

8. 探索

実践プログラミング I
情報工学科 鈴木雅人

探索とは

大量のデータの中から，ある条件を満たすデータを探し出すことを「探索」と言う

| 学籍番号 | 氏 名 |
|--------|-------|
| S07102 | 高専 太郎 |
| S07423 | 高専 花子 |
| S07274 | 東京 次郎 |
| S07518 | 梶田 三郎 |
| ⋮ | ⋮ |

学籍番号が
S07462の
学生は？

- ①最初から順番に探す
→線形探索
- ②ある規則に従ってデータを予め
整列しておき，効率よく探す
→二分探索

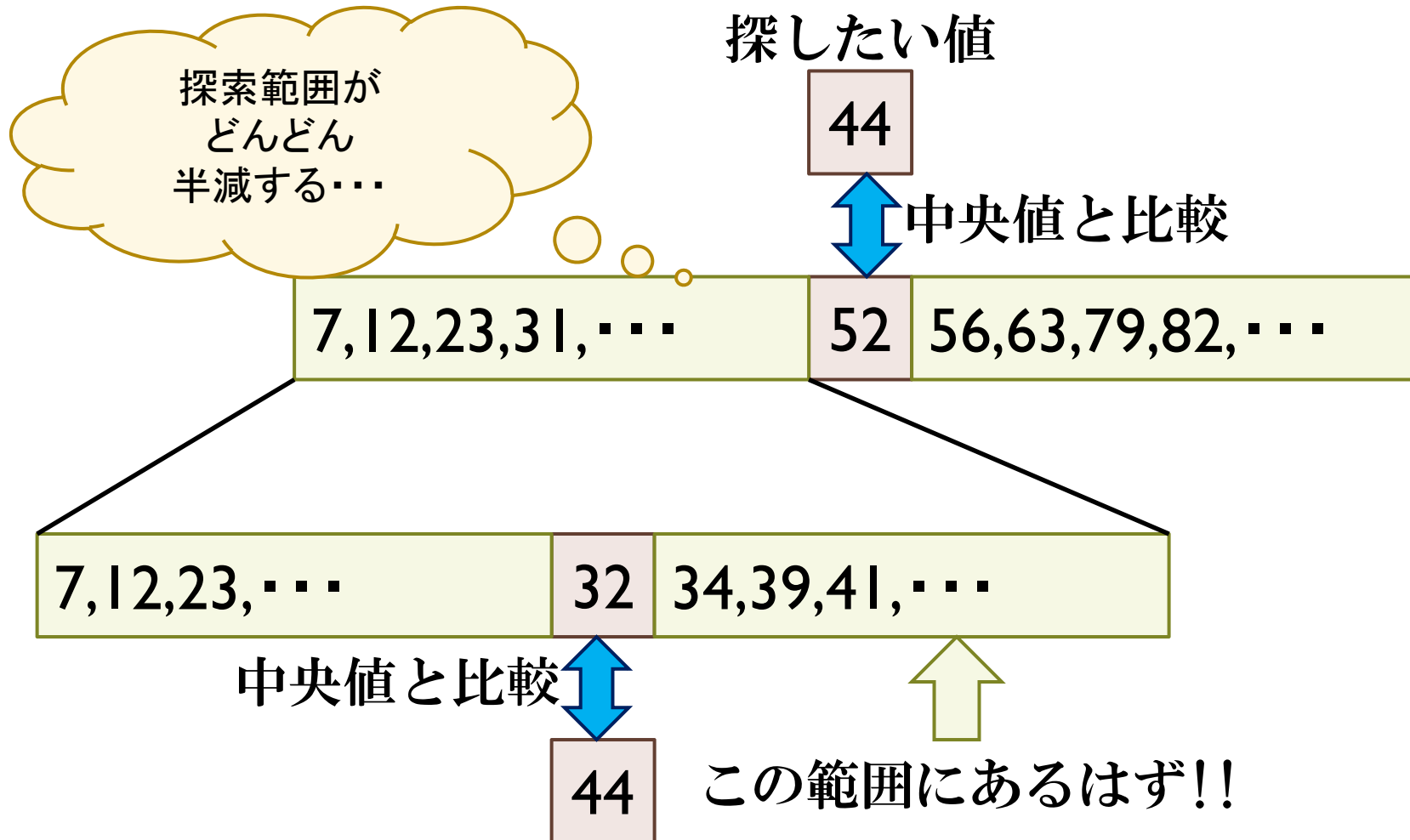
二分探索の考え方

- ① データはソート済み (ここでは小さい順) で横一列に並んでいると仮定する
- ② 探すデータ X とデータ列の中央の値 M とを比較する.
 - $X = M$ なら終了.
 - $X < M$ なら, M より左側の部分データ列に対して②を再度行う.
 - $X > M$ なら, M より右側の部分データ列に対して②を再度行う.

部分データ列内の要素数が
1個または2個の場合要注意



二分探索の手順



二分探索の計算量

厳密な計算は自分でやってみよう！

データが $N = 2^n$ 個あるものと仮定する

1 回の比較で探索対象データ数は約 $\frac{N}{2} = 2^{n-1}$ になる

探索対象データ数が 1 になるまでに n 回の比較を行う

つまり、 $n = \log N$ なので、計算量は $O(\log N)$ である。



【課題8-1】

ファイルに書き込まれている1000万件のデータを読み込み、キーボードから入力した値を二分探索によって探し出し、その位置(行数)を表示するプログラムを作成しなさい。
ただし、位置は先頭を0行目と数えるものとし、また、該当データが見つからない場合は、その旨表示するものとする。

(次ページに続く)



【課題8-1】（続き）

次の手順によって動作確認すること。

- ① 課題5で示した1000万件のデータを，自作のクイックソートによって小さい順に並べ替えてファイルに保存する。
（ファイル名をsorted.datとする）
- ② sorted.datを作成したプログラムに読み込ませ，下記テストデータを用いて動作確認を行う。

| | | |
|-----|-----------|-----------|
| (1) | 644104147 | 6000000番目 |
| (2) | 816070184 | 7600000番目 |
| (3) | 454691142 | 該当なし |

