CSCI2100 Data Structures
Maximum Subsequence Sum Problem

Irwin King

king@cse.cuhk.edu.hk
http://www.cse.cuhk.edu.hk/~king

Department of Computer Science & Engineering
The Chinese University of Hong Kong

Jill Rides Again

- Jill likes to ride her bicycle, but since the pretty city of Greenhills where she lives has grown, Jill often uses the excellent public bus system for part of her journey.
- She has a folding bicycle which she carries with her when she uses the bus for the first part of her trip. When the bus reaches some pleasant part of the city, Jill gets off and rides her bicycle.
- She follows the bus route until she reaches her destination or she comes to a part of the city she does not like. In the latter event she will board the bus to finish her trip.

- Through years of experience, Jill has rated each road on an integer scale of "niceness."
- Positive niceness values indicate roads Jill likes; negative values are used for roads she does not like.
- Jill plans where to leave the bus and start bicycling, as well as where to stop bicycling and re-join the bus, so that the sum of niceness values of the roads she bicycles on is maximized.
- This means that she will sometimes cycle along a road she does not like, provided that it joins up two other parts of her journey involving roads she likes enough to compensate.

- It may be that no part of the route is suitable for cycling so that Jill takes the bus for its entire route. Conversely, it may be that the whole route is so nice Jill will not use the bus at all.
- Since there are many different bus routes, each with several stops at which Jill could leave or enter the bus, she feels that a computer program could help her identify the best part to cycle for each bus route.

Input

- The input file contains information on several bus routes. The first line of the file is a single integer b representing the number of route descriptions in the file. The identifier for each route (r) is the sequence number within the data file, 1 <= r <= b. Each route description begins with the number of stops on the route: an integer s, 2 <= s <= 20,000 on a line by itself.
- The number of stops is followed by s–1 lines, each line i (1 <= i < s) is an integer n i representing Jill's assessment of the niceness of the road between the two stops i and i+1.

Output

- For each route r in the input file, your program should identify the beginning bus stop i and the ending bus stop j that identify the segment of the route which yields the maximal sum of niceness, m = n i +n i+1 +…+n j–1 . If more than one segment is maximally nice, choose the one with the longest cycle ride (largest j–i).
- To break ties in longest maximal segments, choose the segment that begins with the earliest stop (lowest i).
- For each route r in the input file, print a line in the form:
  - The nicest part of route r is between stops i and j.
- However, if the maximal sum is not positive, your program should print:
  - Route r has no nice parts.

Sample Input

3

3

-1

6

10

4

-5

4

-3

4

4

-4

4

-5

4

-2

-3

-4

Output for the Sample Input

- The nicest part of route 1 is between stops 2 and 3
- The nicest part of route 2 is between stops 3 and 9
- Route 3 has no nice parts

Maximum Subsequence Sum

- Given (possibly negative) integers $A_1, A_2, \ldots, A_N$, find the maximum value of $\sum_{k=1}^{j} A_k$.
  (For convenience, the maximum subsequence sum is 0 if all the integers are negative.)

- Example:
  - -2,11,-4,13,-5,-2
  - The answer is 20 (A2 through A4)

Algorithm 1

```
int
max_subsequence_sum( int a[], unsigned int n )
{ int this_sum, max_sum, best_i, best_j, i, j, k;
/*1*/       max_sum = 0; best_i = best_j = -1;
/*2*/       for( i=0; i<n; i++ )
/*3*/           for( j=i; j<n; j++ ){
/*4*/               this_sum=0;
/*5*/               for( k = i; k<=j; k++ )
/*6*/                   this_sum += a[k];
/*7*/               if( this_sum > max_sum )
{    /* update max_sum, best_i, best_j */
/*8*/                   max_sum = this_sum;
/*9*/                   best_i = i;
/*10*/                  best_j = j;}}
/*11*/      return( max_sum );}
```


Notes

- This brute-force algorithm works in $O(N^3)$.
- Statement 1 takes $O(1)$.
- Statement 5 and 6 take $O(N^3)$.
- Statement 7 and 8 take $O(N^2)$.
- Notice that

$$\sum_{k=1}^{j} A_k = A_j + \sum_{k=1}^{j-1} A_k$$

Algorithm 2

```
int
max_subsequence_sum( int a[], unsigned int n )
{int this_sum, max_sum, best_i, best_j, i, j, k;
/*1*/       max_sum = 0; best_i = best_j = ñ1;
/*2*/       for( i=0; i<n; i++ ){
/*3*/           this_sum = 0;
/*4*/           for( j=i; j<n; j++ ){
/*5*/               this_sum += a[j];
/*6*/               if( this_sum > max_sum )
/* update max_sum, best_i, best_j */;
/*7*/           max_sum = this_sum;
/*8*/                   best_i = i;
/*9*/                   best_j = j;}}
/*10*/      return( max_sum );}
```


Notes

- To avoid the cubic running time by removing a for loop.
- The computation at lines 5 and 6 of Algorithm 1 is expensive.
- This algorithm is clearly $O(N^2)$.

Algorithm 3

- Use Divide-and-conquer to find the solution.
  - Find the max subsequence in the left half.
  - Find the max subsequence in the right half.
  - Find the max subsequence that bridges both parts.
  - The maximum is the max of the three.
- This will run in $O(N \log N)$.

Algorithm 4

```
int
max_subsequence_sum( int a[], unsigned int n )
{int this_sum, max_sum, best_i, best_j, i, j;
/*1*/       i = this_sum = max_sum = 0; best_i = best_j = -1;
/*2*/       for( j=0; j<n; j++ )
{/*3*/            this_sum += a[j];
/*4*/           if( this_sum > max_sum )
{       /* update max_sum, best_i, best_j */
/*5*/                   max_sum = this_sum;
/*6*/                   best_i = i;
/*7*/                   best_j = j;}
/*8*/           else if( this_sum < 0 )
{/*9*/                   i = j + 1;
/*10*/                this_sum = 0;}}
/*11*/      return( max_sum );}
```


Example

Input

1	2	−4	5	6	7	−19	2	3	4

Index

1	2	3	4	5	6	7	8	9	10

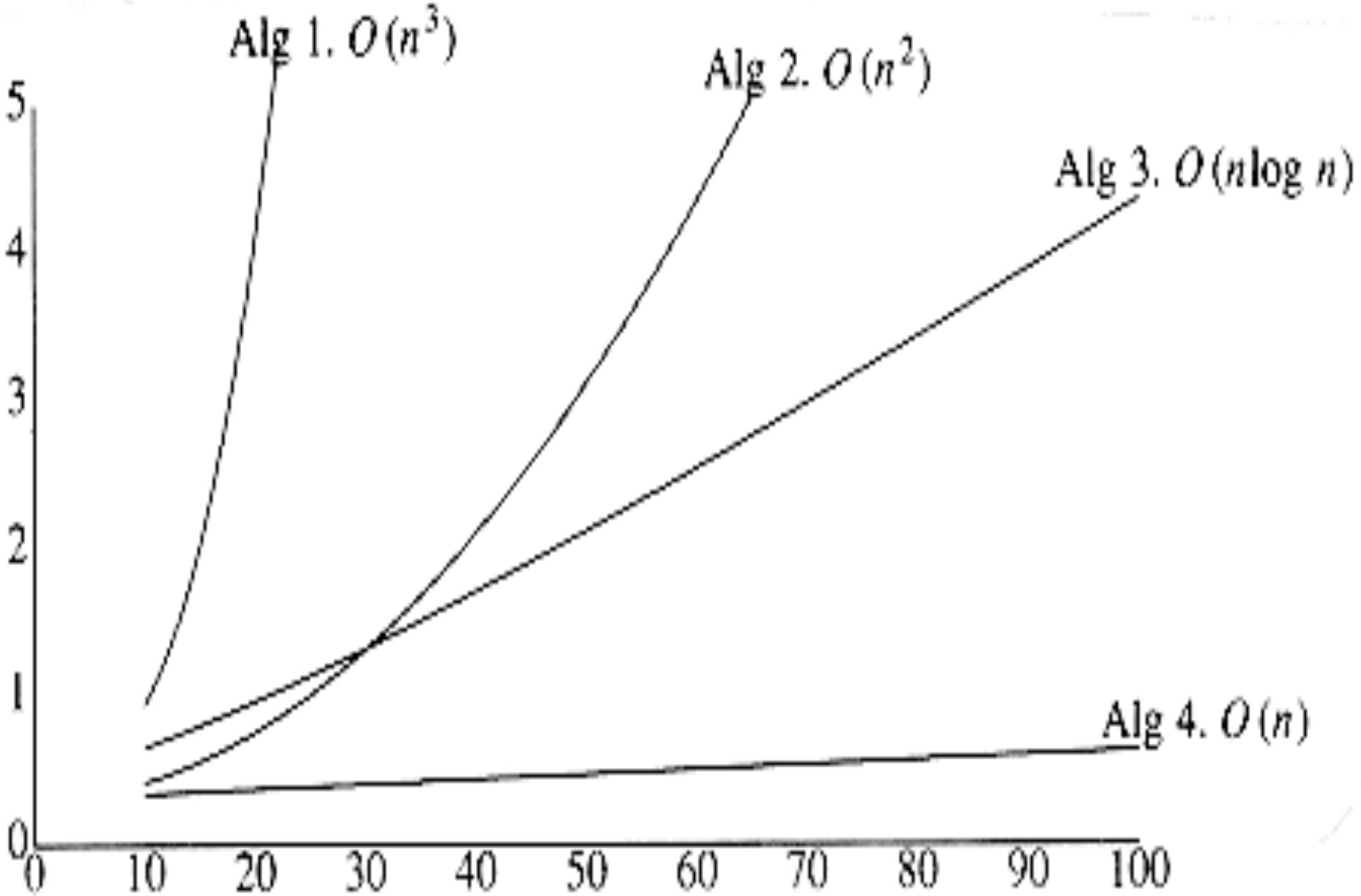
maxsum

1	3	0	5	11	18	0	2	5	9

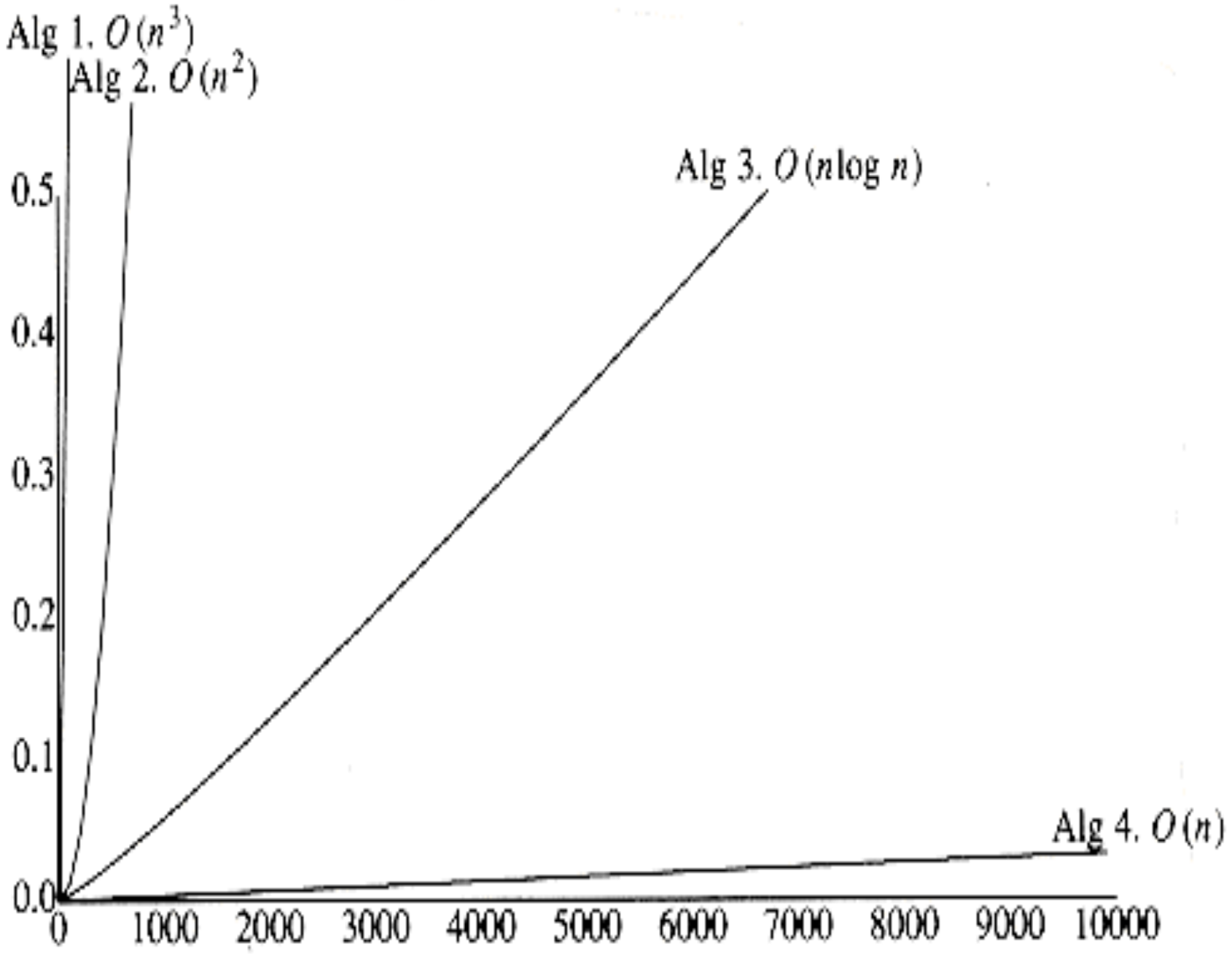
Notes

- This algorithm runs in $O(N)$.
- It makes only one pass through the data.
- Once $A[i]$ is read and processed, it does not need to be remembered.
- At any point in time, the algorithm can correctly give an answer to the subsequence problem for the data it has already read.
- This is an on-line algorithm.

Different Growth Rates



Different Growth Rates



Summary

- How to analyze the complexity of programs.
- Simple programs usually have simple analyses, but this is not always the case.
- Most of the analysis that we will encounter here will be simple and involve counting through loops.