```

THE HEAP

If we weren't about to leave anyway, we should set squares_array_ptr=nullptr



What if we don't know how big the array should be?

We'll build a dynamic list array!

The idea...

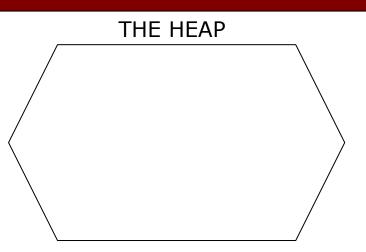
- We will build a service that can grow our arrays as necessary
- We can change our minds and ask for more each time, but...
- C++ arrays can't grow!
- The function will reallocate a bigger array if necessary. If so it will:
 - Copy all the data from the old to the new
 - Always return the pointer to the latest array...the caller uses that!

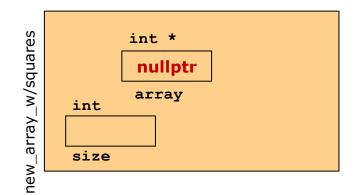
Calling a function that can grow arrays

```
/*
 * Get new array on heap and fill each
 * arr[] with i squared
 */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

    return array;
}</pre>
```



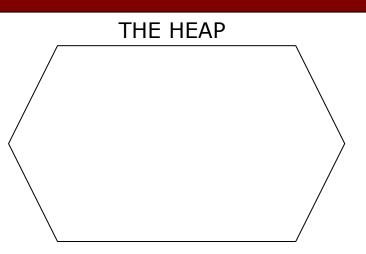


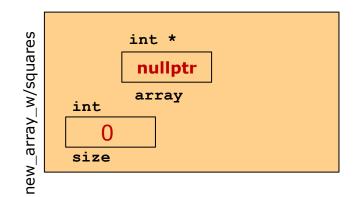
Calling a function that can grow arrays

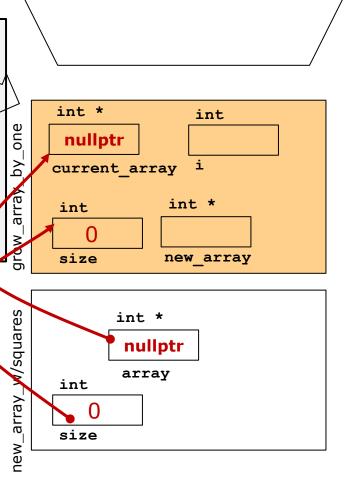
```
/*
 * Get new array on heap and fill each
 * arr[] with i squared
 */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

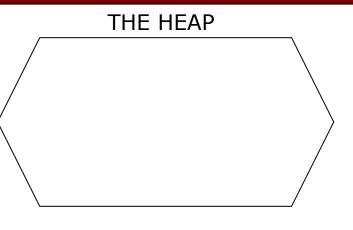
    return array;
}</pre>
```

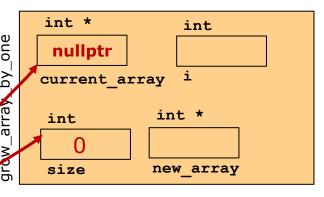


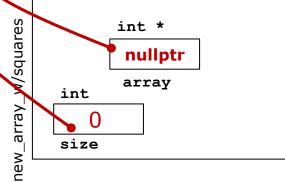


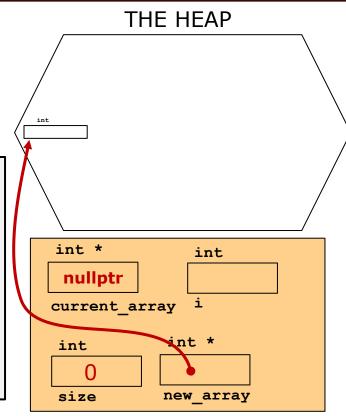


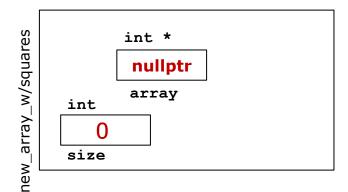
THE HEAP





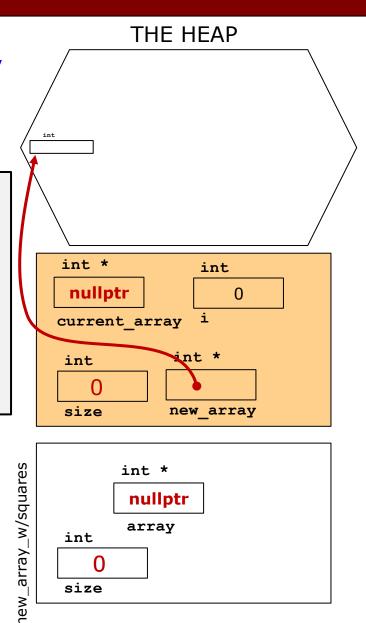






Does nothing this time

(size is 0)



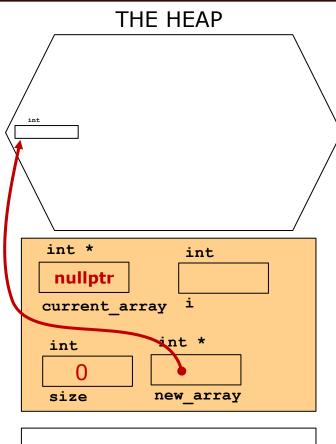
```
int *grow_array_by_one(int *current_array, int size)
{
   int i;
   int *new_array = new int [size + 1];

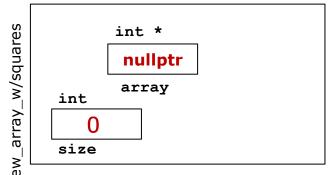
   for (i = 0; i < size; i++) {
      new_array[i] = current_array[i];
   }

   delete [] current_array;  // OK to use delete on nullptr
      // the first time through
   return new_array;
}</pre>
```

Does nothing this time

```
(current_array is nullptr)
```

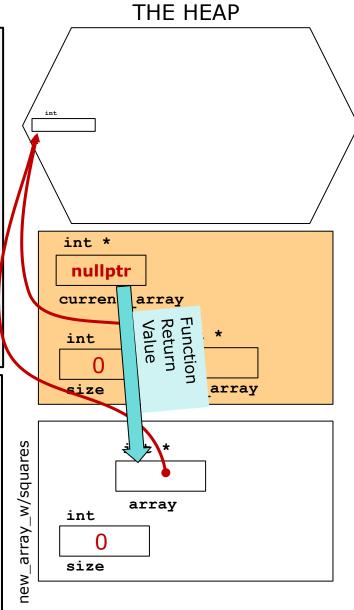




```
/*
  * Get new array on heap and fill each
  * arr[] with i squared
  */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

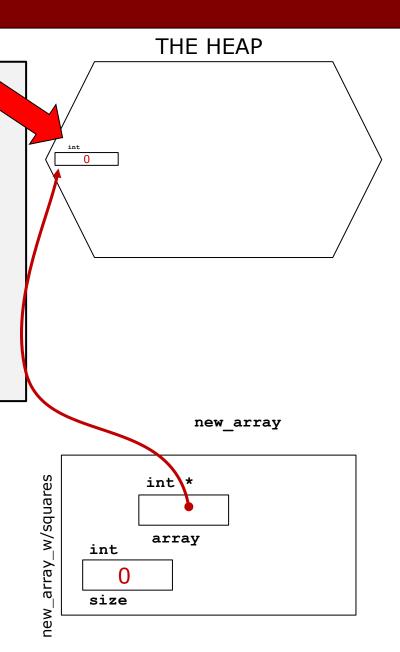
    return array;
}</pre>
```



```
/*
 * Get new array on heap and fill each
 * arr[] with i squared
 */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow array by one(array, size);
        array[size] = size * size; }

    return array;
}</pre>
```



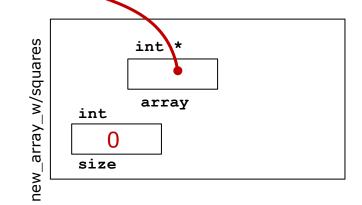
```
/*
 * Get new array on heap and fill each
 * arr[] with i squared
 */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

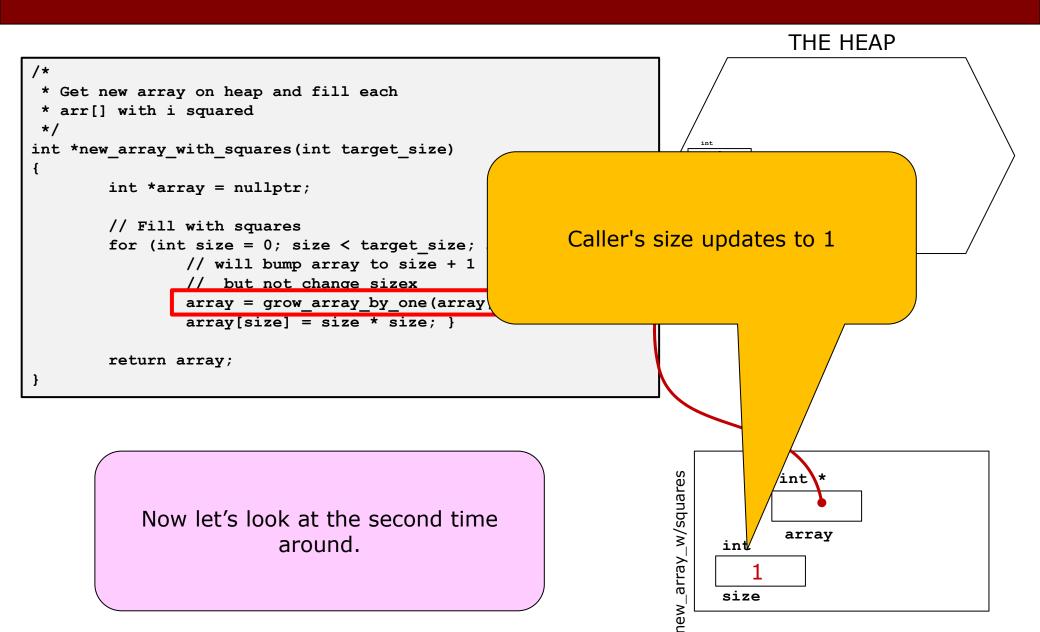
    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

    return array;
}</pre>
```

THE HEAP

Now let's look at the second time around.





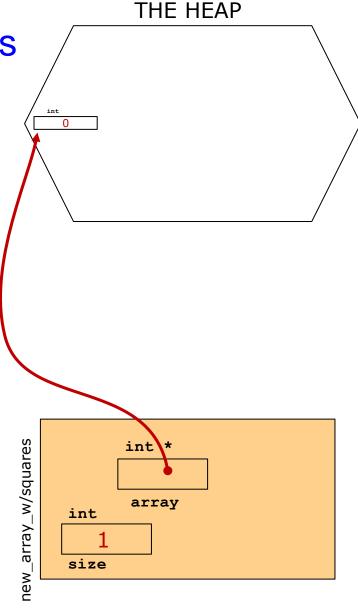
size

Calling a function that can grow arrays

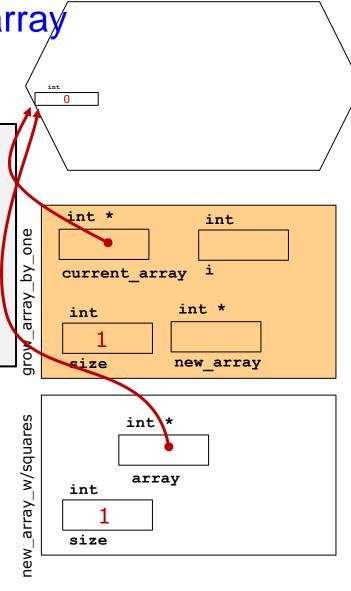
```
/*
 * Get new array on heap and fill each
 * arr[] with i squared
 */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

    return array;
}</pre>
```



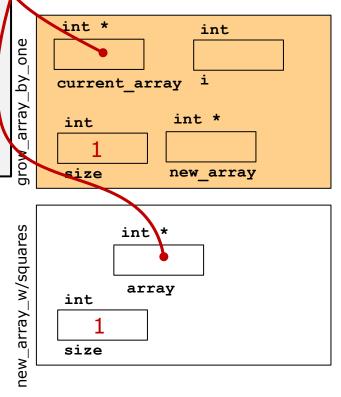
Allocating the second version of the array



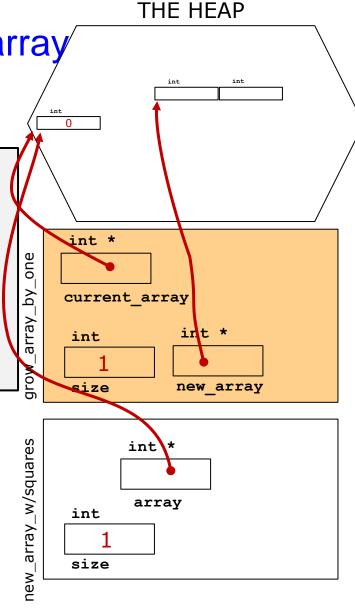
THE HEAP

THE HEAP

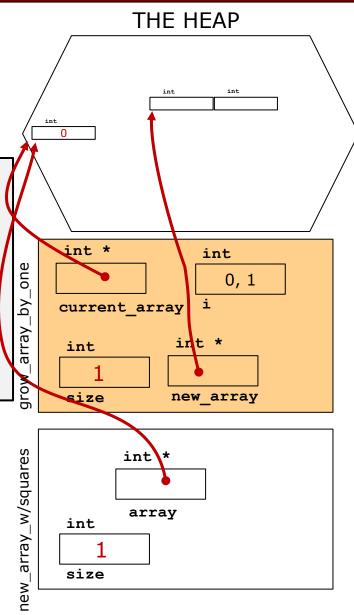
Allocating the *second* version of the array



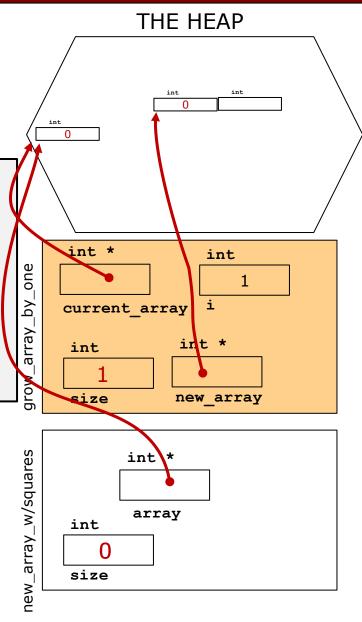
Allocating the second version of the array



Copying the old to the new



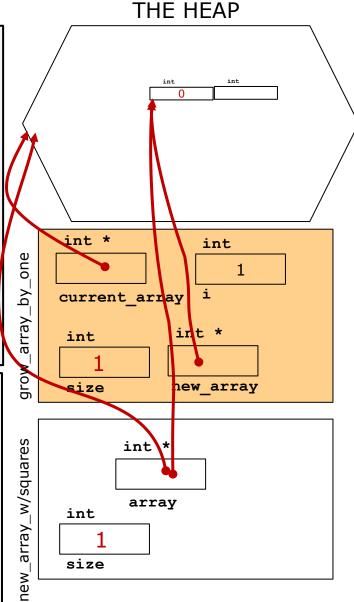
Deleting the one that's been replaced



```
/*
  * Get new array on heap and fill each
  * arr[] with i squared
  */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

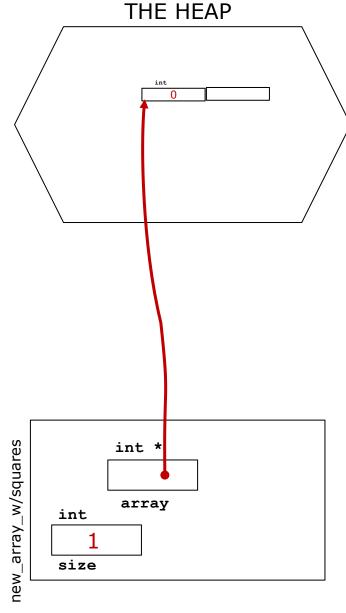
    return array;
}</pre>
```



```
/*
 * Get new array on heap and fill each
 * arr[] with i squared
 */
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1
        // but not change sizex
        array = grow_array_by_one(array, size);
        array[size] = size * size; }

    return array;
}</pre>
```



How does main get the array?

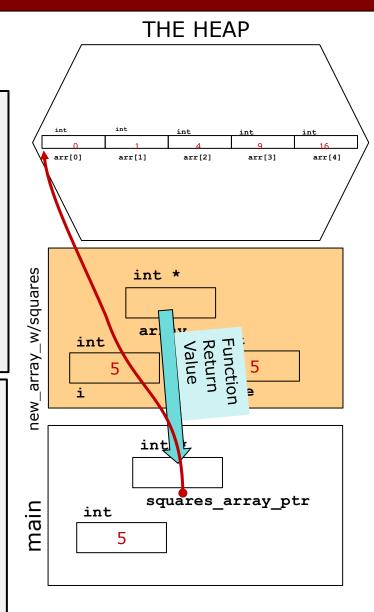
```
int *new_array_with_squares(int size);
void print_array(int *array, int length);
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

```
int *new_array_with_squares(int size)
{
    int i;
    int *array = new int [size];

    // An "identity array" sets each slot
    // value to the index of the slot
    for (i = 0; i < size; i++) {
        array[i] = i * i; // array[i] gets i squared
    }
    return array;
}</pre>
```



Modularity!!

Can We Hide The Details of the Dynamic Array from its users?

We can make this code much more modular

- We'll put all the variables for each dynamic array into a structure
- Our squares application will use functions to perform all manipulations of the dynamic array.
- These functions make the application cleaner and more modular
- They abstract away (hide!) array details from the squares application
- We will sketch the details today...
- ...on Tuesday this will lead us to a more robust approach using classes!

The old squares program

```
int main ()
{
    int *squares_array_ptr;
    int len;

    cout << "How many squares do you want to store? ";
    cin >> len;

    squares_array_ptr = new_array_with_squares(len);
    print_array(squares_array_ptr, len);
    delete [] squares_array_ptr;
}
```

The old squares program

Our new_array_with_squares_function had to manage the underlying array, and its length

```
int *new_array_with_squares(int target_size)
{
    int *array = nullptr;

    // Fill with squares
    for (int size = 0; size < target_size; size++) {
        // will bump array to size + 1 but not change size
        array = grow_array_by_one(array, size);
        array[size] = size * size; // array[i] gets i squared
    }

    return array;
}</pre>
```

The new squares program

Details of the array implementation are in a Smart_Array struct.

```
int main ()
{
         Smart_Array squares_array;
         int len;

         cout << "How many squares do you want to store? ";
         cin >> len;

         squares_array = new_array_with_squares(len);
         print_array(squares_array);
         delete_smart_array(&squares_array);
}
```

The new_array_with_squares fund

Smart set_at function is like:

```
arr[i] = ...
```

...but automatically grows the array as needed.

The Smart_Array Struct

This struct encapsulates both the pointer to the current allocated array, and its size.

```
struct Smart_Array {
    int *current_array;
    int size; // highest set index+1
};
```

```
struct Smart_Array {
    int *current_array;
    int size; // highest set index+1
};
```

```
int get_at(Smart_Array arr, int i)
{
    // Here we assume that all requests are
    // properly within the size of the array
    return arr.current_array[i];
}
```

Example of two of the Smart array

This struct encapsulates both the pointer to the current allocated array, and its size.

```
struct Smart_Array {
    int *current_array;
    int size; // highest index+1
};
```

```
void set_at(Smart_Array *arr_ptr, int i, int new_value)
{
     while(arr_ptr->size <= i) {
          grow_array(arr_ptr);
     }
     arr_ptr->current_array[i] = new_value;
}
```

Example of two of the Smart array

In the end we'll update the array in the obvious way, but...

```
struct Smart_Array {
    int *current_array;
    int size; // highest set
};
```

```
void set_at(Smart_Array *arr_, int i, int new_value)
{
    while(arr_ptr->size = i) {
        grow_array arr_ptr);
    }
    arr_ptr->current_array[i] = new_value;
}
```

```
struct Smart_Array {
    int *current_array;
    int size; // highest set
};
```

We automatically grow the array for the caller *if* it's not already big enough.

```
void set_at(Smart_Array *arr r, int i, int new_value)
{
    while(arr_ptr->size <= i) {
        grow_array(arr_ptr);
    }
    arr_ptr->current_array[i] = new_value;
}
```

struct Smart_Array {
 int *current_array;
 int size; // highest
};

Grow_array may update the struct by allocating a new array and changing the size, thus we must pass it into this function by reference,

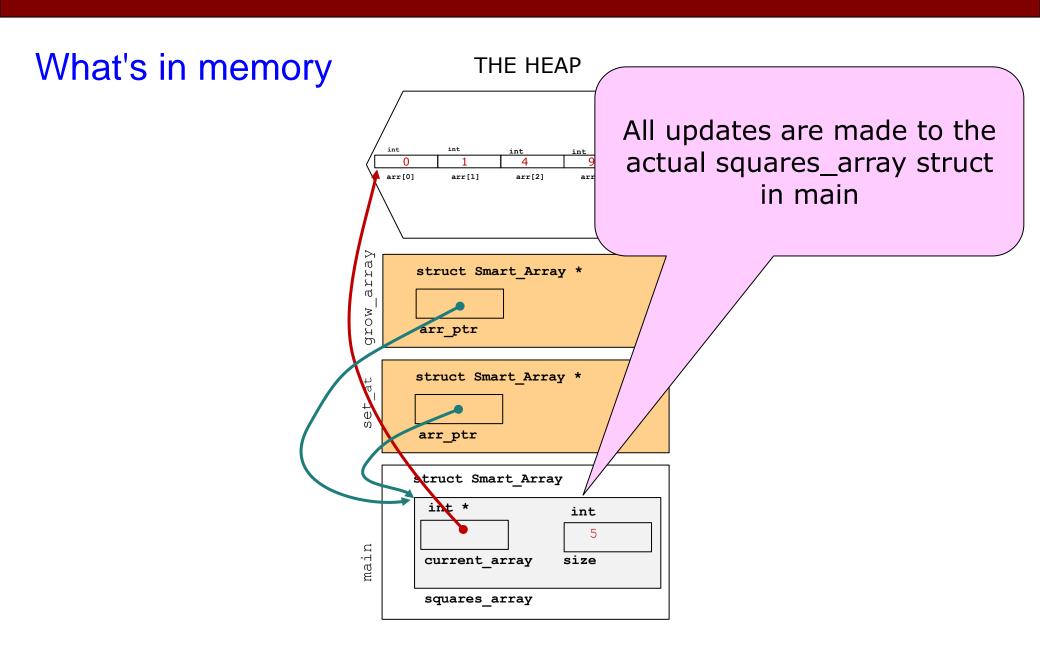
I.e. we receive here the *address* of the caller's Smart_Array struct, not a copy.

```
void set_at(Smart_Array *arr_ptr, int i, int new_value)
{
    while(arr_ptr->size <= i) {
        grow_array(arr_ptr);
    }
    arr_ptr->current_array[i] = new_value;
}
```

```
struct Smart_Array {
    int *current_array;
    int size; // highest
};
```

...and for the same reason we pass on the pointer to grow_array, which is actually manipulating the struct in main.

```
void set_at(Smart_Array *arr_pt int i, int new_value)
{
    while(arr_ptr->size <= i) {
        grow_array(arr_ptr);
    }
    arr_ptr->current_array[i] = new_value;
}
```



The grow array function

```
struct Smart_Array {
    int *current_array;
    int size; // highes
};
```

Using the arr_ptr argument, grow_array is updating the struct that's on the stack in *main's* activation record.

```
void grow array(Smart Array *arr ptr)
        int i;
                                 r ptr->size + 1];
        int *new array = new in/
                               size; i++) {
        for (i = 0; i < arr p*
                new array[i/
                             /arr ptr->current array[i];
        delete [] arr // tr->current_array;
        arr ptr->size = arr ptr->size + 1;
        arr ptr->current array = new array;
```

Review: Using structs and functions for modularit

- We've seen how the application code got much simpler
- Details of managing the arrays were hidden...
- ...e.g. we can change the array growing strategy without changing the code that makes and prints lists of squared numbers
- The pointer stuff looks a little tricky at first, but these are classic techniques...they're not hard once you get the hang of it!
- In our next lesson we will see how by turning our structs into classes, we can make the story even more powerful
- ...a little later we will learn to package classes like Smart_Array into separate, shareable C++ sources files so we can reuse them in lots of programs!

Summary