Linked Lists

Towards Queues



bool queue_empty(queue_t S) // O(1)

Queue Interface

// typedef _____* queue_t;

/*@requires S != NULL; @*/

- queue t queue new() // O(1) /*@ensures \result != NULL: @*/
- /*@ensures queue empty(\result); @*/;
- void enq(queue_t S, int x) // O(1) /*@requires 3 != NULL; @*/
 - /*@crisures !queue_empty(S); @*/
- int deg(queue t S) // O(1) /*@requires S != NULL;
- @*/ **@***/; /*@requires !queue empty(S);

Implementation-side type struct queue_header { int[] data;

typedef struct queue_header queue

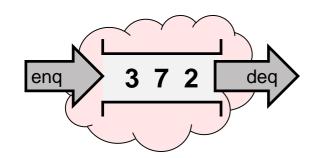
// Client type typedef queue* queue_t;

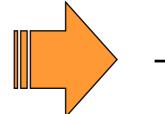
We want to **implement** the queue library

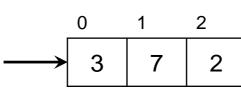
 So far we only wrote client code using its interface

A queue stores a bunch of elements of the same type—say int for a change

Idea: represent a queue as an array







But ...

> arrays have fixed length yet queues are unbounded

how would we add and remove elements?

> can we achieve the complexity goals?

create new arrays each time?

where is the front? the back?

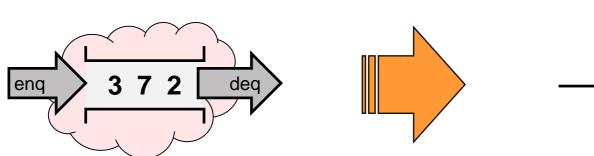
move elements around?

Toward Queues

- A queue stores a bunch of elements of the same type
 - Represent a queue as an array



- We want something like an array but where
 - we can add/remove elements at the beginning and end
 - have it grow and shrink as needed
- Some kind of disembodied array ...



Queue Interface // typedef _____* queue_t; bool queue_empty(queue_t S) // O(1) /*@requires S != NULL; @*/: queue t queue new() // O(1) /*@ensures \result != NULL: @*/ /*@ensures queue empty(\result); @*/; void enq(queue_t S, int x) // O(1) /*@requires S != NULL; @*/ /*@ensures!queue_empty(S); @*/ int deg(queue t S) // O(1) /*@requires S != NULL; @*/ /*@requires !queue empty(S);

removing an element removes a cell

7
2

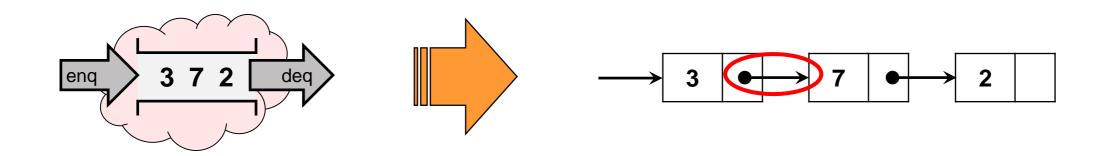
Adding an element adds a cell,

But how to reach elements after the first?

Toward Queues

- A disembodied array
 - o how to reach the elements after the first?
- Use pointers to go to the next element

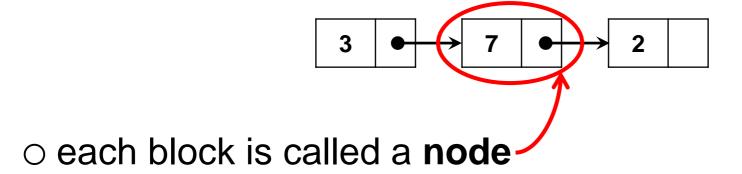
```
Queue Interface
// typedef _____* queue_t;
bool queue_empty(queue_t S)
                              // O(1)
/*@requires S != NULL;
                                 @*/:
queue t queue new()
                              // O(1)
/*@ensures \result != NULL:
                                  @*/
/*@ensures queue_empty(\result); @*/;
void enq(queue_t S, int x)
                              // O(1)
/*@requires S != NULL;
                                 @*/
/*@ensures!queue_empty(S);
                                 @*/
int deq(queue_t S)
                              // O(1)
/*@requires S != NULL;
                                 @*/
/*@requires !queue empty(S);
```



This is called a linked list

Linked Lists

Linked lists use pointers to go to the next element



Let's implement it:

- a node consists of
 - a data element an int here
 - o a pointer to the next node

```
struct list_node {
  int data;
  struct list_node* next;
};
```

The whole list is a pointer to its first node

```
struct list_node {
  int data;
  struct list_node* next;
};
```

- Linked lists are a recursive type
 - a struct list_node is defined in terms of itself
- What if we don't have this pointer?

```
a node that contains an int and
a node that contains an int and
a node that contains an int and
```

- It would take an infinite amount of memory!
- The C0 compiler disallows this
 - > recursion can only occur behind a pointer (or an array)

3

2

•

```
struct list_node {
  int data;
  struct list_node* next;
};
```

Let's make it more readable

```
typedef struct list_node list; // ADDED

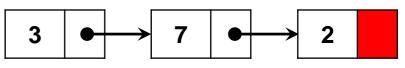
struct list_node {
  int data;
  list* next;
};
This can go before
or after the struct
```

Implementing this linked list

```
list* L = alloc(list);
L->data = 3;
L->next = alloc(list);
L->next->data = 7;
L->next->next = alloc(list);
L->next->next->data = 2;
```

- Does this help us implement queues?
 - Linked lists can be arbitrarily large or small
 - > use just the nodes we need
 - > size is not fixed like arrays
 - It's easy to insert an element at the beginning
 - > allocate a new node and point its next field to the list
 - In fact, it's easy to insert an element between any two nodes
 - > allocate a new node and move pointers around
- What about inserting an element at the end?
 - O How do we indicate the end of a linked list?

So far we just drew an empty box ...

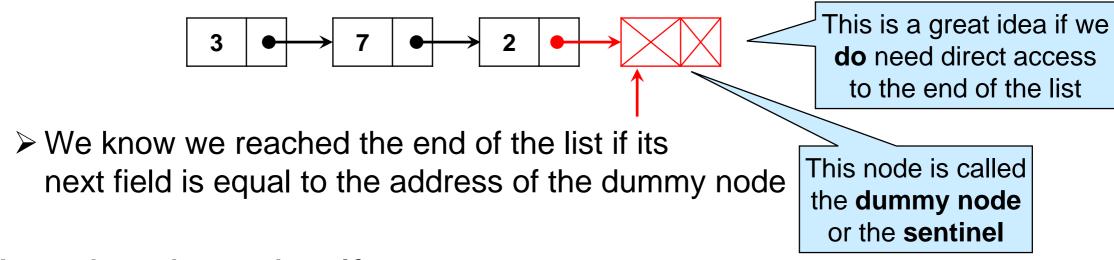


The End of a List

We need to make the pointer in the last node special

Use the NULL pointer
 This is a great idea if we don't need direct access to the end of the list

- > This is a **NULL-terminated list**
- Point it to a special node we keep track of somewhere



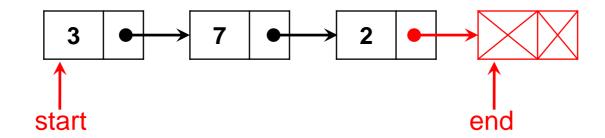
Have it point to itself

This works too, but nobody does that

List Segments

Lists with a Dummy Node

We need to keep track of two pointers



- o start: where the first node is
- o end: the address in the next field of the last node
 - > the address of the dummy node
- What's in the dummy node?
 - o some values that are not important to us
 - some number and some pointer —

These values are not special in any way:

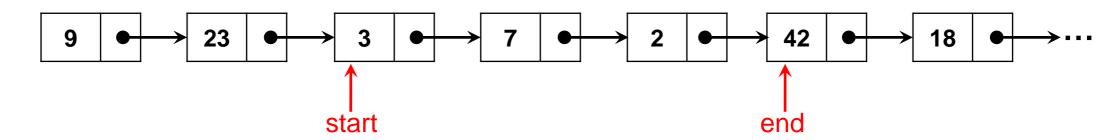
- data could be any element
- next may or may not be NULL

- we say its fields are unspecified
 - no way to test for "unspecified"

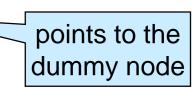
A dummy value is a value we don't care what it is

List Segments

There may be more nodes before and after



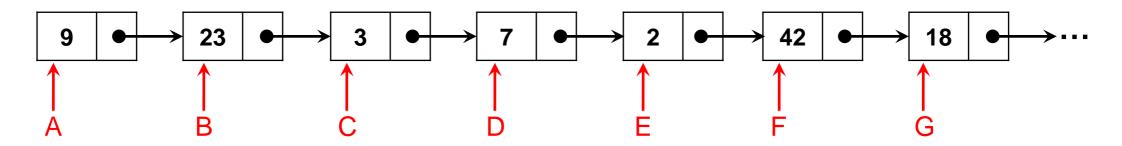
- The pair of pointers start and end identify our list exactly
 - > start is inclusive (the first node of the list)
 - > end is exclusive (one past the last node of the list)



- They identify the list segment [start, end)
 - □ here it contain values 3, 7 and 2
 - > similar to array segments A[lo, hi)

List Segments

There are many list segments in a list

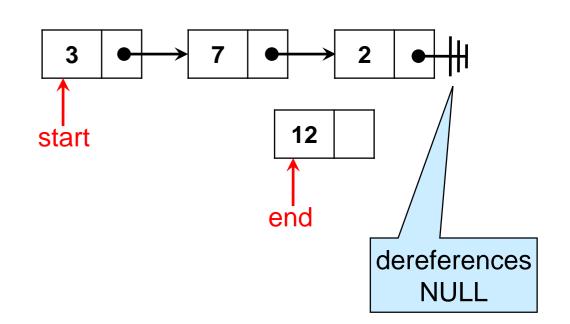


- The list segment [C, F) contains elements 3, 7, 2
 - ☐ its dummy node contains 42 and the pointer G
- The list segment [A, G) contains 9, 23, 3, 7, 2, 42
 - ☐ its dummy node contains 18 and the some pointer
- The list segment [B, D) contains 23, 3
 - □ its dummy node contains 7 and the pointer E
- The list segment [C, C) contains no elements
 - □ its dummy node contains 3 and the pointer D
 - > this is the **empty segment**
 - > any segment where start is the same as end
 - □ [A, A), [B, B), ...

```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

- We want to write a specification function that checks that two pointers start and end form a list segment
 - Follow the next pointer from start until we reach end

```
bool is_segment(list* start, list* end) {
    list* I = start;
    while (I != end) {
        I = I->next;
    }
    return true;
}
```

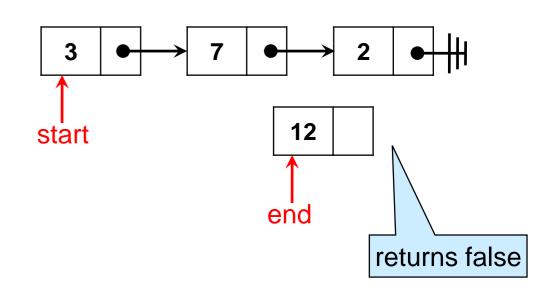


- O Does this work?
 - > the dereference I->next may not be safe
 - we need NULL-checks!
 - > we never return false



```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

- We want to write a specification function that checks that two pointers start and end form a list segment
 - Follow the next pointer from start until we reach end



- O Does this work?
 - > if there is a list segment from start to end, it will return true
 - > if it returns false, there is no list segment from start to end
- It works then ...

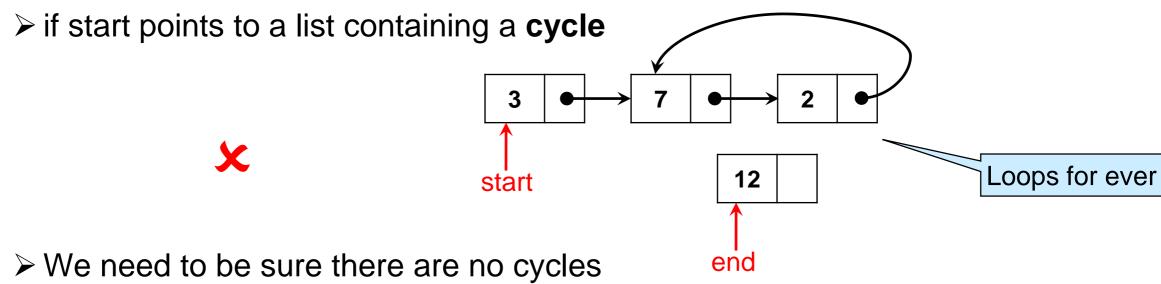
```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

A function that checks that start and end form a list segment

```
bool is_segment(list* start, list* end) {
    list* I = start;
    while (I != NULL) {
        if (I == end) return true;
        I = I->next;
    }
    return false;
}
```

- if there is a list segment from start to end, it will return true
- if it returns false, there is no list segment from start to end

Can there be no list segment but it does not return false



```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

- A function that checks that start and end form a list segment
 - We need to be sure there are no cycles

```
bool is_segment(list* start, list* end)
   //@requires is_acyclic(start); // ADDED
                                                        We will implement it later
    list* | = start;
    while (I != NULL) {
      if (I == end) return true;
      I = I - \text{next};
     return false;
O Does this work?
   > Yes!
                                                     3
                                                                          12
                                                   start
                                                                                               Fails
                                                                                           precondition
                                                                        end
```

```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

A function that checks that start and end form a list segment

```
bool is_segment(list* start, list* end)
//@requires is_acyclic(start);
{
    list* I = start;
    while (I != NULL) {
        if (I == end) return true;
        I = I->next;
    }
    return false;
}
```

O Notes:

- > returns false if start == NULL
- > or if end == NULL
 - NULL is not a pointer to a list node
 - subsumes NULL-check for both start and end

```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

- We can also write it more succinctly
 - using a for loop

```
bool is_segment(list* start, list* end)
//@requires is_acyclic(start);
{
  for (list* I = start; I != NULL; I = I->next) {
    if (I == end) return true;
  }
  return false;
}
```

All 3 versions are equivalent

recursively

Detecting Cycles

- How to check if a list is cyclic?
 Use a counter and look for overflows
 very inefficient!
 also, C0 pointers are 64 bits but ints are 32 bits

 Keep track of visited nodes somewhere

 in an array?
 in another list?

 how big to make it?

 array indices are 32 bits

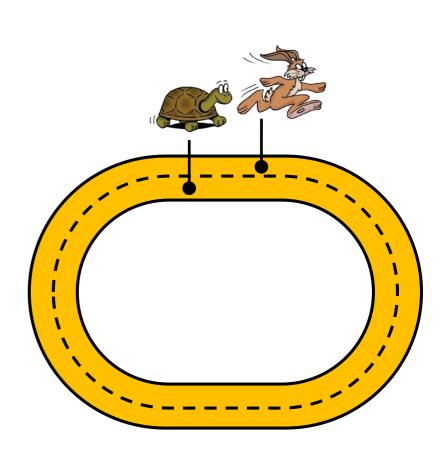
 how do we check it has no cycles?
 - Add a "visited" field to the nodes (a boolean)
- we need to know the list is acyclic to initialize it to false!
 - O What then?

Detecting Cycles

The tortoise and hare algorithm by this dude

- Traverse the list using two pointers
 - > the tortoise starts at the beginning and moves by 1 step
 - > the hare starts just ahead of the tortoise and moves by 2 steps
- If the hare ever overtakes the tortoise, there is a cycle

```
bool is_acyclic(list* start) {
 if (start == NULL) return true;
 list* t = start;
                       // tortoise
 list* h = start->next; // hare
 while (h != t) {
  if (h == NULL || h->next == NULL) return true;
  //@assert t != NULL; // hare hits NULL quicker
               // tortoise moves by 1 step
  t = t->next;
  h = h->next->next; // hare moves by 2 steps
//@assert h == t; // hare has overtaken tortoise
 return false;
```





Robert W. Floyd

Detecting Cycles

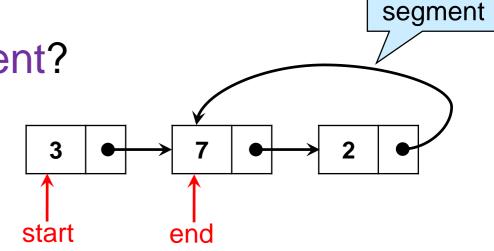
The tortoise and hare algorithm

- Returns
 - > true if there is no cycle
 - ➤ false if there is a cycle

Does it fix our problem with is_segment?

- Too aggressive
- Exercise: fix it!

Hint: you need to account for end

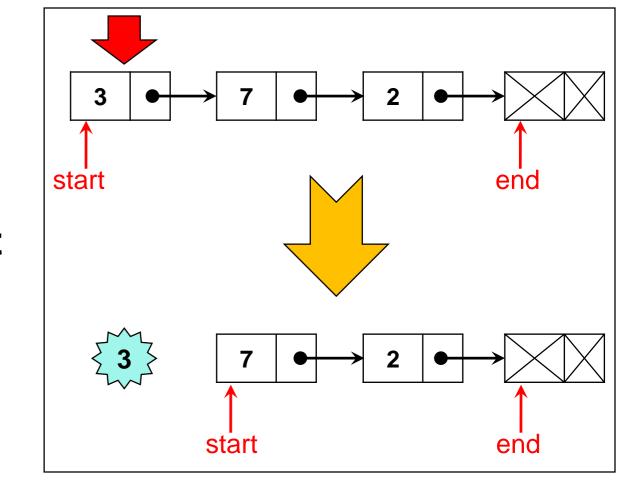


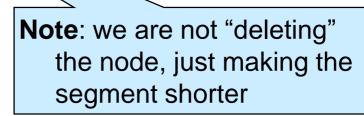
cycle after

Manipulating List Segments

Deleting an Element

- How do we remove the node at the beginning of a non-empty list segment [start, end)?
 - > and return the value in there
 - 1. grab the value in the start node
 - 2. move start to point to the next node
 - 3. return the value
 - int x = start->data;
 start = start->next;
 return x;
 - Complexity: O(1)



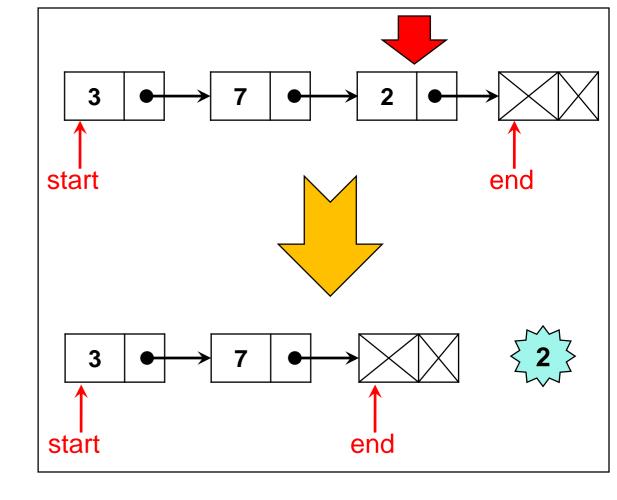


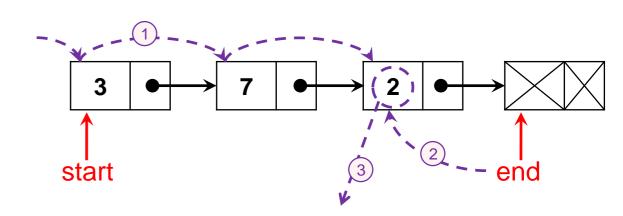
Deleting an Element

- How do we remove the last node of a non-empty list segment [start, end)?
 - > and return the value in there
 - we must go from start
 - end is one node too far
 - 1. follow next until just before end
 - 2. move end to that node
 - 3. return its value

Complexity: O(n)





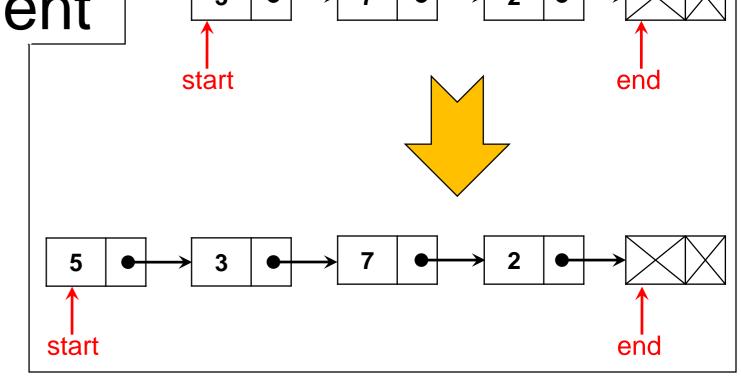


Notes:

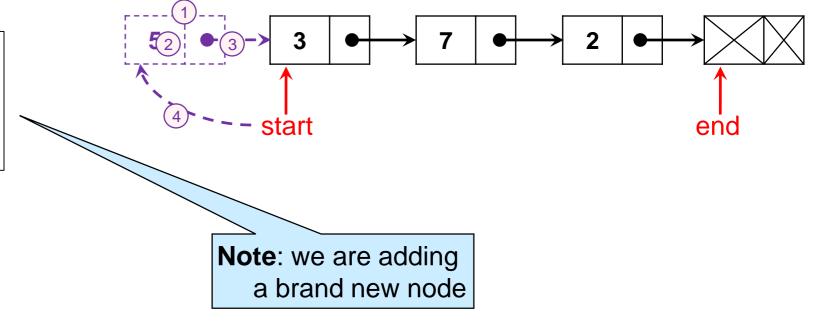
- The old last node becomes the new dummy node
- We are not "deleting" anything, just making the segment shorter

Inserting an Element

• How do we add a node at the beginning of a list segment [start, end)?



- 1. create a new node
- 2. set its data field to the value to add
- 3. set its next field to start
- 4. set start to it
- 1 list* l = alloc(list);
 2 l->data = x;
 3 l->next = start;
 4 start = l;
 - Complexity: O(1)



Inserting an Element

• How do we add a node as the last node of a list segment [start, end)?

3 • 7 • 2 • 5 • dend

end

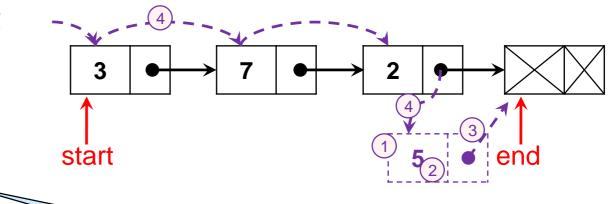
- 1. create a new node
- 2. set its data field to the value to add

Expensive!

- 3. set its next field to end
- 4. point the old last node to it

```
list* new_last = alloc(list);
new_last->data = x;
new_last->next = end;
list* I = start;
while (I->next!= end)
I = I->next;
I->next = new_last;
```

○ Complexity: O(n)



start

Note: we are adding a new last node, but we modify the next pointer of the old last node

Inserting an Element

• How do we add a node as the last node of a list segment [start, end)?

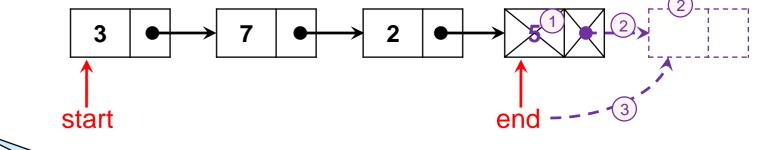
3 • 7 • 2 • 5 • ond

end

start

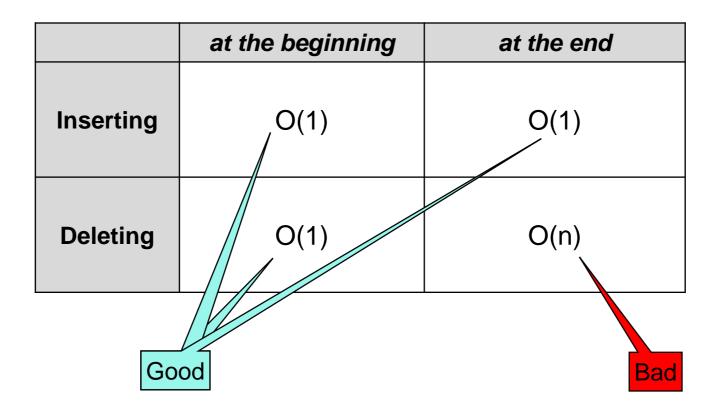
- Can we do better?
 - 1. set the data field of end to the value to add
 - 2. set its next field to a new dummy node
 - 3. set end to it
- 1 end->data = x; 2 end->next = alloc(list);
- end = end->next;





Note: we are using the old dummy node as the new last node, and creating a new dummy

Summary



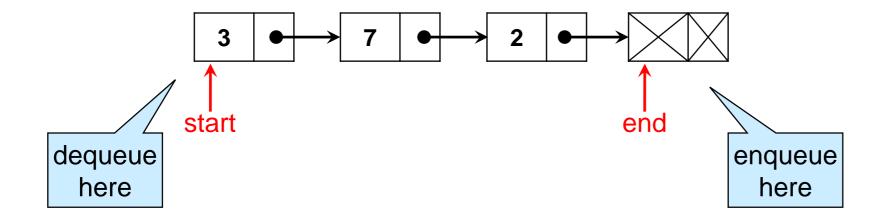
 We will use this as a guide when implementing queues (and stacks) to achieve their complexity goals

Implementing Queues

- Implementing queues
 - We add and remove from opposite ends
 - Cost must be O(1)

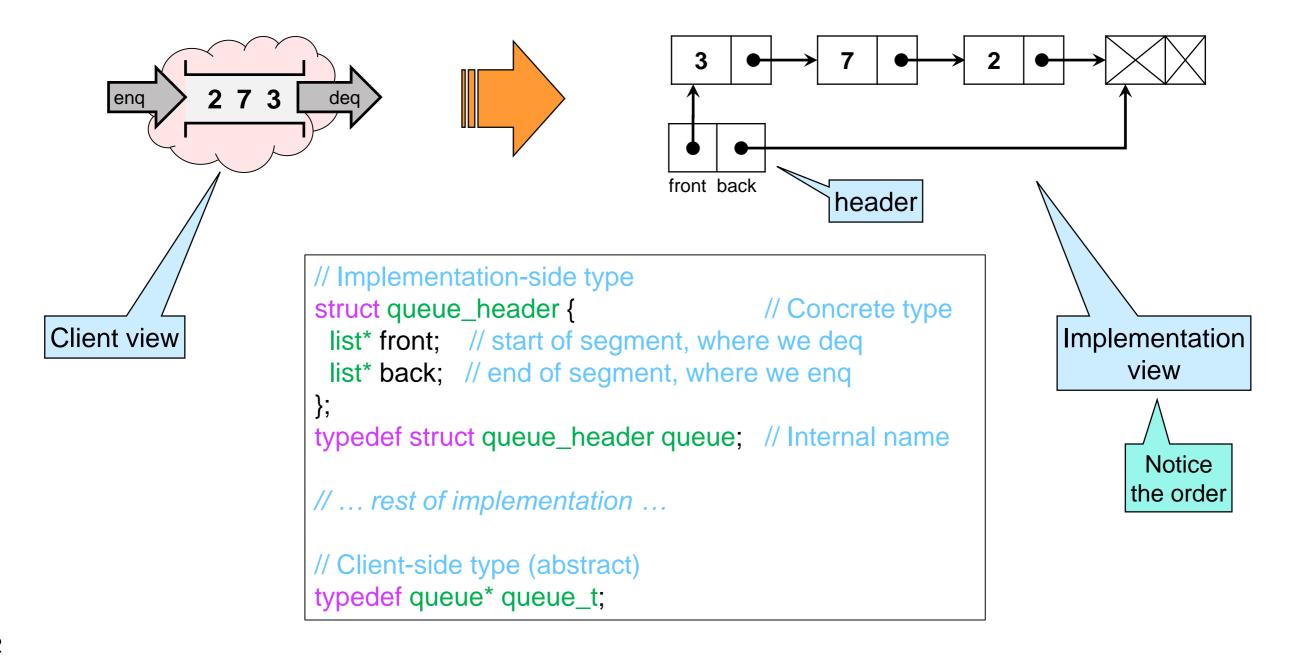
at the end		he beginning	at tl	
	O(1)	O(1)		Inserting
)	O(n)	O(1)	(Deleting
•				

- The front of the queue is the start of the segment
 - because that's where we remove elements from
 - > choosing the end would give deq cost O(n)
- The back of the queue is the end of the segment
 - > the dummy node



```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

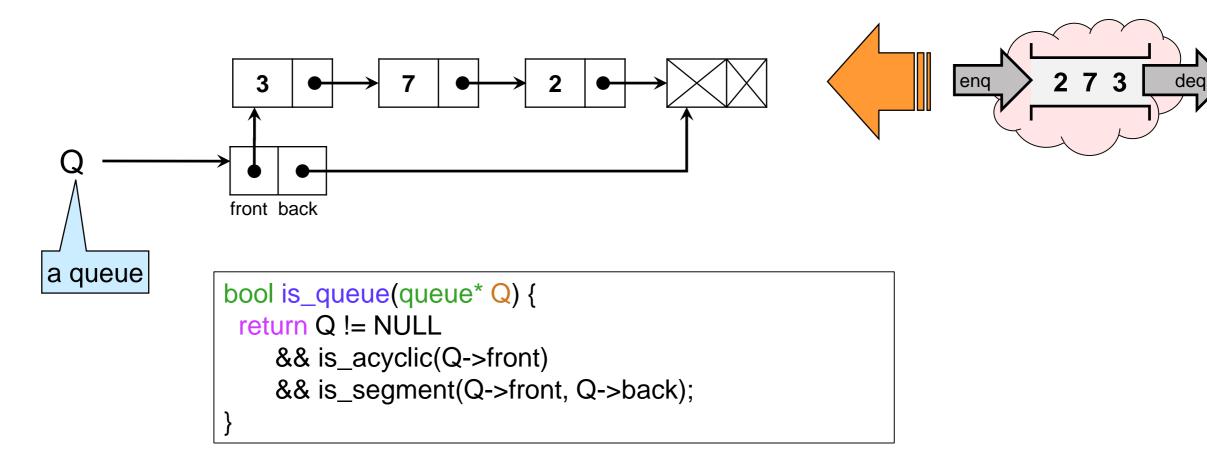
- The front of the queue is the start of the segment
- The back of the queue is the end of the segment



```
typedef struct list_node list;
struct list_node {
   int data;
   list* next;
};

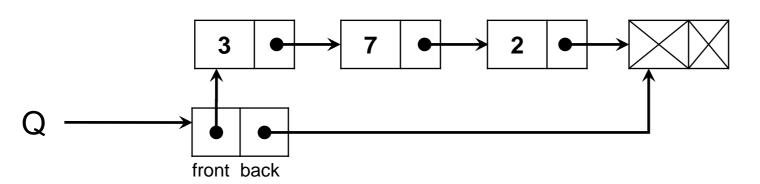
struct queue_header {
   list* front;
   list* back;
};
typedef struct queue_header queue;
```

- Internally, queues are values of type queue*
 - o must be non-NULL
 - front and back fields must bracket a valid list segment



Next we implement the operations exported by the interface

```
Queue Interface
// typedef _____* queue_t;
bool queue_empty(queue_t S) // O(1)
/*@requires S != NULL;
                                 @*/:
                             // O(1)
queue_t queue_new()
/*@ensures \result != NULL;
                                 @*/
/*@ensures queue_empty(\result); @*/;
void eng(queue_t S, int x)
                             // O(1)
/*@requires S != NULL;
                                 @*/
                                @*/;
/*@ensures!queue_empty(S);
int deq(queue_t S)
                             // O(1)
/*@requires S != NULL;
                                 @*/
/*@requires !queue_empty(S);
```



```
typedef struct list_node list;
struct list_node {
   int data;
   list* next;
};

struct queue_header {
   list* front;
   list* back;
};
typedef struct queue_header queue;
```

- Enqueuing
 - o add at the back

```
void enq(queue* Q, int x)
//@requires is_queue(Q);
//@ensures is_queue(Q);
//@ensures !queue_empty(Q);
{
    Q->back->data = x;
    Q->back->next = alloc(list);
    Q->back = Q->back->next;
}
```

- Dequeueing
 - o remove from the front

```
int deq(queue* Q)
//@requires is_queue(Q);
//@requires !queue_empty(Q);
//@ensures is_queue(Q);
{
  int x = Q->front->data;
  Q->front = Q->front->next;
  return x;
}
```

- This is the code we wrote earlier with
 - start changed to Q->front
 - ➤ end changed to Q->back

Cost is **O(1)**

Queues as List Segments

```
typedef struct list_node list;
struct list_node {
    int data;
    list* next;
};

struct queue_header {
    list* front;
    list* back;
};

typedef struct queue_header queue;
```

- Q front back
- The empty queue
 empty segment has start equal to end

```
bool queue_empty(queue* Q)
//@requires is_queue(Q);
{
   return Q->front == Q->back;
}

Cost is O(1)
```

Creating a queueo we create an empty queue

```
queue* queue_new()
//@ensures is_queue(\result);
//@ensures queue_empty(\result);
{
   queue* Q = alloc(queue);
   Q->front = alloc(list);
   Q->back = Q->front;
   return Q;
}
```

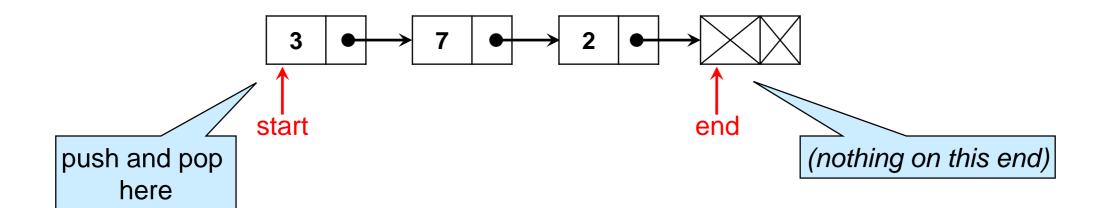
Implementing Stacks

Stacks as List Segments

- Implementing stacks
 - We add and remove from the same end
 - Cost must be O(1)

	at the beginning		at the end	
Inserting		O(1)		O(1)
Deleting		O(1)		O(n)

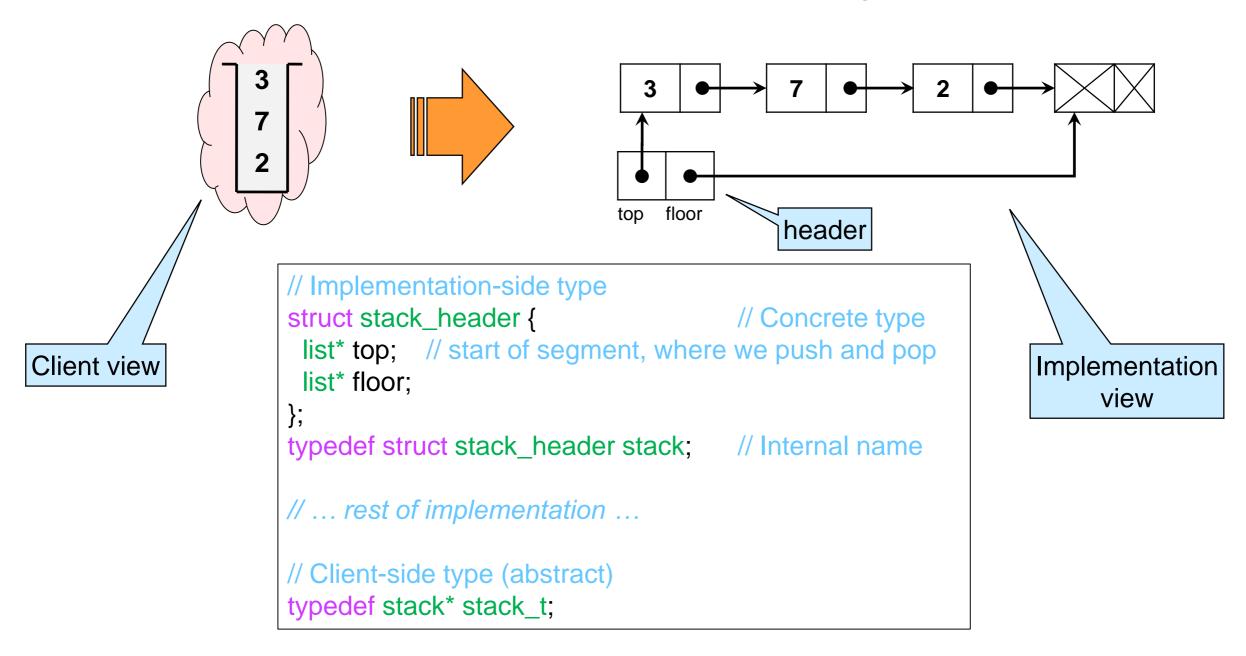
- The top of the stack is the start of the segment
 - because that's where we add and remove elements
 - > choosing the end would give pop cost O(n)
- The floor of the stack is the end of the segment
 - > the dummy node



Stack as List Segments

```
typedef struct list_node list;
struct list_node {
  int data;
  list* next;
};
```

The top of the stack is the start of the segment



The representation invariant is_stack is just like is_queue

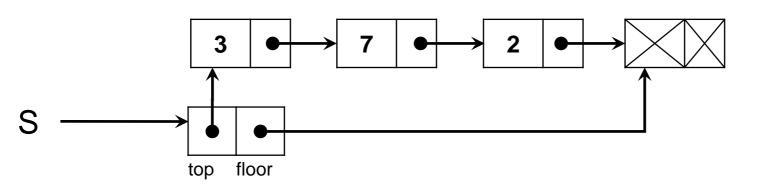
Stacks as List Segments

Next we implement the operations exported by the interface

```
Stack Interface
// typedef _____* stack_t;
bool stack_empty(stack_t S) // O(1)
/*@requires S != NULL;
                                 @*/:
stack_t stack_new()
                             // O(1)
/*@ensures \result != NULL;
                                 @*/
/*@ensures stack_empty(\result); @*/;
void push(stack t S, int x)
                            // O(1)
/*@requires S != NULL;
                                 @*/
/*@ensures!stack_empty(S);
                                 @*/:
int pop(stack_t S)
                            // O(1)
/*@requires S != NULL;
                                 @*/
                                 @*/;
/*@requires !stack_empty(S);
```

Also updated to int elements

Stacks as List Segments



```
typedef struct list_node list;
    struct list_node {
        int data;
        list* next;
     };

struct stack_header {
        list* top;
        list* floor;
     };

typedef struct stack_header stack;
```

```
bool stack_empty(stack* S)
//@requires is_stack(S);
{
   return S->top == S->floor;
}
```

Same code we wrote for queues with front/back replaced with top/floor

```
int pop(stack* S)
//@requires is_stack(S);
//@requires !stack_empty(S);
//@ensures is_stack(S);
{
  int x = S->top->data;
  S->top = S->top->next;
  return x;
}
```

```
stack* stack_new()
//@ensures is_stack(\result);
//@ensures stack_empty(\result);
{
    stack* S = alloc(stack);
    S->top = alloc(list);
    S->floor = S->top;
    return S;
}
```

```
void push(stack* S, int x)
//@requires is_stack(S);
//@ensures is_stack(S);
//@ensures !stack_empty(S);
{

list* I = alloc(list);
I->data = x;
I->next = S->top
S->top = I;
}
```



Code we wrote earlier with start replaced with S->top

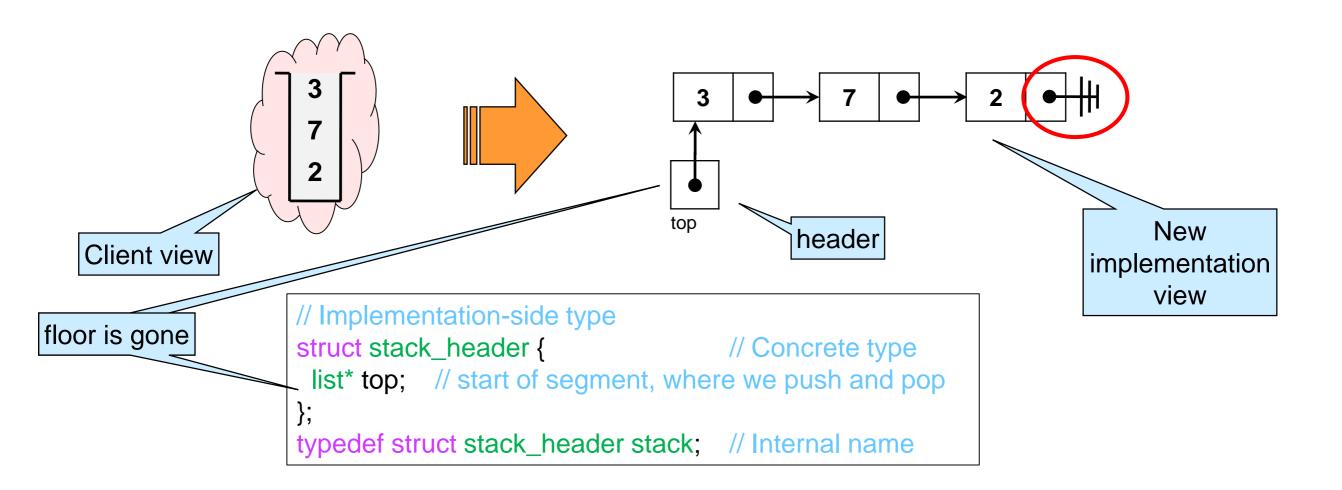
Another Implementation of Stacks

This is a great idea if we

don't need direct access

to the end of the list

- The floor field goes mostly unused
 only to check that a stack is empty
- We can get rid of it ...
 - o ... if we represent stacks as NULL-terminated lists



Another Implementation of Stacks

- Valid stacks are
 - non-NULL and
 - the top field is a NULL-terminated list
 ➤ i.e., is acyclic

```
bool is_stack(stack* S) {
  return S != NULL
    && is_acyclic(S->top);
}
```

 The empty stack has NULL in the top field

```
S \longrightarrow \underbrace{\bullet}_{top}
```

```
bool stack_empty(stack* S)
//@requires is_stack(S);
{
   return S->top == NULL;
}
```

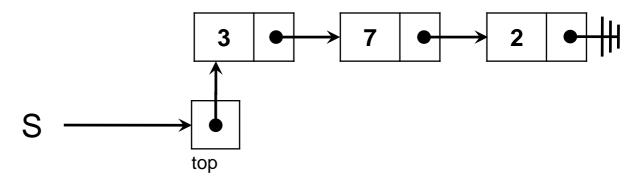
```
stack* stack_new()
//@ensures is_stack(\result);
//@ensures stack_empty(\result);
{
    stack* S = alloc(stack);
    S->top = NULL:
    return S;
}
```

Nothing else changes!

Stacks without Headers

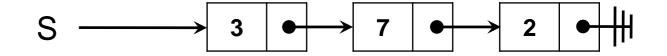
Since the header contains just one field,

```
struct stack_header {
   list* top;
};
typedef struct stack_header stack;
```



o why not get rid of it?

```
typedef list* stack;
```



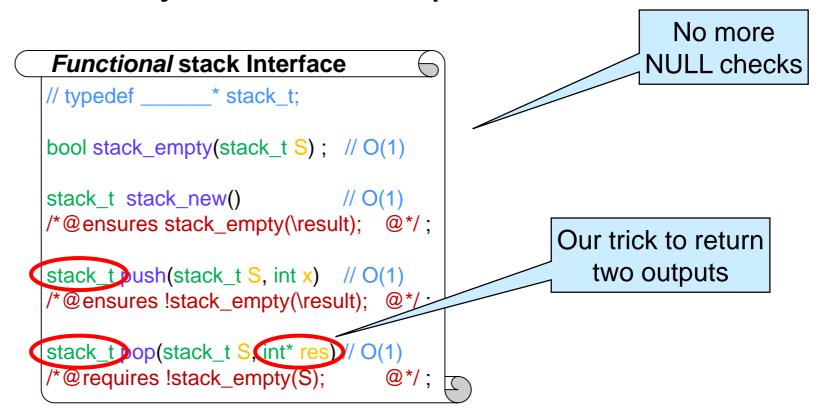
- push and pop are now incorrect
 - □ they modify the local stack variable but not the caller's
 - □ aliasing!
- it breaks the interface: NULL is now the empty stack



Stacks without Headers

typedef list* stack; S → 3 → 7 → 2 → H

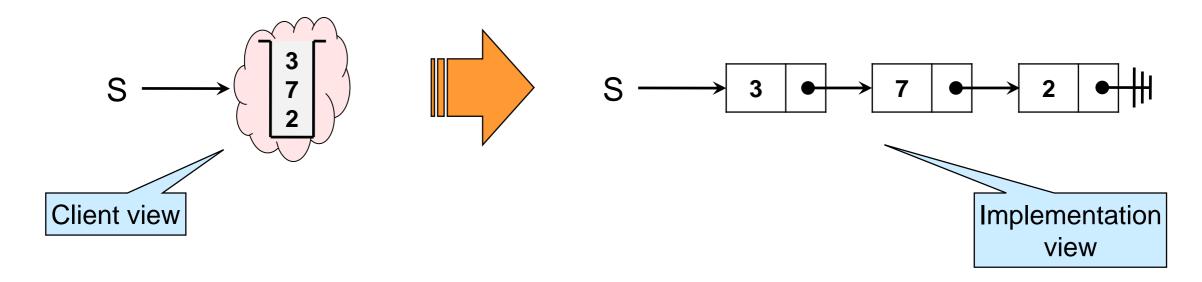
But we're fine if we always return the updated stack



- Functions transform an input stack into an output stack
 - > this is a functional interface

Functional Stacks

• How to create this stack?



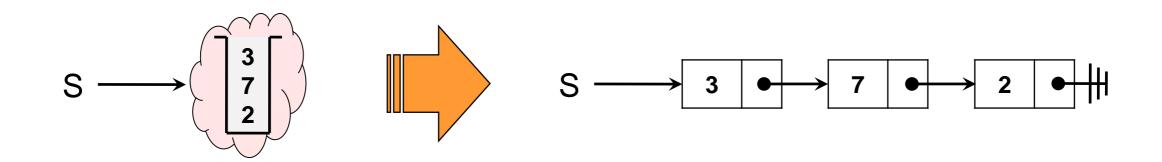
```
stack_t S = stack_new();
S = push(S, 2);
S = push(S, 7);
S = push(S, 3);
```

> equivalently

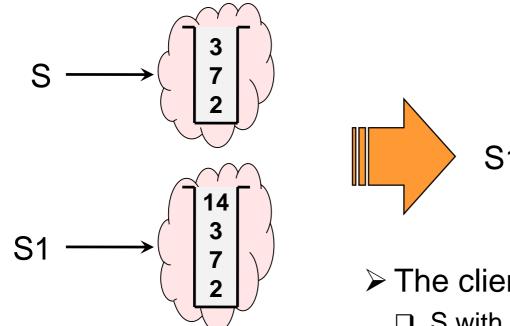
```
stack_t S = push(push(push(stack_new(), 2), 7), 3);
```

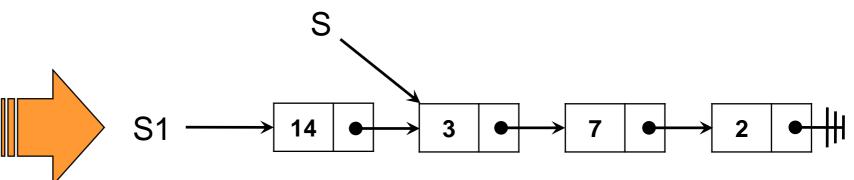
□ but harder to read

Functional Stacks



What if now we do





- > The client has two stacks
 - □ S with 3, 7, 2
 - □ S1 with 14, 3, 7, 2
- ➤ In the implementation, they **share** a suffix
 - □ the linked list 3, 7, 2 is shared

- A functional stack library supports sharing list suffixes
 - This takes up much less space than our earlier implementation!
 - The client has no idea

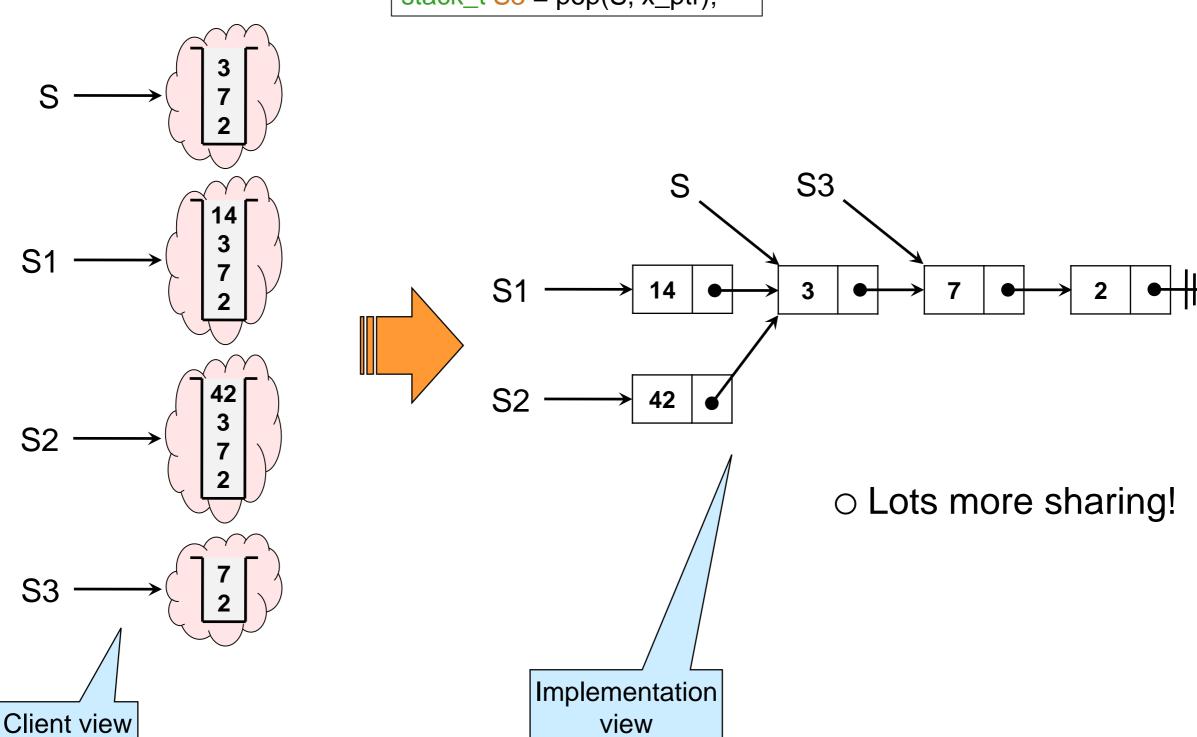
• What if we now do this?

```
stack_t S2 = push(S, 42);
stack_t S3 = pop(S, x_ptr);

The variable S is still around
```

What if we now do

stack_t S2 = push(S, 42); stack_t S3 = pop(S, x_ptr);



- If sharing is so great, why don't our libraries always use it?
 - It takes a change of mindset
 - > using functions that don't modify data structures in place
 - A lot of code we write uses one instance of a data structure
 - ➤ So what? Sharing wouldn't hurt anyway
 - □ Good point
 - It doesn't work for all data structures
 - > Try it on queues!
- Functional programming languages rely heavily on sharing

Wrap Up

What have we done?

- We introduced linked lists
 and two common ways to use them
 - NULL-terminated linked lists
 - list segments
- We learned about list manipulations and their complexity
- We used them to implement stacks and queues
- We talked about sharing

Linked Lists vs. Arrays

• How do they compare?

	Arrays (unsorted)	Linked lists
Pros	O(1) accessbuilt-in	 self-resizing O(1) insertion* O(1) deletion*
Cons	fixed sizeO(n) insertion	Given the right pointersO(n) accessno special syntax

- Question to help decide which one to use:
 - Ocan we anticipate the size we need?
 - O Do they allow us to achieve our target complexity?