

# CSE3081 (2반): 알고리즘 설계와 분석

## <프로그래밍 숙제 1>

담당 교수: 임 인 성

2017년 9월 27일

마감: 10월 16일 월요일 오전 11시 정각

제출물, 제출 방법, LATE 처리 방법 등: 조교가 과목 게시판에 공고할 예정임.

**목표:** 이번 숙제는 두 가지 문제에 대하여 각각을 해결해주는 알고리즘을 설계하고 구현해봄을 목적으로 한다. 특히, 첫 번째 문제 (A)의 경우 다양한 크기의 입력 데이터를 사용하여 수행시간을 측정한 후, 이론적인 시간 복잡도와 실제 수행시간 간의 연관 관계를 분석하여 본다.

### A. Maximum Subsequence Sum Problem (최대 부분 수열 합 문제)

다음과 같은 문제를 고려하자.

Given  $n$  integers  $a_0, a_1, \dots, a_{n-1}$ , find the maximum value of  $\sum_{k=i}^j a_k$  for  $0 \leq i \leq j \leq n - 1$ . In this problem, the result must be 0 if all input integers are negative.

이 문제에 대하여 세 가지 서로 다른 시간 복잡도를 가지는 알고리즘을 구현하라.

- **Algorithm A1:** 시간 복잡도  $O(n^2)$
- **Algorithm A2:** 시간 복잡도  $O(n \log n)$
- **Algorithm A3:** 시간 복잡도  $O(n)$

프로그램 수행 결과 기본적으로 자신이 구한 부분 수열의 합과 인덱스  $(i, j)$  정보를 출력해주어야 한다. 만약 동일한 최대값을 가지는 부분 수열이 두 개 이상 존재하면, 입력 수열에서 가장 앞에 있는 것의 정보를 출력해주어야 한다.

1. 실험에 사용할 입출력 파일에 대한 정보는 다음과 같다.

(a) 입력 파일 (binary file)

- 입력 파일에는  $n, a_0, a_1, \dots, a_{n-1}$  순서대로  $n+1$ 개의 정수 값들이 binary 형식으로 저장되어 있다. (각 정수 값은 4바이트를 소요함)

(b) 출력 파일 (text file)

- 프로그램 수행 결과는 다음과 같이 .txt 타입의 파일에 정확히 다섯 줄의 텍스트 파일 형태로 출력하라.

1024

1

371

125 172

19.24

이때 각 줄에 출력할 내용을 다음과 같다.

- i. 첫째 줄에는 문제의 크기, 즉  $n$  값을 정수 형태로 출력하라.

- ii. 둘째 줄에는 문제를 해결하기 위하여 사용한 알고리즘의 번호를 정수 형태로 출력하라.  
(1: **Algorithm A1**, 2: **Algorithm A2**, 3: **Algorithm A3**)
  - iii. 셋째 줄에는 자신이 구한 최대 부분 수열 합을 정수 형태로 출력하라.
  - iv. 넷째 줄에는 자신이 찾은 최대 부분 수열의 가장 앞과 가장 뒤의 원소에 대한 두 인덱스를 정수 형태로 출력하라. 만약 최대 부분 수열 값이 0이라면 -1을 두 번 출력해야 한다.
  - v. 다섯째 줄에는 이 문제를 해결하는데 소요한 시간(단위: millisecond)를 소수점 이하 자리 수는 적절히 정하여 출력하라.
2. 이제 여러분이 작성한 프로그램은 이름이 정확히 HW1\_MSS\_config.txt 인 텍스트 파일에서 한 줄에 한 번씩 입력 파일과 출력 파일 이름, 그리고 사용할 알고리즘에 대한 정보를 읽어들이고, 그에 대한 계산을 수행하라. 다음은 HW1\_MSS\_config.txt 파일의 예인데, 이 파일과 입출력 파일은 동일한 디렉토리에 존재해야 한다(조교는 정확히 어디에 존재하는 디렉토리인지를 공지할 예정임).

```
MSS_09.input MSS_09.output.txt A1
MSS_19.input MSS_19.output.txt A3
MSS_02.input MSS_02.output.txt A2
MSS_07.input MSS_07.output.txt A2
```

3. 입력 데이터 크기에 따른 수행 시간과 시간 복잡도의 관계를 충실히 분석할 수 있도록, 다양한 입력 크기  $n$ 에 대해 상기 세 가지 알고리즘에 대한 실험을 수행하라. 가급적 측정된 시간의 오차를 줄이기 위하여 다음과 같은 예를 비롯하여 적절한 실험 방법을 적용하라. 보고서 파일에 실험 결과의 신빙성을 높이기 위하여 어떤 노력을 기했는지 반드시 기술하라.

- 동일한  $n$ 에 대해 최소한 다섯 개 이상의 서로 다른 데이터에 대하여 실험을 수행한 후 평균을 낸다.
- 주어진 동일한 데이터에 대하여 다섯 번 이상 실험을 한 후 평균을 낸다.
- 기타

4. 상기 실험을 통하여 산출한 실험 결과로부터 각 알고리즘의 이론적인 시간 복잡도와 수행 시간 간의 관계를 분석하고, 그로부터 자신이 발견한 사실과 그에 대한 의견을 보고서 파일에 정확히 기술하라.

- 본 숙제의 중요한 목적 중의 하나는 이론적인 시간 복잡도와 실제 프로그램의 수행 시간과의 연관성을 찾아내는 것이다. 따라서 각 함수에 대하여 그러한 연관성을 가장 잘 나타낼 수 있는  $n$  값들을 신중히 선택할 것. 특히 시간 측정이 가능한 범위에서 매우 큰  $n$  값에 대해서도 실험할 것 (반복적인 함수 호출에 따른 stack overflow 문제를 완화하는 방법은 조교가 공지할 예정임).
- 각 함수에 대한 실험 결과를 적절한 형식의 테이블로 요약하라. 실제 측정된 시간 값과 이론적으로 구한 시간 복잡도의 연관성을 증명하기 위하여 그래프를 그린다던가 또는 수식으로 함수 관계를 보인다던가 하는 등의 공학적인 방법을 사용하여 자신의 주장을 논리적으로 기술할 것.
- 보고서에 실험 결과를 기술하기 전에 본인이 사용한 컴퓨터의 실험 환경을 정확히 기술할 것 (아래의 예를 참조할 것).

```
OS: Window 10 Education
CPU: Intel(R) Core(TM) i7-2600 CPU @ 3.40GHz 3.39GHz
RAM: 8.00GB
Compiler: Visual Studio 13.0 Release Mode
```

5. 이 문제 채점 시 입력 크기  $n$ 은 최대  $2^{20}$ 까지만 고려하나, 자신의 실험에서는 가능하다면 더 큰 값에 대해서도 실험해볼 것.

## B. Inversion Counting Problem

다음과 같은 문제를 고려하자.

In an  $n$ -integer sequence  $a_0, a_1, \dots, a_{n-1}$ , two element  $a_i$  and  $a_j$  ( $i < j$ ) are called to form an *inversion* if  $a_i > a_j$ . Find the total number of inversions that exist in the sequence. (예를 들어, 1, 5, 4, 8, 10, 2, 6, 9, 3, 7에는 총 17개의 inversion이 존재함)

이 문제를  $O(n \log n)$ 의 시간 복잡도를 가지는 알고리즘을 설계한 후, 강의자료 57, 58쪽에 주어진 **merge sort** 코드를 기반으로 최소한의 변경을 통하여 구현하라.

1. 실험에 사용할 입출력 파일에 대한 정보는 다음과 같다.

(a) 입력 파일 (binary file)

- 입력 파일의 형식과 내용은 A번 문제와 동일하다.

(b) 출력 파일 (text file)

- 프로그램 수행 결과는 다음과 같이 .txt 타입의 출력 파일에 정확히 두 줄의 텍스트 파일 형태로 저장하라.

1024

592

이때 각 줄에 출력할 내용은 다음과 같다.

- 첫째 줄에는 문제의 크기, 즉  $n$ 값을 정수 형태로 출력하라.
- 둘째 줄에는 자신이 구한 inversion의 총 개수를 정수 형태로 출력하라.

2. 이제 여러분의 프로그램은 이름이 정확히 HW1\_IC.config.txt인 텍스트 파일에서 한 줄에 한 개씩 입력 파일과 출력 파일의 이름을 읽어들이, 그에 대한 계산을 수행 한다. 다음은 HW1\_IC.config.txt 파일의 예인데, 이 파일과 입출력 파일은 동일한 디렉토리에 존재해야 한다(조교는 정확히 어디에 존재하는 디렉토리인지를 공지할 예정임).

```
IC_09.input IC_09.output.txt
IC_19.input IC_19.output.txt
IC_02.input IC_02.output.txt
IC_07.input IC_07.output.txt
```

3. 이 문제 채점 시 입력 크기  $n$ 은 최대  $2^{20}$ 까지만 고려한다.

### • [주의]

1. 본인의 결과를 HW1\_S2016\*\*\*\*. {hwp, doc, txt}와 같은 이름의 보고서에 기술하여 프로그램 및 데이터와 함께 제출하라. (파일명에 학번을 기술할 것)
2. 시간 측정은 조교가 공지하는 방식을 사용할 것.
3. 조교는 채점 시 여러분의 원시 코드를 확인한 후 자신의 명령어 파일과 입력 데이터를 사용하여 여러분의 프로그램이 정확한 값을 계산하는지 기계적으로 확인할 예정임.
4. 숙제 제출 기간 동안 조교가 본 숙제와 관련하여 중요한 공지 사항을 게시판에 올릴 수 있으니 항상 수업 게시판을 확인하기 바람.
5. 제출 화일에서 바이러스 발견 시 본인 점수 **X (-1)**이고, 다른 사람의 숙제를 복사할 경우 **관련된 사람 모두에 대하여 만점 X (-10)**임.