CS 577 - Data Structures & Amortized Analysis

Manolis Vlatakis

Department of Computer Sciences University of Wisconsin – Madison

Fall 2024



DATA STRUCTURES

Static problems

Given an input, produce an output.

Examples

Sorting, Compute Fourier Transform, shortest paths, ...

Data structures

Static problems

Given an input, produce an output.

Examples

Sorting, Compute Fourier Transform, shortest paths, ...

Dynamic problems

Given a sequence of operations (given one at a time), produce a sequence of outputs.

Examples

Dynamic Median Maintenance of a growing list, Incremental Convex Hull of a online updated set of points, ...

DATA STRUCTURES

Algorithm Vs Data structure

Algorithm. Step-by-step procedure to solve a problem. **Data structure.** Way to store and organize data.

Data structures

Algorithm Vs Data structure

Algorithm. Step-by-step procedure to solve a problem. **Data structure.** Way to store and organize data.

What data structures have you learned about in your courses so far?

DATA STRUCTURES



Ex. Array, linked list, binary search tree, hash table, ...

What data structures have you learned about in your courses so far?



What data structures have you learned about in your courses so far?

Let's see an example of a Data Structure Problem.

DATA STRUCTURE PROBLEMS

Goal. Design a data structure to support **all** operations in O(1) time.

- **INIT**(*n*): create and return an initialized array (all zero) of length *n*.
- **READ**(*A*,*i*): return element *i* in array.
- WRITE(*A*, *i*, **value**): set element *i* in array to **value**.

Assumptions.

- Can MALLOC an uninitialized array of length *n* in *O*(1) time. true in C or C++, but not Java
- Given an array, can read or write element i in O(1) time.

Appetizer.

Efficiently null-initialized array

Goal. Design a data structure to support **all** operations in O(1) time.

- **INIT**(*n*): create and return an initialized array (all zero) of length *n*.
- **READ**(*A*,*i*): return element *i* in array.
- WRITE(*A*, *i*, **value**): set element *i* in array to **value**.

Assumptions.

• Can MALLOC an uninitialized array of length *n* in *O*(1) time. true in C or C++, but not Java

Remark. A standard array does

INIT in $\Theta(n)$ time **READ** and **WRITE** in $\Theta(1)$ time

Appetizer.

Efficiently null-initialized array

Goal. Design a data structure to support **all** operations in *O*(1) time.

- **INIT**(*n*): create and return an initialized array (all zero) of length *n*.
- **READ**(*A*,*i*): return element *i* in array.
- WRITE(*A*, *i*, **value**): set element *i* in array to **value**.

How can we build such a structure using standard arrays like Lego blocks?

true in C or C++, but not Java

Remark. A standard array does

INIT in $\Theta(n)$ time **READ** and **WRITE** in $\Theta(1)$ time









EFFICIENTLY NULL-INITIALIZED ARRAY IMPLEMENTATION (EFFORT #1) Initialization Stamp



Is-Valid(A, i)

If $(1 \le B[i] \le k)$ Return true. Else Return false.

EFECTENTER NITE TATE AT THE AT

IM i. k keeps track of how many cells have been safely modified. ii. A[:] is used to store and retrieve the values.

iii. B[i] indicates when A[i] was initially modified.

INIT(A, n)

 $k \leftarrow 0$.

 $A \leftarrow Malloc(n)$.

 $B \leftarrow Malloc(n)$.

$\operatorname{Read}(A, i)$

If (Is-VALID(A[i]))
Return A[i].
Else

-

Return 0.

WRITE(A, i, value)
<pre>If (Is-Valid(A[i]))</pre>
$A[i] \leftarrow value.$
Else
$k \leftarrow k + 1$.
$A[i] \leftarrow value.$
$B[i] \leftarrow k.$

Is-Valid(A, i)

If $(1 \leq B[i] \leq k)$

Return true.

Else

Return false.

EFECTENTERY NITTER TRADE A DDAY

i. k keeps track of how many cells have been safely modified.
 ii. A[:] is used to store and retrieve the values.

iii. B[i] indicates when A[i] was initially modified.

Write(A,4,99), Write(A,6,33), Write(A,2,22), Write(A,3,55), Read(A,5)											
l	1	2	3	4	5	6	7	8			
A[]	?	22	55	99	?	33	?	?			
	k = 4										
B[]	?	3	4	1	?	2	?	?			

What if $1 \le B[5] \le 4$ is true by coincidence?

Ef	\mathbf{F} i. k keeps track of how many cells have been safely modified.												
Ιм	ii. A[:] is used to store and retrieve the values.												
	iii. B[i] indicates when A[i] was initially modified.												
	iv C[i] = index of i-th initialized element												
I			, , , , ,			<u>, </u>	, <u> </u>	lititui		cici			Je)
k	← 0.	(in) IATari		1 00	TA 7	mito (167	2) I	Aluito	(1)	22)		i1))
Δ	← Mai.i.		ie(A,	,4,99), vv	rne(1	ч,0,3 `	55), V	vrite	(A, Z)	,22),		
R		Write(A,3,5	5), I	Read	l(A,5)					Lue.	
C			1	2	3	4	5	6	7	8			
U	- MALL	A[]	?	22	55	99	?	33	?	?		1110	
												Lue.	
		B[]	?	3	4	1	?	2	?	?			
	Is-VAT												
	13 VAL												
	If (1	C[]	4	6	2	3	?	?	?	?			
	Retu					1. 4							
	Else					к = 4							
	Retu	What	if 1	≤ B [5] ≤	4 is	true	e by	coir	cide	ence?		
	_							J					

Efficiently Null-Initialized Array Implementation (Effort #2)

If $(1 \le B[i] \le k)$, then we still can't have (C[B[i]] = i) because $C[1..k] \ne i$ if A[i] not initialized. (Induction at $k \odot$)



Efficiently Null-Initialized Array Implementation (Effort #2)



Is-Valid(A, i)

```
If (1 \le B[i] \le k) and (C[B[i]] = i)
Return true.
Else
Return false.
```

Amortized Analysis

Definition

Worst-case analysis

Determine worst-case running time of a data structure operation as a function of the input size *n*.

Pessimistic if expensive operations follow many cheap ones.

Amortized analysis

Determine worst-case running time of a sequence of *n* data structure operations.

Amortized - Cost =
$$\frac{\sum_{i \in [n]} actual - cost[move_i]}{n}$$

Goal

Increment *n* times a *k*-bit binary counter (mod 2^k) where *k* is super-large $k = \omega(2^{2^n})$.

 $A[j] = j^{th}$ least significant bit of the counter.

Counter value	A[7]	A[6]	A[5]	A[4]	A[3]	A[2]	A[1]	A[0]
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1	0
3	0	0	0	0	0	0	1	1
4	0	0	0	0	0	1	0	0
5	0	0	0	0	0	1	0	1
6	0	0	0	0	0	1	1	0
7	0	0	0	0	0	1	1	1
8	0	0	0	0	1	0	0	0
9	0	0	0	0	1	0	0	1
10	0	0	0	0	1	0	1	0
11	0	0	0	0	1	0	1	1

Table: Binary Counter Table

Cost model

Number of bits flipped.

How many bit flips are needed to increment the counter *n* times, starting from zero?

Counter value	A[7]	A[6]	A[5]	A[4]	A[3]	A[2]	A[1]	A[0]
0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1
2	0	0	0	0	0	0	1	0
3	0	0	0	0	0	0	1	1
4	0	0	0	0	0	1	0	0
5	0	0	0	0	0	1	0	1
6	0	0	0	0	0	1	1	0
7	0	0	0	0	0	1	1	1
8	0	0	0	0	1	0	0	0
9	0	0	0	0	1	0	0	1
10	0	0	0	0	1	0	1	0
11	0	0	0	0	1	0	1	1

Table: Binary Counter Table

Cost model

Number of bits flipped.

⁽²⁾How many bit flips are needed to increment the counter n times, starting from zero?

In worst-case, at a	most	k bits	s flipp	ped p	er in	creme	ent. S	bo O(nk	:).
1	0	0	0	0	0	0	U	1	
2		0	0	0	0	0	1	0	
3	0	0	0	0	0	0	1	1	
4	0	0	0	0	0	1	0	0	
5	0	0	0	0	0	1	0	1	
6	0	0	0	0	0	1	1	0	
7	0	0	0	0	0	1	1	1	
8	0	0	0	0	1	0	0	0	
9	0	0	0	0	1	0	0	1	
10	0	0	0	0	1	0	1	0	
11	0	0	0	0	1	0	1	1	

Cost model

Number of bits flipped.

How many bit flips are needed to increment the counter *n* times, starting from zero?

In worst-case, at most k bits flipped per increment. So $O(n\dot{k})$.

But k bits together would only flip during the last increment.

-			-	-		-			
3	0	0	0	0	0	1	0	1	
6	0	0	0	0	0	1	1	0	
7	0	0	0	0	0	1	1	1	
8	0	0	0	0	1	0	0	0	
9	0	0	0	0	1	0	0	1	
10	0	0	0	0	1	0	1	0	
11	0	0	0	0	1	0	1	1	

Intuition

Measure running time in terms of credits (time = money).

Intuition

Measure running time in terms of credits (time = money).

Key Idea

- Expensive operations: Prepay for these by storing up credits during cheaper operations.
- <u>Cheap operations</u>: Use the already stored credits (no extra cost at the time).

Intuition

Measure running time in terms of credits (time = money).

Key Idea

- Expensive operations: Prepay for these by storing up credits during cheaper operations.
- <u>Cheap operations</u>: Use the already stored credits (no extra cost at the time).

How It Works

- For every operation, we consume one credit.
- Assign each operation a cost that includes both the actual cost and some extra "savings."
- Use these "savings" to cover the cost of future cheap operations.

AN EXAMPLE IS WORTH A THOUSAND WORDS

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

 $\# \text{ operations} = \underbrace{(\text{Dollars we consumed})}_{\text{actual cost}} \\ \# \text{ operations} = \underbrace{(\text{Dollars we borrowed})}_{\text{bank cost}} - \underbrace{(\text{Dollars left in our pocket})}_{\text{Potential overcharging}}.$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

 $# \text{ operations} = \underbrace{(\text{Dollars we consumed})}_{\text{actual cost}}$ $# \text{ operations} = \underbrace{(\text{Dollars we borrowed})}_{\text{bank cost}} - \underbrace{(\text{Dollars left in our pocket})}_{\text{Potential overcharging}}.$ $\underbrace{\text{Operation Bank}}_{\text{Expensive 5\$}} + \underbrace{\text{Pocket}}_{\text{Expensive 5\$}}$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

 $# \text{ operations} = \underbrace{(\text{Dollars we consumed})}_{\text{actual cost}}$ $# \text{ operations} = \underbrace{(\text{Dollars we borrowed})}_{\text{bank cost}} - \underbrace{(\text{Dollars left in our pocket})}_{\text{Potential overcharging}}.$ $\underbrace{\text{Operation Bank Pocket}}_{\text{Expensive 5\$}} \underbrace{4\$}$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

 $# \text{ operations} = \underbrace{(\text{Dollars we consumed})}_{actual cost}$ $# \text{ operations} = \underbrace{(\text{Dollars we borrowed})}_{bank cost} - \underbrace{(\text{Dollars left in our pocket})}_{Potential overcharging}.$ $\boxed{\begin{array}{c} Operation & Bank & Pocket \\ \hline Expensive & 5\$ & 4\$ \\ \hline Cheap & & \\ \end{array}}$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

 $# \text{ operations} = \underbrace{(\text{Dollars we consumed})}_{\text{actual cost}}$ $# \text{ operations} = \underbrace{(\text{Dollars we borrowed})}_{\text{bank cost}} - \underbrace{(\text{Dollars left in our pocket})}_{\text{Potential overcharging}}.$ $\underbrace{ \begin{array}{c|c} \mathbf{Operation} & \mathbf{Bank} & \mathbf{Pocket} \\ \hline \text{Expensive} & \underline{5\$} & \underline{4\$} \\ \hline \text{Cheap} & \underline{3\$} \\ \end{array}}$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

operations = (Dollars we consumed)

actual cost
operations = (Dollars we borrowed) - (Dollars left in our pocket).

bank cost

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

operations = (Dollars we consumed)

actual cost

operations = (Dollars we borrowed) – (Dollars left in our pocket).

bank cost

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		2\$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

operations = (Dollars we consumed)

actual cost

operations = (Dollars we borrowed) – (Dollars left in our pocket).

bank cost

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		2\$
Expensive		

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

operations = (Dollars we consumed)

actual cost

operations = (Dollars we borrowed) – (Dollars left in our pocket).

bank cost

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		2\$
Expensive	5\$	

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

operations = (Dollars we consumed)

actual cost

operations = (Dollars we borrowed) – (Dollars left in our pocket).

bank cost

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		2\$
Expensive	5\$	6\$

AN EXAMPLE IS WORTH A THOUSAND WORDS Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank.

- For expensive operations, we borrow \$5. The leftover money stays in our pocket for future use.
- For cheap operations, we borrow only \$1 if our pocket is empty.

At the end,

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		2\$
Expensive	5\$	6\$

operations = $\underbrace{(\text{Dollars we borrowed})}_{10\$} - \underbrace{(\text{Dollars left in our pocket})}_{6\$} = 4.$

AN EXAMPLE IS WORTH A THOUSAND WORDS

Every operation costs \$1 to execute. If we don't have money in our pocket, we borrow from the bank

Aesop's Moral

- Bank: Our worst-case outlook on expensive operations.
- Pocket: Our amortized hope that costly operations are rare.

At the end,

Operation	Bank	Pocket
Expensive	5\$	4\$
Cheap		3\$
Cheap		2\$
Expensive	5\$	6\$

operations = $\underbrace{(\text{Dollars we borrowed})}_{10\$} - \underbrace{(\text{Dollars left in our pocket})}_{6\$} = 4.$

Amortized Analysis

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

Amortized Analysis

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

Increment		A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0		0	0	0	0	0		
Ļ							2\$	1\$
1		0	0	0	0	1		
	1							

Amortized Analysis

I

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

$0 \qquad \cdots \qquad 0 \qquad $	
↓ 2\$	1\$
$1 \cdots 0 0 0 0 1$	
↓ 2\$	1\$
2 0 0 0 1 0	

Amortized Analysis

Ŀ

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

ncrement	 A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0	 0	0	0	0	0		
Ļ						2\$	1\$
1	 0	0	0	0	1		
↓						2\$	1\$
2	 0	0	0	1	0		
↓						2\$	2\$
3	 0	0	0	1	1		

Amortized Analysis

Ŀ

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

ncrement		A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0		0	0	0	0	0		
Ļ							2\$	1\$
1		0	0	0	0	1		
Ļ							2\$	1\$
2		0	0	0	1	0		
Ļ							2\$	2\$
3		0	0	0	1	1		
Ļ							2\$	1\$
4		0	0	1	0	0		

Amortized Analysis

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

Increment		A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0		0	0	0	0	0		
Ļ							2\$	1\$
1		0	0	0	0	1		
Ļ							2\$	1\$
2		0	0	0	1	0		
Ļ							2\$	2\$
3		0	0	0	1	1		
Ļ							2\$	1\$
4		0	0	1	0	0		
Ļ							2\$	2\$
5		0	0	1	0	1		

Amortized Analysis

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

Increment		A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0		0	0	0	0	0		
\downarrow							2\$	1\$
1		0	0	0	0	1		
Ļ							2\$	1\$
2		0	0	0	1	0		
Ļ							2\$	2\$
3		0	0	0	1	1		
↓							2\$	1\$
4		0	0	1	0	0		
Ļ							2\$	2\$
5		0	0	1	0	1		
↓							2\$	2\$
6		0	0	1	1	0		

Amortized Analysis

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

Increment		A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0		0	0	0	0	0		
\downarrow							2\$	1\$
1		0	0	0	0	1		
\downarrow							2\$	1\$
2		0	0	0	1	0		
\downarrow							2\$	2\$
3		0	0	0	1	1		
\downarrow							2\$	1\$
4		0	0	1	0	0		
\downarrow							2\$	2\$
5		0	0	1	0	1		
\downarrow							2\$	2\$
6		0	0	1	1	0		
\downarrow							2\$	3\$
7		0	0	1	1	1		

Amortized Analysis

- If we don't have enough money in our pocket, we borrow \$1 per necessary flip.
- Any leftover money stays in our pocket for future use.

Increment		A[4]	A[3]	A[2]	A[1]	A[0]	Bank	Pocket
0		0	0	0	0	0		
\downarrow							2\$	1\$
1		0	0	0	0	1		
\downarrow							2\$	1\$
2		0	0	0	1	0		
\downarrow							2\$	2\$
3		0	0	0	1	1		
\downarrow							2\$	1\$
4		0	0	1	0	0		
Ļ							2\$	2\$
5		0	0	1	0	1		
Ļ							2\$	2\$
6		0	0	1	1	0		
\downarrow							2\$	3\$
7		0	0	1	1	1		
\downarrow							2\$	1\$
8		0	1	0	0	0		

Amortized Analysis

Rules: The cost of any flip is \$1. We always borrow \$2 from the bank.

• If we don't have enough money in our pocket, we borrow \$1 per necessary flip.

Observations

- #1 Pocket dollars = Aces on the counter register
- #2 Bank loan of 2\$ is always sufficient.

(Proof by induction ☺)

↓ 3		0	0	0	1	1	ZΦ	Zφ	
Ļ		0	0	1	0	0	2\$	1\$	
4 ↓		0	0	1	0	0	2\$	2\$	
5		0	0	1	0	1			
↓ 6		0	0	1	1	0	2\$	2\$	
Ļ		0	0	1	1	Ū	2\$	3\$	
7		0	0	1	1	1	20	10	
↓ 8		0	1	0	0	0	∠⊅	1\$	

Amortized Analysis

Rules: The cost of any flip is \$1. We always borrow \$2 from the bank.

• If we don't have enough money in our pocket, we borrow \$1 per necessary flip.

Observations

- #1 Pocket dollars = Aces on the counter register
- #2 Bank loan of 2\$ is always sufficient.

0

(Proof by induction ⁽²⁾)

Amortized Analysis

operations = (Dollars we borrowed) - (Dollars left in our pocket)

bank cost

Potential overcharging

operations = $2n - (\text{Dollars left in our pocket}) \le 2n$

Appendix

References

Image Sources I



https://en.wikipedia.org/wiki/Bijection