

# 탐색 알고리즘의 이해

---

송기태 ([kitae040522@gmail.com](mailto:kitae040522@gmail.com))  
Soongsil Univ. (Computer Science and Engineering)

# Content

선형 탐색

이분 탐색

파이썬으로 탐색 알고리즘 구현

과제

## 선형 탐색 알고리즘의 개념

- 시퀀스 데이터 구조에서 원하는 항목을 찾는 과정을 나타내는 검색 알고리즘이다. 처음부터 끝까지 하나씩 요소를 검사하여 목표 항목을 찾을 때까지 반복한다.

## 선형 탐색 알고리즘의 작동 과정

- a. 시작부터 데이터 구조의 첫 번째 항목부터 시작한다.
- b. 현재 항목이 목표 항목과 일치하는지 확인한다.
- c. 만약 일치한다면 검색을 종료한다.
- d. 만약 일치하지 않는다면 다음 항목으로 이동하고 다시 비교한다.
- e. 이 과정을 목표 항목을 찾거나 데이터 구조의 끝까지 진행한다.

## 선형 탐색 알고리즘의 작동 과정

탐색할 배열이 주어진다

8	31	48	73	3	11
---	----	----	----	---	----

# 선형 탐색 알고리즘의 작동 과정

73을 찾기 위해 선형 탐색 알고리즘을 사용한다

8	31	48	73	3	11
---	----	----	----	---	----

# 선형 탐색 알고리즘의 작동 과정

1번째 요소인 8이 73인지 비교한다

8	31	48	73	3	11
---	----	----	----	---	----

## 선형 탐색 알고리즘의 작동 과정

73이 아니므로, 다음 요소(31)와 73을 비교한다

8	31	48	73	3	11
---	----	----	----	---	----



## 선형 탐색 알고리즘의 작동 과정

73이 아니므로, 다음 요소(48)와 73을 비교한다

8	31	48	73	3	11
---	----	----	----	---	----

# 선형 탐색 알고리즘의 작동 과정

73을 찾았기 때문에, 탐색을 종료한다

8	31	48	73	3	11
---	----	----	----	---	----

## 선형 탐색 알고리즘의 장점/단점

- 장점
  - 간단하고 이해하기 쉬운 알고리즘
  - 정렬되지 않은 데이터에서도 사용 가능
- 단점
  - 대량의 데이터에 대해 비효율적
  - 다른 고급 검색 알고리즘에 비해 느린 속도

## 숫자 게임

- 철수는 1 이상 N 이하인 자연수  $X$ 를 하나 선택한다.
- 영희는 철수가 생각한  $X$ 를 맞춰야 한다.
- 영희가 어떤 수  $Y$ 를 말하면, 철수는 아래와 같은 대답을 한다.
  - $X > Y$ 라면: Up
  - $X < Y$ 라면: Down
  - $X = Y$ 라면: Correct
- 영희의 질문 횟수를 최소화 시키려면?

## 이분 탐색 알고리즘의 개념

- 데이터를 절반으로 나누어 검색 범위를 반으로 줄여가며 목표 항목을 찾는 알고리즘이다.
- 정렬 된 배열에서만 사용할 수 있지만, 매우 빠른 속도로 검색을 수행할 수 있다.
- 업/다운 게임으로 생각하면 편하다!

# 이분 탐색 알고리즘의 작동 과정

- a. 데이터의 중간 항목을 선택한다.
- b. 중간 항목과 목표 항목을 비교한다.
- c. 중간 항목이 목표 항목과 일치한다면 검색을 종료한다.
- d. 중간 항목이 목표 항목보다 크다면, 검색 범위를 왼쪽 절반으로 좁힌다.
- e. 중간 항목이 목표 항목보다 작다면, 검색 범위를 오른쪽 절반으로 좁힌다.
- f. 이 과정을 목표 항목을 찾거나 검색 범위가 더 이상 줄어들지 않을 때까지 반복한다.

# 이분 탐색 알고리즘의 작동 과정

탐색할 정렬 된 배열이 주어진다

1	3	8	11	31	48	73
---	---	---	----	----	----	----

# 이분 탐색 알고리즘의 작동 과정

31을 찾기 위해 이분 탐색 알고리즘을 사용한다

1	3	8	11	31	48	73
---	---	---	----	----	----	----



## 이분 탐색 알고리즘의 작동 과정

중간 값을 정한다 (중간 값: 배열의 크기 / 2)

1	3	8	11	31	48	73
---	---	---	----	----	----	----

# 이분 탐색 알고리즘의 작동 과정

중간 값이 31보다 작으므로, 앞에 있는 값은 무시한다.

1	3	8	11	31	48	73
---	---	---	----	----	----	----

# 이분 탐색 알고리즘의 작동 과정

중간 값을 정한다 (중간 값: 배열의 크기 / 2)

1	3	8	11	31	48	73
---	---	---	----	----	----	----

# 이분 탐색 알고리즘의 작동 과정

중간 값이 31보다 크므로, 뒤에 있는 값은 무시한다.

1	3	8	11	31	48	73
---	---	---	----	----	----	----

## 이분 탐색 알고리즘의 작동 과정

중간 값을 정한다 (중간 값: 배열의 크기 / 2)

1	3	8	11	31	48	73
---	---	---	----	----	----	----

# 이분 탐색 알고리즘의 작동 과정

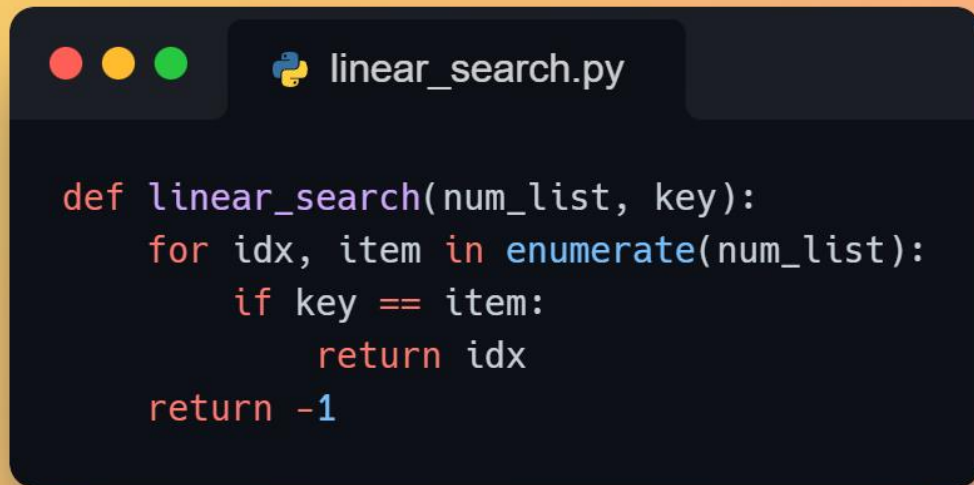
중간 값이 31이므로, 탐색을 종료한다

1	3	8	11	31	48	73
---	---	---	----	----	----	----

# 이분 탐색 알고리즘의 장점/단점

- 장점
  - 매우 빠른 검색 속도
  - 정렬된 데이터에서 사용 가능
  - 대용량 데이터에도 효과적
- 단점
  - 데이터가 정렬되어 있어야 함
  - 삽입 또는 삭제 연산 시에는 추가 작업 필요

# 파이썬으로 검색 알고리즘 구현

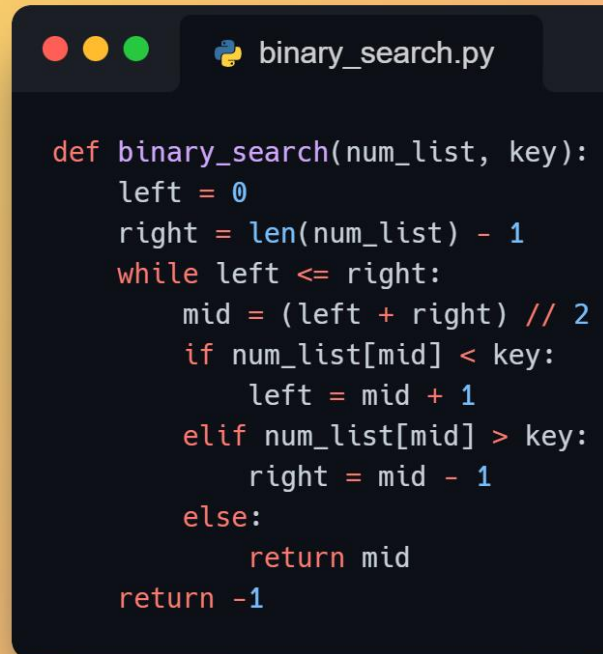


```
def linear_search(num_list, key):  
    for idx, item in enumerate(num_list):  
        if key == item:  
            return idx  
    return -1
```

snappify.com



# 파이썬으로 검색 알고리즘 구현



```
def binary_search(num_list, key):  
    left = 0  
    right = len(num_list) - 1  
    while left <= right:  
        mid = (left + right) // 2  
        if num_list[mid] < key:  
            left = mid + 1  
        elif num_list[mid] > key:  
            right = mid - 1  
        else:  
            return mid  
    return -1
```

snappify.com

# 과제

## 입력

첫째 줄에 문자열이 주어진다.

## 출력

주어진 문자열에 포함된 알파벳 대문자들을 이용해 **MOBIS**를 만들 수 있으면 "YES", 그렇지 않으면 "NO"를 출력한다.

## 제한

- $1 \leq \text{문자열의 길이} \leq 100$
- 문자열은 알파벳 대문자로만 이루어져 있다.

### 예제 입력 1 복사

MOIISB

### 예제 출력 1 복사

YES

### 예제 입력 2 복사

MOBI

### 예제 출력 2 복사

NO

### 예제 입력 3 복사

INHACTP

### 예제 출력 3 복사

NO

4 1920번

제출

맞힌 사람

숏코딩

재채점 결과

채점 현황

내 제출

난이도 기여

질문 게시판

수 찾기

성공

☆

4 실버 IV

시간 제한	메모리 제한	제출	정답	맞힌 사람	정답 비율
1 초	128 MB	221956	67258	44786	29.723%

## 문제

N개의 정수  $A[1], A[2], \dots, A[N]$ 이 주어져 있을 때, 이 안에 X라는 정수가 존재하는지 알아내는 프로그램을 작성하시오.

## 입력

첫째 줄에 자연수  $N(1 \leq N \leq 100,000)$ 이 주어진다. 다음 줄에는 N개의 정수  $A[1], A[2], \dots, A[N]$ 이 주어진다. 다음 줄에는  $M(1 \leq M \leq 100,000)$ 이 주어진다. 다음 줄에는 M개의 수들이 주어지는데, 이 수들이 A안에 존재하는지 알아내면 된다. 모든 정수의 범위는  $-2^{31}$  보다 크거나 같고  $2^{31}$  보다 작다.

## 출력

M개의 줄에 답을 출력한다. 존재하면 1을, 존재하지 않으면 0을 출력한다.

### 예제 입력 1 복사

```
5
4 1 5 2 3
5
1 3 7 9 5
```

### 예제 출력 1 복사

```
1
1
0
0
1
```

## 영단어 암기는 괴로워 성공



3 실버 III

시간 제한	메모리 제한	제출	정답	맞힌 사람	정답 비율
1 초 (추가 시간 없음)	1024 MB	9523	4297	3367	45.432%

### 문제

화은이는 이번 영어 시험에서 틀린 문제를 바탕으로 영어 단어 암기를 하려고 한다. 그 과정에서 효율적으로 영어 단어를 외우기 위해 영어 단어장을 만들려 하고 있다. 화은이가 만들고자 하는 단어장의 단어 순서는 다음과 같은 우선순위를 차례로 적용하여 만들어진다.

1. 자주 나오는 단어일수록 앞에 배치한다.
2. 해당 단어의 길이가 길수록 앞에 배치한다.
3. 알파벳 사전 순으로 앞에 있는 단어일수록 앞에 배치한다

$M$ 보다 짧은 길이의 단어의 경우 읽는 것만으로도 외울 수 있기 때문에 길이가  $M$ 이상인 단어들만 외운다고 한다. 화은이가 괴로운 영단어 암기를 효율적으로 할 수 있도록 단어장을 만들어 주자.

### 입력

첫째 줄에는 영어 지문에 나오는 단어의 개수  $N$ 과 외울 단어의 길이 기준이 되는  $M$ 이 공백으로 구분되어 주어진다. ( $1 \leq N \leq 100\,000$ ,  $1 \leq M \leq 10$ )

둘째 줄부터  $N + 1$ 번째 줄까지 외울 단어를 입력받는다. 이때의 입력은 알파벳 소문자로만 주어지며 단어의 길이는 10을 넘지 않는다.

단어장에 단어가 반드시 1개 이상 존재하는 입력만 주어진다.

### 출력

화은이의 단어장에 들어 있는 단어를 단어장의 앞에 위치한 단어부터 한 줄에 한 단어씩 순서대로 출력한다.

#### 예제 입력 1 복사

```
7 4
apple
ant
sand
apple
append
sand
sand
```

#### 예제 출력 1 복사

```
sand
apple
append
```

# Thank You!

---

송기태 ([kitae040522@gmail.com](mailto:kitae040522@gmail.com))  
Soongsil Univ. (Computer Science and Engineering)