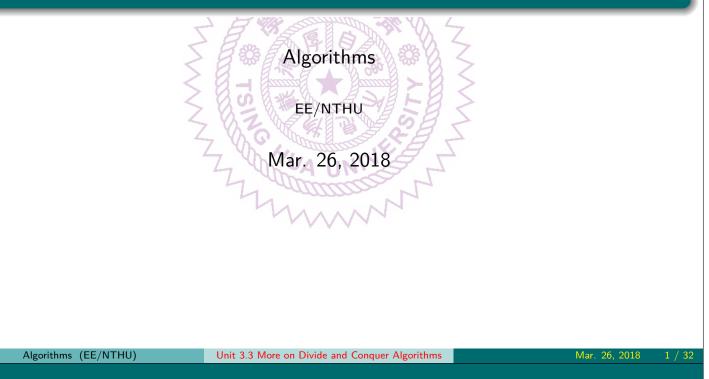
### Unit 3.3 More on Divide and Conquer Algorithms



## Selection Algorithm

- The selection problem is to find the k'th element of an array A and place all elements less than or equal to A[k] in A[1:k-1] and the rest in A[k+1:n].
- Divide and conquer can be applied to this problem as well.
  - The Partition algorithm can be effective in this selection problem.

#### Algorithm 3.3.1. Selection

```
1 Algorithm Select1(A, n, k)
 2 // Partition the array into A[1:k-1] \leq A[k] \leq A[k+1:n].
 3 {
        low := 1; high := n + 1; A[n + 1] := \infty; j := k - 1;
 4
        while (j \neq k) do {
 5
             j := Partition(A, low, high);
 6
 7
             if (k < j) then high := j;
             else if (k > j) then low := j + 1;
 8
 9
        }
10 }
```

• After completing Select1, A[k] is the k'th element.

#### Selection Algorithm – Complexity

- Note that the Partition (A, low, high) decrease the range of the array A by at least 1.
- Thus, the worst-case complexity of the Select1 algorithm is  $\mathcal{O}(n^2)$ .
- Let  $T_A^k(n)$  be the average time to find the k'th smallest element in A[1:n]• The average is taken over all *n*! different permutations.
- Define

$$T_{A}(n) = \frac{1}{n} \sum_{k=1}^{n} T_{A}^{k}(n)$$
(3.3.1)  
$$R(n) = \max_{k} T_{A}^{k}(n)$$
(3.3.2)

• 
$$T_A(n)$$
 is the average execution time of Select1 algorithm.

• And it is obvious that  $T_A(n) \leq R(n)$ .

Algorithms (EE/NTHU)

Unit 3.3 More on Divide and Conquer Algorithms

## Selection Algorithm – Complexity, II

#### Theorem 3.3.2.

The average execution time  $T_A(n)$  of Select1 algorithm is  $\mathcal{O}(n)$ .

**Proof.** The complexity of Partition algorithm is  $\mathcal{O}(n)$  and hence there is a constant c such that B. S. G. G. S. S. S.

$$\begin{aligned} T_A^k(n) &\leq c \cdot n + \frac{1}{n} \Big( \sum_{i=1}^{k-1} T_A^k(n-i) + \sum_{i=k+1}^n T_A^k(i-1) \Big), \\ R(n) &\leq c \cdot n + \frac{1}{n} \max_k \Big( \sum_{i=1}^{k-1} R(n-i) + \sum_{i=k+1}^n R(i-1) \Big), \\ R(n) &\leq c \cdot n + \frac{1}{n} \max_k \Big( \sum_{i=n-k+1}^{n-1} R(i) + \sum_{i=k}^{n-1} R(i) \Big). \end{aligned}$$

To show by induction that  $R(n) \leq 4c \cdot n$ .

(3.3.2)

# Selection Algorithm – Complexity, III

For 
$$n = 2$$
,  

$$R(n) \leq 2 \cdot c + \frac{1}{2} \max \left( R(1), R(1) \right)$$

$$\leq 2.5 \cdot c \leq 4 \cdot c \cdot n$$
Next assume  $R(n) \leq 4 \cdot c \cdot n$  for  $2 \leq n < m$ .  
For  $n = m$ ,  

$$R(m) \leq c \cdot m + \frac{1}{m} \max_{R} \left( \sum_{k=m-k+1}^{m-1} R(i) + \sum_{i=k}^{m-1} R(i) \right)$$
Since  $R(n)$  is a nondecreasing function of  $n$ .  $\sum_{m=m+k+1}^{m-1} R(i) + \sum_{i=k}^{m-1} R(i)$  is maximum when  $k = m/2$  when  $m$  is even, and  $k = (m+1)/2$  when  $m$  is odd.  
When  $m$  is even
$$R(m) \leq c \cdot m + \frac{2}{m} \sum_{i=m/2}^{m-1} R(i)$$

$$\leq c \cdot m + \frac{8c}{m} \sum_{i=m/2}^{m-1} i$$

$$\leq 4 \cdot c \cdot m$$
When  $m$  is odd
$$R(m) \leq c \cdot m + \frac{2}{m} \sum_{i=(m+1)/2}^{m-1} R(i)$$

$$\leq c \cdot m + \frac{8c}{m} \sum_{i=(m+1)/2}^{m-1} R(i)$$

$$\leq c \cdot m + \frac{8c}{m} \sum_{i=(m+1)/2}^{m-1} R(i)$$

$$\leq c \cdot m + \frac{8c}{m} \sum_{i=(m+1)/2}^{m-1} R(i)$$
Since  $T_A(n) \leq R(n)$ , therefore  $T_A(n) \leq 4 \cdot c \cdot n$  and  $T_A(n)$  is  $\mathcal{O}(n)$ .

#### Selection Algorithm – Complexity, V

- The execution time of Select1 in the worst-case is  $\mathcal{O}(n^2)$ .
  - The worst-case can happen if the partition element, A[low], is close to extreme.
  - If the partition element is close to the median, A[(low + high)/2], then the number of iterations can be reduced significantly.
  - Using this argument, the selection algorithm is modified to have worst-case linear time complexity.
- The array A is divided into subarrays each has r elements
  - $\lceil n/r \rceil$  groups
  - A small r is usually preferred (r = 5, for example).
- Then the median of each group is found and move to the front array A
- The median of the medians, *mm*, is then found using Partition function
- Now, this mm can be used to partition array A.
- Since *mm* is used for each partition step in the selection algorithm, worst-case linear time can be guaranteed.
- Note that though Select2 is worst-case linear, it has a much larger coefficient, as compared to Select1, thus for small to median-size problems, Select2 may not be faster in execution.
  - Select2 returns the position j such that A[j] is the k'th element.

Algorithms (EE/NTHU)

Unit 3.3 More on Divide and Conquer Algorithms

## Selection Algorithm – Worst-case Linear

#### Algorithm 3.3.3. Selection – Worst-case Linear

1 Algorithm Select2( $A, k, low, high, r$ )	
2 // P 3 {	Partition the array into $A[1:k-1] \leq A[k] \leq A[k+1:n].$
4	j := k + low - 1;
5	while $(k  eq j - low)$ do {
6	n := high - low + 1;
7	if $(n \leq r)$ then $\{ \ // \ { m small array}$
8	InsertionSort(A, low, high);
9	$return \ low + k;$
10	}
11	for $i:=1$ to $\left\lceil \ n/r  ight ceil$ do $\left\{ \ // \  ext{find} \  ext{median} \  ext{of} \  ext{group} \  ext{and} \  ext{move} \  ext{to} \  ext{front}$
12	InsertionSort(A, low + (i-1) * r, low + i * r - 1);
13	$\mathbf{Swap}(low + i - 1, low + (i - 1) * r + \lceil r/2 \rceil - 1);$
14	}
15	$j := \mathtt{Select2}(A, \lceil \lfloor n/r \rfloor/2 \rceil, low, low + \lceil n/r \rceil - 1); // find median of medians$
16	$\mathtt{Swap}(\mathit{low},j);//$ move median of median to $A[\mathit{low}]$
17	j := Partition(A, low, high + 1);
18	if $(k < j - low)$ $high := j; //$ reduce to $A[low : j]$
19	else if $(k>j-low)$ { $//$ reduce to $A[j+1:high]$
20	$k:=k-\left(j-low+1\right);$
21	low := j + 1;
22	}
23	}
24	$\texttt{return} \ j;$
25 }	

• Given two  $n \times n$  matrix **A** and **B**,  $\mathbf{A}[i, j] \in \mathbb{R}$ ,  $\mathbf{B}[i, j] \in \mathbb{R}$ ,  $1 \le i, j, \le n$ , then  $n \times n$  matrix **C** is the product of **A** and **B**,  $(\mathbf{C} = \mathbf{A} \cdot \mathbf{B})$ ,

$$\mathbf{C}[i,j] = \sum_{k=1}^{n} \mathbf{A}[i,k] \times \mathbf{B}[k,j], \qquad 1 \le i,j \le n.$$
(3.3.3)

- Note that to calculate C[i, j], one needs n multiplications and n-1 additions.
- Thus to calculate C, which has  $n^2$  elements, the time complexity is  $\Theta(n^3)$ .

Algorithms (EE/NTHU

Unit 3.3 More on Divide and Conquer Algorithms

## Matrix Multiplication – Divide and Conquer

- Suppose  $n = 2^k$ , we can apply divide and conquer approach to matrix multiplication problem.
- Divide each matrix into 4 submatrices with  $\frac{n}{2} imes \frac{n}{2}$  dimensions each, then

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix}$$
(3.3.4)

where

$$C_{11} = A_{11}B_{11} + A_{12}B_{21}$$

$$C_{12} = A_{11}B_{12} + A_{12}B_{22}$$

$$C_{21} = A_{21}B_{11} + A_{22}B_{21}$$

$$C_{22} = A_{21}B_{12} + A_{22}B_{22}$$
(3.3.5)

- To calculate matrix  $\mathbf{C}$ , we need n
  - Eight matrix multiplications  $(\frac{n}{2} \times \frac{n}{2})$ ,
  - Four matrix additions ( $\mathcal{O}(n^2)$  complexity due to  $n^2$  elements in C).
- Let T(n) be the complexity, then

$$T(n) = \begin{cases} b, & n \le 2\\ 8 \cdot T(n/2) + c \cdot n^2, & n > 2. \end{cases}$$
(3.3.6)

quer Algorithms

where b and c are constants.

Mar. 26, 2018 10 / 32

#### Matrix Multiplication – Divide and Conquer, II

• If 
$$n = 2^k$$

$$T(n) = 8T(n/2) + c \cdot n^{2}$$

$$= 8\left[8T(n/4) + c \cdot \left(\frac{n}{2}\right)^{2}\right] + c \cdot n^{2}$$

$$= 8^{2}T(n/4) + c \cdot n^{2} (2+1)$$

$$= 8^{3}T(n/8) + c \cdot n^{2} (4+2+1)$$

$$= 8^{k-1}T(n/2^{k-1}) + c \cdot n^{2} \sum_{i=0}^{k-2} 2^{i}$$

$$= 2^{3k-3}b + c \cdot n^{2} \left(\cdot 2^{k-1}\right)$$

$$= \frac{n^{3}}{8}b + c \cdot n^{2} \left(\frac{n}{2}\right)$$

$$= \left(\frac{b}{8} + \frac{c}{2}\right)n^{3}$$

$$= \mathcal{O}(n^{3})$$

• Thus, this divide and conquer approach does not improve the computational complexity

#### Strassen's Matrix Multiplication

• Given Equations (3.3.4) and (3.3.5), define the following

$$P = (A_{11} + A_{22})(B_{11} + B_{22})$$

$$Q = (A_{21} + A_{22})B_{11}$$

$$R = A_{11}(B_{12} - B_{22})$$

$$S = A_{22}(B_{21} - B_{11})$$

$$T = (A_{11} + A_{12})B_{22}$$

$$U = (A_{21} - A_{11})(B_{11} + B_{12})$$

$$V = (A_{12} - A_{22})(B_{21} + B_{22})$$

$$C_{11} = P + S - T + V$$

$$C_{12} = R + T$$

$$C_{21} = Q + S$$

$$C_{22} = P + R - Q + U$$

$$(3.3.8)$$

Then

- To find matrix C, we need 7 matrix multiplications of  $\frac{n}{2} \times \frac{n}{2}$  and 18 matrix additions.
- Since matrix multiplications,  $\mathcal{O}(n^3)$ , is more expensive than matrix addition,  $\mathcal{O}(n^2)$ , for large n this approach might be more efficient.

11 / 32

#### Strassen's Matrix Multiplication, II

• The recurrence relation for the computation time T(n) is

$$T(n) = \begin{cases} b, & n \le 2, \\ 7 \cdot T(n/2) + c \cdot n^2, & n > 2. \end{cases}$$
(3.3.9)

where b and c are two constants. • If  $n = 2^k$ , then

$$\begin{split} T(n) &= 7 \cdot T(n/2) + c \cdot n^2 \\ &= 7^2 \cdot T(n/4) + 7 \cdot c \cdot (n/2)^2 + c \cdot n^2 \\ &= 7^2 \cdot T(n/4) + c \cdot n^2 (7/4 + 1) \\ &= 7^{k-1} \cdot T(n/2^{k-1}) + c \cdot n^2 \sum_{i=0}^{k-2} (7/4)^i \\ &= 7^{k-1} \cdot b + c \cdot n^2 \left( (7/4)^{k-1} - 1 \right) / (3/4) \\ &\approx n \frac{\lg 7}{7} + c' n^{\lg 4 + \lg 7 - \lg 4} \\ &= \mathcal{O}(n^{\lg 7}) = \mathcal{O}(n^{2.807}) \end{split}$$

Algorithms (EE/NTHU)

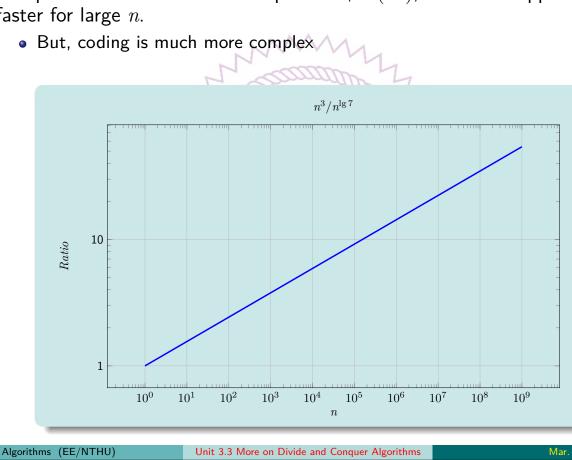
Unit 3.3 More on Divide and Conquer Algorithms

Mar. 26, 2018 13 / 32

26, 2018

#### Strassen's Matrix Multiplication, III

• Compared to direct matrix multiplication,  $\mathcal{O}(n^3)$ , Strassen's approach can be faster for large n.



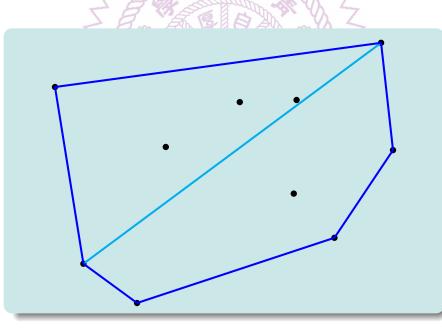
### Convex Hull Problem

- Given a set S that contains points on a 2-D plane, the convex hull is defined as the smallest convex polygon that contains all the points in S.
- A polygon is convex if for any two points  $p_1, p_2$  inside of the polygon, the straight line segment connecting  $p_1$  and  $p_2$  is fully contained in the polygon.
- The vertices of the convex hull of a set S is a subset of S.
  - But, not necessarily a proper subset.

 Market Mar

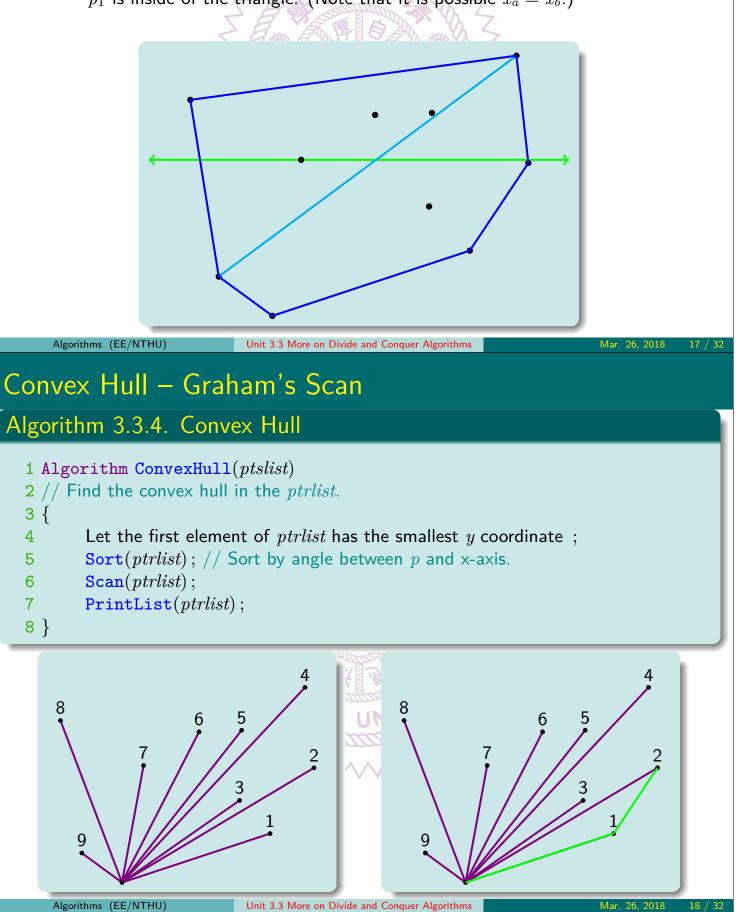
#### Convex Hull – Direct Implementation

- $\bullet\,$  The convex hull of S can be found using the definition above
  - For any  $p_1 \in S$ , if  $p_1$  is inside the triangle formed by  $p_2, p_3, p_4 \in S$ , with  $p_1 \notin \{p_2, p_3, p_4\}$ , then  $p_1$  is not a vertex of the convex hull.
- This direct implementation has the time complexity of  $\mathcal{O}(n^4)$ .
  - n points to be tested,  $n^3$  for all possible triangles.

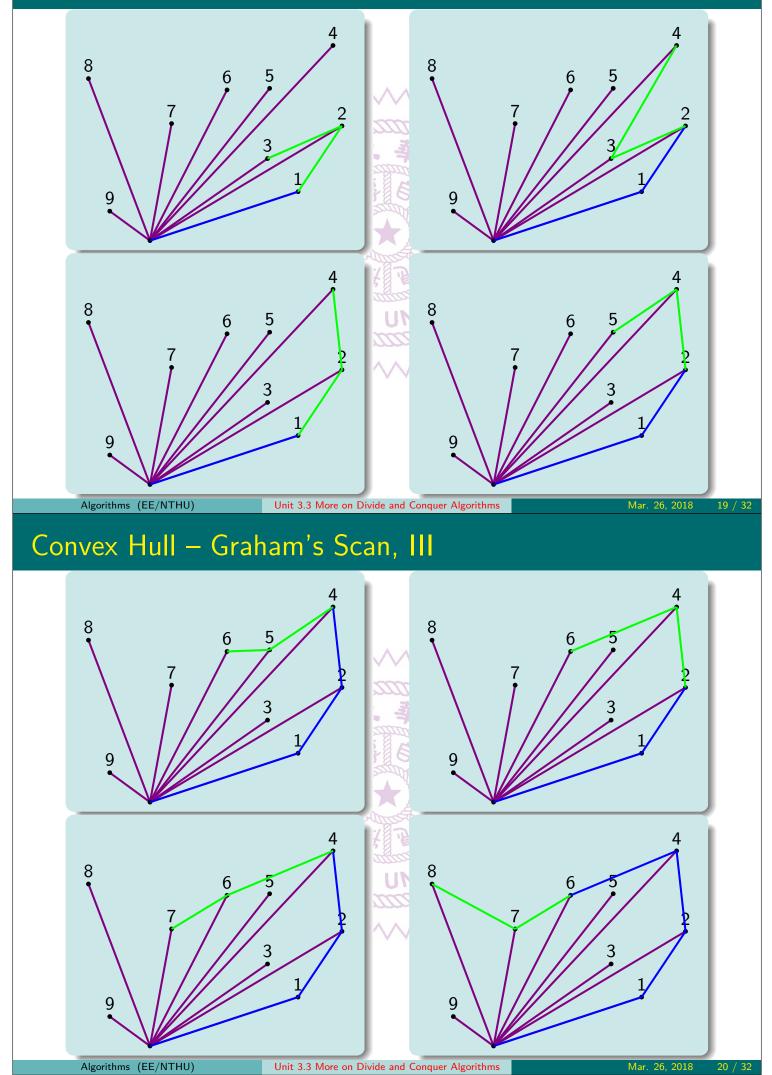


#### Convex Hull – Direct Implementation, II

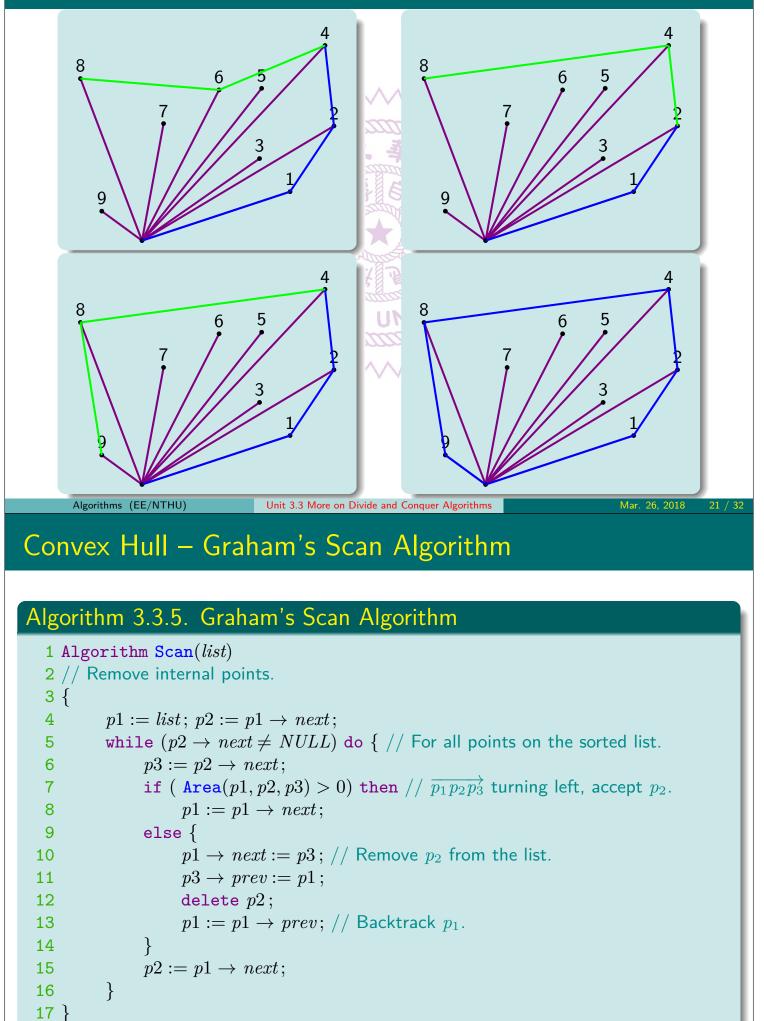
- To test if a point  $p_1$  is inside of a triangle  $riangle p_2 p_3 p_4$ 
  - Let L be the horizontal line passing through p<sub>1</sub> = (x<sub>1</sub>, y<sub>1</sub>), note that L can be described by the linear equation y = y<sub>1</sub>, then check if L intersects with any of the line segments, p<sub>2</sub>p<sub>3</sub>, p<sub>3</sub>p<sub>4</sub>, p<sub>4</sub>p<sub>2</sub>. If not, then p<sub>1</sub> is outside of △p<sub>2</sub>p<sub>3</sub>p<sub>4</sub>. Otherwise let (x<sub>a</sub>, y<sub>1</sub>) and (x<sub>b</sub>, y<sub>1</sub>) be the intersect points, if x<sub>a</sub> ≤ x<sub>1</sub> ≤ x<sub>b</sub> then p<sub>1</sub> is inside of the triangle. (Note that it is possible x<sub>a</sub> = x<sub>b</sub>.)



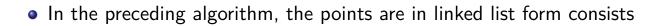
## Convex Hull – Graham's Scan, II

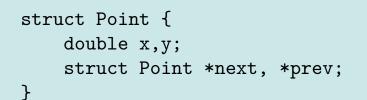


#### Convex Hull – Graham's Scan, IV



#### Convex Hull – Graham's Scan Algorithm, II





- Note that this is a double linked list.
- Let  $p_1(x_1, y_1)$ ,  $p_2(x_2, y_2)$  and  $p_3(x_3, y_3)$  be three points in a plane the function  $Area(p_1, p_2, p_3)$  is defined as

det  $x_2$ 

It can be shown that

Algorithms (EE/NTHU)

• If the area is positive then  $p_3$  is located to the left of the vector  $\overrightarrow{p_1p_2}$ .

Unit 3.3 More on Divide and Conquer Algorithms

• If the area is negative then  $p_3$  is located to the right of the vector  $\overrightarrow{p_1p_2}$ .

 $x_1 \quad y_1$ 

• If the area is zero then  $p_3$  is colinear with  $\overrightarrow{p_1 p_2}$ .

Convex Hull – Graham's Scan Algorithm, Complexity

- The Algorithm (3.3.4) consists of 3 steps
  - (line 4) finding the first element with the smallest y coordinate can be done in  $\mathcal{O}(n)$  time,
  - (line 5) sort by the angle can be done in  $\mathcal{O}(n \lg n)$  time,
  - (line 6) Graham's Scan can be done in  $\mathcal{O}(n)$  time.
- Thus, the time complexity is  $\mathcal{O}(n \lg n)$ .

(3.3.10)

## Quick Hull Algorithm

• Divide and conquer approach can be used to find the convex hull.

#### Algorithm 3.3.6. QuickHull

```
1 Algorithm QuickHull(list, CHull)
 2 // Generate Convex Hull for points in list.
 3 {
 4
         Find p_1 \in list with the smallest x coordinate.
         Find p_2 \in list with the largest x coordinate.
 5
         Let X_1 := \{p | \text{Area}(p_1, p_2, p) > 0\}. // Upper half.
 6
         Let X_2 := \{p | \text{Area}(p_1, p_2, p) < 0\}. // Lower half.
 7
         Hull(p_1, p_2, X_1, UpperHull); // Create upper hull.
 8
         Hull(p_2, p_1, X_2, LowerHull); // Lower hull.
 9
         CHull := Merge(UpperHull, LowerHull); // Merge them.
10
11 }
```

- Finding  $p_1$  and  $p_2$  takes  $\mathcal{O}(n)$  time.
- Finding  $X_1$  and  $X_2$  takes  $\mathcal{O}(n)$  time.
- Merge takes no more than  $\mathcal{O}(n)$  time.
- The time complexity can be dominated by Hull function.

Algorithms (EE/NTHU)

Unit 3.3 More on Divide and Conquer Algorithms

#### Mar. 26, 2018 25 / 3

### Quick Hull Algorithm, II

#### Algorithm 3.3.7. QuickHull

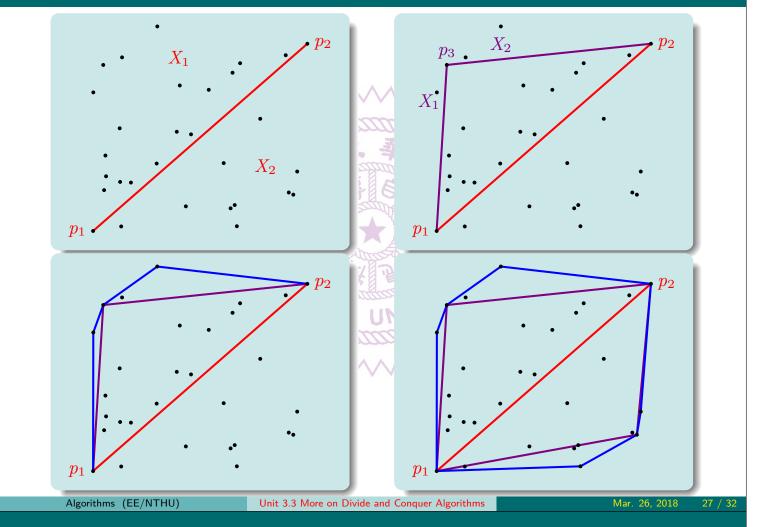
1 Algorithm Hull $(p_1, p_2, list, CHull)$ **2** // Find convex Hull for  $p_1$ ,  $p_2$  and list. 3 { Find  $p_3 \in list$  with the largest  $|\operatorname{Area}(p_1, p_2, p_3)|$ ; 4 Let  $X_1 := \{p | \operatorname{Area}(p_1, p_3, p) > 0\}$ . // All points left to  $\overrightarrow{p_1 p_3}$ / 5 if  $(X_1 = \emptyset)$  then  $H_1 := \{p_1, p_3\}; //$  No more points. 6 7 else HULL $(p_1, p_3, X_1, H_1)$ ; // Recursive call if more points. Let  $X_2 := \{ p | \operatorname{Area}(p_3, p_2, p) > 0 \}.$ 8 9 if  $(X_2 = \emptyset)$  then  $H_2 := \{p_3, p_2\};$ else HULL $(p_3, p_2, X_2, H_2)$ ; 10 11  $CHull := Merge(H_1, H_2); // Combine those two hulls.$ 12 } • Finding  $p_3$ ,  $X_1$ , and  $X_2$  take  $\mathcal{O}(m)$  time, if *list* has m points. • Thus, if T(m) is the time for HULL algorithm we have

$$T(m) = T(m_1) + T(m_2) + O(m),$$
 (3.3.11)

where  $m_1 + m_2 \leq m$ .

- This recurrence relationship is the same as QuickSort.
  - Worst-case complexity is  $\mathcal{O}(m^2)$ , and average-case is  $\mathcal{O}(m\lg m)$ .

### Quick Hull Example



## Time Complexity of Divide and Conquer Algorithms

- In Algorithm DandC (Algorithm 3.1.1) the problem is divided into k subproblems; each solved recursively; then the results are combined to form the final solution.
- The execution time can be assumed to have a general recurrence equation as

$$T(n) = a \cdot T(n/k) + f(n).$$
 (3.3.12)

where f(n) is the time to divide problem into k subsets and to combine the subsets to form the final solution.

• Let 
$$n = k^{m}$$
, then  $T(n) = a \cdot T(n/k) + f(n)$   
 $= a \cdot \left(a \cdot T(n/k^{2}) + f(n/k)\right) + f(n)$   
 $= a^{2} \cdot T(n/k^{2}) + a \cdot f(n/k) + f(n)$   
 $= a^{m} \cdot T(n/k^{m}) + \sum_{i=0}^{m-1} a^{i} \cdot f(n/k^{i})$   
 $= a^{\log_{k} n} \cdot T(1) + \sum_{i=0}^{m-1} a^{i} \cdot f(n/k^{i})$   
 $= n^{\log_{k} a} + \sum_{i=0}^{m-1} a^{i} \cdot f(n/k^{i})$  (3.3.13)

i=0

#### Master Method

- Note that
  - T(1) is taken out since it is a constant,
  - The summation of the second part has  $m = \log_k n$  terms.
- If there is a positive  $\epsilon$  such that  $n^{\log_k a} = n^{\epsilon} \cdot f(n)$  then for large n

$$T(n) = \Theta(n^{\log_k a}). \tag{3.3.14}$$

2 If there is a positive  $\epsilon$  such that  $n^{\log_k a} = f(n)/n^{\epsilon}$  and if  $a \cdot f(n/k) \le c \cdot f(n)$  for some constant c < 1 for large n

$$T(n) = \Theta(f(n)). \tag{3.3.15}$$

If 
$$f(n) = \Theta(n^{\log_k a})$$
 then  

$$T(n) = \Theta(n^{\log_k a} \lg n).$$
(3.3.16)

This comes from the *m* summation terms.  $m = \log_k n = \Theta(\lg n)$ .

• In general, the time complexity of the divide-and-conquer algorithms fall into one of the three scenarios as shown in Eqs. (3.3.14, 3.3.15, 3.3.16).

Algorithms

• However, exceptions exist.

Algorith

## Master Method – Example

• Example 1, Algorithm MaxMin

$$T(n) = 2T(n/2) + 2$$

$$a = 2, k = 2, n^{\log_k a} = n \text{ and } f(n) = 2.$$
Use Eq. (3.3.14) we have  $T(n) = \Theta(n).$   
• Example 2, Algorithm MaxSubArray  

$$T(n) = 2T(n/2) + n$$

$$a = 2, k = 2, n^{\log_k a} = n \text{ and } f(n) = n.$$
Use Eq. (3.3.16) we have  $T(n) = \Theta(n \lg n).$   
• Example 3, MatrixMultiplication  

$$T(n) = 8T(n/2) + n^2$$

$$a = 8, k = 2, n^{\log_k a} = n^3 \text{ and } f(n) = n^2.$$

Use Eq. (3.3.14) we have  $T(n) = \Theta(n^3)$ .

Algorithms (EE/NTHU)

Mar. 26, 2018

29 / 32

#### Master Method – Example

• Example 4,

$$T(n) = 2T(n/2) + n^2$$

 $= 2\left(2T(n/4) + (n/2)^2\right)$ 

 $0 = 4T(n/4) + n^2(1+1/2)$ 

 $= 2^m T(n/2^m) + n^2 \sum_{k=1}^{m}$ 

 $= n + n^2 \cdot 2 \cdot (1 - 2^-)$ 

 $1/2^{i}$ 

 $T(n) = 2T(n/2) + n^2$ 

 $=\Theta(n^2)$ 

a = 2, k = 2,  $n^{\log_k a} = n$  and  $f(n) = n^2$ . Use Eq. (3.3.15) we have  $T(n) = \Theta(n^2)$ . This can also be derived as follows.

• Thus, the Master method can be effective to find the complexity of divide and conquer algorithms.

Unit 3.3 More on Divide and Conquer Algorithms

#### Summary

• Selection problem

Algorithms (EE/NTHU)

- Matrix multiplication
  - Strassen's matrix multiplication
- Convex hull problem
  - Graham's scan algorithm
  - Quick hull algorithm
- Time complexity of divide and conquer algorithms
  - Master method