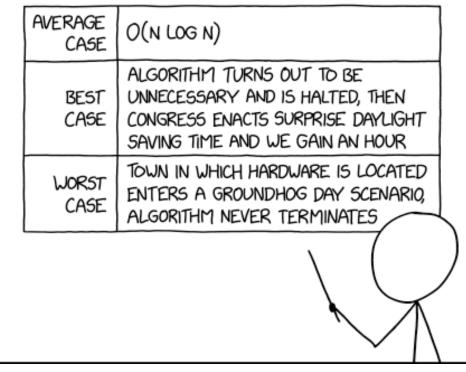
CSE 331 Spring 2025

Arrays I

RESULTS OF ALGORITHM COMPLEXITY ANALYSIS:



xkcd #2939

Matt Wang

& Ali, Alice, Andrew, Anmol, Antonio, Connor, Edison, Helena, Jonathan, Katherine, Lauren, Lawrence, Mayee, Omar, Riva, Saan, and Yusong

Administrivia (05/23)

- HW8 is out!
 - beware: coding portion has a good chunk of math!
 - but also: very fun app :)
- Holiday on Monday...
 - no class
 - no office hours
 - some (reduced) Ed activity
- Implication: please start HW8 early!!
- next Fri: a bit on the final exam

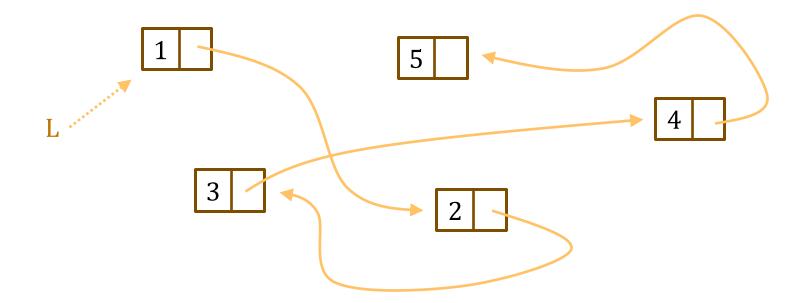
at : (List, N) → Z at(nil, n) := undefined at(x :: L, 0) := x at(x :: L, n+1) := at(L, n)

• Retrieve an element of the list by <u>index</u>

– use "L[j]" as an abbreviation for at(j, L)

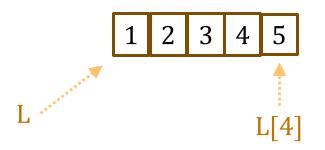
• Not an efficient operation on lists...

Linked Lists in Memory



- Must follow the "next" pointers to find elements
 - at(L, n) is an O(n) operation
 - no faster way to do this

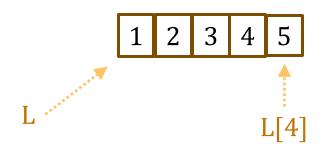
Faster Implementation of at



- Alternative: store the elements next to each other
 - can find the $n\mbox{-th}$ entry by arithmetic:

location of L[4] = (location of L) + 4 * sizeof(data)

- Resulting data structure is an **array**
 - consider: arrays can be an implementation of the List ADT



- Resulting data structure is an array
- Efficient to read L[i]
- Inefficient to...
 - insert elements anywhere but the end
 - write operations with an immutable ADT
 - trees can do all of this in $O(\log n)$ time

- Easily access both L[0] and L[n-1], where n = len(L)
 - can process a list in either direction
- "With great power, comes great responsibility"

— the Peter Parker Principle

- Whenever we write "A[j]", we must check $0 \le j < n$
 - new bug just dropped!

with list, we only need to worry about nil and non-nil once we know L is non-nil, we know L.hd exists

– TypeScript will not help us with this!

type checker does catch "could be nil" bugs, but not this

sum-acc(nil, r) := r sum-acc(x :: L, r) := sum-acc(L, x + r)

Tail recursive version is a loop

```
const sum = (S: List<bigint>): bigint => {
  let r = 0;
  // Inv: sum(S<sub>0</sub>) = r + sum(S)
  while (S.kind !== "nil") {
    r = S.hd + r;
    S = S.tl;
  }
  return r;
};
```

Change to a version that uses indexes...

• Change to using an array and accessing by index

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  // Inv: ...
  while (j !== S.length) { // ... S.kind !== "nil"
    r = S[j] + r; // ... r = S.hd + r
    j = j + 1; // ... S = S.tl
  }
  return r;
};
Note that S is no longer changing
```

Sum Array by Index: compared to sum-acc

```
\begin{aligned} \text{sum-acc} : (\text{List}, \mathbb{N}, \mathbb{Z}) &\to \mathbb{Z} \\ \text{sum-acc}(S, j, r) & := r & \text{if } j = \text{len}(S) \\ \text{sum-acc}(S, j, r) & := \text{sum-acc}(S, j+1, S[j] + r) & \text{if } j \neq \text{len}(S) \end{aligned}
```

• Change to using an array and accessing by index

```
const sum = (S: Array<bigint>): bigint => {
    let r = 0;
    let j = 0;
    // Inv: ...
    while (j !== S.length) {
        r = S[j] + r;
        j = j + 1;
    }
    return r;
};
```

• Use indexes to refer to a section of a list (a "sublist"):

sublist : (List< \mathbb{Z} >, N, Z) → List< \mathbb{Z} >sublist(L, i, j):= nilif j < i</td>sublist(L, i, j):= L[i] :: sublist(L, i + 1, j)

- Useful for *reasoning* about lists and indexes
- This includes <u>both</u> L[i] and L[j]

$$\begin{aligned} \text{sublist}(L, 0, 2) &= L[0] :: \text{sublist}(L, 1, 2) & \text{def} \\ &= L[0] :: L[1] :: \text{sublist}(L, 2, 2) & \text{def} \\ &= L[0] :: L[1] :: L[2] :: \text{sublist}(L, 3, 2) & \text{def} \\ &= L[0] :: L[1] :: L[2] :: \text{nil} & \text{def} \\ &= [L[0], L[1], L[2]] \end{aligned}$$

lef of sublist (since $0 \le 2$) lef of sublist (since $1 \le 2$) lef of sublist (since $2 \le 2$)

ef of sublist (since 3 < 2)

• Use indexes to refer to a section of a list (a "sublist"):

sublist : (List $< \mathbb{Z} >$, \mathbb{N} , \mathbb{Z}) \rightarrow List $< \mathbb{Z} >$

sublist(L, i, j):= nilif j < isublist(L, i, j):= L[i] :: sublist(L, i + 1, j)if $i \le j$

• The sublist is empty when the range is empty

sublist(L, 3, 2) = nil

- weird-looking example that comes up a lot:

sublist(L, 0, -1) = nil

not an array out of bonds error! (this is math, not Java)

sublist : (List< \mathbb{Z} >, \mathbb{N} , \mathbb{Z}) \rightarrow List< \mathbb{Z} >sublist(L, i, j):= nilif j < i</td>sublist(L, i, j):= L[i] :: sublist(L, i + 1, j)if i ≤ j

- Will use "L[i .. j]" as shorthand for "sublist(L, i, j)"
 again, using an operator for most common operations
- Some useful facts about sublists:

L = L[0 .. len(L)-1]

L[i ... j] = L[i ... k] + L[k+1 ... j] for any k with $i - 1 \le k \le j$ (and $0 \le i \le j < n$)

Sum Array by Index: sum-acc, in math

sum-acc(S, j, r):= rif j = len(S)sum-acc(S, j, r):= sum-acc(S, j+1, S[j] + r)if $j \neq len(S)$

Change to using an array and accessing by index

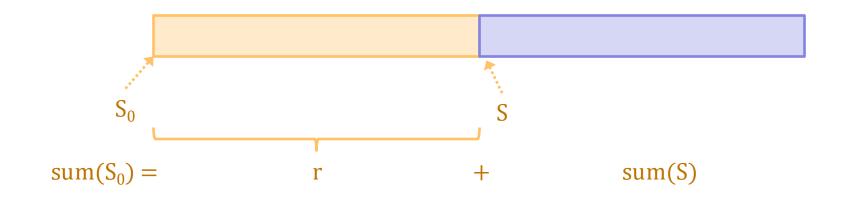
```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  // Inv: ... ?? ...
  while (j != S.length) {
    r = S[j] + r;
    j = j + 1;
  }
  return r;
  Still need to fill in Inv...
};
```

Recall: Sum List With a Loop, with Invariant

Tail recursive version is a loop

```
const sum = (S: List<bigint>): bigint => {
    let r = 0;
    // Inv: sum(S_0) = r + sum(S)
    while (S.kind !== "nil") {
        r = S.hd + r;
        S = S.tl;
    }
    return r;
};
Inv says sum(S_0) is r plus sum of rest (S)
Not the most explicit way of explaining "r"...
```

Visual Intuition for Sum List Loop Invariant



- "r" contains sum of the part of the list seen so far
- Can explain this more simply with indexes...
 - no longer need to move S

Visual Intuition for Index & Sublist Loop Invariant



- Sum is the part in "r" plus the part left in S[j .. n-1]
- What sum is in "r"?

```
\mathbf{r} = \mathbf{sum}(\mathbf{S}[0 \dots \mathbf{j} - 1])
```

- we can use just this as our invariant! (it's all we need)

Sum of an Array: Loop Invariant

• Array version uses access by index

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  // Inv: r = sum(S[0 .. j-1])
  while (j != S.length) {
    r = S[j] + r;
    j = j + 1;
  }
  return r;
};
```

Are we sure this is right? Let's think it through...

Sum of an Array Floyd Logic: Initialization

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  \{\{r = 0 \text{ and } j = 0\}\}
                                     Does Inv hold initially?
  {{ Inv: r = sum(S[0 .. j-1]) }}
  while (j != S.length) {
    r = S[j] + r;
    j = j + 1;
  }
                              sum(S[0 .. j-1])
  return r;
                               = sum(S[0 ... -1]) since j = 0
};
                               = sum([])
                               = 0
                                                   def of sum
                               = r
```

Sum of an Array Floyd Logic: Postcondition

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  {{ Inv: r = sum(S[0 .. j-1]) }}
  while (j != S.length) {
     r = S[j] + r;
    j = j + 1;
  }
  \{\{r = sum(S[0 .. j-1]) \text{ and } j = len(S) \}\}
                                           Does the postcondition hold?
  \{\{r = sum(S)\}\}
  return r;
};
                      r = sum(S[0..j-1])
                        = sum(S[0 .. len(S)-1]) since j = len(S)
                        = sum(S)
```

Sum of an Array Floyd Logic: Loop Body (1/4)

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  {{ Inv: r = sum(S[0 .. j-1]) }}
  while (j != S.length) {
     {{ r = sum(S[0 .. j-1]) and j \neq len(S) }}
     \mathbf{r} = \mathbf{S}[\mathbf{j}] + \mathbf{r};
     j = j + 1;
     \{\{r = sum(S[0..j-1])\}\}
   }
  return r;
};
```

Sum of an Array Floyd Logic: Loop Body (2/4)

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  {{ Inv: r = sum(S[0 .. j-1]) }}
  while (j != S.length) {
     {{ r = sum(S[0 .. j-1]) and j \neq len(S) }}
     \mathbf{r} = \mathbf{S}[\mathbf{j}] + \mathbf{r};
 {{ r = sum(S[0..j]) }}
j = j + 1;
    \{\{r = sum(S[0.j-1])\}\}
  }
  return r;
};
```

Sum of an Array Floyd Logic: Loop Body (3/4)

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  {{ Inv: r = sum(S[0 .. j-1]) }}
  while (j != S.length) {
     \{\{ r = sum(S[0 .. j-1]) and j \neq len(S) \}\}
    {{ S[j] + r = sum(S[0.j]) }}
 {{ sum(S[0..j]) }}
{{ r=sum(S[0..j]) }}
     j = j + 1;
     \{\{r = sum(S[0 .. j-1])\}\}
  }
  return r;
};
```

Sum of an Array Floyd Logic: Loop Body (4/4)

```
const sum = (S: Array<bigint>): bigint => {
  let r = 0;
  let j = 0;
  {{ Inv: r = sum(S[0 .. j-1]) }}
  while (j != S.length) {
    {{ r = sum(S[0 .. j-1]) and j \neq len(S) }} Is this valid?
     \{\{S[j] + r = sum(S[0., j])\}\}
     r = S[j] + r;
    \{\{r = sum(S[0 .. j])\}\}
     j = j + 1;
    \{\{r = sum(S[0 .. j-1])\}\}
  }
  return r;
};
```

Proving Loop Body "Preservation" (1/3)

```
{{ r = sum(S[0 .. j-1]) and j \neq len(S) }}
{{ S[j] + r = sum(S[0 .. j]) }}
```

```
S[j] + r
= S[j] + sum(S[0 .. j-1]) since r = sum(S[0 .. j-1])
= sum(S[0 .. j-1]) + S[j]
= sum(S[0 .. j-1]) + sum([S[j]]) def of sum
= sum(S[0 .. j-1]) + sum(S[j .. j])
= ...
= sum(S[0 .. j])
```

Proving Loop Body "Preservation" (2/3)

```
\{\{r = sum(S[0 .. j-1]) \text{ and } j \neq len(S) \}\}
   \{\{ S[j] + r = sum(S[0.j]) \}\}
S[i] + r
 = S[j] + sum(S[0..j-1])
                                          since r = sum(S[0 ... j-1])
 = sum(S[0 .. j-1]) + S[j]
 = sum(S[0 .. j-1]) + sum([S[j]])
                                          def of sum
 = sum(S[0 .. j-1]) + sum(S[j .. j])
 = ....
 = sum(S[0 .. j-1] # S[j .. j])
 = sum(S[0 ... j])
```

- We saw that len(L # R) = len(L) + len(R)
- **Does** sum(L # R) = sum(L) + sum(R)?
 - Yes! Very similar proof by structural induction. (Call this Lemma 3)

Proving Loop Body "Preservation" (3/3)

```
{{ r = sum(S[0 .. j-1]) and j \neq len(S) }}
{{ S[j] + r = sum(S[0 .. j]) }}
```

```
\begin{split} S[j] + r \\ &= S[j] + sum(S[0 .. j-1]) & since r = sum(S[0 .. j-1]) \\ &= sum(S[0 .. j-1]) + S[j] \\ &= sum(S[0 .. j-1]) + sum([S[j]]) & def of sum \\ &= sum(S[0 .. j-1]) + sum(S[j .. j]) \\ &= sum(S[0 .. j-1] + S[j .. j]) & by Lemma 3 \\ &= sum(S[0 .. j]) \end{split}
```

(The need to reason by induction comes up all the time.)

contains(nil, y):= falsecontains(x :: L, y):= trueif x = ycontains(x :: L, y):= contains(L, y)if $x \neq y$

Tail-recursive definition

```
const contains =
   (S: List<bigint>, y: bigint): bigint => {
    // Inv: contains(S<sub>0</sub>, y) = contains(S, y)
   while (S.kind !== "nil" && S.hd !== y) {
      S = S.tl;
    }
   return S.kind !== "nil"; // implies S.hd === y
};
```

Change to a version that uses indexes...

contains(nil, y):= falsecontains(x :: L, y):= trueif x = ycontains(x :: L, y):= contains(L, y)if $x \neq y$

• Change to using an array and accessing by index

```
const contains =
  (S: Array<bigint>, y: bigint): bigint => {
  let j = 0;
  // Inv: ...
  while (j !== S.length && S[j] !== y) {
    j = j + 1;
  }
    s.hd with s changing becomes
  return j !== S.length;    S[j] with j changing
};
What is the invariant now?
```

Linear Search of an Array: Loop Invariant

contains(nil, y):= falsecontains(x :: L, y):= trueif x = ycontains(x :: L, y):= contains(L, y)if $x \neq y$

Change to using an array and accessing by index

```
const contains =
   (S: Array<bigint>, y: bigint): bigint => {
   let j = 0;
   // Inv: contains(S, y) = contains(S[j .. n-1], y)
   while (j !== S.length && S[j] !== y) {
      j = j + 1;
   }
   return j !== S.length; Can we explain this better?
};
```


- What do we know about the left segment?
 - it does not contain "y"
 - that's why we kept searching

S
$$\longrightarrow$$
 \neq y j

Linear Search of an Array: Refined Invariant



Update the invariant to be more informative

```
const contains =
   (S: Array<bigint>, y: bigint): bigint => {
   let j = 0;
   // Inv: S[i] ≠ y for any i = 0 .. j-1
   while (j !== S.length && S[j] !== y) {
      j = j + 1;
   }
   return j !== S.length;
};
```

- "With great power, comes great responsibility"
- Since we can easily access any L[j], may need to keep track of facts about it
 - may need facts about every element in the list applies to preconditions, postconditions, and intermediate assertions
- We can write facts about several elements at once:

– this says that elements at indexes $0 \dots j-1$ are not y

 $S[i] \neq y$ for any $0 \le i < j$

- shorthand for j facts: $S[0] \neq y, ..., S[j-1] \neq y$

Description	Testing	Tools	Reasoning
no mutation	full coverage	type checker	calculation induction
local variable mutation	u	"	Floyd logic
heap state	u	"	rep invariants
arrays	u	"	for-any facts

Sublist "For any" Facts & Pictures

- "With great power, comes great responsibility"
 - since we can easily access any L[j], may need facts about it
- We can write facts about several elements at once:
 - this says that elements at indexes $0 \hfill ... j-1$ are not y

 $S[i] \neq y$ for any $0 \le i < j$

- These facts get hard to write down!
 - we will need to find ways to make this <u>easier</u>
 - a common trick is to draw pictures instead...

Visual Presentation of Facts

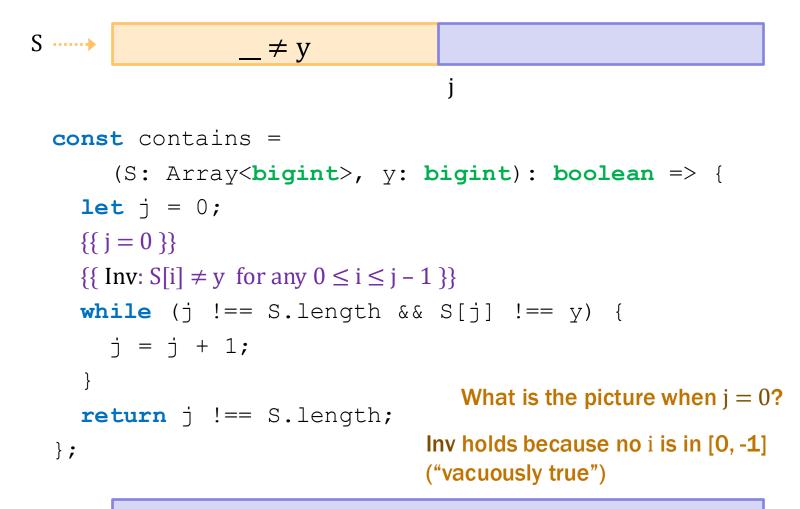


- Just saw this example
- But we have seen "for any" facts with BSTs...

contains-key(y, L) \rightarrow (y < x) contains-key(z, R) \rightarrow (x < z) L R

- "for any" facts are common in more complex code
- drawing pictures is a typical coping mechanism

Proving Linear Search of an Array: Initialization



S ----->

j

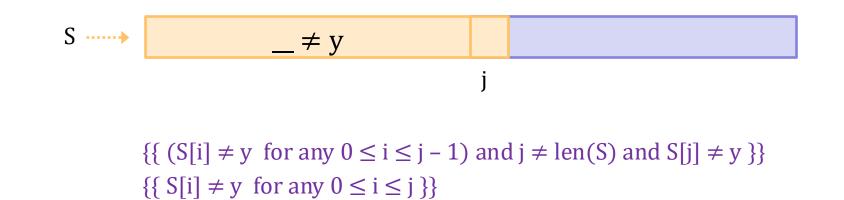
Linear Search of an Array: Preservation (1/4)

```
S .....
                    __ ≠ y
                                         j
  const contains =
        (S: Array<bigint>, y: bigint): boolean => {
     let j = 0;
     {{ Inv: S[i] \neq y for any 0 \le i \le j - 1 }}
     while (j !== S.length && S[j] !== y) {
       {{ (S[i] \neq y for any 0 \leq i \leq j – 1) and j \neq len(S) and S[j] \neq y }}
       j = j + 1;
       \{\{ S[i] \neq y \text{ for any } 0 \le i \le j - 1 \}\}
     }
     return j !== S.length;
  };
```

Linear Search of an Array: Preservation (2/4)

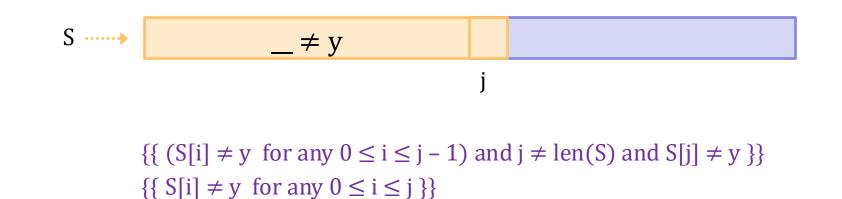
```
S .....
                      __ ≠ y
  const contains =
         (S: Array<bigint>, y: bigint): boolean => {
     let j = 0;
     {{ Inv: S[i] \neq y for any 0 \le i \le j - 1 }}
     while (j !== S.length && S[j] !== y) {
        {{ (S[i] \neq y for any 0 \leq i \leq j – 1) and j \neq len(S) and S[j] \neq y }}
     {{ S[i] \neq y \text{ for any } 0 \le i \le j }}
j = j + 1;
                                                                      Is this valid?
        {{ S[i] \neq y \text{ for any } 0 \leq i \leq j - 1 }}
     }
     return j !== S.length;
  };
```

Linear Search of an Array: Preservation (3/4)



- What does the top assertion say about S[j]?
 - it is not y

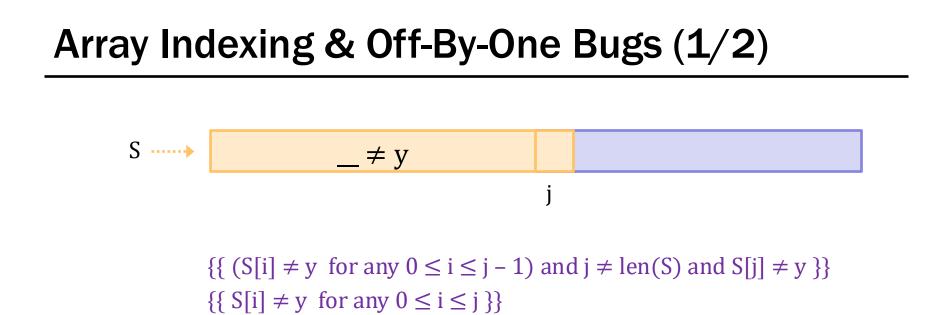




• What is the picture for the bottom assertion?



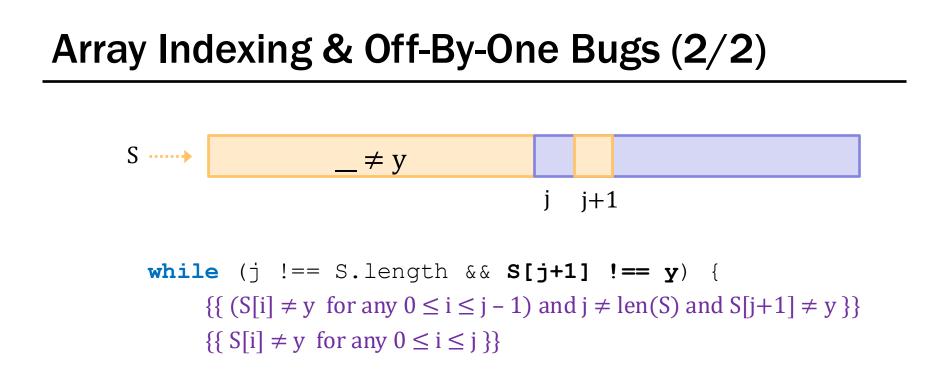
- Do the facts above imply this holds?
 - Yes! It's the same picture



• What is the picture for the bottom assertion?



- Most likely bug is an off-by-one error
 - must check S[j], not S[j-1] or S[j+1]



• What is the picture for the bottom assertion?



Reasoning would verify that this is not correct

Proving Linear Search of an Array: Exit (1/2)

```
S .....
                   __ ≠ y
  const contains =
       (S: Array<bigint>, y: bigint): boolean => {
    let j = 0;
    {{ Inv: S[i] \neq y for any 0 \le i \le j - 1 }}
    while (j !== S.length && S[j] !== y) {
       j = j + 1;
                                         "or" means cases...
    }
                                         Case j \neq len(S):
    {{ Inv and (j = len(S) \text{ or } S[j] = y) }}
                                         Must have S[j] = y.
    {{ contains(S, y) = (j \neq len(S)) }}
                                         What is the picture now?
    return j !== S.length;
  };
                                         Code should and does return true.
```

_______ y ______ y

Proving Linear Search of an Array: Exit (2/2)

```
S .....
                   __ ≠ y
  const contains =
       (S: Array<bigint>, y: bigint): boolean => {
    let j = 0;
    {{ Inv: S[i] \neq y for any 0 \le i \le j - 1 }}
    while (j !== S.length && S[j] !== y) {
       j = j + 1;
                                         "or" means cases...
     }
                                         Case j = len(S):
    {{ Inv and (j = len(S) \text{ or } S[j] = y) }}
                                         What does Inv say now?
    {{ contains(S, y) = (j \neq len(S)) }}
                                         Says y is not in the array!
    return j !== S.length;
                                         Code should and does return false.
  };
```