DESIGN AND ANALYSIS OF ALGORITHMS

CS 4120/5120
MAXIMUM-SUBARRAY PROBLEM

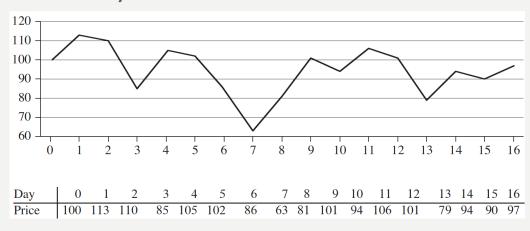
AGENDA

- Maximum-subarray problem definition
- Divide-and-conquer algorithm
- Time complexity analysis

THE STOCK TRADING STORY

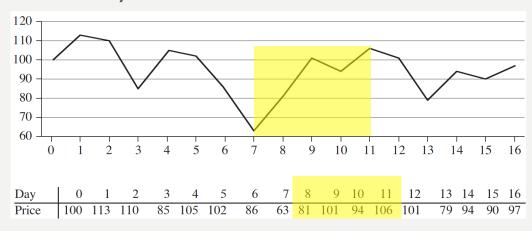
- Suppose that you can learn what the price of the stock will be in the future 17 days.
 - The price of the stock over the 17-day period.
 - You can buy or sell after the close of trading for the day.
 - Using this chart, you can make two trading activities: one buy and one sell.
- What is the best time to buy and the best time to sell?
 - What is the profit?





THE STOCK TRADING STORY MAKING THE MOST MONEY

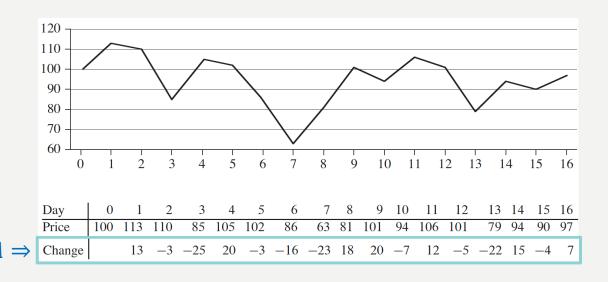
- Suppose that you can learn what the price of the stock will be in the future 17 days.
 - The price of the stock over the 17-day period.
 - You can buy or sell after the close of trading for the day.
 - Using this chart, you can make two trading activities: one buy and one sell.
- The maximum profit can be obtained by buying after the market closes on day _____ and sell after the market closes on day _____.
- Share your thoughts?



THE STOCK TRADING STORY MODEL THE PROBLEM

Transformation

- Compute the change of price of two consecutive days.
- Store the changes in an array.
- Find the longest subarray whose values have the greatest sum.



PLAY WITH THE MODEL

• Given the following input instance.

Index	I	2	3	4	5	6	7	8	9	10	П	12	13	14	15	16
Number	19	-20	8	4	-4	I	7	-15	8	6	-4	6	-1	2	-6	3

• Find the longest subarray whose values have the greatest sum.

MAXIMUM-SUBARRAY PROBLEM PROBLEM DEFINITION

Input

- Array A[1..n] containing both **positive** and **negative** numbers.

Output

- A contiguous subarray A[i..j] of A[1..n], $1 \le i \le j \le n$, such that the sum of values in A[i..j] is the largest of all contiguous subarrays of A[1..n].

- Goal
 - A contiguous subarray A[i..j] of A[1..n], $1 \le i \le j \le n$, such that the sum of values in A[i..j] is the largest of all contiguous subarrays of A[1..n].
- Brute-force solution
 - Step I: List all contiguous subarrays.
 - There are $\binom{n}{2}$ or $\binom{n}{2}$ different subarrays.

- Goal
 - A contiguous subarray A[i..j] of A[1..n], $1 \le i \le j \le n$, such that the sum of values in A[i..j] is the largest of all contiguous subarrays of A[1..n].
- Brute-force solution
 - **Step 2**: Calculate the sum of each subarray A[i..j].
 - For each subarray, the cost of computing the sum is the cost of traversing A[i..j].
 - Therefore, the cost of computing the sum of each subarray is $\underline{j-i}$, where $i \neq j$ and $i, j \in [1, n]$.

- Goal
 - A contiguous subarray A[i..j] of A[1..n], $1 \le i \le j \le n$, such that the sum of values in A[i..j] is the largest of all contiguous subarrays of A[1..n].
- Brute-force solution
 - Step 3: Find the max of all sums.
 - The cost of finding the max of sums is $\binom{n}{2} 1$

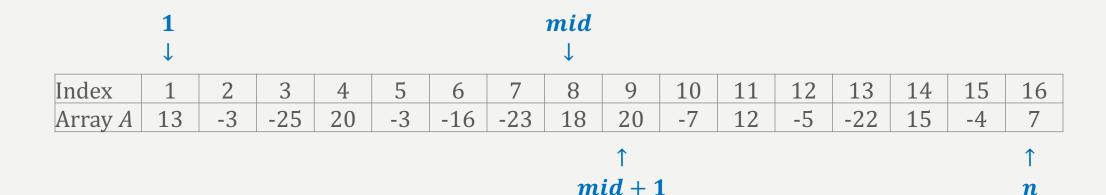
- Goal
 - A contiguous subarray A[i..j] of A[1..n], $1 \le i \le j \le n$, such that the sum of values in A[i..j] is the largest of all contiguous subarrays of A[1..n].
- Brute-force solution
 - The time complexity of brute-force solution is $T(n) = \Omega(n^2)$.
- Better solution?

MAXIMUM-SUBARRAY PROBLEM DIVIDE AND CONQUER

- Apply the **three steps**
 - **Divide** the problem into a number of subproblems that are smaller instances of the same problem.
 - Conquer the subproblems by solving them recursively.
 - If the subproblem sizes are small enough, just solve the subproblems in a straightforward manner.
 - Combine the solutions to the subproblems into the solution to the original problem.

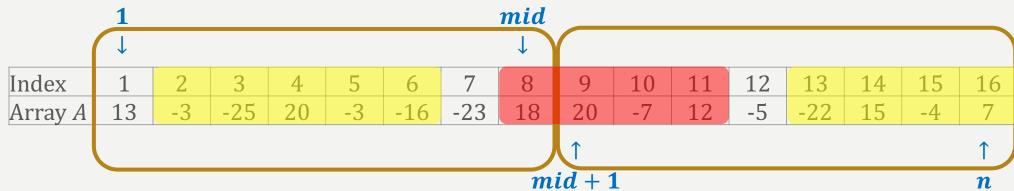
MAXIMUM-SUBARRAY PROBLEM DIVIDE

- Divide the original array A[1..n] into subarrays A[1..mid] and A[mid + 1..n].
 - $mid = \lfloor (1+n)/2 \rfloor$
 - Deal with A[1..mid] and A[mid + 1..n] individually by dividing them into smaller subproblems.



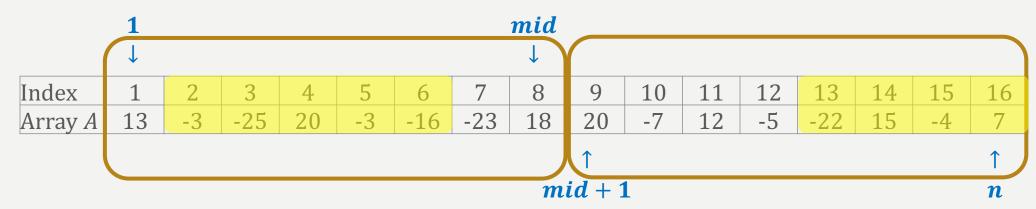
MAXIMUM-SUBARRAY PROBLEM CONQUER

- Conquer the subproblems arisen from the divide step.
 - Two scenarios emerge.
 - Case I: The maximum subarray lies entirely in the left half A[1..mid] or entirely in the right half A[mid + 1..n].
 - Case 2: The maximum subarray A[i...j] happens to cross the midpoint.



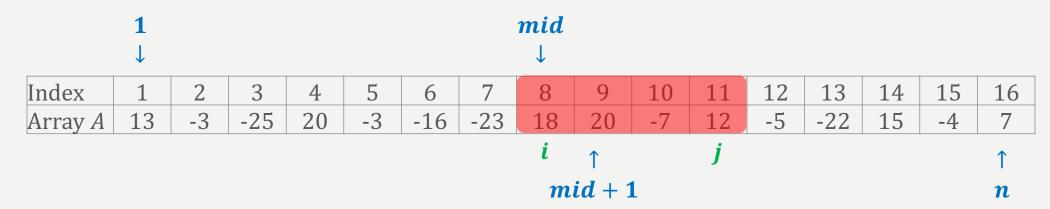
MAXIMUM-SUBARRAY PROBLEM CONQUER CASE 1

- The subproblems are the same as the original problem.
- Generally, the problem can be described as finding the maximum subarray of A[low..high].
- In this case, the indices of maximum subarray A[i..j] will always satisfy $low \le i < j \le high$.
- Solving case I is trivial.



MAXIMUM-SUBARRAY PROBLEM CONQUER CASE 2

- The maximum subarray A[i..j] happens to cross the midpoint.
- In this case, the indices of maximum subarray A[i..j] will be $low \le i \le mid < j \le high$, which is **NOT** exactly the same as the original problem.
- In other words, solution consists of two arrays: A[i..mid] and A[mid + 1..j].



MAXIMUM-SUBARRAY PROBLEM CONQUER CASE 2 (CONT'D)

- The solution A[i..j] crossing the midpoint means A[mid] is included in A[i..j].
- Start at A[mid].
 - Move i to the left side to find subarray A[i..mid] that has the greatest sum, max-left.
 - Move j to the right side to find subarray A[mid + 1..j] that has the greatest sum, max-right.
- The maximum sum is $max-left + max-right_{inid}$

								<u> </u>								
Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Array A	13	-3	-25	20	-3	-16	-23	18	20	-7	12	-5	-22	15	-4	7
						_			_							

 \boldsymbol{n}

MAXIMUM-SUBARRAY PROBLEM COMBINE

- Combine the solutions obtained from case I and case 2. Pick the greatest sum.
 - Case I
 - A[i..j] exists entirely in the left half. The maximum sum is left-sum = 20.
 - A[i..j] exists entirely in the right half. The maximum sum is right-sum=20-7+12=25.
 - Case 2
 - A[i..j] crosses the midpoint. The maximum sum is crosssum = 18 + 20 7 + 12 = 43.

Index	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Array A	13	-3	-25	20	-3	-16	-23	18	20	-7	12	-5	-22	15	-4	7
								↑								↑
	$m{mid}$										n					

MAXIMUM-SUBARRAY PROBLEM THE ALGORITHM

- The FIND-MAXIMUM-SUBARRAY algorithm
 - Use $\Theta(1)$ to denote a constant time cost.
 - Use f (high low +
 1) to denote the cost
 of line 6
 - Use probability P_i to denote the probability of conditions being true at line 7 and 9.

FII	ND-MAXIMUM-SUBARRAY(A, low, high)	Cost	Time
I	$\mathbf{if}\ high == low$	Θ(1)	base
2	return (low, high, A[low]) // base case	$\Theta(1)$	base
3	else $mid = \lfloor (low + high)/2 \rfloor$	$\Theta(1)$	1
4	(left-low, left-high, left-sum) =	T(mid -	1
	$FIND ext{-}MAXIMUM ext{-}SUBARRAY(A, low, mid)$	low + 1)	
5	(right-low, right-high, right-sum) =	T(high	1
	FIND-MAXIMUM-SUBARRAY(A, mid + 1, high)	-mid)	
6	(cross-low, cross-high, cross-sum) =	f(high -	1
	FIND-MAXIMUM-CROSSING-SUBARRAY(A, low, mid, high)	low + 1)	
7	if $left$ - $sum \ge right$ - sum and $left$ - $sum \ge cross$ - sum	$\Theta(1)$	1
8	return (left-low, left-high, left-sum)	$\Theta(1)$	P_1
9	elseif $right$ - $sum \ge left$ - sum and $right$ - $sum \ge cross$ - sum	$\Theta(1)$	P_1
10	return (right-low, right-high, right-sum)	Θ(1)	P_2
11	else return (cross-low, cross-high, cross-sum)	Θ(1)	$\begin{vmatrix} 1 - P_1 \\ -P_2 \end{vmatrix}$

• Derive the running time function of T(high - low + 1)

FII	$ND ext{-}MAXIMUM ext{-}SUBARRAY(A, low, high)$	Cost	Time					
I	$if \ high == low$	$\Theta(1)$	base					
2	return $(low, high, A[low])$ // base case	$\Theta(1)$	base					
3	else $mid = \lfloor (low + high)/2 \rfloor$	$\Theta(1)$	1					
4	(left-low, left-high, left-sum) =	T(mid -	1					
_	FIND-MAXIMUM-SUBARRAY(A, low, mid)	low + 1)	1					
5	(right-low,right-high,right-sum) =	T(high	1					
	FIND-MAXIMUM-SUBARRAY(A, mid + 1, high)	-mid)	1					
6	(cross-low, cross-high, cross-sum) =	f(high -	1					
	FIND-MAXIMUM-CROSSING-SUBARRAY(A, low, mid, high)	low + 1)) 1					
7	if $left$ -sum $\geq right$ -sum and $left$ -sum $\geq cross$ -sum	$\Theta(1)$	1					
8	return (left-low, left-high, left-sum)	$\Theta(1)$	P_1					
9	elseif $right$ - $sum \ge left$ - sum and $right$ - $sum \ge cross$ - sum	$\Theta(1)$	P_1					
10	return (right-low, right-high, right-sum)	Θ(1)	P_2					
11	else return (cross-low, cross-high, cross-sum)	Θ(1)	$ \begin{array}{c} 1 - P_1 \\ -P_2 \end{array} $					

```
• Let low = 1 and let \ high = n.

n = 
T(high - low + 1)
= \Theta(1) + 
T(mid - low + 1) + 
T(high - mid) + 
f(high - low + 1)
```

FII	ND-MAXIMUM-SUBARRAY(A, low, high)	Cost	Time				
I	$\mathbf{if}\ high == low$	$\Theta(1)$	base				
2	return (low, high, A[low]) // base case	$\Theta(1)$	base				
3	else $mid = \lfloor (low + high)/2 \rfloor$	$\Theta(1)$	1				
4	(left-low, left-high, left-sum) =	T(mid -	1				
	$FIND ext{-}MAXIMUM ext{-}SUBARRAY(A, low, mid)$	low + 1)	1				
5	(right-low, right-high, right-sum) =	T(high	1				
	FIND-MAXIMUM-SUBARRAY(A, mid + 1, high)	-mid)	1				
6	(cross-low, cross-high, cross-sum) =	f(high -	1				
	FIND-MAXIMUM-CROSSING-SUBARRAY(A, low, mid, high)	low + 1)	1				
7	if $left$ - $sum \ge right$ - sum and $left$ - $sum \ge cross$ - sum	$\Theta(1)$	1				
8	return (left-low, left-high, left-sum)	$\Theta(1)$	$ P_1 $				
9	elseif $right$ - $sum \ge left$ - sum and $right$ - $sum \ge cross$ - sum	$\Theta(1)$	P_1				
10	return (right-low, right-high, right-sum)	Θ(1)	P_2				
11	else return (cross-low, cross-high, cross-sum)	Θ(1)	$ \begin{array}{c} 1 - P_1 \\ -P_2 \end{array} $				

```
• Let low = 1 and let high = n. mid = _____.
```

• mid = (n+1 is even)

(n+1 is odd)

FI	ND-MAXIMUM-SUBARRAY(A, low, high)	Cost	Time
	$\mathbf{if}\ high == low$	Θ(1)	base
2	return (low, high, A[low]) // base case	$\Theta(1)$	base
3	else $mid = \lfloor (low + high)/2 \rfloor$	$\Theta(1)$	1
4	(left-low, left-high, left-sum) =	T(mid -	1
	FIND-MAXIMUM-SUBARRAY(A, low, mid)	low + 1)	1
5	(right-low,right-high,right-sum) =	T(high	1
	FIND-MAXIMUM-SUBARRAY(A, mid + 1, high)	-mid)	1
6	(cross-low, cross-high, cross-sum) =	f(high -	1
	FIND-MAXIMUM-CROSSING-SUBARRAY(A, low, mid, high)	low + 1)	1
7	if $left$ - $sum \ge right$ - sum and $left$ - $sum \ge cross$ - sum	$\Theta(1)$	1
8	return (left-low, left-high, left-sum)	$\Theta(1)$	P_1
9	elseif $right$ - $sum \ge left$ - sum and $right$ - $sum \ge cross$ - sum	$\Theta(1)$	P_1
10	return (right-low, right-high, right-sum)	Θ(1)	P_2
	else return (cross-low, cross-high, cross-sum)	Θ(1)	$ \begin{array}{c} 1 - P_1 \\ -P_2 \end{array} $

The running time function

$$T(n)$$

$$= \Theta(1) + T\left(\frac{n+1}{2}\right) +$$

$$T(\frac{n-1}{2}) + f(n) \quad (n+1 \text{ is even})$$

$$= \Theta(1) + 2T\left(\frac{n}{2}\right) + f(n) \quad (n+1 \text{ is odd})$$

• Tolerate sloppiness $T(n) = 2T\left(\frac{n}{2}\right) + f(n)$

	FIN	ND-MAXIMUM-SUBARRAY(A, low, high)	Cost	Time
	I	$\mathbf{if}\ high == low$	$\Theta(1)$	base
	2	return (low, high, A[low]) // base case	$\Theta(1)$	base
	3	$else mid = \lfloor (low + high)/2 \rfloor$	$\Theta(1)$	1
	4	(left-low, left-high, left-sum) =	T(mid -	1
		$FIND ext{-}MAXIMUM ext{-}SUBARRAY(A, low, mid)$	low + 1)	1
	5	(right-low,right-high,right-sum) =	T(high	1
ı)	<u> </u>	FIND-MAXIMUM-SUBARRAY(A, $mid + 1$, $high$)	-mid)	1
	6	(cross-low, cross-high, cross-sum) =	f(high -	. 1
	0		low + 1)	1
₁)	7	if $left$ - $sum \ge right$ - sum and $left$ - $sum \ge cross$ - sum	$\Theta(1)$	1
	8	return (left-low, left-high, left-sum)	$\Theta(1)$	P_1
	9	elseif $right$ - $sum \ge left$ - sum and $right$ - $sum \ge cross$ - sum	$\Theta(1)$	P_1
	10	return (right-low, right-high, right-sum)	$\Theta(1)$	P_2
	11	else return (cross-low, cross-high, cross-sum)	Θ(1)	$ \begin{array}{c} 1 - P_1 \\ -P_2 \end{array} $

- Complete the Cost and Time columns.
 - Consider ONLY the worst-case scenario

FI	ND-MAX-CROSSING-SUBARRAY(A, low, mid, high)	Cost	Time (Worst-case)
I	$left$ - $sum = -\infty$	Θ(1)	1
2	sum = 0	$\Theta(1)$	
3	for $i = mid$ downto low	$\Theta(1)$	
4	sum = sum + A[i]	$\Theta(1)$	
5	if $sum > left$ - sum	$\Theta(1)$	
6	left- $sum = sum$	$\Theta(1)$	
7	max- $left = i$	$\Theta(1)$	
8	$right$ - $sum = -\infty$	$\Theta(1)$	
9	sum = 0	$\Theta(1)$	
10	$\mathbf{for} \ j = mid + 1 \ \mathbf{to} \ high$	$\Theta(1)$	
11	sum = sum + A[j]	$\Theta(1)$	
12	if sum > right-sum	Θ(1)	
13	right- $sum = sum$	Θ(1)	
14	max- $right = j$	Θ(1)	
15	return $(max-left, max-right, left-sum + right-sum)$	Θ(1)	

- Complete the Cost and Time columns.
 - Consider ONLY the worst-case scenario

E I I	IND-MAX-CROSSING-SUBARRAY(A, low, mid, high)		Time
	VD-MAX-CROSSING-SOBARRAI (A, tow, mta, mgn)	Cost	(Worst-case)
I	$left$ - $sum = -\infty$	$\Theta(1)$	1
2	sum = 0	$\Theta(1)$	1
3	for i = mid downto low	$\Theta(1)$	(mid - low + 2)
4	sum = sum + A[i]	$\Theta(1)$	(mid - low + 1)
5	if $sum > left$ - sum	$\Theta(1)$	(mid - low + 1)
6	left- $sum = sum$	$\Theta(1)$	(mid - low + 1)
7	max- $left = i$	$\Theta(1)$	(mid - low + 1)
8	$right$ - $sum = -\infty$	$\Theta(1)$	1
9	sum = 0	$\Theta(1)$	1
10	$\mathbf{for}j = mid + 1\mathbf{to}high$	$\Theta(1)$	(high - mid + 1)
П	sum = sum + A[j]	$\Theta(1)$	(high - mid)
12	if $sum > right$ - sum	$\Theta(1)$	(high - mid)
13	right- $sum = sum$	$\Theta(1)$	(high - mid)
14	max- $right = j$	$\Theta(1)$	(high - mid)
15	return $(max-left, max-right, left-sum + right-sum)$	$\Theta(1)$	1

• Running time f(high - low + 1)

E I I	ND-MAX-CROSSING-SUBARRAY (A, low, mid, high)	Cost	Time
1.11	VD-MAX-CROSSING-SOBARRAI (A, tow, mta, ntgn)	Cost	(Worst-case)
	$left$ - $sum = -\infty$	$\Theta(1)$	1
2	sum = 0	$\Theta(1)$	1
3	$\mathbf{for}\ i = mid\ \mathbf{downto}\ low$	$\Theta(1)$	(mid - low + 2)
4	sum = sum + A[i]	$\Theta(1)$	(mid - low + 1)
5	if $sum > left$ - sum	$\Theta(1)$	(mid - low + 1)
6	left- $sum = sum$	$\Theta(1)$	(mid - low + 1)
7	max- $left = i$	$\Theta(1)$	(mid - low + 1)
8	$right$ - $sum = -\infty$	$\Theta(1)$	1
9	sum = 0	$\Theta(1)$	1
10	$\mathbf{for} \ j = mid + 1 \ \mathbf{to} \ high$	$\Theta(1)$	(high - mid + 1)
	sum = sum + A[j]	$\Theta(1)$	(high - mid)
12	if $sum > right$ - sum	$\Theta(1)$	(high - mid)
13	right- $sum = sum$	Θ(1)	(high - mid)
14	max- $right = j$	Θ(1)	(high - mid)
15	return $(max-left, max-right, left-sum + right-sum)$	Θ(1)	1

- Let low = 1, high = n.
- f(high low + 1)= f(n)=

E I I	ND-MAX-CROSSING-SUBARRAY (A, low, mid, high)	Cost	Time
1.11	VD-MAX-CROSSING-SOBARRAI (A, tow, mta, ntgn)	Cost	(Worst-case)
I	$left$ - $sum = -\infty$	$\Theta(1)$	1
2	sum = 0	$\Theta(1)$	1
3	$\mathbf{for}\ i = mid\ \mathbf{downto}\ low$	$\Theta(1)$	(mid - low + 2)
4	sum = sum + A[i]	$\Theta(1)$	(mid - low + 1)
5	if $sum > left$ - sum	$\Theta(1)$	(mid - low + 1)
6	left- $sum = sum$	$\Theta(1)$	(mid - low + 1)
7	max- $left = i$	$\Theta(1)$	(mid - low + 1)
8	$right$ - $sum = -\infty$	$\Theta(1)$	1
9	sum = 0	$\Theta(1)$	1
10	$\mathbf{for} \ j = mid + 1 \ \mathbf{to} \ high$	$\Theta(1)$	(high - mid + 1)
	sum = sum + A[j]	$\Theta(1)$	(high - mid)
12	if $sum > right$ - sum	$\Theta(1)$	(high - mid)
13	right- $sum = sum$	Θ(1)	(high – mid)
14	max- $right = j$	Θ(1)	(high - mid)
15	return $(max-left, max-right, left-sum + right-sum)$	Θ(1)	1

MAXIMUM-SUBARRAY PROBLEM TIME COMPLEXITY

- Combine the recursive function T(n) and f(n).
- The recursive running time of the FIND-MAXIMUM-SUBARRAY algorithm is

$$T(n) = \begin{cases} \Theta(1) & \text{if } n = 1 \\ 2T(n/2) + \Theta(n) & \text{if } n > 1 \end{cases}$$

NEXT UP STRASSEN'S ALGORITHM

REFERENCE

- https://www.vectorstock.com/royalty-free-vector/stacks-of-coins-and-money-bag-vector-log9019
- The stock price chart is a screenshot from the textbook.