

Scope and Lifetime

Declaration is portion of program text to which it applies

Can be contiguous.

Scope is static: independent of data.

Extent of storage is portion of program execution during which it exists.

Can be contiguous

Dynamic: depends on data

Extent:

Lifetime duration of program

Automatic: duration of call or block execution (local variables)

From time of allocation statement (`new`) to deallocation.

27:48 2017

CS61B: Lecture #37 2

Under the Hood: Allocation

References (references) are represented as integer addresses.

Due to machine's own practice.

Do not convert integers ↔ pointers,

Parts of Java runtime implemented in C, or sometimes C++, where you can.

How in C:

```
[STORAGE_SIZE]; // Allocated array
pointer = STORAGE_SIZE;

// Pointer to a block of at least N bytes of storage */
void* malloc(size_t n) { // void*: pointer to anything
    if (n < 0) return ERROR();
    int remainder = n;
    remainder = (remainder - n) & ~0x7; // Make multiple of 8
    return (void*) (store + remainder);
}
```

27:48 2017

CS61B: Lecture #37 4

Explicit Deallocating

Why? We often require explicit deallocation, because of

lack of in-time information about what is array

or of converting pointers to integers.

In-time information about unions:

```
Various {
    int;
    * Pntr;
    double;
} // X is either an int, char*, or double
```

All three problems; automatic collection possible.

Garbage collection can be somewhat faster, but rather error-prone:

Corruption

Leaks

27:48 2017

CS61B: Lecture #37 6

Lecture #37

A side excursion into nitty-gritty stuff: Storage management

27:48 2017

CS61B: Lecture #37 1

Explicit vs. Automatic Freeing

Explicit means to free dynamic storage.

When no expression in any thread can possibly be influenced by an object, it might as well not exist:

```
def cleanup():
```

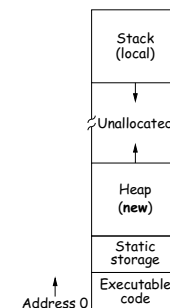
```
    c = new IntList(3, new IntList(4, null));
    tail = c;
    // Note: c is now deallocated, so no way to refer to first cell of list
```

But, Java runtime, like Scheme's, recycles the object c via garbage collection.

27:48 2017

CS61B: Lecture #37 3

Example of Storage Layout: Unix



How to turn chunks of unallocated region into heap.

Automatically for stack.

27:48 2017

CS61B: Lecture #37 5

Free List Strategies

requests generally come in multiple sizes.

requests on the free list are big enough, and one may have to split a chunk and break it up if too big.

strategies to find a chunk that fits have been used:

Best fit:

requests are processed in LIFO or FIFO order, or sorted by address.

requests are adjacent blocks.

requests are processed for **first fit** on list, **best fit** on list, or **next fit** on list to find the first-chosen chunk.

Worst fit: separate free lists for different chunk sizes.

Sticky items: A kind of segregated fit where some newly added blocks of one size are easily detected and combined with other chunks.

requests are processed to reduce **fragmentation** of memory into lots of tiny free chunks.

Free Lists

allocator grabs chunks of storage from OS and gives to program.

program uses recycled storage, when available.

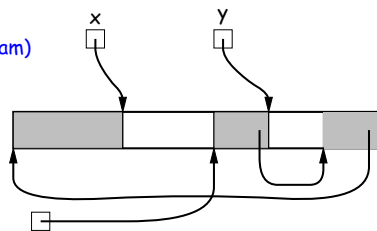
when memory is freed, added to a *free list* data structure to be reused.

program can request explicit freeing and some kinds of automatic garbage collection.

variables (pointers to program)

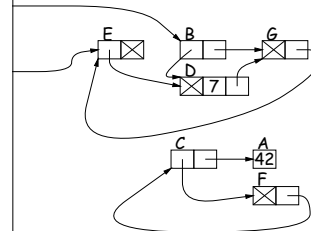
on the Heap

Free List

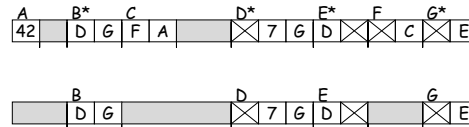


Garbage Collection: Mark and Sweep

(mathematics)



1. Traverse and **mark** the graph of objects.
2. **Sweep** through memory, freeing unmarked objects.



Copying Garbage Collection

approach: **copying garbage collection** takes time proportional to amount of active storage.

traverse the graph of active objects breadth first, copying them to a contiguous area (called "to-space").

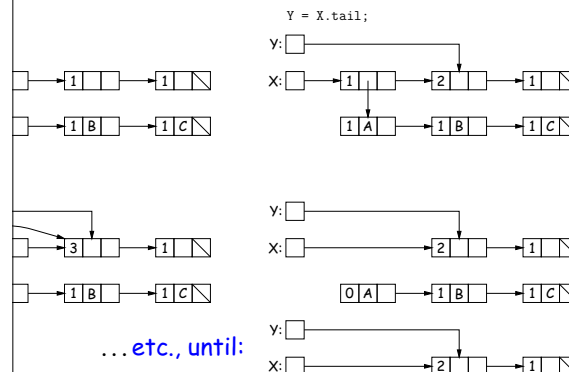
copy each object, mark it and put a *forwarding pointer* that points to where you copied it.

the next time you have to copy an already marked object, just put a forwarding pointer instead.

eventually, the space you copied from ("from-space") becomes to-space; in effect, all its objects are freed in constant time.

Garbage Collection: Reference Counting

keep track of count of number of pointers to each object. Release object when count goes to 0.



Cost of Mark-and-Sweep

garbage collection algorithms don't move any existing objects—pointers are updated.

amount of work depends on the amount of memory swept—proportional to amount of active (non-garbage) storage + amount of garbage. It's not necessarily a big hit: the garbage had to be active at some point, hence there was always some "good" processing in the time spent on each byte of garbage scanned.

Objects Die Young: Generational Collection

Objects stay active, and need not be collected.

Need to avoid copying them over and over.

Generational garbage collection schemes have two (or more) spaces: one for newly created objects (*new space*) and one for objects that have survived garbage collection (*old space*).

Old space garbage collection collects only in new space, ignores pointers pointing to old space, and moves objects to old space.

Old space has usual roots plus pointers in old space that have changed (or might be pointing to new space).

When new space full, collect all spaces.

This leads to much smaller *pause times* in interactive systems.

Copying Garbage Collection Illustrated

from:

A	B	C	D	E	F	G
42	D	G	F	A	7	G

 B: Old object
B': New object
*: marked

to:

--	--	--	--	--	--	--

forwarding pointers

from:

A	B*	C	D	E*	F	G
42	B'	G	F	A	7	G

 Copy roots

to:

B'	E'					
D	G	D				

from:

A	B*	C	D*	E*	F	G*
42	B'	G	F	A	D	7

 Copy from to-space in (b). Only D is new

to:

B'	E'	D'				
D	G	D	7	G	E	

from:

A	B*	C	D*	E*	F	G*
42	B'	G	F	A	D	7

 Copy from to-space in (c). No new objects

There's Much More

Next highlights.

Focus on how to implement these ideas efficiently.

garbage collection: What if objects scattered over many spaces?

Incremental garbage collection: where predictable pause times are important, doing a little at a time.