



이진 탐색

이진 탐색이란?

데이터가 정렬된 배열에서 검색 범위를 줄여가며 특정 값을 찾는 탐색 기법

이진 탐색을 하기 위해서는 반드시 **데이터가 오름차순으로 정렬**되어 있어야 한다!

이진 탐색은 검색을 반복할 때마다 검색 범위가 반으로 줄어들어 탐색 속도가 빠르다는 장점이 있다.

동작 방식

1. 배열의 중간값을 찾는다.
2. 검색값과 중간값을 비교한다.
 - a. 중간값 = 검색값 : 값을 찾았으므로 종료한다.
 - b. 중간값 > 검색값 : 중간값을 기준으로 배열의 왼쪽 구간에서 다시 탐색한다.
 - c. 중간값 < 검색값 : 중간값을 기준으로 배열의 오른쪽 구간에서 다시 탐색한다.
3. 값을 찾거나 검색 범위가 남아있을 때까지 위 과정(1~2)을 반복한다.

예시

다음과 같은 배열이 주어졌을 때, 23이라는 값을 찾는 과정을 살펴보자.

5	12	17	23	35	42	50	59	65
---	----	----	----	----	----	----	----	----

먼저, 이 배열의 중간값 인덱스를 찾아보자

$$\frac{0 + 8}{2} = 4$$

5	12	17	23	35	42	50	59	65
↑				↑				↑
Low				Mid				High

그 다음, 검색값 23과 중간값이 일치하는지 검사했을 때, **23 < 35 (중간값)**으로 검색값이 중간값보다 작다. 이 경우, 중간값을 기준으로 **왼쪽** 구간에서 다시 탐색해야한다.

따라서

- Low = 0
- High = Mid - 1 = 3

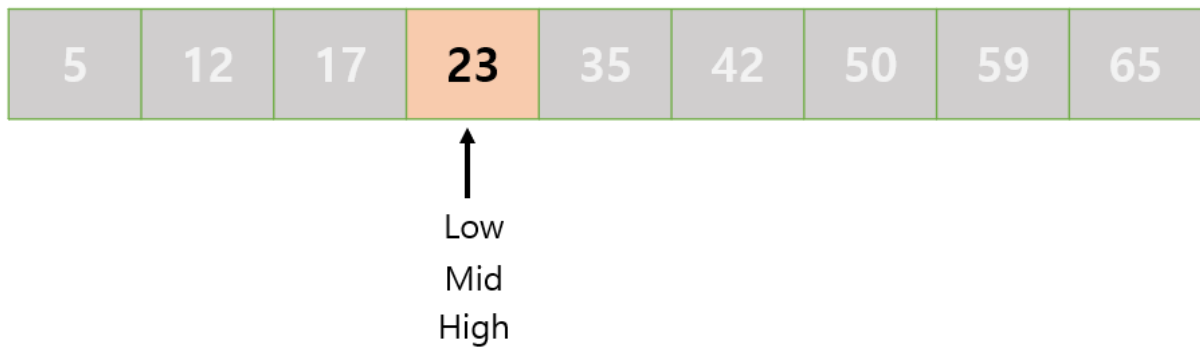
- High = 3

이 구간에서의 중간값은

$$\frac{3 + 3}{2} = 3$$

3이 나온다.

Low = High = Mid = 3이므로, 값은 하나만 남았다. 이 값이 검색값인지 확인해보자.



검색값이 맞다! 탐색을 종료한다.

랜덤으로 만든 배열이었는데 어쩌다보니 끝까지 탐색해버렸다.

시간 복잡도 : $O(\log N)$

간단히 유도를 해보자.

처음 데이터의 개수가 N 이라고 할 때

1. $\frac{N}{2}$

2. $\frac{N}{2} \times \frac{1}{2}$

3.

$$\frac{N}{2} \times \frac{1}{2} \times \frac{1}{2}$$

...

K번째 시행에는

$$\left(\frac{1}{2}\right)^K N$$

개의 데이터가 남게 된다. 이렇게 계속해서 탐색하다보면

$$\left(\frac{1}{2}\right)^K N \approx 1$$

최악의 경우 데이터가 1개 남는 시점까지 도달하게 될 것이다. 따라서 위의 식이 나오게 된다.

위 식의 양 변에 2^K 를 곱해주고 2를 밑으로 하는 로그를 취해주면 다음과 같은 식이 나오게 된다.

$$K \approx \log_2 N$$

여기서 K는 시행횟수를 의미했다.

따라서 자료의 갯수 N에 따른 시행 횟수는

$$\log_2 N$$

가 되고, 시간 복잡도는 Big O 표기법으로 $O(\log N)$ 이 된다.