#### **Runtime Environments**

#### Stephen A. Edwards

**Columbia University** 

Fall 2018



**Storage Classes** 

The Stack and Activation Records

**In-Memory Layout Issues** 

The Heap

**Automatic Garbage Collection** 

Shared Libraries and Dynamic Linking

Objects and Inheritance

**Exceptions** 

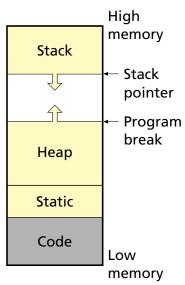
# Storage Classes

#### **Storage Classes and Memory Layout**

Stack: objects created/destroyed in last-in, first-out order

Heap: objects created/destroyed in any order; automatic garbage collection optional

Static: objects allocated at compile time; persist throughout run



#### **Static Objects**

```
class Example {
   public static final int a = 3;

public void hello() {
    System.out.println("Hello");
   }
}
```

#### **Advantages**

Zero-cost memory management

Often faster access (address a constant)

No out-of-memory danger

#### **Examples**

Static class variable
Code for hello method
String constant "Hello"
Information about the
Example class

#### Disadvantages

Size and number must be known beforehand

Wasteful if sharing is possible

# Records

The Stack and Activation

#### **Stack-Allocated Objects**



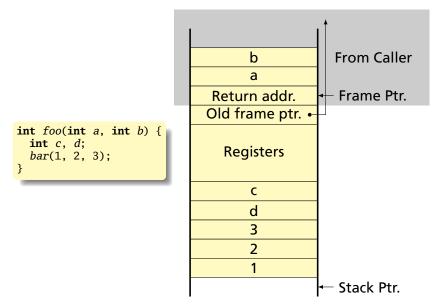
Natural for supporting recursion.

Idea: some objects persist from when a procedure is called to when it returns.

Naturally implemented with a stack: linear array of memory that grows and shrinks at only one boundary.

Each invocation of a procedure gets its own *frame* (*activation record*) where it stores its own local variables and bookkeeping information.

#### An Activation Record: The State Before Calling bar



#### **Recursive Fibonacci**

#### (Real C)

```
int fib(int n) {
   if (n<2)
     return 1;
   else
     return
        fib(n-1)
        +
      fib(n-2);
}</pre>
```

#### (Assembly-like C)

```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1;
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
}</pre>
```

```
fib(3)
fib(2) fib(1)
fib(1) fib(0)
```



```
SP .
```

n = 3

```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
}
```

```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
}
```

return address last frame pointer•
tmp1 = 2
tmp2 =
tmp3 =
n = 2

```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
                                SP
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
```

```
n = 3
return address
last frame pointer.
tmp1 = 2
tmp2 =
tmp3 =
n = 2
return address
last frame pointer.
tmp1=1
tmp2 =
tmp3 =
n = 1
```

```
Executing fib(3)
                                         n = 3
                                         return address
                                         last frame pointer.
                                         tmp1 = 2
                                         tmp2 =
  int fib(int n) {
                                         tmp3 =
      int tmp1, tmp2, tmp3;
                                         n = 2
      tmp1 = n < 2;
                                         return address
      if (!tmp1) goto L1;
                                         last frame pointer.
      return 1:
                                         tmp1=1
 L1: tmp1 = n - 1;
                                         tmp2 =
      tmp2 = fib(tmp1); \Leftarrow
                                         tmp3 =
 L2: tmp1 = n - 2;
                                         n = 1
      tmp3 = fib(tmp1);
                                         return address
 L3: tmp1 = tmp2 + tmp3;
                                         last frame pointer.
      return tmp1;
                                         tmp1 = 1
                                         tmp2 =
                                         tmp3 =
```

```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
```

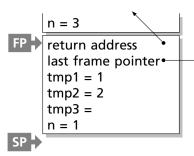
```
n = 3
return address
last frame pointer.
tmp1 = 2
tmp2 =
tmp3 =
n = 2
return address
last frame pointer.
tmp1 = 0
tmp2 = 1
tmp3 =
n = 0
```

```
Executing fib(3)
                                        n = 3
                                        return address
                                        last frame pointer.
                                        tmp1 = 2
                                        tmp2 =
  int fib(int n) {
                                        tmp3 =
      int tmp1, tmp2, tmp3;
                                        n = 2
      tmp1 = n < 2;
                                        return address
      if (!tmp1) goto L1;
                                        last frame pointer•
      return 1;
                                        tmp1 = 0
 L1: tmp1 = n - 1;
                                        tmp2 = 1
      tmp2 = fib(tmp1);
                                        tmp3 =
 L2: tmp1 = n - 2;
                                        n = 0
      tmp3 = fib(tmp1);
                                        return address
 L3: tmp1 = tmp2 + tmp3;
                                        last frame pointer.
      return tmp1;
                                        tmp1 = 1
                                        tmp2 =
                                        tmp3 =
```

```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
}
```

```
n = 3
return address
last frame pointer.
tmp1 = 2
tmp2 =
tmp3 =
n = 2
return address
last frame pointer.
tmp1 = 2
tmp2 = 1
tmp3 = 1
```

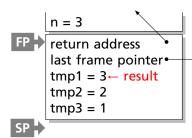
```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
}
```



```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1;
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
```

```
n = 3
return address
last frame pointer.
tmp1 = 1
tmp2 = 2
tmp3 =
n = 1
return address
last frame pointer.
tmp1 =
tmp2/=
tmp3 =
```

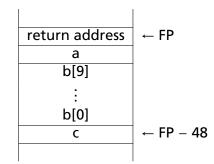
```
int fib(int n) {
    int tmp1, tmp2, tmp3;
    tmp1 = n < 2;
    if (!tmp1) goto L1;
    return 1:
L1: tmp1 = n - 1;
    tmp2 = fib(tmp1);
L2: tmp1 = n - 2;
    tmp3 = fib(tmp1);
L3: tmp1 = tmp2 + tmp3;
    return tmp1;
}
```



#### Allocating Fixed-Size Arrays

Local arrays with fixed size are easy to stack.

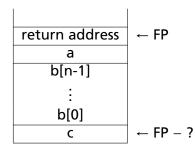
```
void foo()
{
   int a;
   int b[10];
   int c;
}
```



#### Allocating Variable-Sized Arrays

Variable-sized local arrays aren't as easy.

```
void foo(int n)
{
   int a;
   int b[n];
   int c;
}
```

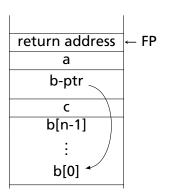


Doesn't work: generated code expects a fixed offset for c. Even worse for multi-dimensional arrays.

#### Allocating Variable-Sized Arrays

As always: add a level of indirection

```
void foo(int n)
{
   int a;
   int b[n];
   int c;
}
```



Variables remain constant offset from frame pointer.

#### **Nesting Function Definitions**

```
let articles words =
                                     let count words w = List.length
                                       (List.filter ((=) w) words) in
 let report w =
                                     let report words w = w ^ ": " ^
   let count = List.length
                                       string_of_int (count words w) in
      (List.filter ((=) w) words)
    in w ^ ": " ^
                                     let articles words =
       string_of_int count
                                       String.concat ", "
                                         (List.map (report words)
 in String.concat ", "
                                          ["a"; "the"]) in
    (List.map report ["a"; "the"])
                                     articles
in articles
                                         ["the": "plt": "class": "is":
    ["the"; "plt"; "class"; "is";
                                          "a"; "pain"; "in";
     "a": "pain": "in":
                                          "the": "butt"]
     "the"; "butt"]
```

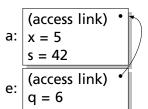
Produces "a: 1, the: 2"

```
let a \times s =
 let b y =
    let c z = z + s in
    let d w = c (w+1) in
    d(y+1) in (* b *)
 let e q = b (q+1) in
e(x+1)(*a*)
```

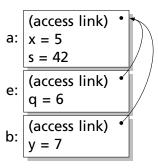
What does "a 5 42" give?

a:  $\begin{cases} (access link) & \bullet \\ x = 5 \\ s = 42 \end{cases}$ 

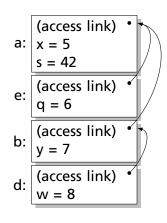
```
let a \times s =
 let b y =
    let c z = z + s in
    let d w = c (w+1) in
    d(y+1) in (* b *)
 let e q = b (q+1) in
e(x+1)(*a*)
```



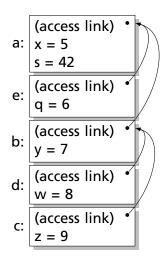
```
let a \times s =
  let b y =
    let c z = z + s in
    let d w = c (w+1) in
    d(y+1) in (* b *)
  let e \ q = b \ (q+1) in
e(x+1)(*a*)
```



```
let a \times s =
  let b y =
    let c z = z + s in
    let d w = c (w+1) in
    d(y+1) in (* b *)
  let e \ q = b \ (q+1) in
e (x+1) (* a *)
```



```
let a \times s =
  let b y =
    let c z = z + s in
    let d w = c (w+1) in
    d(y+1) in (* b *)
  let e \ q = b \ (q+1) in
e (x+1) (* a *)
```



# In-Memory Layout Issues

#### Layout of Records and Unions

Modern processors have byte-addressable memory.

0

1

2



The IBM 360 (c. 1964) helped to popularize byte-addressable memory.

Many data types (integers, addresses, floating-point numbers) are wider than a byte.

16-bit integer:

1 0

32-bit integer:

3

2

2



#### Layout of Records and Unions

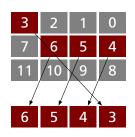
Modern memory systems read data in 32-, 64-, or 128-bit chunks:



Reading an aligned 32-bit value is fast: a single operation.

3	2	1	0
7	6	5	4
11	10	9	8

It is harder to read an unaligned value: two reads plus shifting



SPARC and ARM prohibit unaligned accesses

MIPS has special unaligned load/store instructions

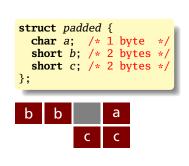
x86, 68k run more slowly with unaligned accesses

#### **Padding**

To avoid unaligned accesses, the C compiler pads the layout of unions and records.

#### Rules:

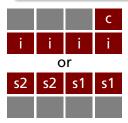
- ► Each *n*-byte object must start on a multiple of *n* bytes (no unaligned accesses).
- Any object containing an n-byte object must be of size mn for some integer m (aligned even when arrayed).



#### Unions

A C struct has a separate space for each field; a C union shares one space among all fields

```
union intchar {
  int i;    /* 4 bytes */
  char c;  /* 1 byte */
};
```



#### **Arrays**

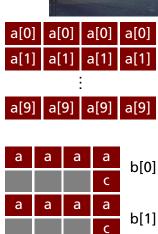


Basic policy in C: an array is just one object after another in memory.

```
int a[10];
```

This is why you need padding at the end of *structs*.

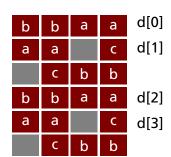
```
struct {
   int a;
   char c;
} b[2];
```



#### Arrays and Aggregate types

The largest primitive type dictates the alignment

```
struct {
    short a;
    short b;
    char c;
} d[4];
```



# **Arrays of Arrays**

char a[4];

a[3] a[2] a[1] a[0]

a[0][3] a[0][2] a[0][1] a[0][0] a[0]

char a[3][4];

a[1][3] a[1][2] a[1][1] a[1][0] a[1]

a[2][3] a[2][2] a[2][1] a[2][0] a[2]

# The Heap

### **Heap-Allocated Storage**

Static works when you know everything beforehand and always need it.

Stack enables, but also requires, recursive behavior.

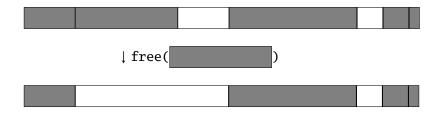
A *heap* is a region of memory where blocks can be allocated and deallocated in any order.

(These heaps are different than those in, e.g., heapsort)

```
struct point {
   int x, y;
}:
int play_with_points(int n)
  int i;
  struct point *points;
 points = malloc(n * sizeof(struct point));
  for (i = 0; i < n; i++) {
    points[i].x = random();
    points[i].y = random();
  /* do something with the array */
  free(points);
```











### Rules:

Each allocated block contiguous (no holes)

Blocks stay fixed once allocated

malloc()

Find an area large enough for requested block

Mark memory as allocated

free()

Mark the block as unallocated

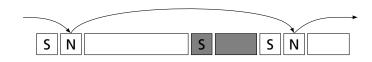


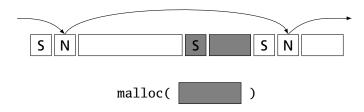
Maintaining information about free memory Simplest: Linked list

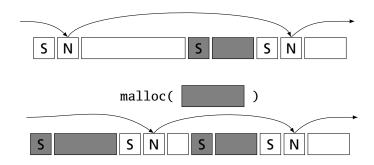
The algorithm for locating a suitable block Simplest: First-fit

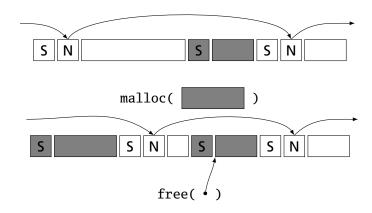
The algorithm for freeing an allocated block

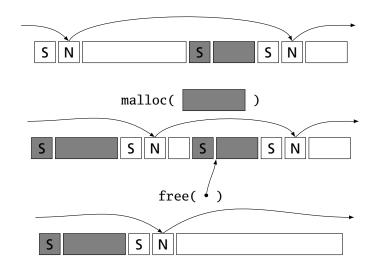
Simplest: Coalesce adjacent free blocks











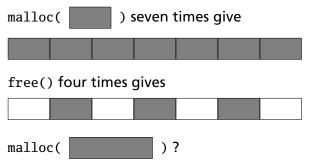
Many, many other approaches.
Other "fit" algorithms
Segregation of objects by size
More clever data structures

### **Heap Variants**

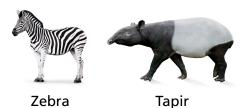
Memory pools: Differently-managed heap areas Stack-based pool: only free whole pool at once Nice for build-once data structures Single-size-object pool:

Fit, allocation, etc. much faster Good for object-oriented programs

### Fragmentation



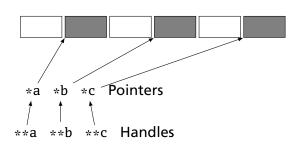
Need more memory; can't use fragmented memory.



### Fragmentation and Handles

Standard CS solution: Add another layer of indirection.

Always reference memory through "handles."



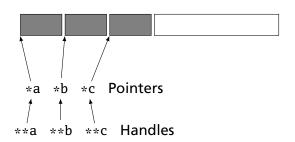


The original Macintosh did this to save memory.

### Fragmentation and Handles

Standard CS solution: Add another layer of indirection.

Always reference memory through "handles."





The original Macintosh did this to save memory.

# Collection

**Automatic Garbage** 

### **Automatic Garbage Collection**

Entrust the runtime system with freeing heap objects

Now common: Java, C#, Javascript, Python, Ruby, OCaml and most functional languages

### **Advantages**

Much easier for the programmer

Greatly improves reliability: no memory leaks, double-freeing, or other memory management errors

### Disadvantages

Slower, sometimes unpredictably so

May consume more memory



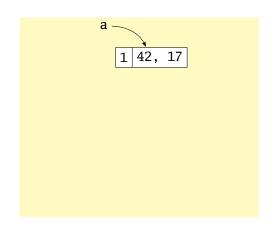
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



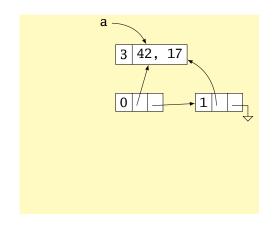
- Maintain count of references to each object
- ▶ Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



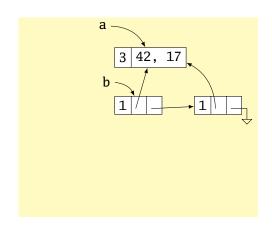
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



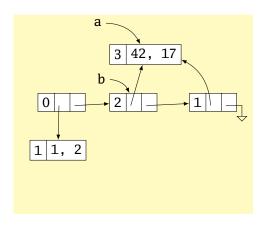
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



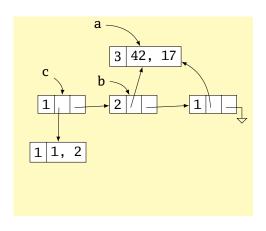
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



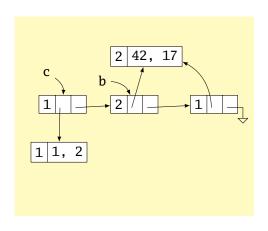
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



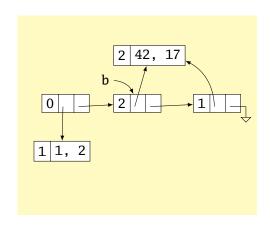
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



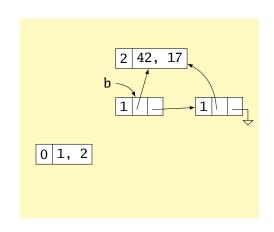
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



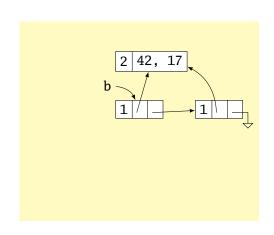
- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



- Maintain count of references to each object
- Free when count reaches zero

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



### **Issues with Reference Counting**

Circular structures defy reference counting:



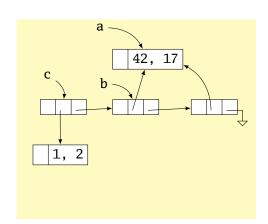
Neither is reachable, yet both have non-zero reference counts.

High overhead (must update counts constantly), although incremental

### Mark-and-Sweep

- Stop-the-world algorithm invoked when memory full
- Breadth-first-search marks all reachable memory
- All unmarked items freed

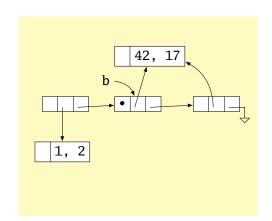
```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



### Mark-and-Sweep

- Stop-the-world algorithm invoked when memory full
- Breadth-first-search marks all reachable memory
- All unmarked items freed

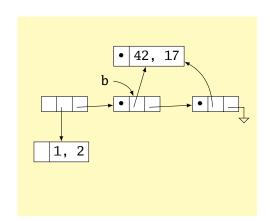
```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



### Mark-and-Sweep

- Stop-the-world algorithm invoked when memory full
- Breadth-first-search marks all reachable memory
- All unmarked items freed

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```

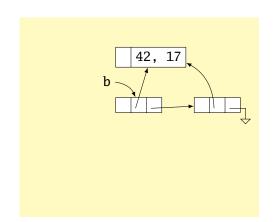


### Mark-and-Sweep

What and when to free?

- Stop-the-world algorithm invoked when memory full
- Breadth-first-search marks all reachable memory
- All unmarked items freed

```
let a = (42, 17) in
let b = [a;a] in
let c = (1,2)::b in
b
```



### Mark-and-Sweep

Mark-and-sweep is faster overall; may induce big pauses

Mark-and-compact variant also moves or copies reachable objects to eliminate fragmentation

Incremental garbage collectors try to avoid doing everything at once

Most objects die young; generational garbage collectors segregate heap objects by age

Parallel garbage collection tricky

Real-time garbage collection tricky

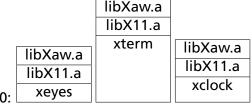
# Shared Libraries and Dynamic Linking



### Shared Libraries and Dynamic Linking

The 1980s GUI/WIMP revolution required many large libraries (the Athena widgets, Motif, etc.)

Under a *static linking* model, each executable using a library gets a copy of that library's code.



Address 0:

### **Shared Libraries and Dynamic Linking**

The 1980s GUI/WIMP revolution required many large libraries (the Athena widgets, Motif, etc.)

Under a *static linking* model, each executable using a library gets a copy of that library's code.



Address 0:

Wasteful: running many GUI programs at once fills memory with nearly identical copies of each library.

Something had to be done: another level of indirection.

### **Shared Libraries: First Attempt**

Most code makes assumptions about its location.

First solution (early Unix System V R3) required each shared library to be located at a unique address:

		libXm.so
libXaw.so	libXaw.so	
libX11.so	libX11.so	libX11.so

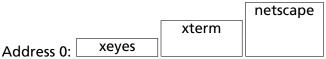
Address 0: xeyes netscape

### **Shared Libraries: First Attempt**

Most code makes assumptions about its location.

First solution (early Unix System V R3) required each shared library to be located at a unique address:



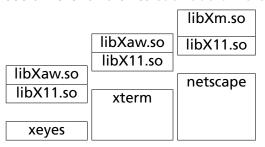


Obvious disadvantage: must ensure each new shared library located at a new address.

Works fine if there are only a few libraries; tended to discourage their use.

### **Shared Libraries**

Problem fundamentally is that each program may need to see different libraries each at a different address.



### Position-Independent Code

Solution: Require the code for libraries to be position-independent. Make it so they can run anywhere in memory.

As always, add another level of indirection:

- All branching is PC-relative
- All data must be addressed relative to a base register.
- All branching to and from this code must go through a jump table.

### Position-Independent Code for bar()

## Normal unlinked code

```
save %sp. -112, %sp
sethi %hi(0), %o0
  R SPARC HI22 .bss
mov %00, %00
  R SPARC LO10 .bss
sethi %hi(0). %o1
  R SPARC HT22 a
mov %o1, %o1
  R SPARC LO10 a
call 14
  R SPARC WDISP30 strcpv
nop
sethi %hi(0), %o0
  R SPARC HI22 .bss
mov %00, %00
  R SPARC LO10 .bss
call 24
  R SPARC WDISP30 baz
nop
ret
restore
```

### gcc -fpic -shared

```
save %sp, -112, %sp
sethi %hi(0x10000), %17
call 8e0
          ! add PC to %17
add %17, 0x198, %17
1d [\%17 + 0x20], \%00
1d
    [\%17 + 0x24].\%01
             Actually just a stub
call
      10a24 ! strcpy
nop
1d [\%17 + 0x20], \%00
      call is PC-relative
call 10a3c | baz
nop
ret
restore
```

**Objects and Inheritance** 

### Single Inheritance

Simple: Add new fields to end of the object

Fields in base class always at same offset in derived class (compiler never reorders)

Consequence: Derived classes can never remove fields

```
C++
class Shape {
  double x, y;
};
class Box : Shape {
  double h, w;
};
class Circle : Shape {
  double r;
};
```

### **Equivalent C**

```
struct Shape {
  double x, y;
};

struct Box {
  double x, y;
  double h, w;
};

struct Circle {
  double x, y;
  double r;
};
```

### **Virtual Functions**

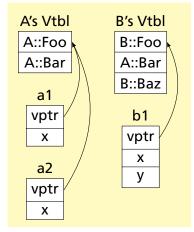
```
class Shape {
  virtual void draw(); // Invoked by object's run-time class
};
                        // not its compile-time type.
class Line : public Shape {
  void draw();
class Arc : public Shape {
  void draw();
};
Shape *s[10];
s[0] = \mathbf{new} \ Line;
s[1] = \text{new } Arc;
s[0]->draw(); // Invoke Line::draw()
s[1]->draw(); // Invoke Arc::draw()
```

### Virtual Functions

Trick: add to each object a pointer to the virtual table for its type, filled with pointers to the virtual functions.

Like the objects themselves, the virtual table for each derived type begins identically.

```
struct A {
  int x;
  virtual void Foo();
  virtual void Bar();
};
struct B : A 
  int v;
  virtual void Foo();
  virtual void Baz();
};
A a1;
A a2;
B b1;
```



# **Exceptions**



### C++'s Exceptions

```
struct Except {} ex; // This struct functions as an exception
void top(void) {
  try {
    child();
  } catch (Except e) { // throw sends control here
    printf("oops\n");
void child() {
  child2();
void child2() {
  throw ex; // Pass control up to the catch block
```





### C's setjmp/longjmp: Idiosyncratic Exceptions

```
#include <setjmp.h>
jmp_buf closure; /* return address, stack & frame ptrs. */
void top(void) {
  switch ( setjmp(closure) ) { /* normal: store closure, return 0 *
                              /* longjmp jumps here, returns 1 */
  case 0: child():
                              /* unexceptional case */
         break:
  case 1: break;
                              /* longjmp( ,1) called */
void child() {
  child2();
void child2() {
  longjmp(closure, 1);
```

### Implementing Exceptions

### One way: maintain a stack of exception handlers

```
try {
                     push(Ex, Handler); // Push handler on stack
  child();
                      child();
                     pop();
                                      // Normal termination
                     goto Exit;
                                    // Jump over "catch"
\} catch (Ex e) {
                    Handler:
  foo();
                     foo();
                                       // Body of "catch"
                    Exit:
void child() {
                   void child() {
  child2();
                      child2();
void child2() {
                    void child2() {
  throw ex;
                     throw(ex); // Unroll stack; find handler
```

Incurs overhead, even when no exceptions thrown

### Implementing Exceptions with Tables

Q: When an exception is thrown, where was the last try?

A: Consult a table: relevant handler or "pop" for every PC

```
void foo() {
23456789
                                     Lines
                                             Action
      try {
        bar();
                                       1–2
                                             Pop stack
      } catch (Ex1 e) {
                                             Handler @ 5 for Ex1
                                       3-5
        a();
    }
10
    void bar() {
11
      baz();
                                     6–15 Pop stack
12
13
14
    void baz() {
15
16
    try {
17
        throw ex1:
                                            Handler @ 18 for Ex2
                                    16–18
18
      } catch (Ex2 e) {
19
        b();
20
                                             Pop stack
                                    19–21
21
```