

operands operator

y = 0+3;

(tatement

primitive 9 of farifilmetic oper

introduces 1 - expression start

privation call start

privation call start

countries flowstart

response of the private oper

structured reference 1 - punction call start

countries flowstart

response of the proceedence 1 - a sq. b 11 c sq. c

private operator

private operator

private operator

privation call start

countries flowstart

private operator

private

# Dynamic memory allocation scenario / motivation 1

· When we define an array, we allocate memory for it

```
int arr[20];
sets aside space for 20 ints (80 bytes)
```

This space is allocated at compile-time (i.e. when the program is compiled)



81

# Dynamic memory allocation scenario / motivation 1

- What if we do not know how large our array should be?
- · length is determined at runtime rather than compile time
- In other words, we need to be able to allocate memory at run-time (i.e. while the program is running)

```
• How?
int n;
printf("How many elements in int array? ");
scanf("%d", &n);
int my_array[n]; /* but not allowed in ANSI-C */
   gcc -ansi -pedantic varArray.c
   gcc -ansi -pedantic-errors varArray.c

**2
   ISO C90 forbids variable length array 'my_array'
```

- Fortunately, C supports dynamic storage allocation: the ability to allocate storage during program execution.
- Using dynamic storage allocation, we can design data structures that grow (and shrink) as needed.
- The <stdlib.h> header declares three memory allocation functions:

```
mallocAllocates a block of memory but doesn't initialize it.CallocAllocates a block of memory and clears it.
```

realloc Resizes a previously allocated block of memory.

- These functions return a value of type void \* (a "generic" pointer).
  - function has no idea what type of data to store in the block.

83

# Common library functions [Appendix of K+R]

```
<stdio.h>
                  <string.h>
                                  <stdlib.h>
                                                     <ctype.h>
printf()
                  strlen(s)
scanf()
                  strcpy(s,s)
                                  double atof(s)
                                                    int islower(int)
getchar()
                  strcat(s,s)
                                  int
                                                    int isupper(int)
                                         atoi(s)
putchar()
                  strcmp(s,s)
                                  long
                                         atol(s)
                                                    int isdigit(int)
                  strtok(s,s)
                                  void rand()
                                                    int isxdigit(int)
                                  void
sscanf()
                                         system()
                                                    int isalpha(int)
                  <math.h>
sprintf()
                                  void
                                         exit()
                  sin() cos()
                                  int
                                         abs(int)
                                                    int tolower(int)
                  exp()
gets() puts()
                                                    int toupper(int)
                  log()
fgets() fputs()
                                  void* malloc()
                                                    <assert.h>
                  () wog
                                  void* calloc()
                                                    assert()
                  sqrt()
fprintf()
                                  void* realloc()
                  ceil()
fscanf()
                                  void free()
                  floor()
```

## malloc()



"stdlib.h" defines:

- allocates memory at run-time
- returns a void pointer to the memory that has at least n bytes available (just allocated for you).
  - Address of first byte e.g., 1000
  - Can be <u>casted</u> to any type

YORK

85

## Summary of pointer operations



- Legal:

  - assignment of pointers of the same type
  - adding or subtracting a pointer with an integer | p++ , p+2, p-2
  - subtracting or comparing two pointers to members of the same p2- p1 if (p1 < p2) while (p1 != p2)
  - assigning or comparing to zero (NULL) (later)

Illegal:

- add two pointers, multiply or divide two pointers, integers
- add or subtract float or double to pointers p + 1.23
- shift or mask pointer variables p << 2
- assign a pointer of one type to a pointer of another (except for void \*) without a cast used in OS course RK

```
pangling Pointers

malloc()

#include <stdlib.h>

int main() {
  int *p; // uninitialized, not point to anywhere

  *p = 52;
  printf("%d\n", *p);
}

segmentation fault
core dump

YORK

YORK

YORK

**TORY
**
```

```
Whenever you need to set a pointer's pointee
e.g.,
    *ptr = var;
    scanf("%s", ptr);
    strcpy(ptr, "hello");
    fgets(ptr, 10, STDIN);
    .....
    *ptrArr[2] = var; // pointer array

Ask yourself: Have you done one of the following
1. ptr = &var. /* direct */
    arr[20]; ptr=&arr[0];
2. ptr = ptr2 /* indirect, assuming ptr2 is good */
3. ptr = (..)malloc(....) /* now */
88
```

88

```
malloc()
#include <stdlib.h>

int main() {
   int *p, x;
   p = &x;
   *p = 52; // x=52
   printf("%d\n", *p);
}
```

malloc()

#include <stdlib.h>

int main() {
 int \*p, x;
 int \*p2 = &x; p = p2;
 \*p = 52; // x=52
 printf("%d\n", \*p);
}

Simple fix2

YORK

\*\*TORY OF THE PLANT OF THE PRODUCT OF THE PLANT OF T

90

```
malloc()
#include <stdlib.h>

int main() {
   int *p;
   p = (int *) malloc(4);
   *p = 52;
   printf("%d\n", *p);
}

• Note: type conversion (cast) on result of malloc
   p = malloc(4); also works. Will convert

YORK

1000 1001 1002 1003

fix3

fix3

Improve?
```

## Improve1 sizeof

· A better approach to ensure portability

```
1000 1001 1002 1003
```

```
int *p;
p = (int *) malloc(4);

p = (int *) malloc( sizeof(int) );
*p = 52;
```

## Improve 2 NULL

- · Allocation not always successful
- malloc() returns NULL when it cannot fulfill the request, i.e., memory allocation fails (e.g. no enough space)

```
int *p;
p = (int *)malloc(100000000);// malloc returns NULL
p = (int *)malloc(-10); // malloc returns NULL
```

93



93

#### **NULL**

- <stdlib.h> <stdio.h> <string.h> ...defines macro
   NULL a special pointer constant with value 0
- 0 (zero) is never a valid address

```
    NULL == "0 as a pointer" == "points to nothing"
    int * p; // p == NULL? Not really
```

```
■ p == 0 ? // better use NULL like EOF
```

```
p = malloc(10000000);
if (p == NULL) {  // an "exception"
   exit(0) /* allocation failed; take appropriate action '
}
else ...

if ( (p = malloc(10000000)) == NULL) {
   exit(0) /* allocation failed; take appropriate action '
}else ....
```

```
1024 1025 1026 1027
                                  p+1 p+2
malloc()
                              1000
                                 1004
                                     1008
                                         1012
                                            1016
#include <stdlib.h>
                               1
                                      12
                               0
                                      2
                                  1
                                          3
int main() {
                                 4n bytes allocated.
  int n;
                                 n=7 28 bytes 1000~1027 allocated
  printf("How many elements in int array? ");
  scanf("%d", &n);
  int * p = (int *)malloc(n * sizeof(int));
  if (p == NULL)
     exit(0);
  // else
  *p = 1;
                   // p[0] = 1 second +1 +4?
  *(p+1) = 2;
                  // p+1 = 1004 p[1] = 2
  *(p+2) = 12;
                   // p+2 = 1008 p[2] = 12
                      pointer arithmetic!!!
```

```
malloc()
#include <stdlib.h>
                                 1000 1001 1002 1003 1004 1005 1006
int main() {
  int n;
  printf("chars in array: "); _{\rm n~bytes~allocated.~Include~for~ \setminus 0}
  scanf("%d", &n);
                                    n=7 7 bytes 1000~1006 allocated
  char * p = (char *)malloc(n * sizeof(char)); //n+1?
  if (p == NULL)
                                                    *(p+0) = 'a';
      exit(0);
                                                    *(p+1) = 'b';
                                                    *(p+2) = 'c';
  strcpy(p, "abc");
                                                    *(p+3) = '\0'
  *(p+1) = 'x';
  printf("%s", p); // axc
  printf("%d", strlen(p));
9 printf("%s", p+1);
```

## calloc()

What if we want to allocate arrays of n element?
 malloc (n \* sizeof(int));
 alternatively,

```
void * calloc(int n, int size);
```

 calloc allocates an array of n elements where each element has size size

```
• e.g.
   int *p;
   p = (int *)calloc(6, sizeof(int));
```

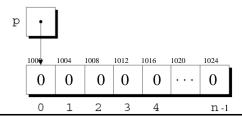
98



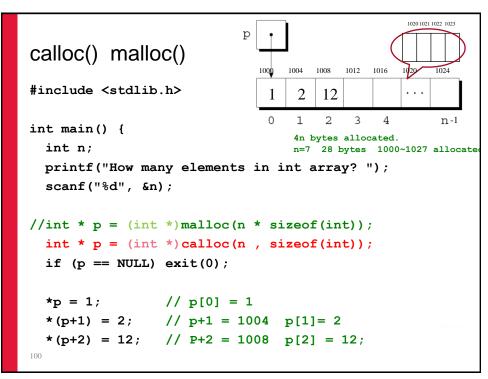
98

## calloc() vs. malloc()

- calloc(x , y) is pretty much the same as
   malloc(x \* y)
- except
  - malloc does not initialize memory
  - calloc initializes memory content to 0 (zero)







## free()

- memory allocation functions malloc, calloc obtain memory blocks from a storage pool known as the heap, where storage is persistent until the programmer explicitly requests that it be deallocated (or program terminates)
- A block of memory that's no longer accessible to a program is said to be garbage.
  - A program that leaves garbage behind has a memory leak.
- Some languages (e.g., Java) provide a garbage collector that automatically locates and recycles garbage, but C doesn't.

YORK

101

### Memory Leaks

- The first memory block is lost "forever" (until program terminates).
- May cause problems (exhaust memory).



102

#### **Memory Leaks**

- What happens if some memory is heap allocated, but never deallocated?
- A program which forgets to deallocate a block is said to have a "memory leak" which may or may not be a serious problem. The result will be that the heap gradually fill up as there continue to be allocation requests, but no deallocation requests to return blocks for re-use.
- For a program which runs, computes something, and exits immediately, memory leaks are not usually a concern. Such a "one shot" program could omit all of its deallocation requests and still mostly work.
- Memory leaks are more of a problem for a program which runs for an indeterminate amount of time. In that case, the memory leaks can gradually fill the heap until allocation requests cannot be satisfied, and the program stops working or crashes.

103

For your information

### free()

 Instead, each C program is responsible for recycling its own garbage by calling the <u>free</u> function to release unneeded memory.

```
void free (void *ptr);
```

- "frees" memory we previously allocated, tells the system we no longer need this memory and that it can be reused
- address in "ptr" must have been returned from either malloc, calloc Or realloc.

```
p = malloc(7*4);
...
104 free(p);
```



```
1024 1025 1026 1027
                                       p+1 p+2
malloc() calloc()
                                      1004
                                          1008 1012 1016
                                                       1020
#include <stdlib.h>
                                   1
                                           12
                                            2
int main() {
                                       4n bytes allocated.
  int n; int *p;
                                      n=7 28 bytes 1000~1027 allocated
  printf("How many elements in int array? ");
  scanf("%d", &n);
  p = (int *)malloc(n * sizeof(int)); //or
  p = (int *)calloc(n , sizeof(int));
  if (p == NULL)
      exit(0);
  *p = 1;
                    // store 1 at address 1000 (1000~1003)
  *(p+1) = 2;
                    // p+1 = 1004 store 2 at address 1004
  *(p+2) = 12;
                    // p+2 = 1008    store 12 at address 1008
                       pointer arithmetic!!!
  free (p);
  p=&i;
```

```
char *ptr;
realloc()
                                         ptr = malloc(20);
                                         ptr = realloc(ptr,50);
 resize a dynamically allocated array.
 void *realloc(void *ptr, int size);
ptr must point to a memory block obtained by a previous call of
 malloc, calloc, Or realloc.
  ptr is NULL, a new block is allocated
 size represents the new size of the block, which may be larger
 or smaller than the original size.
realloc(NULL, n) behaves like malloc(n).
 realloc (ptr, 0) behaves like free (prt), as it frees
 the memory block.
106
                       For your information
```

### More on memory allocation

- We know the syntax
- But when to use it ?????
  - When need to allocate at run time, of course
  - What else?
- Another feature of malloc -- request for heap space!

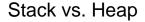


```
#include <stdio.h>
void setArr (int);
int * arr[10]; // global, array of 10 int pointers
int main(int argc, char *argv[])
  setArr(1);
  printf("arr [%d] = %d\n", 1, *arr[1]);
  return 0;
/* set arr[index], which is a pointer,
to point to an integer of value 100 */
void setArr (int index){
                                          0
   *arr[index] = 100;
                                          1
                                                                  100
}
                                          2
                       What is wrong
110
                       here??
```

```
#include <stdio.h>
void setArr (int);
int * arr[10]; // global, array of 10 int pointers
int main(int argc, char *argv[])
  setArr(1);
  printf("arr [%d] = %d\n", 1, *arr[1]);
  return 0;
/* set arr[index], which is a pointer,
to point to an integer of value 100 */
void setArr (int index){
                                          0
  int i = 100;
                                                                  i 100
  arr[index] = &i;
                                          1
                                          2
Compiles but may or
                       What is wrong
 may not work
                       here??
```

```
#include <stdio.h>
 void setArr (int);
 int * arr[10]; // array of 10 int pointers, global variable
 int main(int argc, char *argv[])
                                                          red 396 % a.out
arr[0] -*-> 101
arr[1] -*-> 32706
       setArr(0);
                                                           red 397 % a.out
                                                           arr[0] -*-> 101
arr[1] -*-> 32712
       setArr(1);
       printf("arr[%d] -*->%d\n", 0, *arr[0]);
       printf("arr[%d] -*->%d\n", 1, *arr[1]);
       return 0;
 }
                                                           red 399 %
 /* set arr[index], which is a pointer, to point to an integer of some value */
⊟void setArr (int index){
   int i = index+100;
       arr[index] = &i;
 }
                                                                        YORK
 This will probably not
                           What is wrong
  work
                           here??
```

```
#include <stdio.h>
void setArr (int);
int * arr[10]; // global, array of 10 int pointers
int main(int argc, char *argv[])
   setArr(1);
   printf("arr [%d] = %d\n", 1, *arr[1]);
                                              0
  return 0;
                                                                     i 100
                                              1
}
                                              2
/* set arr[index], which is a pointer,
to point to an integer of value 100 */
void setArr (int index){
                                              i is local variable,
   int i = 100;
                                              lifetime is block/function
   arr[index] = &i;
                                              -- i is in stack, where it is
                                              deallocated when
113
                                              function exits !!!
```



- Local (stack) memory, automatic
  - Allocated on function call, and deallocated automatically when function exits
- Stack Stack Address Space free space Heap free space Address Space Heap **Uninitialized Data** Segment (.bss) Global, static **Initialized Data** Segment (.data) Code Segment (.text)

- Dynamic (heap) memory
  - The heap is an area of memory available to allocate areas ("blocks") of memory for the program.
  - Not deallocated when function exits



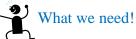
How to allocate in heap then?

115

#### Stack vs. heap

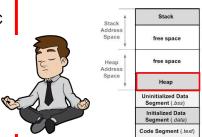
- Local (stack) memory, automatic
  - Allocated on function call, and deallocated automatically when function exits
- Dynamic heap memory
  - The heap is an area of memory available to allocate areas ("blocks")
    of memory for the program.

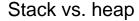
Not deallocated when function exits.



Request a heap memory:

o malloc() / calloc() / realloc() in C



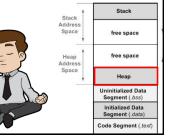


- Local (stack) memory, automatic
  - Allocated on function call, and deallocated automatically when function exits
- Dynamic heap memory
  - The heap is an area of memory available to allocate areas ("blocks") of memory for the program.
  - Not deallocated when function exits.



Request a heap memory:

- malloc() / calloc() / realloc() in Cnew in C++ and Java
  - Student s = new Student();



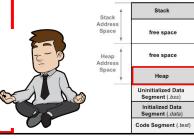
117

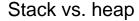
### Stack vs. heap

- · Local (stack) memory, automatic
  - Allocated on function call, and deallocated automatically when function exits
- Dynamic heap memory
  - The heap is an area of memory available to allocate areas ("blocks") of memory for the program.
  - Not deallocated when function exits.

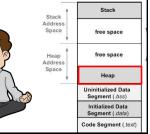


- Request a heap memory:
  - o malloc() / calloc() / realloc() in C
  - o new in C++ and Java
    - Student s = new Student();
- Deallocate from heap memory:





- Local (stack) memory, automatic
  - Allocated on function call, and deallocated automatically when function exits
- Dynamic heap memory
  - The heap is an area of memory available to allocate areas ("blocks") of memory for the program.
  - Not deallocated when function exits.
  - Request a heap memory:
    - o malloc() / calloc() / realloc() in C
    - o new in C++ and Java
      - Student s = new Student();
  - Deallocate from heap memory:
    - o free() in C
    - o delete in C++
    - o garbage collection in Java



What we need!

119

#### When to use malloc?

- When you need to allocate memory in run time, of course
- When you need memory space throughout the program execution
  - 1. ptr = &i. /\* direct \*/
  - o i needs to have persistent lifetime
    - o if i is a local variable in function?



- 2. ptr = ptr2 /\* indirect 1\*/
  - o ptr2 needs to point to persistent address
  - o if ptr2 points to a local variable?



3. ptr = (...)malloc(....)

correct choice!

local variable is in Stack. Not in heap.



120

```
Correct implementation
               #include <stdio.h>
               void setArr (int);
               int * arr[10]; // global, array of 10 int pointers
               int main(int argc, char *argv[])
                  setArr(1);
                  printf("arr [%d] = %d\n", 1, *arr[1]);
                  return 0;
   Heap
               /* set arr[index], which is a pointer,
200
               to point to an integer of value 100 */
               void setArr (int index){
                 ·
                 arr[index] = (int *) malloc(sizeof (int)); // malloc(4)
               }
  121
```

```
Correct implementation
              #include <stdio.h>
              void setArr (int);
              int * arr[10]; // global, array of 10 int pointers
              int main(int argc, char *argv[])
                 setArr(1);
                 printf("arr [%d] = %d\n", 1, *arr[1]);
                                                        // 100
                 return 0;
  Heap
               /* set arr[index], which is a pointer,
               to point to an integer of value 100 */
   100
               void setArr (int index){
                 arr[index] = (int *) malloc(sizeof (int)); // malloc(4)
                 *arr[index] = 100;
               }
                 int i=100; *(arr[index])=i;
```

```
Correct implementation
                 #include <stdio.h>
                 void setArr (int);
                 int * arr[10]; // global, array of 10 int pointers
                 int main(int argc, char *argv[])
                    setArr(1);
                    printf("arr [%d] = %d\n", 1, *arr[1]);
                                                              // 100
                    return 0;
                 }
    Heap
                 /* set arr[index], which is a pointer,
200
                 to point to an integer of value 100 */
    100
                 void setArr (int index){
                   int *p = malloc(sizeof(int));
                   *p = 100;
                   arr[index] = p; // points to heap space
 another way
```

```
#include <stdio.h>
int * arr[10]; // array of 10 int pointers, global variable
int main(int argc, char *argv[])
     int i;
     int a=0, b=100, c=200,d=300,e=400;
     arr[0] = &a;
     arr[1] = &b;
     arr[2] = &c;
     arr[3] = &d;
     arr[4] = &e;
     for(i=0; i<5;i++)</pre>
         printf("arr[%d] -*-> %d\n", i, *arr[i]); /* 0, 100, 200, 300, 400 */
     return 0;
}
This program works (but not practical).
a,b,c,d,e are local variables, in stack, but not deallocated before
program main() terminates/returns
 125
```

- Pointers (Ch5)
  - Basics: Declaration and assignment (5.1)
  - Pointer to Pointer (5.6)
  - Pointer and functions (pass pointer by value) (5.2)
  - Pointer arithmetic +- ++ -- (5.4)
  - Pointers and arrays (5.3)
    - Stored consecutively
    - o Pointer to array elements p + i = &a[i] \*(p+i) = a[i]
    - Array name contains address of 1st element a = &a[0]
    - o Pointer arithmetic on array (extension) p1-p2 p1<>!= p2
    - o Array as function argument "decay"
    - Pass sub\_array
  - Array of pointers (5.6-5.9)
  - Command line arguments (5.10)
  - Memory allocation (extra)
- Structures (Ch6)
  - Pointer to structures (6.4)
  - Self-referential structures (extra)

today

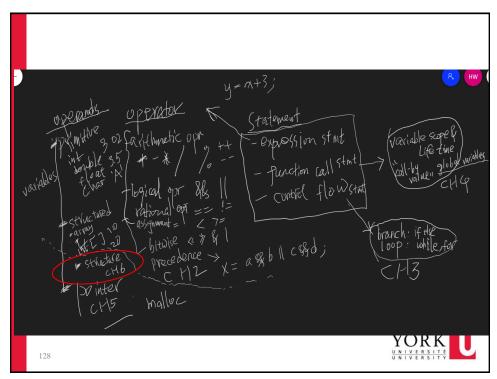


126

### EECS2031 - Software Tools

C - Structures, Unions, Enums & Typedef (K+R Ch.6)





# **Structures**

- A collection of one or more variables grouped under a single name for easy manipulation
- The variables can be of <u>different</u> types
  - Primitive data types, arrays, pointers and other structure
- Encapsulate data

int x;
int y;
int directionX;
int directionY;

• Only contains data (no functions).



129

# Structures

- · Basics: Declaration and assignment
- · Structures and functions
- Pointer to structures
- · Arrays of structures
- Self-referential structures (e.g., linked list, binary trees)

YORK

130

130

```
Structures
struct {
  float width;
  float height;
} chair;
                          - width
                          - height
 struct {
                         is the type
                                      // like int a;
    float width;
                                       // Student s; 🎉 Java
    float height
                         chair is variable name.
struct {
  float width;
                       Need to repeat
  float height;
  table;
```

#### Structure Names

• Give a name (tag) to a struct, so we can reuse it:

132

#### Structures

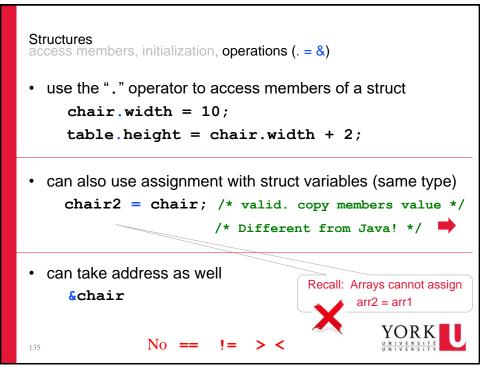
access members, initialization, operations (. = &)

• use the "." operator to access members of a struct
 chair.width = 10;
 float f = chair.height;
 table.height = chair.width + 2;

Operator Type	Operato		Associativity
Primary Expression Operators	()[]>		left-to-right
Unary Operators	* & + - !~ (typecast) siz		right-to-left
	* / %	arithmetic	
	+-	arithmetic	
	>> <<	bitwise	
	<><=>=	relational	
Rinary Operators	== !=	relational	left-to-right

```
Structures
access members, initialization, operations (. = \&)
struct shape {
    float width;
    float height;
struct shape chair = {2,4}; // approach 1
                             width
                                    height
                                                       width 2
                                                       height 4
struct shape chair;
chair.width = 2;
                                          approach 2
chair.height = 4;
                                      Size of struct not necessarily the
                                      sum of its elements. Use sizeof()
struct myshape {
                              chair
  int data;
                               data 2
  float arr[3];
                               arr
                                    1.5 2.5
struct myshape s = {2, {1.5, 2.5}}; //approach 1
(s.arr)[2] = 3.3; // approach 2 set directly

    associativity
```



```
Struct shape chair = {2,4};

struct shape chair2 = chair; // copy members values only // different from Java chair2.width = chair.width chair2.height = chair.width, chair.height);

printf("%d %d", chair.width, chair2.height);

printf("%d %d", chair2.width, chair2.height);

chair2.width = 20; // does not affect chair

printf("%d %d", chair.width, chair2.width);

? What if an element is a pointer? deep/shallow copy?
```

```
Structures
access members, initialization, operations (. = &)
                                                       chair
                                                          width 2
                              width
                                        height
                                                          height 4
 struct shape chair = {2,4};
 struct shape chair2 = chair; // copy members values
                                                          width 2
                                                          height 4
                                // different from Java
     chair2.width = chair.width
     chair2.height = chair.height
 struct shape2 {
                                          data 2
     float data;
                                                          800
                                           800
                                                          5 | i
     int * p;
 };
                                          data 2
                                         p 800
 int i = 5;
 struct shape2 s ={2, &i};
 struct shape2 s2 = s1; // s2.p = s.p
                                                   "shallow copy"
 *(s2.p) = 20;
                                                     YORK
 printf("%d %d", *(s.p), *(s2.p)); 20 20
```

Operator Type	Operato		
Primary Expression Operators	()[]>	associativity Left to right	*s.p = 3;
Unary Operators	* & + - (typecast)		scanf("%f", &chair2.wi
	* / %	arithmetic	
	+ -	arithmetic	&(chair2.wi
	>> <<	bitwise	
	<><=>=	relational	s2.arr[2] = 3
Binary Operators	== !=	relational	<b>─</b>
Diliary Operators	&	bitwise	
	^	bitwise	No () needed
	1	bitwise	
	&&	logical	
	II	logical	
Ternary Operator	?:		(* ptr).width
Assignment Operators	= += -= *= /= ^=  =	= %= >>= <<= &=	later
Comma	, '		Ť

```
Nested Structures
struct point {
 int x;
  int y; };
                             \boldsymbol{y}
                                          pt2
struct rect {
 struct point pt1;
 struct point pt2;
                                pt1
};
struct rect screen;
screen.pt1.x = 1;
screen.pt2.x = 8;
                        Associativity
(screen.pt2).y = 7;
                       left to right
```

### Structures vs. Arrays (so far)

- Both are aggregate (non-scalar) types in C -- type of data that can be referenced as a single entity, and yet consists of more than one piece of data.
- Both cannot be compared using == != X
- Array: elements are of same type
   Structure: elements can be of different type
- Array: element accessed by [index/position] arr[1] = 3;
   Structure: element accessed by .name chair.width = 4;
- Array: cannot assign as a whole arr2 = arr1 X
  Structure: can assign/copy as a whole chair2 = chair1
- Array: size is the sum of size of elements
   Structure: size not necessarily the sum of size of elements

use sizeof

140

## Structures

- · Basics: Declaration and assignment
- Structures and functions
- Pointer to structures
- Arrays of structures
- Self-referential structures (e.g., linked list, binary trees)

