

## 2024년 1학기 알고리즘 과제

- 과제의 프로그램 소스와 보고서를 작성하여 **e-campus**에 업로드

▮ 프로그램 소스: 알고리즘 A,B를 한 번에 수행될 수 있도록 하나의 파이썬 프로그램으로 만들어 이름+학번+hw1.py 저장. 저장된 프로그램 소스는  $n=5,000$ ,  $10,000$ 에 대해 수행되도록 작성

▮ 보고서: 이름+학번+hw1.pdf로 저장

- 보고서에는 과제 내용의 알고리즘 A,B 및 (2) ~ (6)의 답변을 작성. 시간 및 문제 크기의 추정 근거를 서술

- 두 파일 (.py, .pdf)을 **e-campus**에 기한 내에 업로드

동일한 과제를 제출한 모든 학생들에게 페널티 부과

1일 지연 시 만점의 20% 감점

$n$ 개의 데이터 (키값은  $1 \sim 1,000$  사이의 자연수를 **random**으로 생성)를 비내림차순으로 정렬하는 문제에 대해

(1)  $O(n^2)$  알고리즘인 **selection sort**(알고리즘 A)와  $O(n \log n)$  알고리즘 **merge sort**(알고리즘 B)를 python으로 구현한다. **selection sort**의 시간복잡도 분석, **merge sort**의 시간복잡도 분석은 본 강의 12주차 3차시, 3주차 3차시에 각각 설명되어 있다.

(2) 다음의 문제 크기  $n$ 에 대해 알고리즘 A, B가 종료될 때까지의 시간을 측정하여 다음 테이블에 작성하라.

n	알고리즘 A	알고리즘 B
5,000	0.35936665534973145 초	0.01554417610168457 초
10,000	1.4229905605316162 초	0.015623807907104492 초
15,000	3.2037556171417236 초	0.015623807907104492 초
20,000	5.7207701206207275 초	0.031244754791259766 초

30,000	13.05269169807434 초	0.06249284744262695 초
40,000	22.97588586807251 초	0.06248950958251953 초
80,000	x	0.1562361717224121 초

(3)  $n$ 개의 데이터에 대해 알고리즘 A의 수행시간을  $f_A(n)$ , 알고리즘 B의 수행시간을  $f_B(n)$ 로 표현한다. (2)항의 테이블 값을 이용해서  $\frac{n'}{n}=2, 3, 4$  일 때  $f_A(n')/f_A(n)$ ,  $f_B(n')/f_B(n)$ 의 평균값을 계산하라. 즉, 데이터의 크기가 2배, 3배, 4배 될 때 수행시간의 비율을 구하는 것이다.

$n'/n$	평균 $f_A(n')/f_A(n)$	평균 $f_B(n')/f_B(n)$
2	4.017592389261092	2.2999816093269643
3	9.043861822412133	2.492222764975862
4	16.032615642832496	3.6631895531819603

(4) (3)의 결과에서 관찰한 내용과  $n=40,000$ 일 때의 결과를 이용하여  $n=50,000,000$  일 때의 알고리즘 A의 수행시간을 추정한다. 추정 결과를 year 단위로 표시하라. 추정 방법에 대해 설명한다.

$n=40,000$ 일 때 A 알고리즘 수행시간: 22.97588586807251

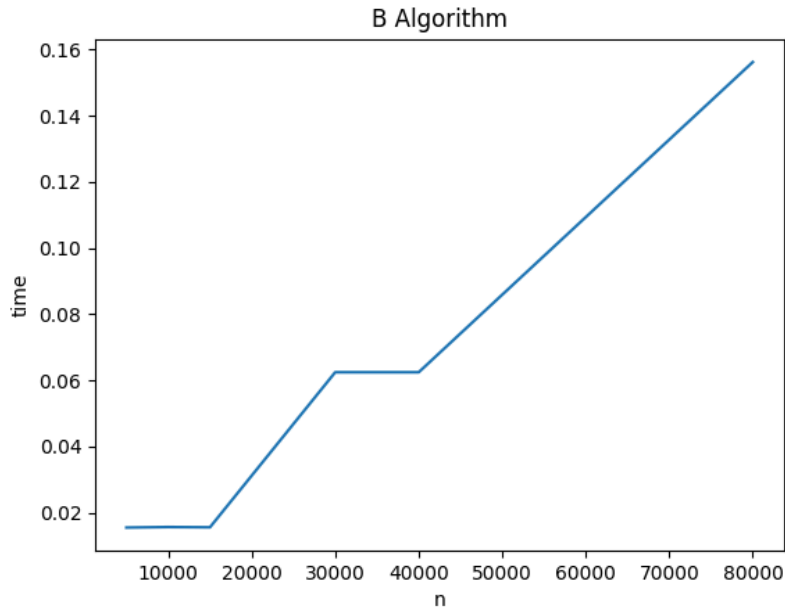
현재 데이터가  $x$ 배 증가했을 때 수행시간 비율은 약  $x^2$ 의 값을 가진다.

이를 이용하여 40,000의 1250배에 해당하는  $n=50,000,000$ 일때의 수행시간을 추정하면,

$(1250)^2 = (n=50,000,000\text{일때의 수행시간}) / (n=40,000\text{일때의 수행시간})$ 이므로,

$n=50,000,000$ 일때의 수행시간은  $22.9758 * (1250)^2$  (초) = 1.14(년)이다.

(5) (2)의 결과를 이용하여 가로축이  $n$ , 세로축이 시간인 그래프에  $f_B(n)$  를 표시하라. 이 그래프를  $a * n \log_2 n$  함수로 표시할 때  $a$ 값을 추정하라. 추정 방법을 설명한다.



$$a * n \log_2 n = f_B(n)$$

$$a = f_B(n) / n \log_2 n$$

위 공식에 맞추어 각  $n$ 값마다의  $a$  값을 구하고 그 값들의 평균값을  $a$ 로 추정함

$$a = 1.3297753217957335e-07$$

(6) (5)의 결과를 이용하여 알고리즘 B를 컴퓨터로 1분간 수행할 때 해결할 수 있는 문제의 크기  $n'$ 를 추정하라. 추정 방법을 설명한다.

$$a * n \log_2 n \leq 60$$

$$n \log_2 n \leq 60/a$$

이때의  $n$ 값을 추정하기 위해  $a * n \log_2 n$  값이 60을 넘길때까지  $n$ 값을 1,000,000부터 1,000,000씩 증가시켰고,  $n=19000000$ 일때 61.09초가 소요된다는 추정 결과가 나왔다.

이후 12000000로부터 1씩 줄여가며  $n$ 값을 추정했을 때  $n=18679562$ 일 때 소요 시간이 59.99초를 확인했다.

**=> n'=18666647**